

Winter Arctic polynyas in CMIP6 models

Céline Heuzé¹, Jonathan W. Rheinländer², Tian Tian³, and Carmen Hau Man Wong¹

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

²Nansen Environmental and Remote Sensing Center, Bergen, Norway

³National Center for Climate Research, Danish Meteorological Institute, Copenhagen, Denmark

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

Abstract. Winter Arctic polynyas, openings in the pack ice, play a crucial role for the climate from sea ice production to cloud formation and are hotspots for the ecosystem and human activity. Their area has significantly increased since satellite records began. Yet their representation has yet to be evaluated in any generation of global climate models, most likely because their automatic multi-model retrieval is challenging. We here use a newly-developed machine-learning based method and evaluate polynya activity against the satellite-derived one over 1979-2024 in the 18 CMIP6 models with daily sea ice concentration available. We find that models overestimate winter Arctic polynya area but underestimate its frequency, and limit their opening to the seasonally ice-covered regions. Polynya area is increasing in most models, but the bias of these trends are inconsistent. Although the model with the highest resolution has both the highest areas and frequencies, the sea ice model component is a more robust predictor of polynya activity, with most activity in the models whose thermodynamics scheme enhances ice growth/melt. Accordingly, we found more polynyas in the models with a larger seasonal cycle, in particular those with a warmer autumn that would delay ice growth or melt early. Finally, we confirm preliminary findings that polynya activity does not seem to impact the representation of the water column; if anything, we find less-dense water at the bottom of the continental shelf following larger polynya activity. Overall, our results suggest that in the Arctic, CMIP6 models unrealistically open only sensible-heat polynyas.

1 Introduction

Polynyas, non-linear openings enclosed in sea ice (World Meteorological Organization, 2014), are crucial for the climate, ecosystem, and the exploitation of polar regions. Winter polynyas expose the comparatively warm ocean to the frigid atmosphere, acting as “ice factories” (Smith Jr and Barber, 2007) responsible for a non-negligible part of the total sea ice production in the Arctic (Preußner et al., 2019). 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). Sea ice production results in brine rejection at the surface of the ocean; in the Arctic, the dense water thus produced overflows off the shelf, modifying and ventilating the deep ocean (Aagaard et al., 1981; Martin and Cavalieri, 1989; Ohshima et al., 2016; Vivier et al., 2024). The combined effect of ocean mixing and air-sea exchanges further leads to intense gas exchanges and nutrient upwelling, making polynyas hotspots

25 for marine life (Moore et al., 2023), which indigenous communities and whalers have long known and utilised (Hastrup et al., 2018).

Most Arctic polynyas are so-called “latent heat polynyas” that open in response to strong winds (Smith Jr and Barber, 2007), often after a pre-conditioning period with warm air temperatures that thin the ice (Wong et al., 2026). They also open primarily on the continental shelf, close to the shelf break (Morales Maqueda et al., 2004). Winter Arctic polynya areas have
30 been increasing since satellite record began in the 1970s (Wong et al., 2026), which is to be expected as the Arctic warms (Rantanen et al., 2022). However, the increase in polynya area appears to have slowed down over the last decade (Preußer et al., 2019; Wong et al., 2026), which Wong et al. (2026) explain as the result of inconsistent trends in winds in the Arctic, in particular their direction. It may also be that as polynyas grow and reach the shelf break, the oceanic forcing can no longer be ignored, and the mechanisms leading to polynya openings are changing (Morales Maqueda et al., 2004); Wong et al. (2026)
35 did find that half of the Arctic regions had become hybrid polynyas, where thermodynamics also plays a role. Owing to the crucial role polynyas play in the climate system, understanding and projecting their changes is urgent. Before we can study future modelled polynyas, one needs to know how well historical polynyas are represented in climate models.

The latest generation of state-of-the-art global climate models are those who participated in the Climate Model Intercomparison Project phase 6 or “CMIP6” (Eyring et al., 2016). Their representation of the Arctic sea ice is acceptable but suffers
40 many biases (Notz and SIMIP Community, 2020); it is common for authors to have to reject a large number of CMIP6 models whose sea ice concentration (Heuzé and Jahn, 2024) or thickness (Athanasé et al., 2025) is unrealistic. Besides, both modelled Arctic air temperature (Tian et al., 2024) and wind speed (Zapponini and Goessling, 2024) are biased, and so is the upper Arctic Ocean (Muilwijk et al., 2023). We therefore expect the modelled Arctic polynyas to be biased as well. Besides, in the Southern Ocean, Mohrmann et al. (2021) and more recently Landrum et al. (2024) found large biases in the area and frequency
45 of open water and coastal polynyas in CMIP6 models. To the best of our knowledge, nobody has assessed Arctic polynyas, in any CMIP phase, aside from an appendix figure showing the polynya frequency on the Siberian shelf in Heuzé et al. (2023), to try and explain biases in modelled deep Arctic Ocean properties in CMIP6. This is probably because Arctic polynyas are particularly tedious to detect in a multi-model framework, owing to the models’ different grid geometries, resolutions and therefore representations of the coastline and archipelagos, and their treatments of the North Pole. Recently, Heuzé and Wong
50 (2025) presented a new machine-learning based algorithm specifically designed to address these challenges, thus facilitating the detection of polynyas across different sea-ice models; we use their algorithm in this study.

Here, we evaluate the representation of winter Arctic polynyas in CMIP6 models. We quantify their biases in area, frequency and trends when compared to observations in section 3, including the effect of using monthly output instead of daily output in subsection 3.2. We then try to determine the causes for these biases in section 4, looking at sea ice thickness, discussing the
55 sea ice model components used, before moving to winds and air temperature. Finally, in section 5, we go back to Heuzé et al. (2023) and the question that originally motivated this paper: do modelled Arctic polynyas impact the modelled deep Arctic Ocean?

2 Data and Methods

2.1 Polynya detection and observational polynya masks

60 We here use the new polynya detection algorithm of Heuzé and Wong (2025), developed with the explicit purpose to detect polynyas in CMIP6 models. As reviewed by e.g. Landrum et al. (2024) for the Southern Ocean and Wong et al. (2026) for the Arctic, there is no consensus in the literature regarding the sea ice concentration or thickness threshold that should be used to detect a polynya, be it for observations or models. Besides, prior to applying that threshold one has to perform a flood-fill to separate polynyas and open ocean (e.g. Mohrmann et al., 2021; Wong et al., 2026). Flood-fill algorithms change adjacent values to a common value, starting from a “seed” point in the open ocean, effectively detecting the ice edge. This is near-impossible to implement in a multi-model comparison framework in the Arctic owing to the models’ different grids, in particular their treatment of the North Pole, their choice of reference longitude, and how their resolution represents the Canadian archipelago. Heuzé and Wong (2025) therefore took a new approach and developed, on observations and for CMIP6 models, a neural network-based method that detects patterns, gradients in the sea ice rather than absolute concentration or thickness values, and requires no flood-fill.

We now briefly describe how the algorithm works and re-direct the interested reader to the full publication (Heuzé and Wong, 2025) for the details. As summarised on Fig. 1, the method consists of four steps:

1. The daily sea ice concentration (SIC) input image is prepared: All pixels with SIC lower than 15% are set to 0 to increase the sharpness of the ice edge, and all pixels south of 68°N are cut out. A second version is created by interpolating onto the National Snow and Ice Data Center (NSIDC) Northern Hemisphere Equal-Area Scalable Earth Grid 2.0 (Brodzik et al., 2012).
2. The interpolated daily SIC are passed through the “Unet” system, i.e. two convolutional neural networks that “zoom in” and “zoom back out” of the image, respectively, so that the result is a pixel-wise classification. Unet is particularly well suited to so-called anomaly detection (Ronneberger et al., 2015), in our case a small opening in the pack ice. The Unet system used here was trained by Heuzé and Wong (2025) using the daily NSIDC observational sea ice concentration as input (more information after this list) and the labelled daily polynya masks of Wong et al. (2026), which used a SIC threshold of 50%, as target. As the model is already trained, this step takes barely a few seconds per model in our current study.
3. The sea ice on the native grid is passed through a Gaussian Mixture Model (GMM) system, an unsupervised classification method that Heuzé and Wong (2025) optimised to detect the open ocean and marginal ice zone (MIZ) pixels as one class. As the model is not trained, this step takes more than 1 hour per model and ensemble member. Note that the output of the AWI-CM-1-1-MR CMIP6 model was too heavy on the native grid, so we had to use its interpolated version (the one used for step 2).

4. The GMM mask is interpolated onto the same grid as the Unet result to produce a filtered version of the Unet polynya / non polynya classification. Only pixels that are detected as polynyas in step 2 and not detected as open ocean + MIZ in step 3 are kept as polynya pixels in this step.

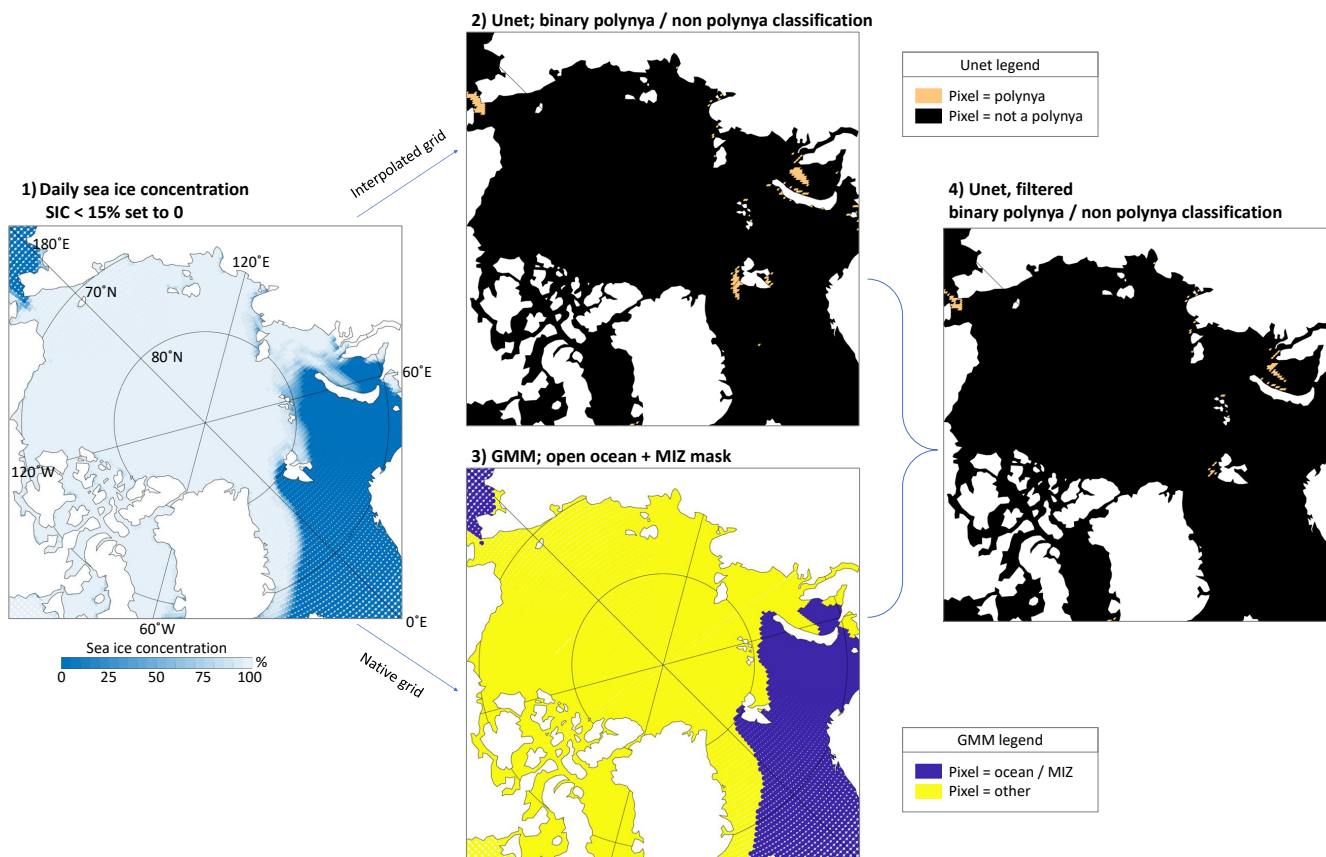


Figure 1. Steps of the polynya detection method, illustrated using the CMIP6 model MRI-ESM2-0 on 1st January 2015.

We leave the settings of Heuzé and Wong (2025) unchanged for Unet. For the GMM filtering, following their recommendation, we treat the input of step 3 differently for the observations and the CMIP6 models: For the observations, all daily sea ice concentration values lower than 40% were set to 0; for the CMIP6 models, all values lower than 15%. We tested also a 30% threshold for CMIP6 but found no significant difference in the performance; since running the GMM on all models took nearly two weeks, we did not test other thresholds.

The observed daily NSIDC sea ice concentration from 1st December to 31st March, from 1978 to 2024, at a 25 km resolution, has already been classified at the pixel level as polynya / non polynya by Heuzé and Wong (2025). Rather than re-classifying the sea ice data, we here re-use the observed polynya masks of Heuzé and Wong (2025) as reference. The NSIDC sea ice concentration is retrieved from passive microwave satellites, processed with the NASA Team algorithm (Cavalieri et al., 1999).

For section 3.2 we generated a monthly observational reference by first computing the monthly mean of the NSIDC sea ice concentration from their daily data, and then passing these monthly sea ice concentration through the Unet+GMM algorithm.

2.2 CMIP6 data: ice, wind, and ocean

To compare models to the full observational time series (1978 - 2024), we use both the historical run (Eyring et al., 2016) that finishes on 31st December 2014 and the Shared Socio-economic Pathway “SSP” climate change run (Tebaldi et al., 2021) of the Climate Model Intercomparison Project phase 6. Our analysis centres on the 18 models from 10 different model families that had daily sea ice concentration “siconc” available at the time of download (summer 2025), listed in Table 1. They cover nearly three orders of magnitude in horizontal resolution in the Arctic, from just above 100 km² per grid cell area for AWI-CM-1-1-MR to nearly 10 000 km² for IPSL-CM5A2-INCA. Most models are of the order of 1000 to 2000 km², or about 1° longitude/latitude. Their sea ice and ocean components will be further discussed in the analysis. We acknowledge that real-world Arctic polynyas are preferably detected using sea ice thickness rather than concentration (e.g. Tamura and Ohshima, 2011; Preußner et al., 2019). However, the daily sea ice thickness in CMIP6 is often not available, and it has been reported to have large errors compared to observations on multiple occasions (Mohrmann et al., 2021; Athanase et al., 2025), so we use the more reliable and widely available daily sea ice concentration. We also use the monthly sea ice concentration for a brief comparison; see Appendix Table A2 for the ensemble members and SSPs.

When possible, we used up to 10 ensemble members per model. The ensemble spread is the first thing analysed in section 3. We tried to use only SSP2.4-5 but it was not available for some models. In the absence thereof, we used SSP3.7-0, and SSP5-8.5 when it was the only one available; see Appendix Table A1. As shown in Tebaldi et al. (2021), using different SSPs should have no impact on our results as the scenarios do not diverge before approximately year 2050 whereas we stop in 2024. We nonetheless redid the analyses with the historical run only and find no significant difference (not shown): the values were obviously different, but the correlation between the two was near-perfect, and most importantly, the relative model biases were unchanged.

For investigating the potential causes of polynya biases, we use the monthly sea ice thickness “sivol” (in m, computed as the total volume of sea ice in m³ divided by grid-cell area in m²) since the daily thickness has large errors, as we described above, or is not available. Note that the variable “sithick” exists but is ambiguously defined as the floe thickness, and is less widely available than “sivol”. For CanESM5 and the EC-Earth3 models, “sivol” is not available; we used “simass” (total mass of sea ice divided by grid-cell area) that we converted to sivol by dividing it by their ice density of 917 kg m⁻³ (Vancoppenolle et al., 2023). Because of missing files, only 14 models can be used for this analysis; the other four are marked with “no thickness” in Table 1. We also used the daily surface zonal and meridional wind speeds “uas” and “vas”, respectively; they were not available for IPSL-CM5A2-INCA, as indicated in Table 1.

For investigating potential consequences of polynya activity on the ocean, we used the monthly full-depth potential temperature “thetao” and practical salinity “so” on the models’ native grid. Full depth variables being very heavy, we use only the three models with the most and three models with the least polynya activity (° next to the model name in Table 1). As we will show, there is a large agreement in polynya activity across ensemble members, so we use only the first ensemble member

135 of each model: r1i1p1f1 for ACCESS-CM2, AWI-CM-1-1-MR, CanESM5, IPSL-CM6A-LR, and NESM3, and r1i1p1f2 for MIROC-ES2L.

Model	Res. >70°N	Sea ice	Ocean	Comment	Model reference
b) ACCESS-CM2 °	965	CICE5	MOM5		Bi et al. (2020)
c) AWI-CM-1-1-MR °	138	FESOM	FESOM		Semmler et al. (2020)
d) CanESM5 °	2543	LIM2	NEMO3.4.1	simass	Swart et al. (2019a)
e) EC-Earth3-AerChem	2526	LIM3	NEMO3.6	simass	Döscher et al. (2021)
f) EC-Earth3-CC	2526	LIM3	NEMO3.6	simass	Döscher et al. (2021)
g) EC-Earth3-Veg-LR	2526	LIM3	NEMO3.6	simass	Döscher et al. (2021)
h) EC-Earth3-Veg	2526	LIM3	NEMO3.6	simass	Döscher et al. (2021)
i) IPSL-CM5A2-INCA	8681	LIM2	NEMO3.6	no thickness no uas and vas	Sepulchre et al. (2020)
j) IPSL-CM6A-LR °	2543	LIM3	NEMO3.6		Boucher et al. (2020b)
k) MIROC-ES2L °	1100	COCO	COCO	no thickness	Hajima et al. (2020)
l) MIROC6	1100	COCO	COCO		Tatebe et al. (2019)
m) MPI-ESM-1-2-HAM	2074	MPIOM1.6.3	MPIOM1.6.3		-
n) MPI-ESM1-2-HR	1406	MPIOM1.6.3	MPIOM1.6.3		Müller et al. (2018)
o) MPI-ESM1-2-LR	2074	MPIOM1.6.3	MPIOM1.6.3		Mauritsen et al. (2019)
p) MRI-ESM2-0	1098	~CICE	MRI.COMv4		Yukimoto et al. (2019a)
q) NESM3 °	2490	CICE	NEMO3.4	no thickness	Yang et al. (2020)
r) NorESM2-LM	1472	CICE5	BLOM	no thickness	Seland et al. (2020)
s) NorESM2-MM	1472	CICE5	BLOM		Seland et al. (2020)

Table 1. For the models with daily sea ice concentration output: Model name and letter used on all figures and tables; Horizontal resolution “Res.” north of 70°N as grid cell area in km² (total area north of 70°N, in km², divided by number of grid cells north of 70°N); Sea ice component and version, when indicated in the reference (~ means similar to, according to the reference); Ocean component and version, when indicated in the reference; Comments; and reference of the model description paper. ° next to the model name indicates a model used in the ocean analysis of section 5. See also Appendix Table A1 for the ensemble members, simulations, and dataset DOIs.

2.3 Polynya statistics and other analyses

140 In line with e.g. the “real” Arctic polynya assessment of Preußner et al. (2019) or that of polynyas in CMIP6 in Southern Ocean of Mohrmann et al. (2021), our evaluation metric is the yearly/winter mean polynya area. We define winter for polynyas as the period from 1st December to 31st March to reduce the risk of mis-detecting late freeze-up and early melt onset as polynyas. The yearly/winter mean polynya area is then the mean of the daily polynya areas, which are for each day the sum of the grid cell areas of all pixels detected as a polynya. For some diagnostics, we keep the 45-winter time series. For most of the manuscript however, we then compute the 45-winter mean of the mean polynya area. We tested using instead the historical run only (i.e.

1979 to 2014 instead of 1979 to 2024) and found no significant differences, with correlations exceeding 0.99 (not shown); we
145 found the same result if using instead 1982-2014 as Tian et al. (2024) did.

Two other diagnostics were used by Wong et al. (2026): the cumulative polynya area, i.e. sum of the daily values instead
of the mean (described above), and the total area of each winter, i.e. the sum of the grid cell areas of all pixels detected as
a polynya at least once each winter. We computed them both for all CMIP6 models (not shown) and found a near-perfect
correlation (>0.9) between our yearly/winter means and their diagnostics. In particular, the relative biases of the models did
150 not change, i.e. the models with the largest values with one diagnostic were still the largest with another one. We therefore use
the yearly mean polynya area only. We define the polynya frequency of each pixel as the number of winters where that pixel is
detected as a polynya at least once.

As is common in observational studies (e.g. Tamura and Ohshima, 2011; Preußer et al., 2019; Wong et al., 2026), we study
the polynya behaviour of different regions of the Arctic separately. We use the same region definition as Wong et al. (2026) as
155 it is optimised to reduce noise from incorrectly detected marginal ice zone and is adapted to the same time period as our study,
including the emergence of polynyas in north-east Greenland. The coordinates of each region are given in Appendix Table A3
and shown on the first pan-Arctic map of section 3.

Unless stated otherwise, all across-model correlations shown are significant at 95% and conducted on the ensemble-average
values. We show linear regressions only if the corresponding regression is significant. For the trend analysis, linear trends
160 (with 95% significance) are calculated on the individual ensemble members first and the ensemble-average trend is the value
presented. As for the yearly means, changing the time period for the trend calculation to 1979-2014 or even 1982-2014 made
no significant difference (not shown), and in particular did not change the models' relative biases.

For the beginning of the analysis, where we determine the pan-Arctic behaviour of the models, the sea ice area is computed
as the sum of all concentration \times grid cell area, using the concentration on the interpolated grid cut at 68°N . Using the same
165 grid, the sea ice extent is computed as the sum of the grid cell areas for all grid cells where the concentration is strictly larger
than 15%. Along with the 1979-2024 mean of the whole-year means we also compute the mean seasonal cycle: for each
year, we compute the difference between the yearly maximum and the yearly minimum SIA or SIE, respectively, and take the
45-year mean.

The objective of this study is to try and find a link between polynya biases and model biases across the CMIP6 ensemble,
170 not to conduct a process study in individual models. Therefore, our priority is to find across-model relationships. Only in the
absence of such relationship do we resort to time series correlations at the individual model level. Since polynya area and
frequency have regularly been linked to wind speed, including its extremes, and air temperature (see Wong et al., 2026, and
references therein), we use these two parameters in our study.

The daily wind speed is computed from the daily "uas" and "vas". In each region, we extract the mean of the three variables,
175 but also their 90th percentile as the most extreme value. For the two components of the wind "uas" and "vas", we have to
extract both the 10th percentile (most extreme negative values, i.e. west- or southward) and the 90th percentile (most extreme
positive values, i.e. east- or northward).

We use the air temperature biases of Tian et al. (2024). They use only the first ensemble member of each model (r1i1p1f*), only the time period 1982-2014 (historical run), and before computation regridded all the models to a common 1° grid. For across-model correlations, we therefore also use the first ensemble member only and the 1982-2014 time period for our polynya areas. Compared to our list of models, the only model missing from their study is AWI-CM-1-1-MR. All their values are biases, i.e. model value minus a common satellite-based reference; see Tian et al. (2024) for more information. Since it is a common offset for all models, it has no impact on our across-model correlations.

Finally for the ocean variables, we limit our investigation to a comparison of the Kara and Laptev seas. We extend the geographical limit of the Kara Sea northward to 86°N compared to the rest of the study (coordinates in Appendix Table A3) to capture the potential effect of the polynya not only on the water column below, but in the deep basin offshore via potential dense water overflows and effects on the halocline; the original definition of the Laptev Sea already includes the relevant deep ocean area. We further split each sea into three regions based on their bathymetry

- continental shelf: shallower than 500 m;
- shelf break: bathymetry between 500 and 1000 m;
- abyssal plain: deeper than 3000 m.

The bathymetry itself was computed as the depth of the last level with salinity data, since the modelled bathymetry “depth” was not available for all models at the time of download. We compute the bottom properties as the properties T and S in that last level. Given the large range of depths considered, we use the density anomaly referred to 2000 m σ_2 , computed following TEOS10 as implemented in the Gibbs SeaWater toolbox (Roquet et al., 2015). Finally, following Heuzé et al. (2023), we extract each model’s properties of the Atlantic Water core, defined as the depth of the maximum temperature deeper than 200 m.

3 Results: Assessment of biases in winter Arctic polynya

3.1 Polynyas are too large but do not open often enough

At the pan-Arctic scale, 10 out of 18 models overestimate the winter polynya area for all their ensemble members (Fig. 2). The largest overestimation is from AWI-CM1-1-MR, which is nearly 10 times too high. This is also the model with the highest horizontal resolution by far (Table 1), which is consistent with both Wong et al. (2026) and Landrum et al. (2024) who showed that higher resolution in the product leads to larger polynya area detected. The bias does not seem related to the ocean model: both CanESM5 and NESM3 use the same model NEMO (both version NEMO3.4), yet the polynya area in CanESM5 is twice as large as the observations, while that of NESM3 is only a third of the observed area. We will show in section 4.1 that this is probably a consequence of their different sea ice models. The intermodel spread is much larger than the spread across ensemble members of each model (Fig. 2). This result still holds when looking at the individual regions separately (Appendix Figure A1). The rest of the bias analysis is therefore done with the ensemble mean of each model, unless specified otherwise.

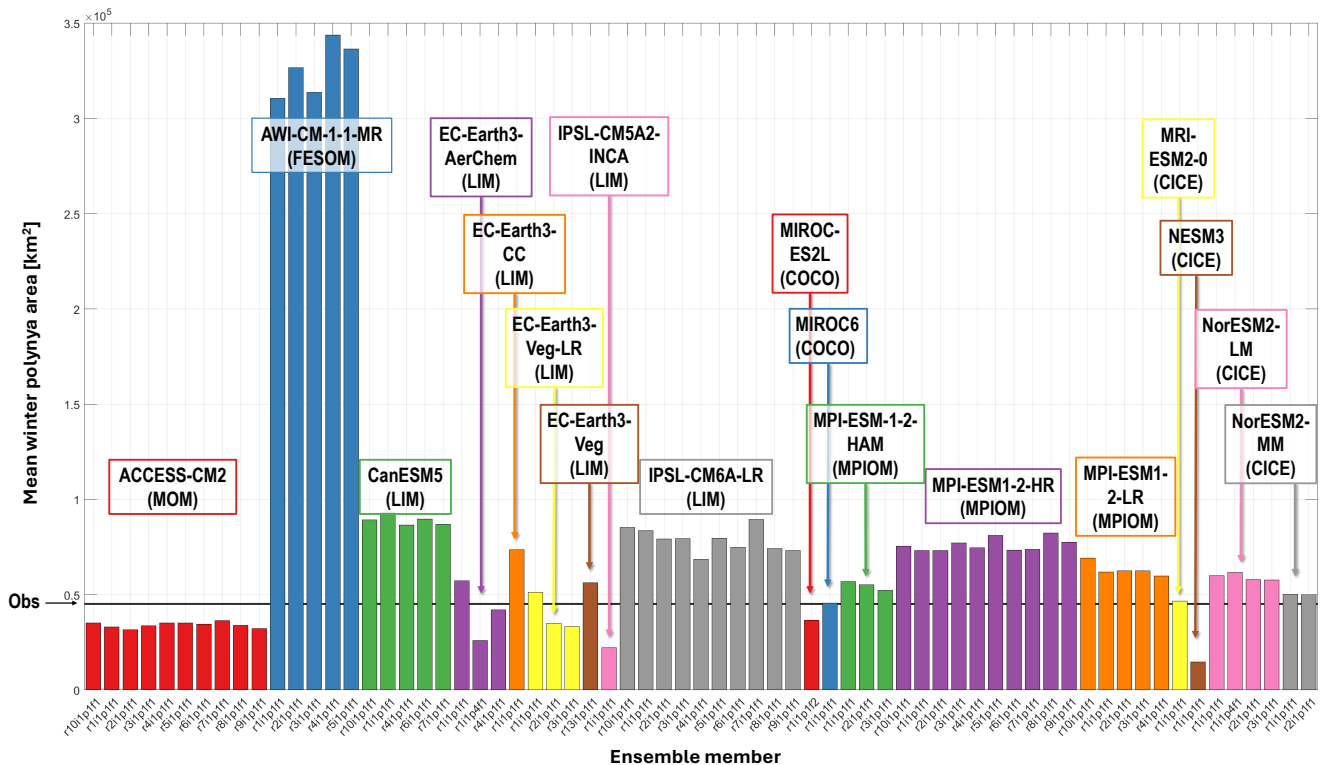


Figure 2. 1979-2024 mean of the winter (Dec - Mar) mean polynya area for the entire Arctic, for the observations (black horizontal line) and for each ensemble member of the models with daily sea ice concentration. Individual regions shown on Appendix Fig. A1. Acronym in parentheses is the sea ice model family, as discussed in the rest of this section and in section 4.1.

At the pan-Arctic scale, models not only overestimate the 45-year mean polynya area, they also overestimate how often polynyas open (Fig. 3). Most models have more polynya activity than the observations, either because low-activity areas extend further north or south (red colours), or because the regions are more active than the observations (yellow colours). In observations, an increase in polynya activity as sea ice area/extent declines has been reported (e.g. Preußner et al., 2019; Heuzé and Wong, 2025; Wong et al., 2026). At the multi-model level, we find the opposite relationship. We find an across-model correlation between pan-Arctic sea ice area / extent and mean winter polynya area of 0.75 / 0.84, i.e. the models with most sea ice also have the most polynyas. This is most likely because, with the exception of CanESM5 and the EC-Earth3 models, polynya activity in models is mostly limited to the seasonally ice covered area (between the summer extent in cyan and the winter extent in dark blue, Fig. 3), whereas in observations there are polynyas in the inner ice pack (north of the cyan line). The across-model correlation between polynya activity and the model's mean seasonal cycle in SIA / SIE is 0.50 / 0.53: the larger the seasonally ice-covered region, the more polynyas. This suggests a link with sea ice thickness, which we will investigate in section 4.

220 Again, the highest resolution model AWI-CM1-1-MR has the most activity. Note that we visually verified its daily sea
ice concentration output and can confirm that these are indeed polynyas. There are however many pixels with a low polynya
opening frequency out of the main, predefined regions, for most models: this is noise from the marginal ice zone in the North
Atlantic, Barents Sea and on the Pacific side. As discussed both for observations by Preußner et al. (2016) and for models by
Heuzé and Wong (2025), this is unavoidable in the Arctic, even with traditional threshold + flood-fill methods, owing to the
225 complex shape of the coastline and the MIZ itself. To reduce this noise, we conduct the rest of the bias analysis on the nine
regions highlighted on Fig. 3a (see also Appendix Table A3).

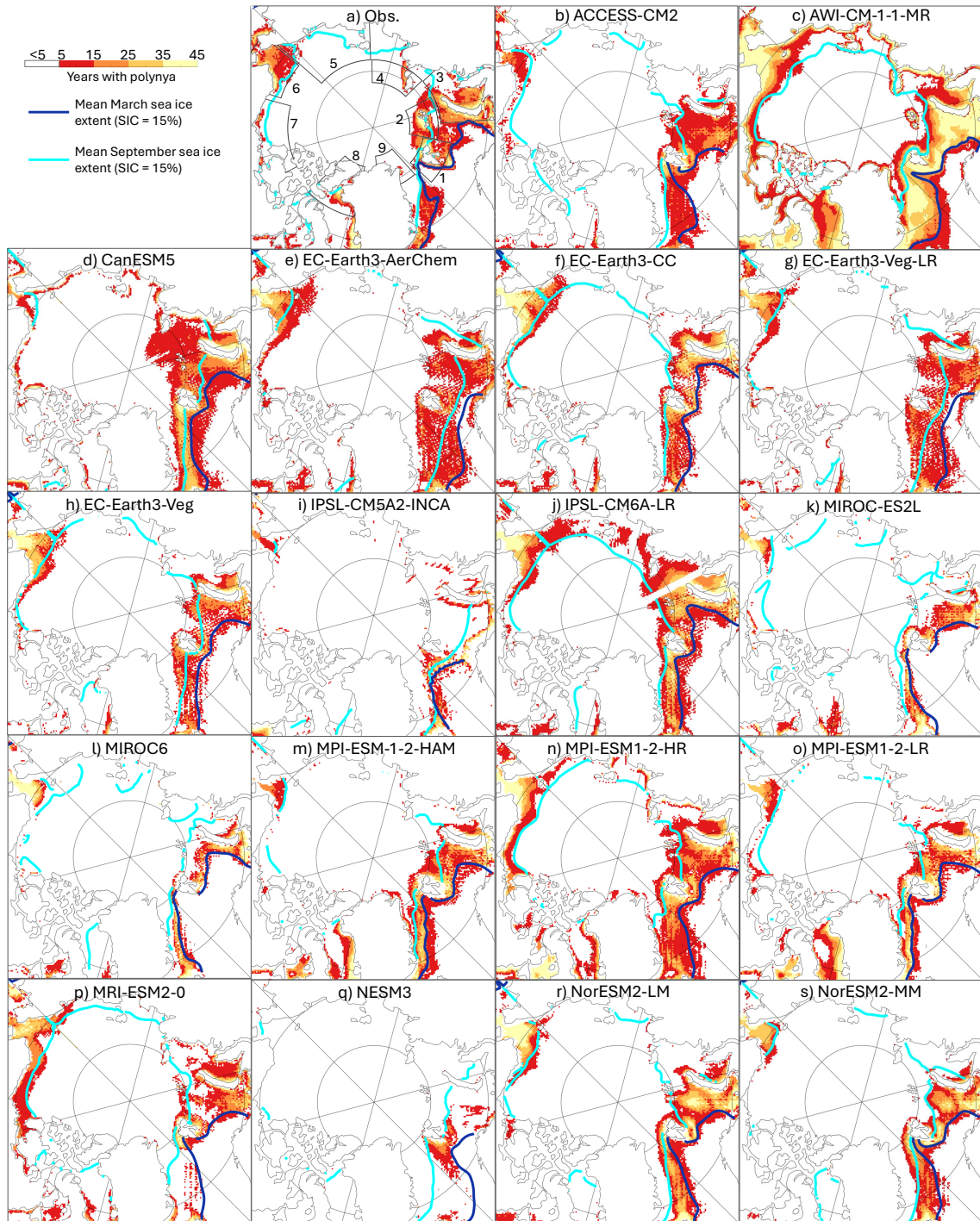


Figure 3. For each pixel, for the observations (a) and the ensemble-average of the CMIP6 models with daily output, number of winters over 1979-2024 where a polynya opened there at least once between 1st December and 31st March. Note that polynyas are detected using sea ice concentration only. Dark blue and cyan are the mean sea ice extent in March and September, respectively, over the same period 1979-2024; the “tail” of sea ice (cyan line) in the top-left corner along 180° , seen for some models, is a plotting artefact. Thick black lines on a) are the regions used in this study, as per Wong et al. (2026), also defined in [Appendix Table A3](#).

To quantify the biases in polynya frequency, we compare the percentage of each region where polynyas have opened for more than 25 out of the 45 years (not necessarily continuously), in both models and observations (pale orange and pale yellow on Fig. 3). In the regions Svalbard (column 1 in Table 2), East Siberian (5) and Chukchi (6), a majority of models overestimate
230 locations that have a high opening frequency (10, 12 and 12 models, respectively). The overestimation is most dramatic in the East Siberian Sea where the observational value is 0. Most models show very little activity in the North Water Polynya (NOW, column 8). This region is one of the most active polynya regions in the Arctic but is difficult to model owing to the narrow straits and complex bathymetry (e.g. Moore et al., 2023). It is therefore surprising that the models' behaviour in the North Water Polynya region is not related to the model's horizontal resolution. For example, the higher resolution MPI-ESM1-2-HR
235 is lower than observations (3.9 instead of 6.9%, column 8 in Table 2), but its low resolution counterpart MPI-ESM1-2-LR is higher than the observations (15.0 instead of 6.9%).

The largest overestimation of the frequent-opening locations is from AWI-CM-1-1-MR in the Beaufort Sea, with a value 48 times that of the observations. In fact, this model largely overestimates the area with frequent polynyas in all regions except Franz Josef Land, where it underestimates it by nearly half (8.0% instead of 13.1%). MRI-ESM2-0 is the only other model
240 that is quite consistent in overestimating the frequent-opening areas: it does so in all regions except Franz Josef Land, Kara Sea, and the North Water polynya. It is more common for CMIP6 models to underestimate the frequent-opening areas: seven of the models do so in at least 7 out of 9 regions (ACCESS-CM2, EC-Earth3-AerChem, EC-Earth3-Veg, IPSL-CM5A2-INCA, MIROC6, MPI-ESM-1-2-HAM, NESM3), and 13 in at least 6 (same + EC-Earth3-CC, EC-Earth3-Veg-LR, MIROC-ES2L, MPI-ESM1-2-LR, NorESM2-LM, NorESM2-MM). In summary, even though Fig. 3 suggests that models open polynyas too
245 often, what they do is to open with a low frequency at too many locations. The modelled polynyas' spatial distribution is mostly confined to the marginal seas and rarely form in the central ice pack.

Finally, winter Arctic polynya area has been increasing in observations (e.g. Preußer et al., 2016; Preußer et al., 2019; Wong et al., 2026), with Wong et al. (2026) in particular finding positive trends since satellite records began in all regions except Svalbard. Do the CMIP6 models reproduce these trends? In most regions, even in Svalbard, a majority of CMIP6 models
250 exhibit a positive trend in polynya area over 1979-2024 (Fig. 4). The trend is however inaccurate when compared to the observed one, but there is no consistency in whether the models over- or underestimate it (compare symbols to the horizontal black line). We find rather that there often is a relationship between the 45-year mean bias in polynya area and the bias in trend of the polynya area (Fig. 4), with five regions exhibiting large significant positive correlations between the two, albeit with three different behaviours. In Franz Josef Land (region 2, $R=0.81$) and the Laptev Sea (region 4, $R=0.84$), most models have
255 a low bias in both 45-year mean and trend, except for CanESM5 (d) and IPSL-CM6A-LR (j) that have a large bias in both. In the Kara (3, 0.67) and Beaufort (7, 0.77) seas, all models overestimate the trend, but the trend is greater for larger polynya areas. In the North Water polynya (8, 0.91) in contrast all models underestimate the trend, but likewise the trend is larger for larger polynya areas. Note that AWI-CM-1-1-MR is excluded, as it is a clear outlier; when it is included (see Appendix Fig. A2), all regions except the MIZ-dominated Svalbard have an extremely strong correlation, exceeding 0.9 in five regions.
260 Note that although Heuzé and Wong (2025) found, in observations, an increased performance of the detection algorithm as

	1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
a) Obs.	17.9	13.1	21.5	0.5	0.0	11.2	0.4	6.9	0.3
b) ACCESS-CM2	28.1	1.1	2.0	0.0	0.0	1.1	0.0	0.0	0.0
c) AWI-CM-1-1-MR	66.7	8.0	53.7	17.7	14.7	74.3	17.7	25.8	15.6
d) CanESM5	23.7	0.0	18.0	1.8	1.2	3.3	3.2	0.2	0.3
e) EC-Earth3-AerChem	0.2	0.0	2.4	0.0	0.6	28.3	0.0	0.0	0.0
f) EC-Earth3-CC	17.0	4.2	21.4	0.0	3.7	54.9	1.0	0.0	0.0
g) EC-Earth3-Veg-LR	0.4	0.0	3.8	0.0	0.6	23.3	0.7	0.2	0.0
h) EC-Earth3-Veg	6.6	0.0	8.9	0.0	1.3	45.5	0.2	0.0	0.0
i) IPSL-CM5A2-INCA	1.9	0.0	1.3	0.0	0.0	0.0	0.1	0.0	0.0
j) IPSL-CM6A-LR	23.7	20.5	28.7	0.1	0.7	42.1	0.1	0.0	0.0
k) MIROC-ES2L	8.7	0.2	3.5	0.2	0.1	16.5	1.0	0.0	0.0
l) MIROC6	6.7	0.2	9.4	0.1	0.0	28.9	0.1	0.0	0.0
m) MPI-ESM-1-2-HAM	23.5	0.0	7.8	0.1	0.0	0.1	0.0	4.3	0.0
n) MPI-ESM1-2-HR	35.8	0.7	18.4	2.6	0.8	37.8	5.5	3.9	0.0
o) MPI-ESM1-2-LR	23.5	3.6	11.0	0.1	0.0	10.4	0.0	15.0	0.0
p) MRI-ESM2-0	23.1	0.9	10.7	0.5	1.9	27.5	0.8	0.4	0.6
q) NESM3	5.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
r) NorESM2-LM	24.5	1.6	14.6	0.0	0.4	45.1	0.0	0.0	0.0
s) NorESM2-MM	23.5	0.0	4.8	0.0	0.2	26.0	0.0	0.0	0.0

Table 2. For each region, percentage of the region where there is a polynya more than 25 out of 45 winters over 1979-2024, for the observations (top row) and the ensemble-average of the models. Only non-land pixels considered.

SIA/SIE decreased, we find no across-model correlation between the trends in SIA/SIE and either the trend or mean values of the polynya area.

In summary, there is no across-model agreement on the polynya area trend, rather different clusterings in the different regions. We see no link with the model resolution in any of the regions (shading on Fig. 4), but there appears to be some clustering by sea ice model. The LIM models (stars on Fig. 4) in particular are often found together, most visibly by the Pacific inflow regions of the East Siberian (all of them are to the right) and Beaufort seas (all of them are to the bottom centre/left). We will investigate why that might be in section 4, but first, since the number of models with daily sea ice is way less than that with monthly sea ice, let us investigate the impact of using monthly CMIP6 data for polynya detection.

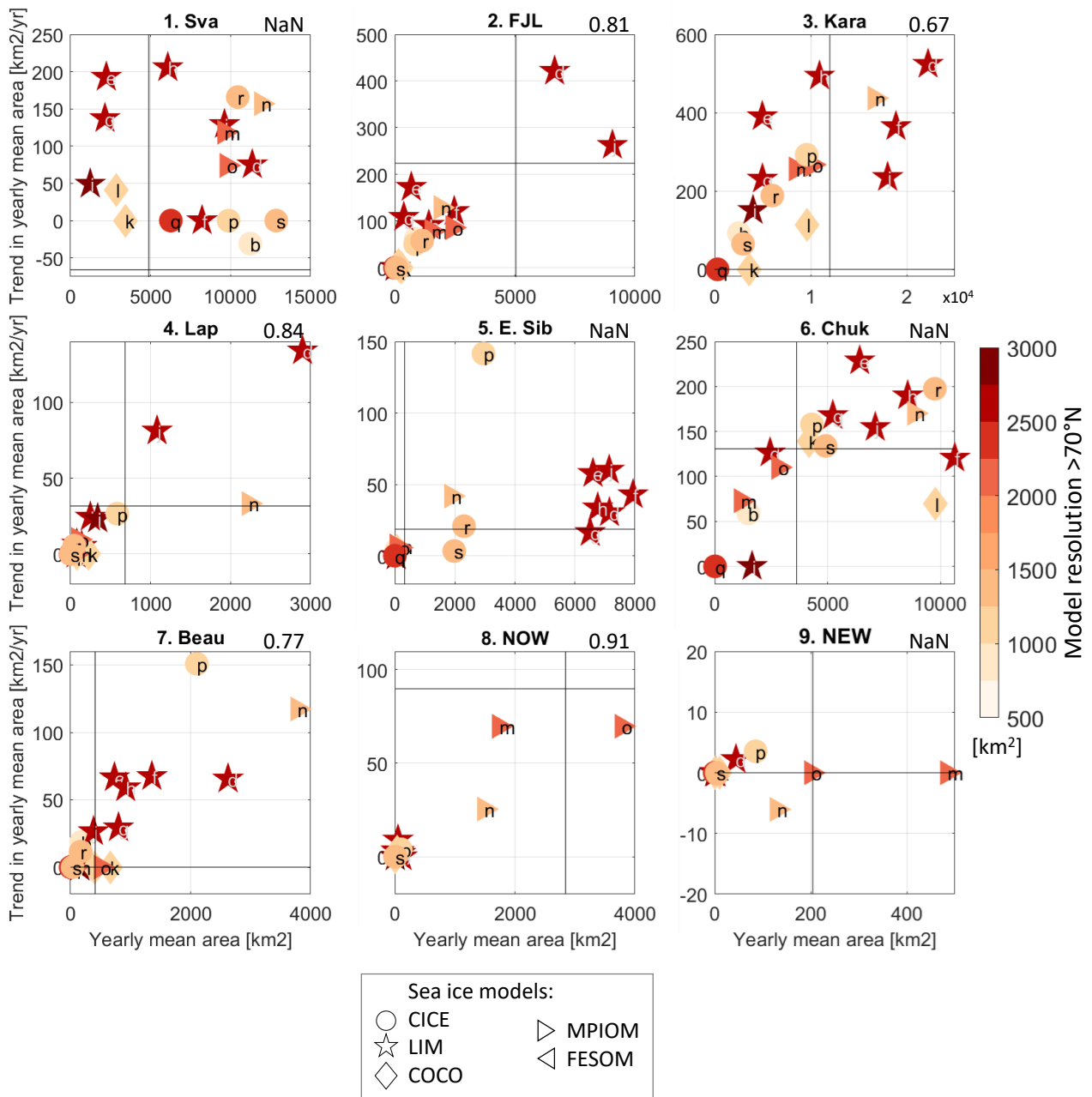


Figure 4. Across-model relationship between the 1979-2024 mean of the winter-mean polynya area (x-axis) and the 1979-2024 trend (95% significance) in winter mean polynya area (y-axis) for each region as defined in Appendix Table A3. If the trend is insignificant, it is shown here as 0. Models referred to with same letters as on Fig. 3. Numbers in the top right corner of each panel is the correlation between the two (95% significance), considering only the models with a significant trend; ‘NaN’ means no significant correlation. Black lines are the observational values. Shading is each model’s nominal resolution north of 70°N; symbol, its sea ice component. AWI-CM-1-1-HR is excluded for clarity; see Appendix Figure A2 for the version with that model included. Mean area and trend values are also given in Appendix Table A4.

3.2 A linear relationship between values from daily and monthly output

270 There are nearly twice as many CMIP6 models with monthly sea ice concentration than with daily sea ice concentration (35 vs 18). It is therefore tempting to try and detect polynyas in CMIP6 models using monthly data, although polynyas usually last but a few days (Smith Jr and Barber, 2007). We can immediately see that both areas and frequency are largely underestimated when using monthly output (Fig. 5), including for the monthly-version of the observations where only a handful of pixels have a polynya 5 to 10 out of the 45 years, but none more often than that (i.e. only red pixels on panel a). The same holds for 275 the CMIP6 models, with most models having very few polynya pixels at the pan-Arctic scale, and nearly no model having frequencies exceeding 15 years (i.e. another colour than red) anywhere.

The relationship with the sea ice area / extent and its seasonal cycle have also been lost, probably because of the extreme behaviours in the monthly-only models. Without the need for any calculation, we can flag several models as having clearly inaccurate sea ice concentration, models that are therefore not trustworthy in terms of their polynya representation. For example, 280 CIESM and the two CMCC models have virtually no summer sea ice on average over 1979-2024 (cyan contours around the Canadian archipelago on Fig. 5). CAMS-CSM1-0, CAS-ESM2-0, and the CNRM models in contrast have an immense winter ice cover (dark blue contours), including the entire Greenland Sea all the way to Iceland. CAS-ESM2-0 also has one of the most extreme seasonal cycles (distance between the cyan and dark blue lines), while CESM2 has virtually none. Should the reader consider using monthly output for a polynya analysis, we strongly recommend that they reject these models with unrealistic 285 sea ice cover.

AWI-CM-1-1-MR and IPSL-CM5A2-INCA are not available as monthly output, and EC-Earth3-Veg had no ensemble member in common for the daily and monthly output at the time of download. We therefore momentarily limit our study to the 15 remaining models and compare the mean polynya area obtained when using their daily or monthly output. Note that we used only the ensemble members common to both time resolutions, so the polynya area values presented on Fig. 6 differ slightly 290 from those on Fig. 4 and in Appendix Table A4.

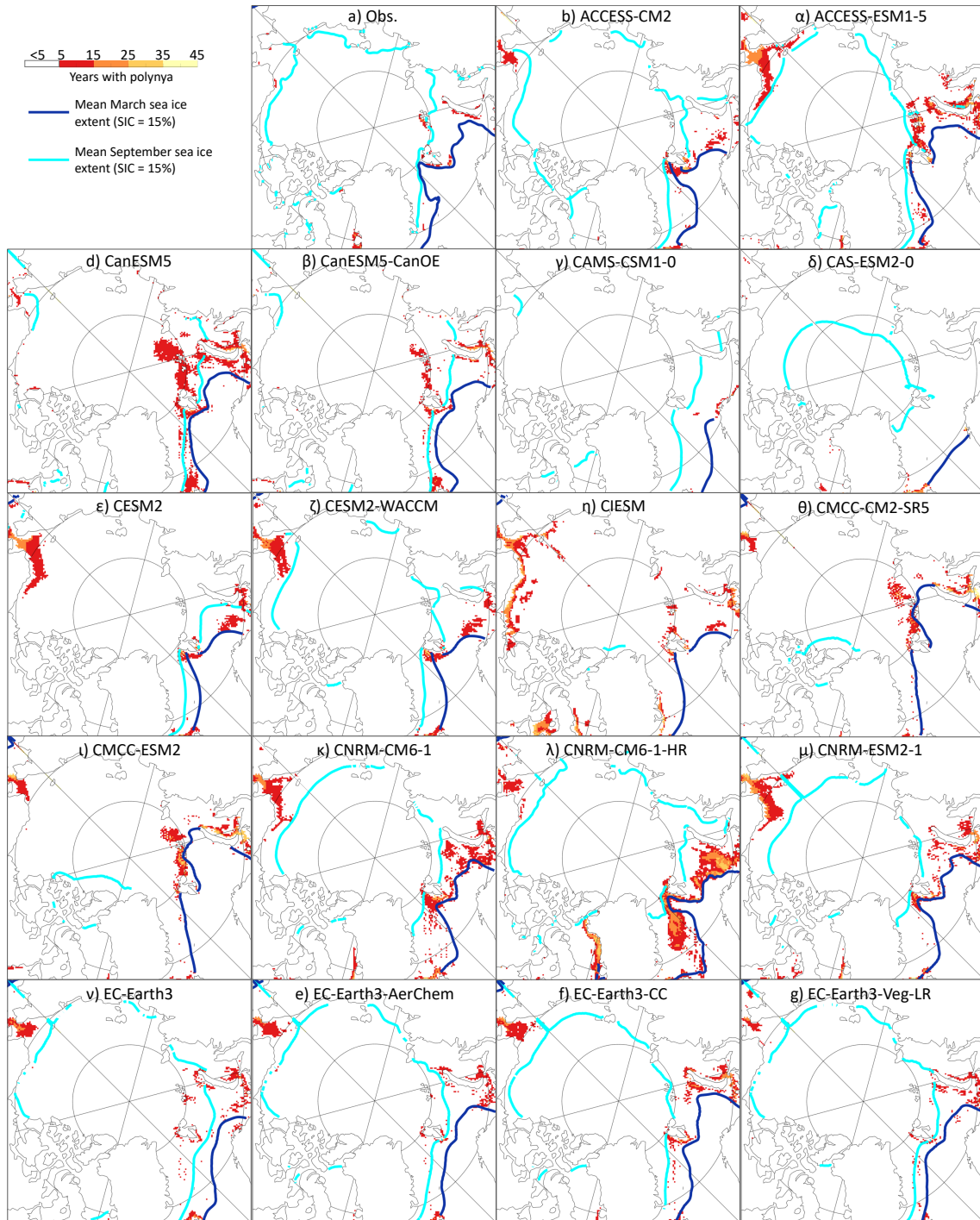


Figure 5. Same as Fig. 3 but for monthly output: For each pixel, for the observations (a) and the ensemble-average of the CMIP6 models with monthly sea ice output, number of winters over 1979–2024 where a polynya opened there at least once between 1st December and 31st March. Dark blue and cyan are the mean sea ice extent in March and September, respectively, over the same period 1979–2024; the “tail” of sea ice in the top-left corner of some models is a plotting artefact.

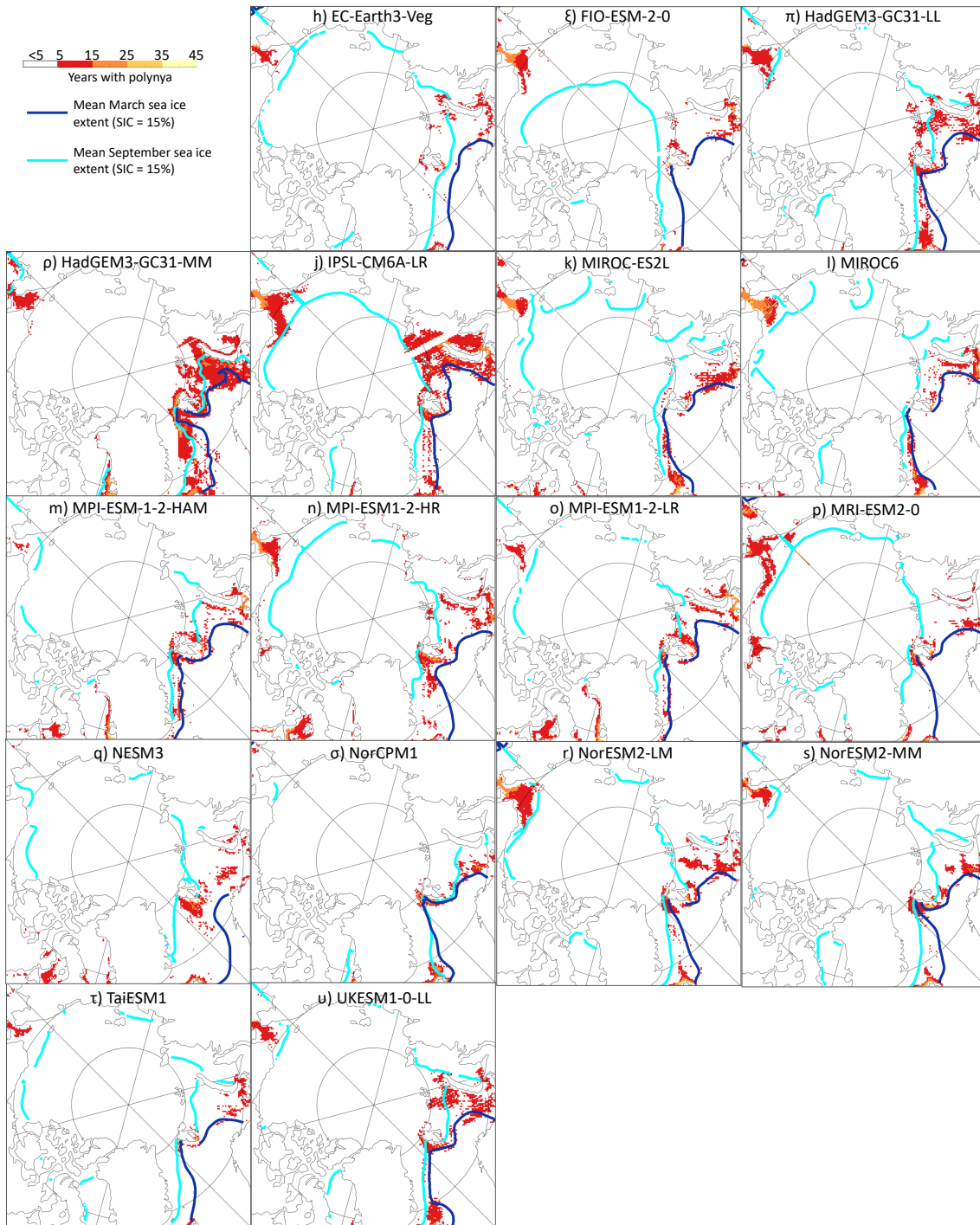


Figure 5. continued.

We find a very strong significant across-model correlation in all regions between polynya area from monthly and that from daily values (Fig. 6). In four regions, as expected, the monthly value is much smaller than daily: half as small in the Beaufort Sea (region 7), 0.4 times in the North Water polynya (region 8), 5 times smaller in the Laptev Sea (region 4), and up to 10 times smaller in the North East Water polynya (region 9). In four more regions, the regression shows a nearly one-to-one relationship between the monthly and daily results (0.82 in Svalbard, 0.93 in Franz Josef Land, 0.72 in the Kara Sea, and 0.86 in the East Siberian Sea). Because detection in these regions is impacted by noise from the MIZ, this agreement suggests that monthly and daily results are similarly impacted. In the Chukchi Sea (region 6), the monthly results are larger than the daily results for 6 of the models, which toggles the regression slope above 1. That is the only such region. Visual inspection (not shown) reveals that the daily sea ice there is very “binary” in these models: small areas are clearly open (concentration of 15% or less) next to a very sharp transition to the consolidated ice pack (concentration above 80%). So even though this transitional zone varies a lot from day to day, this results in a large region with sea ice concentration around 40% as monthly mean, which the algorithm detects as a large polynya. The same happens to MRI-ESM2-0 (model p) in the Beaufort Gyre (not shown): It experiences a very late freeze up in December 2015 and 2017, during which the daily ice edge moves a lot from daily to daily images. The daily polynya area is not affected, since these regions are filtered out as open ocean for the daily results, but the resulting monthly mean sea ice concentration is low over a large region and erroneously interpreted as many polynya pixels for the monthly result.

Note that standard threshold methods would be affected by these behaviours as well, unless the authors select a high sea ice concentration as threshold, which then further reduces their ability to remove the marginal ice zone. This could be fixed in our method by fine-tuning the thresholds and settings of either Unet or the GMM per model and/or region, but we want to keep the same values for all models, at the pan-Arctic scale. One could also use the daily sea ice concentration to generate the MIZ filter and apply it to the monthly classification result, although the computing time saved would be insignificant - if daily output are available, we recommend their consistent use. A dedicated study of these specific models in the Chukchi and Beaufort seas is beyond the scope of this paper. Note also that even though the observations are not included in these correlation and regression calculations, they fall close to the regression line but have systematically lower results from monthly than daily values, as is to be expected.

In summary, one can mix and match monthly and daily results if one is careful, since the relationship between the two differs depending on the Arctic region. We recommend against doing that and would rather encourage the widespread release of daily sea ice data, as polynyas are short-lived events that open and close on daily to sub-daily time scales. For that reason, in the next section, we use only the daily outputs. We would also like to re-iterate that sea ice concentration alone is not optimum for detecting polynyas in the Arctic and would welcome a widespread release of daily sea ice thickness output.

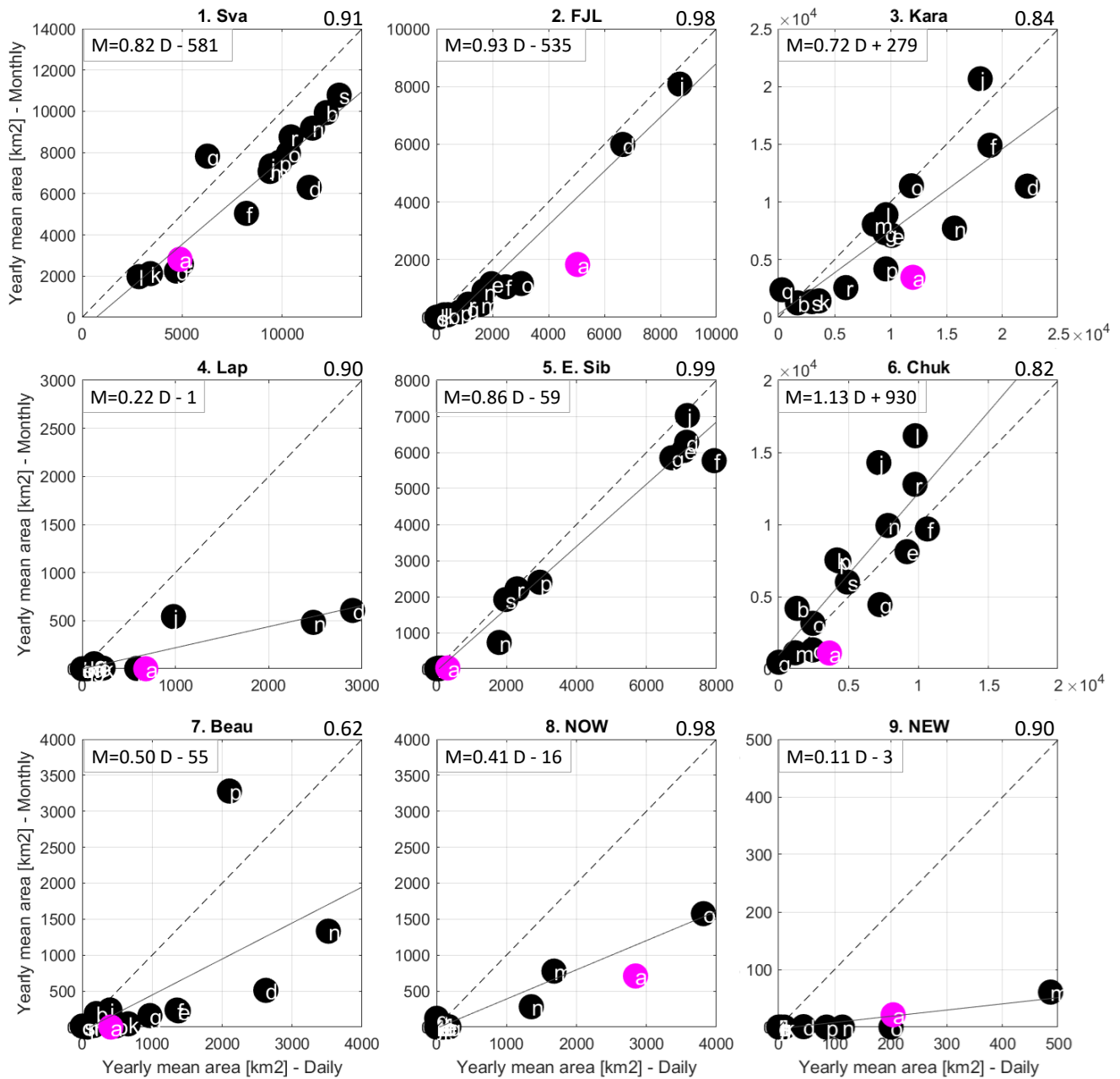


Figure 6. For the models that have both daily and monthly sea ice concentration, across-model relationship between the 1979-2024 mean of the winter mean polynya area computed from the daily output (x-axis) and that from the monthly output (y-axis) for each region as defined in Appendix Table A3. Models referred to with same letters as on Fig. 3. Numbers in the top right corner of each panel is the correlation between the two (95% significance), considering only the models with a significant trend. Equation in the top left corner is the linear relationship between monthly “M” and daily “D” values, also shown as the plain black line. Dashed black line is the unit $x=y$ line. Observational values shown as magenta dot but not included in the correlation or regression calculation.

4 Potential causes of biases in winter Arctic polynyas

In this section, we highlight processes or features that might explain the biases in winter Arctic polynya activity in CMIP6 models. The reader must bear in mind that a proper mechanistic understanding is impossible as we do not have access to the high-resolution model output O(10 min), only their daily average. Likewise, we do not have access to the model code and cannot run sensitivity experiments, only point out similarities. Therefore this section relies only on correlations, at the individual model level and across-models, and carefully discusses potential causalities.

4.1 Sea ice thickness and sea ice model

One expects a negative correlation between sea ice thickness and polynya area in winter, since sea ice thickness is commonly used to detect polynyas (e.g., Preußer et al., 2016). For each model and each region, we correlate the time series of their region-average, season-average sea ice thickness with the time series of the same region's DJFM polynya area. We indeed find negative correlations between winter thickness and winter polynya area in most regions for most models (Appendix Table A5): the thinner the winter sea ice, the larger the polynyas that same winter. The occasional positive correlation in the Svalbard region in our opinion is an indirect one, which rather indicates that by definition where there is sea ice (an exception in that region, the norm in all the others), polynyas can open. We also find many significant negative correlations with the previous summer (JJA) and autumn (SON) ice thicknesses. However, since there is a strong autocorrelation between different seasonal sea ice thicknesses (around 0.5 in all regions and models, not shown), and bearing in mind the link between sea ice seasonal cycle and polynya activity of Fig. 3, these many correlations suggest rather that models that tend to have thin ice in general have large polynya activity.

We do find an across-model relationship between 45-year mean sea ice thickness and 45-year mean polynya activity, but not everywhere nor in all regions (Table 3). In the Kara, Laptev and Beaufort seas and the North Water polynya, the correlation is strongest with the summer sea ice thickness; in the Chukchi sea, it is strongest with the winter thickness; in Franz Josef Land, it is nearly equally strong with both seasons. Since few models had correlations between their time series of autumn thickness and polynya area, it is not surprising that we find across-model correlations between their 45-year means only in the Kara and NOW regions, and these are of lesser magnitude than for the other seasons (-0.66 in Kara, compared to -0.91 with summer and -0.77 in winter). Many of the correlations depend on whether the AWI-CM-1-1-MR model is included, as exemplified on Appendix Fig. A3: in the Kara Sea where it has an average thickness, including its very high polynya area breaks the correlations; in the Laptev Sea where it has a low thickness, it makes them.

In observations, Sumata et al. (2023) identified a regime shift in sea ice when the sea ice draft passes below the 1.3 m threshold, which corresponds to sea ice becoming thinner than approximately 1.5 m. This 1.5 m thickness threshold has also been seen in models: Athanase et al. (2025) argued that it is because it matches the sea ice concentration threshold at which the free drift parameterisation is switched on or off (approximately 80%), while in some sea ice models at 1.5 m thickness the sea ice albedo parameterisation changes (Rousset et al., 2015). Looking only at the regions with a significant across-model correlation between thickness and polynya activity, we do see a suggestion of larger polynya activity for models with mean

thickness below 1.5 m (Appendix Figure A4, 1.5 m is the black horizontal line). This is however not a clean cut: in all regions, some models exhibit polynya activity even at thicknesses larger than 2 m while others still have no polynya below 1 m thickness. The thickness bias alone cannot explain the polynya activity bias.

Table 3. Across-model correlations between the 1979-2024 (*1982-2014) mean winter (DJFM) polynya area and the mean of the sea ice thickness “sivol” and of the air temperature “T2m” in summer (JJA), autumn (SON) and winter (DJF) for each region, when the polynya-area outlier AWI-CM-1-1-MR is excluded (plain text) and included (text in brackets); AWI-CM-1-1-MR is not available for T2m. Last rows: correlation between the sea ice thickness and the air temperature. ‘-’ indicates correlation non significant at 95%.

		1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
sivol vs area	JJA	- (-)	-0.66 (-0.66)	-0.91 (-)	-0.77 (-0.60)	- (-)	- (-0.58)	-0.58 (-0.65)	-0.71 (-)	- (-)
	SON	- (-)	- (-)	-0.66 (-)	- (-)	- (-)	- (-0.53)	- (-)	-0.68 (-)	- (-)
	DJF	- (-)	-0.69 (-0.69)	-0.77 (-)	- (-0.55)	- (-)	-0.76 (-0.73)	- (-0.60)	-0.67 (-)	- (-)
T2m* vs area	JJA	-	-	0.55	-	-	0.43	-	-	-
	SON	0.60	0.72	0.76	-	-	0.74	0.42	0.49	-
	DJF	-	0.76	0.83	-	0.46	0.75	-	-	-
sivol* vs T2m*	JJA	-0.55	-	-0.60	-	-	-0.79	-	-	-
	SON	-0.87	-0.93	-0.79	-0.91	-0.84	-0.79	-0.82	-0.72	-0.78
	DJF	-0.75	-0.87	-0.92	-0.70	-0.67	-0.89	-0.59	-	-0.50

To summarise, 1) there is no consistent across-model relationship between polynya area and sea ice thickness biases, not even in winter; and 2) there seems to be some clustering of the biases by sea ice model (Fig. 4). A clue that the polynya bias is not linked to one specific, physical process, but rather caused by a deep model-dependent bias, e.g. in its treatment of thermodynamics, would be an across-model link between polynya bias and the model’s equilibrium climate sensitivity (ECS). We do find a significant correlation of 0.76 between the ECS of Zelinka et al. (2020) and the pan-Arctic polynya area, with the caveat that there are only 9 models in common in our studies. That is, this link may not be robust, but it is encouraging. Let us therefore investigate more closely the impact of the sea ice component of the models. Two sea ice models are over-represented in our study: the Los Alamos sea ice model CICE (Hunke et al., 2017) and the Louvain-La-Neuve sea ice model LIM (Rousset et al., 2015), with the latter typically simulating larger polynya areas (see section 3). Many parameters can be relevant for polynyas, sometimes with opposite results: Chevallier et al. (2017) found for example that, in ocean reanalyses, having several thickness categories (as CICE does) resulted in a thinner ice, but so did having a low “P*” ice strength (as LIM does). One crucial point is that the two models have completely different thermodynamics treatments and surface processes included, with CICE having long included a scheme for melt ponds. As a result, CICE has a rather realistic sea ice growth / melt (Sun and Solomon, 2024) when compared to observations, whereas in LIM both growth and melt are faster and with a larger amplitude than expected (Uotila et al., 2017). This would explain LIM’s large and frequent polynyas.

The dynamics treatments also differ between LIM and CICE. Winter Arctic polynyas typically form in response to strong winds (Smith Jr and Barber, 2007), making an accurate representation of wind forcing on sea ice crucial for simulating polynya

openings. In sea ice models, it is the rheological formulation and roughness (or drag) that control how efficiently wind forcing is transmitted into sea ice motion. Ocean reanalyses with a low air-ice drag coefficient (such as LIM) have a slower sea ice drift (Chevallier et al., 2017), i.e. should be harder to open by wind divergence. There is also an ice-ocean drag coefficient; ocean reanalyses with a low ice-ocean drag coefficient (such as CICE) have a thinner sea ice (Chevallier et al., 2017), i.e. should be easier to open by wind divergence. Besides, the sea ice models use different rheologies: elastic-viscous-plastic for CICE and LIM3 but viscous-plastic for LIM2. This could impact sea ice deformation on short time scales and its sensitivity to wind stress. All these hypotheses however assume that models, as in reality, form Arctic polynyas in response to winds rather than thermodynamics forcing; we verify this assumption in the next subsection.

4.2 Winds and air temperature

There is no across-model correlation, even when excluding the polynya area outlier AWI-CM-1-MR, between the 45-year mean wind characteristics and the polynya area in any region. However, most models have strong significant correlations between the time series of the polynya area and that of the wind speed and/or the wind components (Appendix tables A6 to A8). Correlations with the wind speed, both its mean and extreme, are widespread in the Eurasian Arctic, from Svalbard to the East Siberian sea, whereby stronger wind speeds are associated with larger polynya areas. This is in agreement with e.g. Smith Jr and Barber (2007); Wong et al. (2026).

In the other regions, fewer models have correlations between the winter wind and polynya time series. In the Chukchi and Beaufort seas and in the North East Water polynya, the models that have such correlation tend to have a negative correlation with mean U and the 10th percentile of U, and positive with mean V and the 90th percentile of V 90th or negative with the 10th percentile of V (Appendix Tables A7 and A8), i.e. the stronger the northwestward winds, the larger the polynya area. The correlation with the northward winds is in agreement with real-world polynya observations; Lee et al. (2023) in particular showed that the few observed winter polynya events in NEW were entirely driven by northward (southerly) winds mechanically redistributing the thick ice. The correlation with the westward winds both for NEW and the Pacific side seems counterintuitive, but in observations Peralta-Ferriz and Woodgate (2017) found that enhanced westward winds lead to a stronger inflow of Pacific Water through Bering Strait, i.e. a larger supply of upper-ocean heat that can melt the sea ice in the Chukchi and Beaufort regions. Lavoie et al. (2022) found that some CMIP6 models do not represent any Pacific Water; this may explain why the correlation is not more widespread among our models. Similarly, westward winds would impact the models' (re)circulation of Atlantic Water north of Fram Strait, which may explain the occasional model's correlation with the NEW polynya area. Finally, as in the real world, the North Water Polynya is associated with southward winds, albeit only seen for the high horizontal resolution models (Appendix Tables A8). The other models most likely are too coarse to properly represent the Nares Strait geography or because the resolution of the atmosphere component directly impacts the representation of the most extreme winds (Huot et al., 2021), crucial for Greenland polynyas (Smith Jr and Barber, 2007). Moore (2021) explicitly states that a resolution of at least 9 km in the atmosphere is required; for comparison, the high resolution version of the CMIP6 models that participated in HighResMIP is of the order of 25 km (Haarsma et al., 2016).

In contrast, we find strong across-model correlations between air temperature and polynya area in all regions except the Laptev Sea and North East Water polynya (Table 3). Correlations with the same winter are to be expected, since e.g. Wong et al. (2026) found a similar relationships in observations in the Arctic, which they explained as warm air thinning the ice, preconditioning the region for polynya opening. This is consistent with the strong across-model negative correlation between air temperature and sea ice thickness in winter, and that between thickness and polynya area discussed in the previous section (Table 3). We however find hardly any across-model relationship in summer, even though summer thickness and polynya area were strongly associated, which further suggests that a generally thin summer ice is not a driver but a consequence of large polynya activity in a model.

We rather find large across-model relationships between autumn air temperature and both winter polynya area and autumn sea ice thickness throughout the Arctic (Table 3, middle), with mean autumn air temperature being the only one correlated with mean polynya area in Svalbard ($R=0.60$), the Beaufort Sea ($R=0.42$) and the North Water Polynya ($R=0.49$). In all regions, the absolute value of the correlation between autumn air temperature and autumn sea ice thickness exceeds 0.7. Interestingly, this is similar to what Zhou et al. (2022) found in observations for open water polynyas in the Southern Ocean: the subsequent opening of the polynya was detectable already in the autumn as a thin ice anomaly, which pre-conditioned the region for opening with the winds acting as final trigger.

To summarise, as in observations, we find significant relationships in each model between polynya area and wind speed or direction. The most important wind component depends on the location, but also on the horizontal resolution of the model in narrow locations such as the Canadian archipelago. We also find a large contribution of the model's overall temperature bias via its impact on sea ice thickness, with warmer models having thinner ice in autumn. Determining whether individual models are actively melting sea ice in autumn or simply delaying its growth is beyond the scope of this study. The outcome is the same: a sea ice in which it is easier to open polynyas in the subsequent winter. Across-model correlations between polynya area and sea ice thickness year-round, along with the autocorrelation of sea ice, perpetuate the polynya cycle, resulting in some models being more polynya-prone than others. We finish this manuscript with an analysis of the potential impacts of this difference in overall polynya behaviour on the ocean.

5 Do winter Arctic polynyas impact the modelled ocean?

One of the motivations for this study was the finding of Heuzé et al. (2023) that most CMIP6 models have no dense overflow representation, and that the few models that do have overflows have them only in Saint Anna Trough, in the Kara Sea, even though they seemed to also have polynyas in the Laptev Sea. They further found that few models have dense water at the bottom of the shelf break, below the expected polynya area, and that the properties of this dense water were inconsistently biased. They therefore suspected that something had gone wrong in the ocean - sea ice coupling, suggesting a disconnect between the representation of polynyas in models and the simulated ocean properties. We finish this study by briefly expanding on their analysis.

5.1 Polynya activity does not seem to impact the mean hydrographic biases

440 We contrast the three models with most polynya activity (red on Figs 7 and 8) to the three with least activity (blue). We limit our analysis to the Kara and Laptev seas, two regions that are regularly cited in observational and modelling studies as being highly productive in terms of sea ice formation and brine rejection, which impact dense water formation and deep ocean ventilation (e.g. Tamura and Ohshima, 2011; Itkin et al., 2015; Preußer et al., 2016). In the Kara Sea, the mean temperature at the surface of the shelf break is colder in the models with least activity: all three are below -1°C (blue on Fig. 7), whereas the surface of the most active models is between -0.8 to $+1.2^{\circ}\text{C}$. That they are not at freezing temperature, whether because of warm water upwelling or warm air, suggests that brine rejection will be hindered. This surface temperature difference is also consistent with our earlier finding of thinner ice and stronger seasonal cycle being associated with higher polynya activity, although whether the upper ocean is systematically warmer because sea ice melts in summer, or sea ice melts in summer because the ocean is warmer, we cannot tell from these data. It may also be a coincidence in this subsample of models, since the other depth levels, properties and regions show no consistent behaviour. In fact, considering the temperature on the shelf, the lines of the least (blue) and most active (red) models cross already at 40 m depth. And since even at the surface, the salinity biases are inconsistent between these models on the continental shelf, so is their density (Fig. 7, left). In particular, the models with most polynya activity do not have denser water at the bottom of the shelf than the models with least activity: of the six models, the densest is MIROC-ES2L, among the least active models (36.88 kg m^{-3} , Appendix Table A9). Therefore, it comes as no surprise that there is no consistency at the shelf break, in any of the properties, nor in the deep basin. Although it seems at first glance that the Atlantic Water core is deeper off-shelf in the least active models (blue in Fig. 7, right), we find that the core depth varies from 358 to 876 m in the most active models, and 667 to 917 m in the least active, i.e. their ranges overlap (Appendix Table A9).

CMIP6 models have 10 times smaller polynya areas in the Laptev Sea than in the Kara Sea, and they are more sporadic (see section 3). These differences appear to have no impact on the ocean profiles (Fig. 8). We find again that the only consistent clustering is in the temperature at the very surface of the continental shelf, with the least active models being colder than the most active ones (blue, below -1°C ; red, -0.8 to 0°C). It is worth noting again that in the most active models, these temperatures are above freezing, i.e. not compatible with sea ice production and brine rejection. The bias order is not maintained with depth and there is no consistent bias in salinity, so the density at the bottom of the shelf does not significantly differ between the three most polynya-active and three least active models; again, the densest model is in fact MIROC-ES2L, among the least active models (35.97 kg m^{-3} , Appendix Table A9). We find no consistency in the Atlantic Water core or bottom properties, either at the shelf break or in the deep basin (Fig. 8).

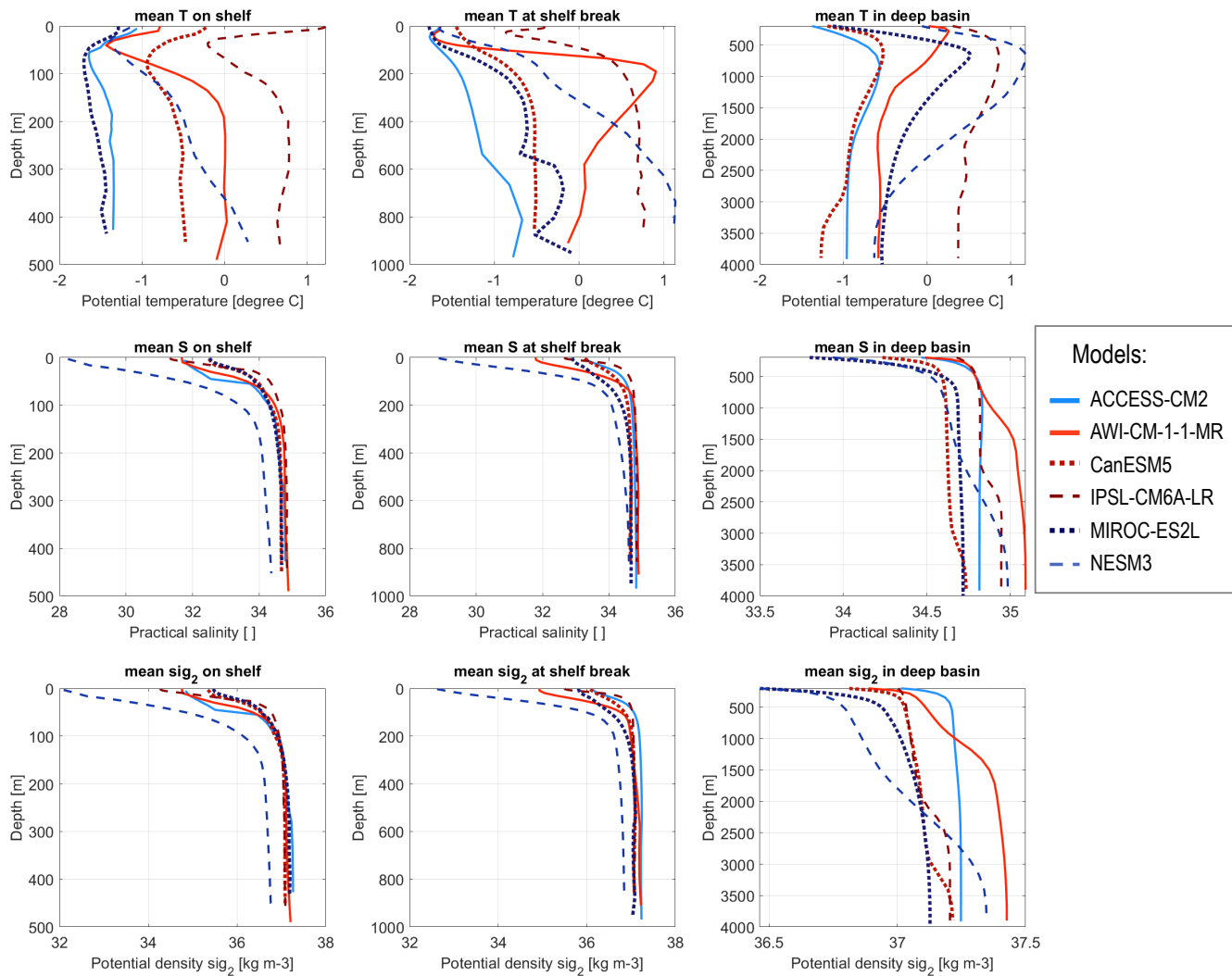


Figure 7. In the Kara Sea, for the three models with most (red) and least (blue) polynya activity, profiles of potential temperature (T, top), practical salinity (S, middle) and potential density anomaly σ_2 (sig_2 , bottom) on the continental shelf (left), at the shelf break (centre), and in the deep basin (right). For the values and identifying which model is which, see Appendix Table A9. Note that in the deep basin, the depth axis starts at 200 m depth.

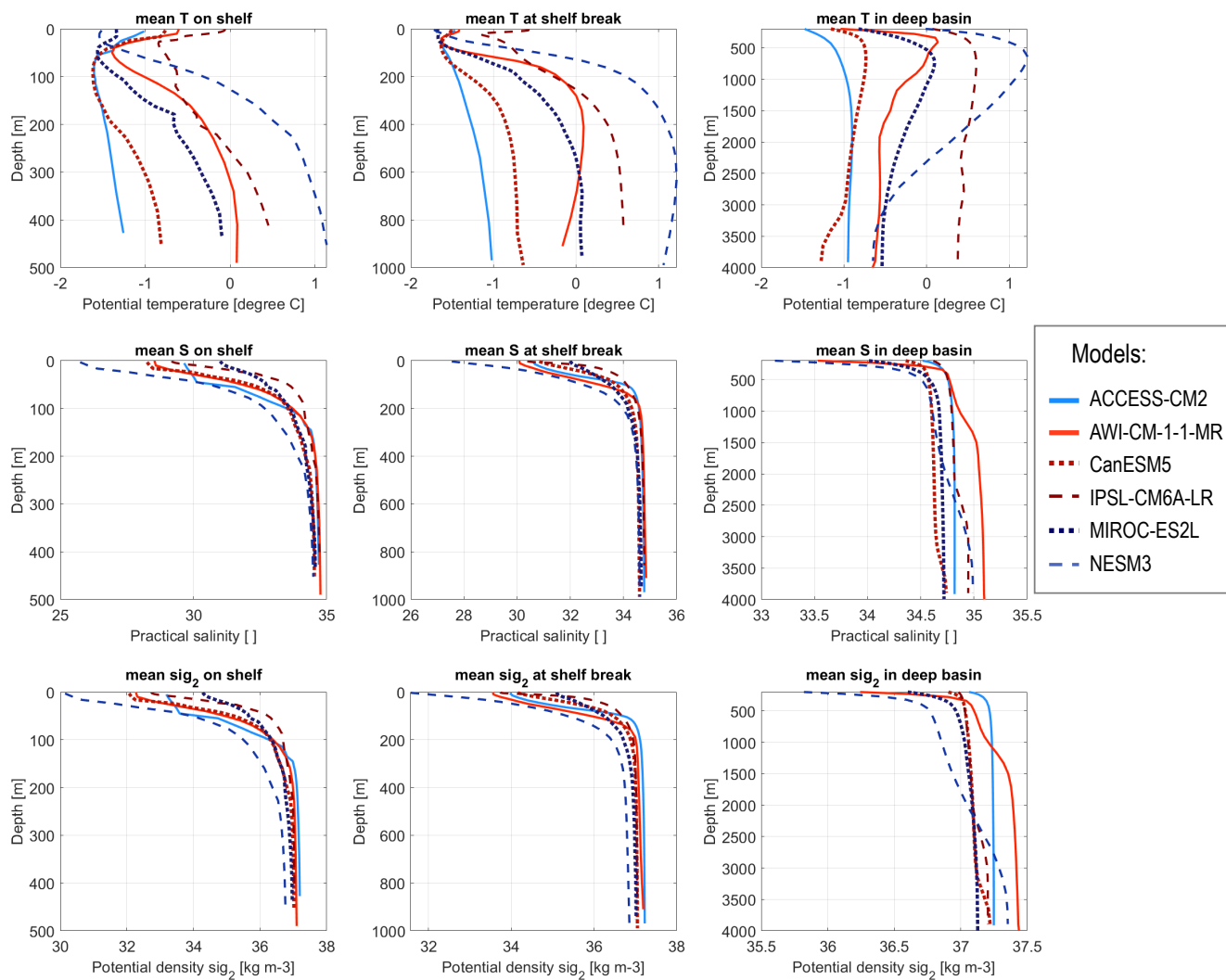


Figure 8. Same as Fig. 7 but for the Laptev Sea.

5.2 Polynya activity can have a short-lived impact on ocean properties in individual models

In particular in the Laptev Sea, where polynyas are simulated less frequently, it may be that taking the 45-year average masks
470 the effect of polynyas on the water column. We therefore correlate the time series of polynya area and ocean properties, considering both the yearly-mean properties and the different seasons to account for the time the dense water may need to travel from the shelf to the deep basin. In observations, this takes 3 to 6 months (Ivanov and Golovin, 2007; Luneva et al., 2020). In the two active models CanESM5 and IPSL-CM6A-LR, this takes up to one year, with maximum correlations found between polynya activity and properties at the bottom of the deep ocean in both Kara and Laptev in the subsequent autumn
475 (Table 4). In AWI-CM-1-1-MR, correlations are rather with the yearly value or that of the same winter; it is not impossible that in that very high resolution model, jets of dense water reach the deep basin shortly after their formation. The active models tend to agree on the sign of the correlations, with more polynya activity being associated with warmer and fresher deep basins. AWI-CM-1-1-MR and CanESM5 also agree that more polynya activity is associated with warmer and less-dense waters at the bottom of the shelf and at the bottom of the shelf break, in both seas. This is unexpected since polynya activity is supposed to
480 generate dense water (Morales Maqueda et al., 2004). Even when polynyas have been observed to result in warmer deep waters, it was in conjunction with high salinities (Vivier et al., 2024). Models do not systematically have a correlation between polynya activity and the bottom salinity there, but when they do, this correlation is negative, i.e. more polynya activity is surprisingly associated with fresher waters (Table 4, bottom S). This would be consistent with polynyas opening by melting the ice, i.e. sensible heat polynyas, rather opening dynamically in response to strong winds, i.e. latent heat polynyas (Morales Maqueda
485 et al., 2004). Note that the third of the most active polynya models, IPSL-CM6A-LR, has no such correlation between polynya activity and bottom properties (Table 4), even though it uses the same ocean and sea ice model as CanESM5, albeit a newer version of it. In the “real” Arctic, there are no purely sensible heat polynyas, but there are hybrid ones that switch modes while open from latent heat to sensible heat; after the switch, dense water production has been observed to stop (Hirano et al., 2016). Models simulating sensible heat polynyas would be consistent with the lack of connection between polynya activity and bottom
490 properties.

Finally, both most and least active models have negative correlations at the Kara Sea shelf break between the Atlantic Water core depth and polynya activity, i.e. the more polynya activity the shallower the core (Table 4, AW core depth). In the real Eurasian Arctic, the dense water cascade entrains Atlantic Water as it flows off the shelf there, which results in a cooling and deepening of the Atlantic Water core (e.g. Woodgate et al., 2001; Dmitrenko et al., 2014). Heuzé et al. (2023) found that
495 CMIP6 models systematically had too deep an Atlantic Water core, often sitting below the shelf break depth, and that this bias was the result of a biased Atlantic Water inflow through Fram Strait that was not corrected by ventilation in the Arctic, since there was virtually no ventilation. It may therefore be that polynya activity momentarily corrects that advected Atlantic Water bias. However, the lack of consistency between polynya activity and both temperature and salinity, most likely a result of how models open their polynya, suggests that polynyas cannot correct the other ocean biases.

Table 4. For the models with most (left) and least (right) polynya activity, in the Kara (top) and Laptev (bottom) seas, maximum correlation between the 1979-2024 winter polynya area time series and that of the ocean properties on the continental shelf, at the shelf break, and in the deep basin, in the Atlantic Water core and at the bottom cell. ‘-’ indicates that no significant correlation (at 95%) could be found. We show only the maximum correlation, in absolute value, and indicate whether this correlation corresponds to the yearly mean ocean property (Year), to that of the same winter as the polynya (DJF), or to the subsequent spring (MAM), summer (JJA) or autumn (SON).

			c) AWI-CM-1-1-MR	d) CanESM5	j) IPSL-CM6A-LR	k) MIROC-ES2L	q) NESM3	b) ACCESS-CM2
Shelf	bottom	T	0.74 (DJF)	0.48 (Year)	-	-0.36 (MAM)	-	-
		S	-	-0.49 (Year)	-	-	-	-
		σ_2	-0.43 (SON)	-0.59 (JJA)	-	-	-	-
Break	AW core	T	0.73 (SON)	0.45 (JJA)	-	-	-0.31 (Year)	-
		S	0.40 (Year)	-0.34 (SON)	-	-	-0.29 (Year)	-
		σ_2	-0.69 (SON)	-0.48 (JJA)	-	-	-	-
	bottom	depth	-0.56 (Year)	-0.30 (JJA)	-	-0.40 (Year)	-	-0.38 (Year)
		T	0.64 (Year)	0.47 (SON)	-	-	-0.31 (Year)	-
		S	-	-0.38 (Year)	-	-	-0.29 (Year)	-
σ_2	-0.49 (MAM)	-0.42 (DJF)	-	-	-	-		
Deep	AW core	T	0.59 (Year)	0.40 (SON)	-	-	-	-
		S	-	0.35 (Year)	-	-	-	-
		σ_2	-0.62 (Year)	-0.38 (SON)	-	-	-	-
	bottom	depth	-0.59 (Year)	-0.46 (SON)	-	-0.28 (Year)	-	-
		T	0.54 (DJF)	0.38 (Year)	0.43 (SON)	-	-	-
		S	-0.64 (Year)	-0.41 (SON)	0.41 (SON)	-	-	-
σ_2	-0.61 (DJF)	-0.39 (Year)	-0.43 (SON)	-	-	-		
Shelf	bottom	T	0.39 (DJF)	0.38 (Year)	-	-	-	-
		S	-0.34 (SON)	-	-	-	-	-
		σ_2	-0.36 (SON)	-	-	-	-	-
Break	AW core	T	0.37 (Year)	0.60 (SON)	-	-	-	-
		S	-	-0.28 (Year)	-	-	-	-
		σ_2	-0.37 (Year)	-0.60 (SON)	-	-	-	-
	bottom	depth	-	-0.48 (Year)	-	-	-	-
		T	0.55 (DJF)	0.57 (Year)	-	-	-	-
		S	-0.37 (DJF)	-	-	-	-	-
σ_2	-0.50 (DJF)	-0.57 (Year)	-	-	-	-		
Deep	AW core	T	-	0.58 (Year)	-	-	-	-
		S	0.28 (SON)	-0.35 (JJA)	-	-	-	-
		σ_2	-	-0.59 (Year)	-	-	-	-
	bottom	depth	-	-0.47 (Year)	-	-	-	-
		T	0.55 (Year)	-0.38 (SON)	-	-	-	-
		S	-	0.31 (Year)	0.44 (SON)	-	-	-
σ_2	-0.43 (DJF)	0.34 (Year)	-	-	-	-		

Although winter Arctic polynyas play a crucial role for all components of the Arctic climate and its ecosystem (Smith Jr and Barber, 2007), and that satellite observations have revealed a strong increase in their area since 1970s (Wong et al., 2026), to the best of our knowledge, this study was first to evaluate the representation of winter Arctic polynyas in any CMIP generation. We detected polynyas using their daily sea ice concentration since daily sea ice thickness is not widely available in CMIP6, and found that the majority of models overestimate the polynya area while they underestimate their frequencies. The exception was the very high resolution model AWI-CM-1-1-MR ($\sim 100 \text{ km}^2$ grid cells, Semmler et al., 2020) which was overestimating both, in all regions. The spatial distribution matches that of observational studies relying on sea ice concentration, but obviously misses some thin-ice polynyas that are better detected via sea ice thickness or surface fluxes (e.g. Tamura and Ohshima, 2011; Preußer et al., 2016). Unfortunately in CMIP6, only daily sea ice concentration was both widely available and reliable. The temporal resolution obviously mattered as well, with polynya areas computed from monthly sea ice concentration being lower than those from the daily, yet usually correlated with each other. That is, it would technically be possible to infer daily results from monthly output in order to increase the amount of models available; we however advise against this, as the relationship was region-dependent, and are rather looking forward to the planned more widespread release of daily output in CMIP7 (Fox-Kemper et al., 2025).

Despite the apparent link with both temporal and spatial resolution, we find that models tend to cluster by sea ice model. This is not that surprising, since polynya dynamics are strongly influenced by the sea ice physics and their rheologies. Besides, although most winter Arctic polynyas are supposed to be wind-driven (Morales Maqueda et al., 2004), we found no across-model relationship between wind bias and polynya bias. The two most common sea ice models, CICE and LIM, have different drag coefficients and thermodynamics schemes that make it harder to open polynyas in response to winds but easier with thermodynamics, respectively (Hunke et al., 2017; Uotila et al., 2017; Chevallier et al., 2017). We find that the LIM models are among those with the largest polynyas, but also the largest sea ice seasonal cycle. In fact, across models, we find that the larger polynya areas are associated with warmer autumns, leading to thinner ice / delayed ice growth in the large seasonally ice free regions, which are thus preconditioned for opening polynya in the winter. Thermodynamics do matter in observations, even for wind-driven polynyas, as a warmer atmosphere preconditions the region for polynya opening (Wong et al., 2026), especially in the autumn (Zhou et al., 2022). Our results however show a lack of across-model relationship between wind and polynya activity, i.e. thermodynamics does not only matter for preconditioning but is also the trigger. As CMIP6 models project an increase in wind speed in the Arctic and in the transfer of wind stress to the ice, notably in winter (Mulwijk et al., 2024), one would expect a continued increase in “real” polynya area in the future. Given the models’ lack of agreement on recent polynya trends (Fig. 4) and suspected incorrect opening mechanism, we prefer leaving future projections out of this study.

Finally, rather disappointingly, we find that polynya activity does not seem to affect the representation of the water column in CMIP6 models, in agreement with the preliminary findings of Heuzé et al. (2023). Even by limiting the water column analysis to the three most and three least polynya-active models, we found no consistent relationship between polynya activity and properties on the continental shelf, at the shelf break, or in the deep basin. In particular, even though real-world polynyas are

hotspots of sea ice production (Tamura and Ohshima, 2011) and therefore result in dense water at the bottom of the shelf (Vivier et al., 2024), our densest model was one of the least polynya-active. At the individual model level, we even found negative correlations between polynya activity and bottom density, i.e. the larger the polynya, the less-dense the shelf water of that year, further suggesting that polynyas in CMIP6 models are the result of sea ice melt rather than wind-driven dynamics opening. That is, CMIP6 models seem to be opening sensible heat polynyas in the Arctic, where all observed polynyas are primarily wind-driven latent-heat or hybrid (Wong et al., 2026). Another surprising result was widespread negative correlations between polynya activity in the Eurasian Arctic and Atlantic Water core depth, i.e. the larger the polynya, the shallower the core, i.e. the opposite of the real-world (e.g. Dmitrenko et al., 2014). Our optimistic interpretation is that polynyas momentarily reduce the deep bias detected by Heuzé et al. (2023), but due to polynya's intermittent nature and their apparent wrong formation process, polynyas cannot fix the overall model bias, caused by advected errors through Fram Strait.

How to fix these biases? A higher resolution is the obvious answer, both in the ocean to represent the complex coastline geometry and in the atmosphere for the strong winds (Huot et al., 2021; Moore, 2021), but it is unrealistic. In fact, some modelling centres have even decided to not increase their model resolution beyond $\sim 1^\circ$, as the marginal improvement in the global climate representation cannot justify the increased climate footprint of the simulation (OMIP protocol for CMIP7, to be submitted soon). Besides, the highest resolution model in our study was a clear outlier. One option that seems promising is the better representation of land-fast ice (Itkin et al., 2015), which is crucial for polynya development in the real Arctic (Morales Maqueda et al., 2004) but usually not captured in models. It is also possible that the biases will fix themselves: as polynya area increases (Wong et al., 2026) and the shelf break keeps warming (Ingvaldsen et al., 2021), real-world winter Arctic polynyas may soon become open-ocean polynyas that open through thermodynamics, which models already do, making CMIP polynya representation and projection more trustworthy.

Code and data availability. The scripts specific to this manuscript will be published on Zenodo closer to manuscript acceptance. The machine-learning models and scripts as well as the reference dataset, all from Heuzé and Wong (2025), are freely available on Zenodo at <https://doi.org/10.5281/zenodo.17540456> (last access: 9 February 2026). The observed sea ice concentration data from the National Snow and Ice Data Centre (dataset DOI <https://doi.org/10.5067/MPYG15WAA4WX>, DiGirolamo et al., 2022) are freely available at <https://nsidc.org/data/nsidc-0051/versions/2> (last access: 9 February 2026). The CMIP6 dataset DOIs and corresponding citations are listed in Appendix Table A1 and A2.

Table A1. For the models with daily sea ice concentration output: Model name, Ensemble members used in this study and their corresponding Shared Socio-economic Pathway (SSP) run, and the Dataset DOI of the CMIP6 runs along with its reference.

Model	Ensemble members	SSP used	Dataset DOI	Reference
ACCESS-CM2	r1:10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.2281	Dix et al. (2019)
AWI-CM-1-1-MR	r1:5i1p1f1	3-7.0	10.22033/ESGF/CMIP6.359	Semmler et al. (2018)
CanESM5	r1,4,6:7,10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1303	Swart et al. (2019c)
EC-Earth3-AerChem	r1i1p4f1	2-4.5	10.22033/ESGF/CMIP6.639	EC-Earth Consortium (EC-Earth) (2020a)
	r1,4i1p1f1	3-7.0		
EC-Earth3-CC	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.640	EC-Earth Consortium (EC-Earth) (2020b)
EC-Earth3-Veg-LR	r1:3i1p1f1	5-8.5	10.22033/ESGF/CMIP6.643	EC-Earth Consortium (EC-Earth) (2020c)
EC-Earth3-Veg	r13i1p1f1	2-4.5	10.22033/ESGF/CMIP6.642	EC-Earth Consortium (EC-Earth) (2019b)
IPSL-CM5A2-INCA	r1i1p1f1	N/A	10.22033/ESGF/CMIP6.13642	Boucher et al. (2020a)
IPSL-CM6A-LR	r1:6,10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1534	Boucher et al. (2018)
	r7:9i1p1f1	3-7.0		
MIROC-ES2L	r1i1p1f2	2-4.5	10.22033/ESGF/CMIP6.902	Hajima et al. (2019)
MIROC6	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.881	Tatebe and Watanabe (2018)
MPI-ESM-1-2-HAM	r1:3i1p1f1	3-7.0	10.22033/ESGF/CMIP6.1622	Neubauer et al. (2019)
MPI-ESM1-2-HR	r1:2i1p1f1	2-4.5	10.22033/ESGF/CMIP6.741	Jungclaus et al. (2019)
	r3:10i1p1f1	3-7.0		
MPI-ESM1-2-LR	r1:4,10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.742	Wieners et al. (2019)
MRI-ESM2-0	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.621	Yukimoto et al. (2019b)
NESM3	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.2021	Cao and Wang (2019)
NorESM2-LM	r1:3i1p1f1; r1i1p4f1	2-4.5	10.22033/ESGF/CMIP6.502	Seland et al. (2019)
NorESM2-MM	r1:2i1p1f1	2-4.5	10.22033/ESGF/CMIP6.506	Bentsen et al. (2019)

Table A2. Same as table A1 but for the models with monthly sea ice concentration output.

Model	Ensemble members	SSP used	Dataset DOI	Reference
ACCESS-CM2	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.2281	Dix et al. (2019)
ACCESS-ESM1-5	r1,10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.2288	Ziehn et al. (2019)
CAMS-CSM1-0	r1:2i1p1f1	3-7.0	10.22033/ESGF/CMIP6.1399	Rong (2019)
CAS-ESM2-0	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1944	Chai (2020)
CESM2-WACCM	r1:3i1p1f1	2-4.5	10.22033/ESGF/CMIP6.10024	Danabasoglu (2019a)
CESM2	r4,10:11i1p1f1	2-4.5	10.22033/ESGF/CMIP6.2185	Danabasoglu (2019b)
CIESM	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1352	Huang (2019)
CMCC-CM2-SR5	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1362	Lovato and Peano (2020)
CMCC-ESM2	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.13164	Lovato et al. (2021)
CNRM-CM6-1-HR	r1i1p1f2	2-4.5	10.22033/ESGF/CMIP6.1385	Voltaire (2019)
CNRM-CM6-1	r1:6i1p1f2	2-4.5	10.22033/ESGF/CMIP6.1375	Voltaire (2018)
CNRM-ESM2-1	r1i1p1f2	2-4.5	10.22033/ESGF/CMIP6.1391	Seferian (2018)
CanESM5-CanOE	r1i1p2f1	2-4.5	10.22033/ESGF/CMIP6.10205	Swart et al. (2019b)
CanESM5	r1:10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1303	Swart et al. (2019c)
EC-Earth3-AerChem	r1i1p1f1	3-7.0	10.22033/ESGF/CMIP6.639	EC-Earth Consortium (EC-Earth) (2020a)
EC-Earth3-CC	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.640	EC-Earth Consortium (EC-Earth) (2020b)
EC-Earth3-Veg-LR	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.643	EC-Earth Consortium (EC-Earth) (2020c)
EC-Earth3-Veg	r4i1p1f1	5-8.5	10.22033/ESGF/CMIP6.642	EC-Earth Consortium (EC-Earth) (2019b)
EC-Earth3	r1,10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.181	EC-Earth Consortium (EC-Earth) (2019a)
FIO-ESM-2-0	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.9047	Song et al. (2019)
HadGEM3-GC31-LL	r1:5i1p1f3	2-4.5	10.22033/ESGF/CMIP6.419	Ridley et al. (2018)
HadGEM3-GC31-MM	r1:4i1p1f3	5-8.5	10.22033/ESGF/CMIP6.420	Ridley et al. (2019)
IPSL-CM6A-LR	r1:6i1p1f1	2-4.5	10.22033/ESGF/CMIP6.1534	Boucher et al. (2018)
	r7:9i1p1f1	3-7.0		
MIROC-ES2L	r1,5,8:10i1p1f2	2-4.5	10.22033/ESGF/CMIP6.902	Hajima et al. (2019)
	r6i1p1f2	5-8.5		
MIROC6	r1:3i1p1f1	2-4.5	10.22033/ESGF/CMIP6.881	Tatebe and Watanabe (2018)
	r5,7i1p1f1	5-8.5		
MPI-ESM-1-2-HAM	r1,3i1p1f1	3-7.0	10.22033/ESGF/CMIP6.1622	Neubauer et al. (2019)
MPI-ESM1-2-HR	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.741	Jungclaus et al. (2019)
	r10i1p1f1	3-7.0		
MPI-ESM1-2-LR	r1,10i1p1f1	2-4.5	10.22033/ESGF/CMIP6.742	Wieners et al. (2019)
MRI-ESM2-0	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.621	Yukimoto et al. (2019b)
	r3,5i1p1f1	3-7.0		
NESM3	r1:2i1p1f1	2-4.5	10.22033/ESGF/CMIP6.2021	Cao and Wang (2019)
NorCPM1	r10i1p1f1	N/A	10.22033/ESGF/CMIP6.10843	Bethke et al. (2019)
NorESM2-LM	r1:3i1p1f1, r1i1p4f1	2-4.5	10.22033/ESGF/CMIP6.502	Seland et al. (2019)
NorESM2-MM	r1:2i1p1f1	2-4.5	10.22033/ESGF/CMIP6.506	Bentsen et al. (2019)
TaiESM1	r1i1p1f1	2-4.5	10.22033/ESGF/CMIP6.9684	Lee and Liang (2019)
UKESM1-0-LL	r1:4,8:10i1p1f2	2-4.5	10.22033/ESGF/CMIP6.1569	Tang et al. (2019)

Table A3. Coordinates of the nine regions used in this study, as per Wong et al. (2026). Also shown on Fig. 3a.

Region	Longitude	Latitude
1. Svalbard	2°-34°E	76°-81°N
2. Franz Josef Land	35°-73°E	78.5°-83.5°N
3. Kara Sea	49°-88°E	68°-78°N
4. Laptev Sea	89°-136°E	70°-82°N
5. East Siberian Sea	137°E-176°W	68°-78°N
6. Chukchi Sea	153°-175°W	65°-73°N
7. Beaufort Sea	108°-150°W	67°-75°N
8. North Water polynya (NOW)	57°-90°W	74°-83°N
9. North East Water polynya (NEW)	3°-40°W	79°-85°N

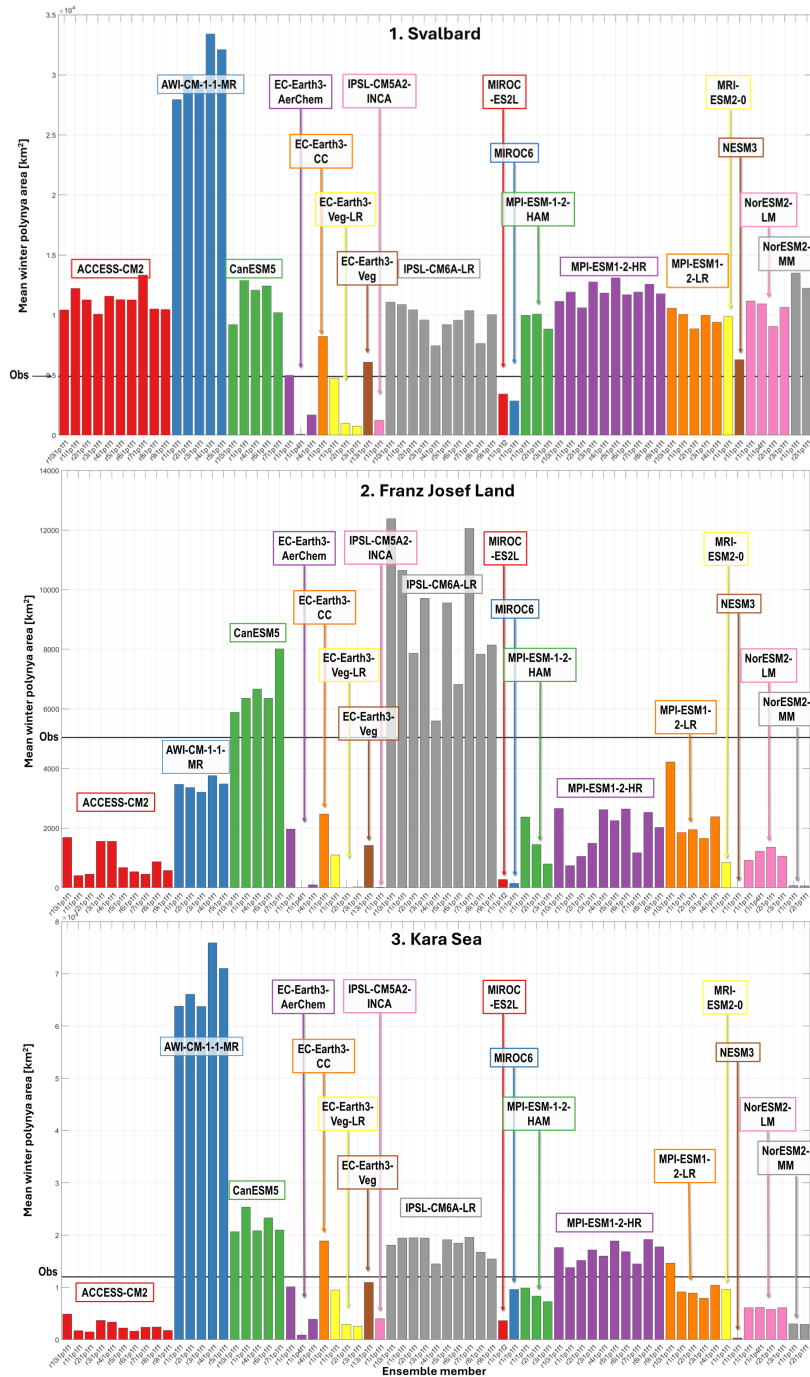


Figure A1. Same as Fig. 2 but for the nine regions analysed in this paper: 1. Svalbard, 2. Franz Josef Land, 3. the Kara Sea, 4. the Laptev Sea, 5. the East Siberian Sea, 6. the Chukchi Sea, 7. the Beaufort Sea, 8. the North Water Polynya and 9. the North East Water Polynya. Regions shown on Fig. 3a. and defined in Appendix Table A3.

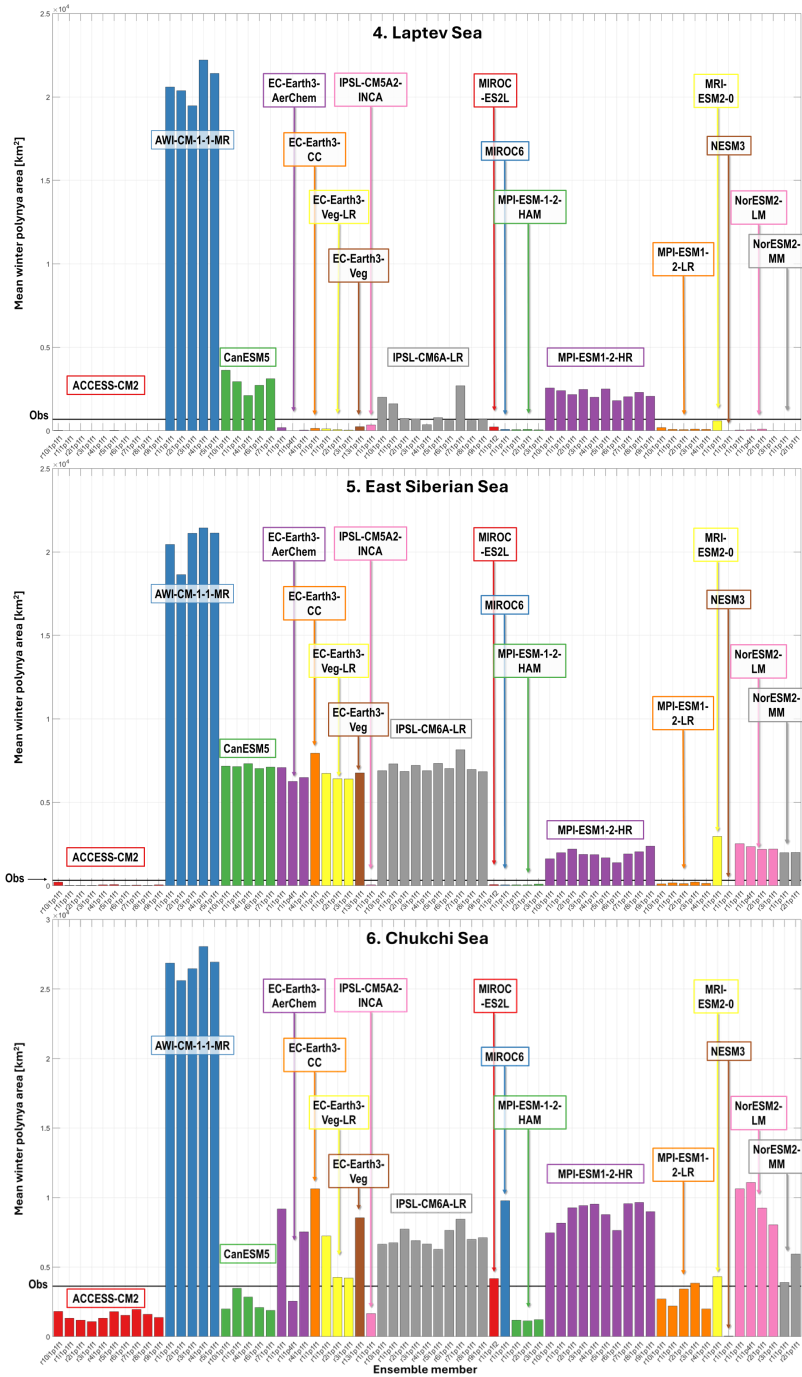


Figure A1. continued.

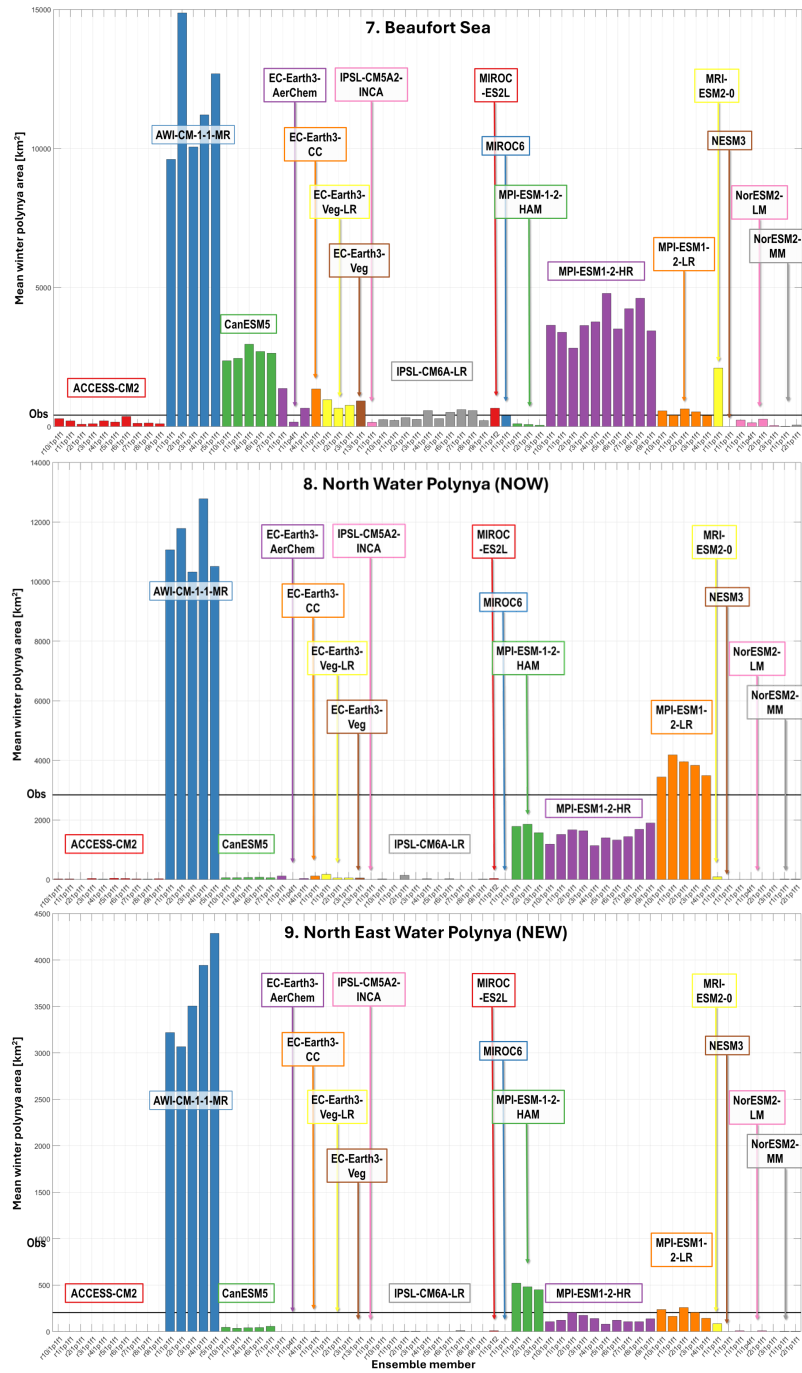


Figure A1. continued.

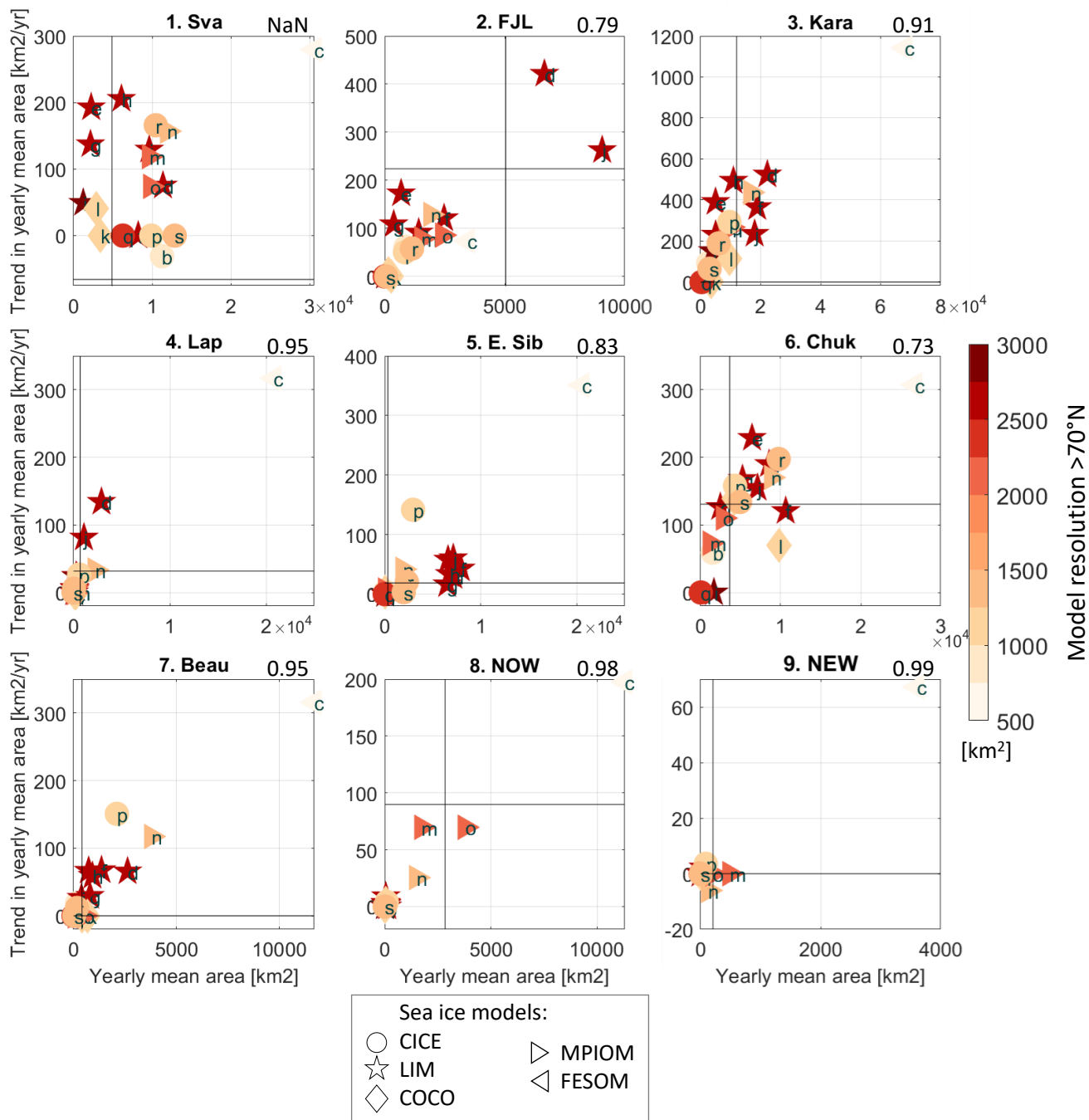


Figure A2. Same as Fig. 4 but including AWI-CM-1-1-MR.

Table A4. Values shown on Fig. 4. Top rows: 1979-2024 mean of the yearly winter (Dec-Mar) mean polynya area, in km². Bottom rows: 1979-2024 trends in yearly winter mean polynya area, in km²/year, at 95% significance; ‘NaN’ if insignificant.

	1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
a) Obs.	4889	5040	11997	681	323	3627	410	2843	204
	-66	224	NaN	32	19	131	NaN	90	NaN
b) ACCESS-CM2	11232	875	2537	8	51	1496	175	24	0
	-31	62	93	1	4	59	19	0	0
c) AWI-CM-1-1-MR	30567	3452	68106	20810	20552	26776	11685	11293	3602
	280	74	1142	317	351	307	316	198	67
d) CanESM5	11351	6653	22224	2898	7148	2457	2624	66	44
	75	421	524	134	30	126	65	3	2
e) EC-Earth3-AerChem	2235	685	4966	78	6599	6420	731	52	1
	193	173	390	7	58	229	66	9	NaN
f) EC-Earth3-CC	8216	2473	18879	140	7934	10630	1355	119	3
	NaN	121	365	5	43	120	67	NaN	NaN
g) EC-Earth3-Veg-LR	2148	375	4982	69	6507	5232	798	95	0
	137	108	231	NaN	16	168	29	2	0
h) EC-Earth3-Veg	6078	1417	10924	247	6757	8545	918	54	3
	205	91	494	24	34	190	59	2	NaN
i) IPSL-CM5A2-INCA	1231	13	4000	339	44	1658	157	7	0
	49	1	151	23	NaN	NaN	NaN	NaN	NaN
j) IPSL-CM6A-LR	9616	9061	18016	1081	7141	7110	382	25	3
	129	263	235	81	60	155	26	3	NaN
k) MIROC-ES2L	3414	270	3612	223	68	4168	659	40	10
	NaN	NaN	NaN	NaN	3	139	NaN	3	NaN
l) MIROC6	2844	142	9589	80	52	9761	397	0	1
	41	6	114	NaN	3	70	NaN	NaN	NaN
m) MPI-ESM-1-2-HAM	9621	1539	8475	64	65	1170	77	1737	484
	119	79	256	NaN	7	73	NaN	70	NaN
n) MPI-ESM1-2-HR	11920	1915	16678	2229	1884	8842	3785	1491	130
	157	129	437	33	42	170	117	25	-6
o) MPI-ESM1-2-LR	9767	2408	10177	97	148	2827	505	3781	202
	74	86	267	10	6	110	NaN	70	NaN
p) MRI-ESM2-0	9875	845	9586	584	2955	4306	2104	98	85
	NaN	51	293	27	142	158	151	4	4
q) NESM3	6266	6	305	0	0	18	0	3	0
	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
r) NorESM2-LM	10447	1138	6009	47	2303	9747	168	1	5
	166	58	188	6	21	198	12	NaN	0
s) NorESM2-MM	12853	64	2967	4	1979	4914	32	13	4
	NaN	NaN	65	NaN	3	134	NaN	NaN	NaN

Table A5. For each model, in each region, correlation (95% significant) between the time series of the winter (DJFM) mean polynya area and that of the previous summer (JJA), previous autumn (SON), and same winter (DJFM) mean sea ice thickness. ‘-’ indicates non significant correlations.

Model	Season	1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
b) ACCESS-CM2	JJA	-	-0.30	-	-	-0.40	-0.45	-0.47	-0.26	-0.33
	SON	0.48	-0.32	-0.33	-	-0.44	-0.51	-0.36	-	-
	DJFM	0.42	-0.52	-0.49	-	-0.43	-0.74	-0.40	-	-
c) AWI-CM-1-1-MR	JJA	-0.54	-0.44	-0.54	-	-0.25	-0.34	-0.43	-0.42	-0.56
	SON	-0.57	-0.39	-0.71	-0.40	-0.43	-0.49	-0.42	-0.35	-0.49
	DJFM	-0.66	-0.48	-0.78	-0.31	-0.58	-0.68	-0.72	-0.56	-0.66
d) CanESM5	JJA	-	-0.69	-0.53	-0.39	-	-0.29	-0.26	-	-0.40
	SON	-	-0.66	-0.62	-0.53	-	-0.34	-0.34	-	-0.51
	DJFM	-	-0.87	-0.74	-0.71	-0.41	-0.66	-0.50	-0.32	-0.49
e) EC-Earth3-AerChem	JJA	-0.70	-0.68	-0.77	-0.43	-0.63	-0.61	-0.63	-0.39	-
	SON	-0.75	-0.62	-0.69	-0.41	-0.66	-0.62	-0.58	-0.49	-
	DJFM	-0.84	-0.71	-0.91	-0.44	-0.66	-0.79	-0.60	-0.57	-0.29
f) EC-Earth3-CC	JJA	-0.25	-0.45	-0.57	-0.34	-0.54	-0.33	-0.48	-	-0.28
	SON	-0.38	-0.33	-0.54	-	-0.53	-0.32	-0.51	-	-0.38
	DJFM	-0.43	-0.70	-0.87	-0.32	-0.56	-0.65	-0.55	-0.33	-0.35
g) EC-Earth3-Veg-LR	JJA	-0.62	-0.76	-0.56	-0.38	-0.38	-0.32	-0.39	-	-0.29
	SON	-0.72	-0.74	-0.49	-	-0.31	-0.40	-0.34	-	-0.35
	DJFM	-0.81	-0.84	-0.81	-0.30	-0.26	-0.81	-0.46	-0.26	-0.28
h) EC-Earth3-Veg	JJA	-0.76	-0.63	-0.77	-	-0.41	-0.63	-0.48	-0.27	-
	SON	-0.79	-0.52	-0.67	-	-0.40	-0.55	-0.48	-0.38	-
	DJFM	-0.85	-0.65	-0.85	-0.38	-0.41	-0.83	-0.58	-0.40	-
j) IPSL-CM6A-LR	JJA	-	-0.75	-	-0.41	-0.46	-0.53	-	-	-
	SON	-	-0.58	-0.35	-0.34	-0.37	-0.42	-0.30	-0.27	-
	DJFM	-	-0.93	-0.49	-0.65	-0.45	-0.67	-0.41	-0.44	-
l) MIROC6	JJA	-0.66	-0.30	-0.40	-0.42	-0.32	-0.54	-	-	-
	SON	-0.50	-0.36	-0.41	-0.35	-0.31	-0.63	-0.25	-	-
	DJFM	-0.72	-0.50	-0.72	-0.43	-0.35	-0.83	-0.44	-	-
m) MPI-ESM-1-2-HAM	JJA	-0.42	-0.68	-0.45	0.29	-	-0.35	-0.43	-0.43	-
	SON	-0.45	-0.70	-0.57	0.26	-	-0.41	-0.38	-0.58	-
	DJFM	-0.58	-0.75	-0.88	-	-	-0.59	-0.37	-0.69	-0.27
n) MPI-ESM1-2-HR	JJA	-0.48	-0.49	-0.59	-0.26	-0.27	-0.58	-0.49	-0.38	-
	SON	-0.34	-0.49	-0.67	-	-0.40	-0.42	-0.43	-0.44	-
	DJFM	-0.58	-0.59	-0.85	-0.45	-0.44	-0.65	-0.51	-0.36	-
o) MPI-ESM1-2-LR	JJA	-	-0.69	-0.41	-	-0.30	-0.49	-0.32	-0.35	-
	SON	-	-0.59	-0.53	-	-	-0.60	-0.30	-0.37	-
	DJFM	-0.48	-0.66	-0.86	-	-0.29	-0.59	-0.39	-0.56	-
p) MRI-ESM2-0	JJA	0.27	-0.55	-0.46	-0.53	-0.53	-0.51	-0.69	-0.54	-0.32
	SON	-	-0.51	-0.46	-0.47	-0.52	-0.53	-0.66	-0.53	-0.40
	DJFM	-	-0.71	-0.83	-0.65	-0.65	-0.80	-0.63	-0.68	-0.43
s) NorESM2-MM	JJA	-	-0.49	-0.32	-	-0.31	-0.29	-0.36	-	-
	SON	0.34	-0.34	-0.43	-	-0.32	-0.44	-0.29	-	-
	DJFM	0.26	-0.46	-0.71	-0.28	-0.33	-0.67	-0.32	-	-

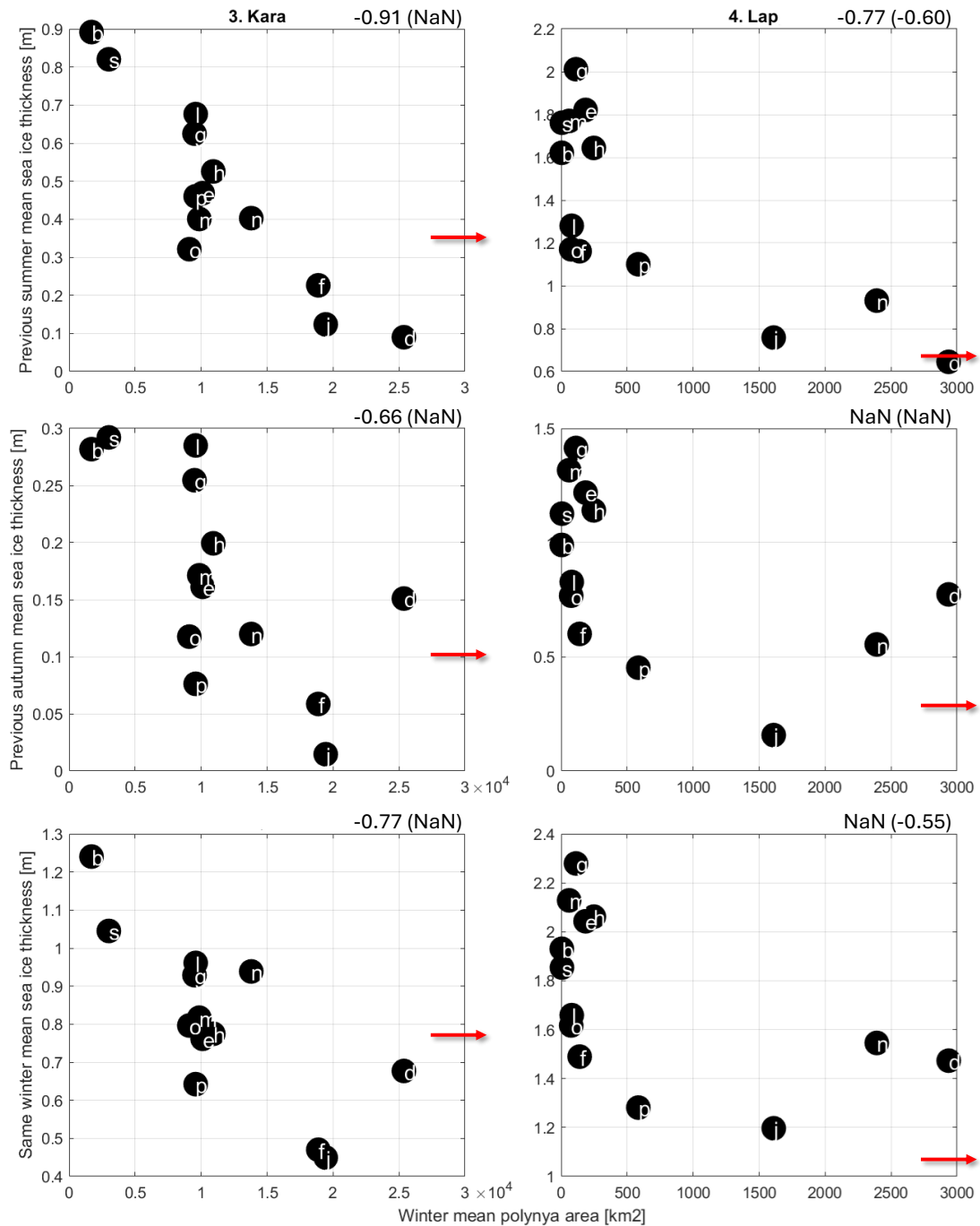


Figure A3. In the Kara (left) and Laptev (right) sea regions, across-model relationship between the 1979-2024 mean of the winter mean polynya area (x-axis) and the 1979-2024 mean sea ice thickness (y-axis) in summer (top), autumn (centre), and winter (bottom). Models referred to with same letters as on Fig. 3. Numbers in the top right corner of each panel is the correlation between the two (95% significance), with that in brackets when the outlier AWI-CM-1-1-MR is included, as per table 3. Red arrow indicates AWI-CM-1-1-MR sea ice thickness value and point in the direction of its polynya area; see area value on previous figures or in e.g. Appendix Table A4.

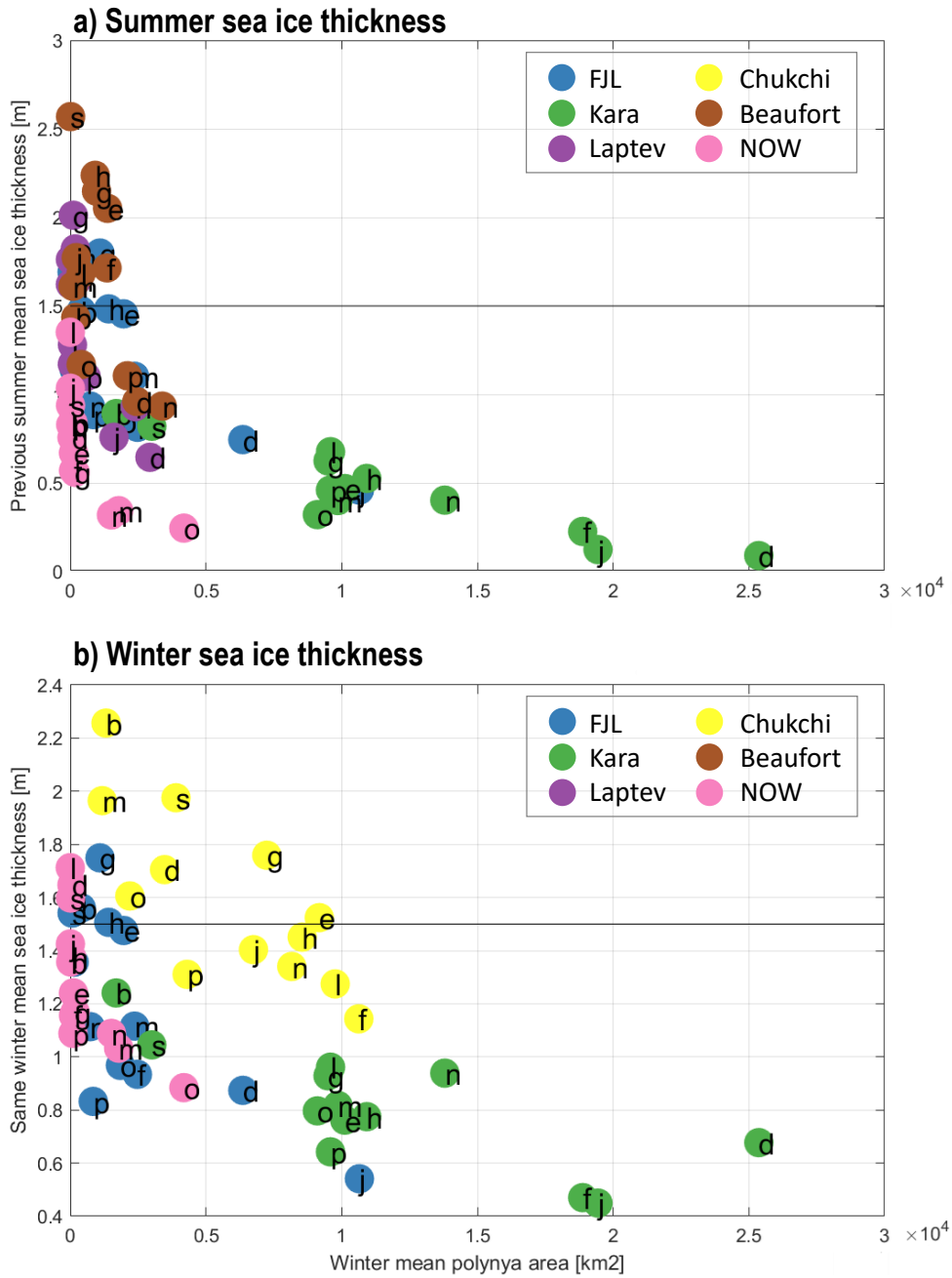


Figure A4. For all regions (colours), across-model relationship between the 1979-2024 mean of the winter mean polynya area (x-axis) and the 1979-2024 mean sea ice thickness (y-axis) in summer (top), and winter (bottom), shown only if there is a significant relationship between the two in the region. See also table 3. Models referred to with same letters as on Fig. 3. Black horizontal line is the 1.5 m sea thickness threshold discussed in the main text.

Table A6. For each model, in each region, correlation (95% significant) between the time series of the winter (DJFM) mean polynya area and that of the same winter mean and 90th percentile wind speed. ‘-’ indicates non significant correlations.

Model	Speed	1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
b) ACCESS-CM2	mean	-0.39	0.54	0.30	-	0.30	-	-	-	-
	90th	-0.33	0.52	-	-	0.29	-	-	-	-
c) AWI-CM-1-1-MR	mean	0.68	0.37	0.52	0.33	0.63	-	0.38	-	0.35
	90th	0.64	0.34	0.58	0.31	0.67	-	0.57	-	0.44
d) CanESM5	mean	-	0.72	0.59	0.46	0.35	0.34	-	-	0.39
	90th	-	0.72	0.62	0.53	0.32	0.27	0.29	-	0.36
e) EC-Earth3-AerChem	mean	0.71	0.58	0.69	-	-	0.32	-	-	-
	90th	0.70	0.74	0.72	-	-	0.30	0.33	-	-
f) EC-Earth3-CC	mean	-	0.70	0.62	0.34	0.27	0.66	-	-	-
	90th	-	0.74	0.67	0.40	0.32	0.65	-	-	-
g) EC-Earth3-Veg-LR	mean	0.60	0.51	0.39	-	0.37	-	-	-	0.25
	90th	0.63	0.57	0.47	-	0.40	-	-	-	0.28
h) EC-Earth3-Veg	mean	0.72	0.55	0.33	0.27	-	-	0.45	-	-
	90th	0.75	0.64	0.43	0.25	-	-	0.43	-	-
j) IPSL-CM6A-LR	mean	-	0.84	0.46	-	-	0.48	-	-	-
	90th	-	0.88	0.44	0.32	0.32	0.49	-	-	-
k) MIROC-ES2L	mean	-	-	-	-	-	-	-	-	-
	90th	-	-	-	-	-	-	-	-	-
l) MIROC6	mean	0.36	0.42	0.75	0.62	-	0.55	-	-	-
	90th	0.30	0.47	0.72	0.60	-	0.60	-	-	-
m) MPI-ESM-1-2-HAM	mean	0.34	0.66	0.71	0.35	0.31	-	0.26	-	-
	90th	-	0.68	0.71	0.36	0.40	-	-	-	-
n) MPI-ESM1-2-HR	mean	0.45	0.25	0.26	0.48	0.30	0.41	-	-	-
	90th	0.41	0.25	0.40	0.50	0.43	0.40	-	-	-
o) MPI-ESM1-2-LR	mean	0.45	0.46	0.63	-	0.37	0.28	-	-	0.37
	90th	0.42	0.55	0.65	0.29	0.35	0.26	-	0.29	0.34
p) MRI-ESM2-0	mean	-	0.73	0.71	-	-	-	-	-	-
	90th	-	0.77	0.70	-	0.30	-	-	-	-
q) NESM3	mean	0.44	-	-	-	-	-	-	-	-
	90th	0.49	-	-	-	-	-	-	-	-
r) NorESM2-LM	mean	-	-	0.46	-	-	-	-	-	-
	90th	-	-	0.36	-	-	-	-	-	-
s) NorESM2-MM	mean	-	-	0.41	-	0.26	0.40	-	-	-
	90th	-	0.31	0.36	-	-	0.36	-	-	-

Table A7. For each model, in each region, correlation (95% significant) between the time series of the winter (DJFM) mean polynya area and that of the same winter mean, 10th and 90th percentiles zonal wind (i.e extreme west- and eastward, respectively). ‘-’ indicates non significant correlations.

Model	U	1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
b) ACCESS-CM2	mean	-	-	-	-	-	-	-	-	-
	10th	-	-	-	-	-0.27	-	-	-	-
	90th	-	0.27	-	0.31	-	-	-	-	-
c) AWI-CM-1-1-MR	mean	-0.25	-	-	-	-	-	-	0.37	-0.33
	10th	-0.29	-	-	-	-	-	-0.34	0.35	-0.47
	90th	-	-	-	-	-	-	-	-0.32	-
d) CanESM5	mean	-	-	-	-	-	-	-0.28	-	0.27
	10th	-	-	-	-	-	-	-0.29	0.31	-
	90th	-	-	0.28	-	0.36	0.41	-	-	-
e) EC-Earth3-AerChem	mean	-0.45	-0.29	0.40	-	-	-	-	-	-
	10th	-0.54	-0.44	-	-	-	-	-0.30	-	-
	90th	-	-	-	-	-	0.44	-	-	-
f) EC-Earth3-CC	mean	-	-0.32	0.55	-	-0.27	-0.56	-	-	-
	10th	-	-0.39	0.47	-	-0.35	-0.59	-	-	-
	90th	-	0.30	-	-	-	-	-	-	-0.33
g) EC-Earth3-Veg-LR	mean	-0.33	-	-	-	-	-	-	-	-
	10th	-0.37	-0.31	-	-	-0.26	-	-	-	-
	90th	-	-	-	0.33	-	-	-	-	-
h) EC-Earth3-Veg	mean	-	-0.26	-	-	-	-	-0.26	0.30	-0.29
	10th	-	-0.41	-	-	-	-	-0.26	-	-0.32
	90th	-	-	-	0.43	-	0.34	-	-	-
j) IPSL-CM6A-LR	mean	-	-	0.41	-	-0.36	-0.32	-0.32	-	-
	10th	-	-0.31	0.36	-	-0.37	-0.41	-0.37	0.28	-
	90th	-	-	-	-	-0.26	-	-	-	-
k) MIROC-ES2L	mean	-	-	-	-	-	-	-	-	-
	10th	-	-	-	-	-	-	-	-	-
	90th	-	-	-	-	-	-	-	-	-
l) MIROC6	mean	-	-	-	0.38	-0.40	-0.56	-	-	-0.38
	10th	-	-	-	0.34	-0.35	-0.61	-	-	-0.42
	90th	-	-	-	0.35	-	-0.32	-	-	-
m) MPI-ESM-1-2-HAM	mean	-	-0.38	0.26	-	-	-	-	-	-0.48
	10th	-	-0.44	-	-	-	-	-	-	-0.48
	90th	0.46	-	0.28	0.42	-	-	-	-0.26	0.54
n) MPI-ESM1-2-HR	mean	-	-	-	-	-0.48	-0.39	-	-	-0.30
	10th	-	-	-	-	-0.51	-0.38	-	-	-0.29
	90th	-	0.31	-	-	-0.27	-	-	-0.40	0.30
o) MPI-ESM1-2-LR	mean	-	-0.35	-	-	-0.34	-	-	-	-
	10th	-	-0.43	-	-	-0.39	-	-	-	-
	90th	0.50	-	0.34	0.26	-0.32	-	-	-	-
p) MRI-ESM2-0	mean	-	-	0.42	-	-	-	-	-	-
	10th	-	-	0.38	-	-	-	-	-	-
	90th	0.27	-	0.58	-	-	-	-	-0.28	-
q) NESM3	mean	-0.27	-	-	-	-	-	-	-	-
	10th	-0.32	-	-	-	-	-	-	-	-
	90th	-	-	-	-	-	-	-	-	-
r) NorESM2-LM	mean	-	-	0.39	-	-	-	-0.37	-	-0.31
	10th	-	-	0.35	-	-	-	-0.36	-	-0.34
	90th	-	-	0.25	-	-	0.26	0.29	-	-
s) NorESM2-MM	mean	-	-	-	-	-0.28	-0.36	-	0.38	-
	10th	-	-	-	-	-0.26	-0.34	-	-	-
	90th	-	-	-	-	-0.34	-	0.34	-0.30	-

Table A8. Same as A7 but for the meridional wind; 10th and 90th percentile are extreme south- and northward, respectively

Model	V	1. Sva	2. FJL	3. Kara	4. Lap	5. E. Sib	6. Chuk	7. Beau	8. NOW	9. NEW
b) ACCESS-CM2	mean	-	0.37	-	0.25	-	-	-	-	-
	10th	-0.25	-	-	-	-	-	-	-	-
	90th	-	0.40	-	-	-	-	-	-	0.26
c) AWI-CM-1-1-MR	mean	0.26	-	-	-	-	0.28	0.32	-0.32	-
	10th	-	-	-	-	-	-	-	0.35	-
	90th	0.32	-	-	-	-	0.33	0.33	-0.26	0.44
d) CanESM5	mean	-	0.38	0.39	0.36	0.49	0.47	-	-	-
	10th	-	-	0.28	-	-	-	-	-	0.38
	90th	-	0.64	0.42	0.49	0.52	0.52	-	-	-
e) EC-Earth3-AerChem	mean	-	-	0.34	-	-	0.47	-	-	-
	10th	-0.29	-	0.53	-	-	-	-	-	-
	90th	0.31	-	0.41	-	-	0.49	0.33	-	-
f) EC-Earth3-CC	mean	-	0.40	-	-	-	-	-	-	-0.27
	10th	-	-	0.61	-	-	-0.50	-	-	-
	90th	-	0.46	0.27	-	-	-	-	-	-0.25
g) EC-Earth3-Veg-LR	mean	-	0.31	0.36	0.29	-	0.37	-	-	-
	10th	-	-	-	-	-	-	-	-	-
	90th	-	0.37	0.44	-	-	0.47	-	-	-
h) EC-Earth3-Veg	mean	-	-	0.33	0.43	-	0.38	-	-	-
	10th	-	-	-	-	-	-	-0.25	0.28	-0.33
	90th	-	0.26	0.43	0.37	-	0.39	-	-	-
j) IPSL-CM6A-LR	mean	-	0.42	-	-	-	-	-	-	-
	10th	-	0.26	0.42	-	-0.31	-	-0.30	-	-
	90th	-	0.72	0.26	-	-	0.32	-	-	-
k) MIROC-ES2L	mean	-	-	-	-	-	-	-	-	-
	10th	-	-	-	-	-	-	-	-	-
	90th	-	-	-	-	-	-	-	-	-
l) MIROC6	mean	0.29	-	0.35	0.34	-	-	-	-	0.28
	10th	-	-	-	0.41	-0.41	-0.46	-	-	-0.29
	90th	0.34	-	0.39	0.32	-	-	-	-	0.30
m) MPI-ESM-1-2-HAM	mean	0.52	-	0.46	0.52	-	-	-	-0.27	0.59
	10th	-	-0.28	0.37	-	-	-	-	-	-0.36
	90th	0.53	0.30	0.56	0.57	-	-	-	-0.26	0.67
n) MPI-ESM1-2-HR	mean	-	0.32	0.29	-	-	-	-	-0.39	0.35
	10th	-	-	-	-	-0.42	-0.38	-	0.28	-
	90th	-	0.28	0.38	-	-	-	-	-0.33	0.41
o) MPI-ESM1-2-LR	mean	0.50	-	0.43	0.32	-	-	-	-	-
	10th	-	-0.26	0.27	-	-	-	-	-	-
	90th	0.49	-	0.50	0.37	-	-	-	-	0.35
p) MRI-ESM2-0	mean	0.32	0.29	0.70	-	-	0.27	-	-	0.25
	10th	-	-	0.42	-	-	-	-	-	-
	90th	0.36	0.42	0.71	-	-	0.33	-	-	0.37
q) NESM3	mean	-	0.29	-	-	-	-	-	-	-
	10th	-	-	-	-	-	-	-	-	-
	90th	-	0.33	-	-	-	-	-	-	-
r) NorESM2-LM	mean	-	-	-	-	-	0.30	0.36	-	-
	10th	-	-	0.38	-	-	-	-0.40	-	-0.28
	90th	-	-	-	-	-	0.32	0.28	-	-
s) NorESM2-MM	mean	-	-	-	-	-0.32	-	0.34	-0.32	-
	10th	-	-	-	-	-0.26	-0.36	-	0.45	-0.25
	90th	-	-	-	-	-	-	-	-0.32	-

Table A9. For the models with most (left) and least (right) polynya activity, in the Kara (top) and Laptev (bottom) seas, 45-year mean polynya area in km² and 45-year mean properties on the continental shelf, at the shelf break, and in the deep basin, in the Atlantic Water core and at the bottom cell. Potential temperature T is in °C; practical salinity S has no unit; potential density anomaly σ_2 is in kg m⁻³; depth is in m.

			c) AWI-CM-1-1-MR	d) CanESM5	j) IPSL-CM6A-LR	k) MIROC-ES2L	q) NESM3	b) ACCESS-CM2
Mean	polynya	area	63801	25366	19442	3612	305	1692
Shelf	bottom	T	-0.81	-0.59	0.31	-1.66	-0.80	-1.42
		S	33.65	33.38	33.81	34.26	31.60	32.92
		σ_2	36.32	36.08	36.32	36.88	34.70	35.80
Break	AW core	T	0.93	-0.47	0.79	-0.38	0.81	-1.06
		S	34.81	34.67	34.84	34.60	34.50	34.79
		σ_2	37.03	37.08	37.07	37.01	36.80	37.24
	bottom	depth	291	448	437	375	575	554
		T	-0.02	-0.53	0.69	-0.90	0.81	-1.08
		S	34.90	34.68	34.85	34.67	34.50	34.80
σ_2	37.21	37.09	37.09	37.13	36.80	37.25		
Deep	AW core	T	0.41	-0.47	0.89	0.53	1.21	-0.53
		S	34.77	34.61	34.82	34.67	34.59	34.83
		σ_2	37.06	37.04	37.04	36.97	36.82	37.22
	bottom	depth	358	876	783	667	726	913
		T	-0.58	-1.25	0.37	-0.53	-0.61	-0.96
		S	35.09	34.72	34.95	34.72	34.98	34.82
σ_2	37.43	37.20	37.21	37.13	37.34	37.25		
Mean	polynya	area	20586	2936	1610	223	0.2	4.5
Shelf	bottom	T	-1.00	-0.96	-0.46	-1.30	-0.96	-1.36
		S	31.89	29.83	32.49	33.15	29.06	30.45
		σ_2	34.95	33.31	35.37	35.97	32.72	33.84
Break	AW core	T	0.14	-0.71	0.57	0.06	1.22	-1.08
		S	34.76	34.58	34.77	34.63	34.56	34.76
		σ_2	37.09	37.04	37.04	36.99	36.80	37.22
	bottom	depth	427	746	707	655	613	744
		T	-0.04	-0.72	0.56	0.04	1.15	-1.08
		S	34.81	34.58	34.77	34.64	34.58	34.76
σ_2	37.15	37.04	37.04	37.01	36.83	37.22		
Deep	AW core	T	0.15	-0.69	0.62	0.11	1.22	-0.90
		S	34.74	34.61	34.79	34.65	34.56	34.81
		σ_2	37.07	37.05	37.06	37.00	36.80	37.24
	bottom	depth	388	1263	987	749	646	1866
		T	-0.59	-1.16	0.39	-0.52	-0.61	-0.94
		S	35.09	34.69	34.95	34.71	34.98	34.82
σ_2	37.43	37.17	37.20	37.12	37.34	37.25		

560 *Author contributions.* J.R. had the original idea for the work. C.H. designed the study and conducted most of the work, with input from C.H.M.W. and J.R. C.H.M.W. and T.T. contributed data. All authors contributed to the writing and reviewing.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. C.H. and C.H.M.W acknowledge funding from the Swedish National Space Agency, grant agreement 2022-00149. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS),
565 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. We thank the two anonymous reviewers whose suggestions improved the manuscript.

570 **References**

- Aagaard, K., Coachman, L., and Carmack, E.: On the halocline of the Arctic Ocean, *Deep Sea Research Part A. Oceanographic Research Papers*, 28, 529–545, [https://doi.org/https://doi.org/10.1016/0198-0149\(81\)90115-1](https://doi.org/https://doi.org/10.1016/0198-0149(81)90115-1), 1981.
- 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.
- 575 Bentsen, M., Olivière, D. J. L., Seland, y., Toniazco, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O. A., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: [dataset] NCC NorESM2-MM model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.506>, 2019.
- Bethke, I., Wang, Y., Counillon, F., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H. R., Chiu, P.-G., Bentsen, M., Guo, C., Tjiputra, J., Kirkevåg, A., Olivière, D. J. L., Seland, y., Fan, Y., Lawrence, P., Eldevik, T., and Keenlyside, N.: [dataset] NCC NorCPM1 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.10843>, 2019.
- 580 Bi, D., Dix, M., Marsland, S., O'farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky, J., Rashid, H. A., et al.: Configuration and spin-up of ACCESS-CM2, the new generation Australian community climate and earth system simulator coupled model, *Journal of Southern Hemisphere Earth Systems Science*, 70, 225–251, <https://doi.org/10.1071/ES19040>, 2020.
- 585 Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Cadule, P., Devilliers, M., Ghattas, J., Lebas, N., Lurton, T., Mellul, L., Musat, I., Mignot, J., and Cheruy, F.: [dataset] IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1534>, 2018.
- Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Balkanski, Y., Checa-Garcia, R., Hauglustaine, D., Bekki, S., and Marchand, M.: [dataset] IPSL IPSL-CM5A2-INCA model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.13642>, 2020a.
- 590 Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., et al.: Presentation and evaluation of the IPSL-CM6A-LR climate model, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002 010, <https://doi.org/10.1029/2019MS002010>, 2020b.
- Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., and Savoie., M. H.: EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets, *ISPRS International Journal of Geo-Information*, 1, 32–45, <https://doi.org/doi:10.3390/ijgi1010032>, 2012.
- 595 Cao, J. and Wang, B.: [dataset] NUIST NESMv3 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.2021>, 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.
- 600 Chai, Z.: [dataset] CAS CAS-ESM2.0 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1944>, 2020.
- Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., et al.: Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project, *Climate Dynamics*, 49, 1107–1136, <https://doi.org/10.1007/s00382-016-2985-y>, 2017.
- 605 Danabasoglu, G.: [dataset] NCAR CESM2-WACCM model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.10024>, 2019a.

- Danabasoglu, G.: [dataset] NCAR CESM2 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.2185>, 2019b.
- 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.
- 610 Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, I., Law, R., Mackallah, C., Marsland, S., O'Farrell, S., Rashid, H., Srbnovsky, J., Sullivan, A., Trenham, C., Vohralik, P., Watterson, I., Williams, G., Woodhouse, M., Bodman, R., Dias, F. B., Domingues, C. M., Hannah, N., Heerdegen, A., Savita, A., Wales, S., Allen, C., Druken, K., Evans, B., Richards, C., Ridzwan, S. M., Roberts, D., Smillie, J., Snow, K., Ward, M., and Yang, R.: [dataset] CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.2281>, 2019.
- 615 Dmitrenko, I. A., Kirillov, S. A., Serra, N., Koldunov, N., Ivanov, V. V., Schauer, U., Polyakov, I. V., Barber, D., Janout, M., Lien, V., et al.: Heat loss from the Atlantic water layer in the northern Kara Sea: Causes and consequences, *Ocean Science*, 10, 719–730, <https://doi.org/10.5194/os-10-719-2014>, 2014.
- Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., Bergmann, T., Bernadello, R., Bousetta, S., Caron, L.-P., et al.: The EC-earth3 Earth system model for the climate model intercomparison project 6, *Geoscientific Model Development Discussions*, 2021, 620 1–90, <https://doi.org/10.5194/gmd-15-2973-2022>, 2021.
- EC-Earth Consortium (EC-Earth): [dataset] EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.181>, 2019a.
- EC-Earth Consortium (EC-Earth): [dataset] EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.642>, 2019b.
- 625 EC-Earth Consortium (EC-Earth): [dataset] EC-Earth-Consortium EC-Earth3-AerChem model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.639>, 2020a.
- EC-Earth Consortium (EC-Earth): [dataset] EC-Earth-Consortium EC-Earth-3-CC model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.640>, 2020b.
- EC-Earth Consortium (EC-Earth): [dataset] EC-Earth-Consortium EC-Earth3-Veg-LR model output prepared for CMIP6 CMIP, 630 <https://doi.org/10.22033/ESGF/CMIP6.643>, 2020c.
- 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.
- Fox-Kemper, B., DeRepentigny, P., Treguier, A. M., Stepanek, C., O'Rourke, E., Mackallah, C., Meucci, A., Aksenov, Y., Du- 635 rack, P. J., Feldl, N., et al.: CMIP7 Data Request: Ocean and Sea Ice Priorities and Opportunities, *EGUsphere*, 2025, 1–58, <https://doi.org/10.5194/egusphere-2025-3083>, 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, <https://doi.org/10.1175/JCLI4046.1>, 2007.
- Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., et al.: 640 High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6, *Geoscientific Model Development*, 9, 4185–4208, <https://doi.org/10.5194/gmd-9-4185-2016>, 2016.
- Hajima, T., Abe, M., Arakawa, O., Suzuki, T., Komuro, Y., Ogura, T., Ogochi, K., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Ohgaito, R., Ito, A., Yamazaki, D., Ito, A., Takata, K., Watanabe, S., Kawamiya, M., and Tachiiri, K.: [dataset] MIROC MIROC-ES2L model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.902>, 2019.

- 645 Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., et al.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, *Geoscientific Model Development*, 13, 2197–2244, <https://doi.org/10.5194/gmd-13-2197-2020>, 2020.
- Hastrup, K., Andersen, A. O., Grønnow, B., and Heide-Jørgensen, M. P.: Life around the North Water ecosystem: Natural and social drivers of change over a millennium, *Ambio*, 47, 213–225, <https://doi.org/10.1007/s13280-018-1028-9>, 2018.
- 650 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. and Wong, C. H. M.: Automatic detection of Arctic polynyas using hybrid supervised-unsupervised deep learning, *The Cryosphere*, 19, 6043–6058, <https://doi.org/10.5194/tc-19-6043-2025>, 2025.
- 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.
- 655 Hirano, D., Fukamachi, Y., Watanabe, E., Ohshima, K. I., Iwamoto, K., Mahoney, A. R., Eicken, H., Simizu, D., and Tamura, T.: A wind-driven, hybrid latent and sensible heat coastal polynya off B arrow, A laska, *Journal of Geophysical Research: Oceans*, 121, 980–997, <https://doi.org/10.1002/2015JC011318>, 2016.
- Huang, W.: [dataset] THU CIESM model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1352>, 2019.
- 660 Hunke, E., Lipscomb, W., Jones, P., Turner, A., Jeffery, N., and Elliott, S.: CICE, The Los Alamos Sea Ice Model, <https://www.osti.gov/biblio/1364126>, 2017.
- Huot, P.-V., Kittel, C., Fichet, T., Jourdain, N. C., Sterlin, J., and Fettweis, X.: Effects of the atmospheric forcing resolution on simulated sea ice and polynyas off Adélie Land, East Antarctica, *Ocean Modelling*, 168, 101 901, <https://doi.org/10.1016/j.ocemod.2021.101901>, 2021.
- 665 Ingvaldsen, R. B., Assmann, K. M., Primicerio, R., Fossheim, M., Polyakov, I. V., and Dolgov, A. V.: Physical manifestations and ecological implications of Arctic Atlantification, *Nature Reviews Earth & Environment*, 2, 874–889, <https://doi.org/10.1038/s43017-021-00228-x>, 2021.
- Itkin, P., Losch, M., and Gerdes, R.: Landfast ice affects the stability of the Arctic halocline: Evidence from a numerical model, *Journal of Geophysical Research: Oceans*, 120, 2622–2635, <https://doi.org/10.1002/2014JC010353>, 2015.
- 670 Ivanov, V. V. and Golovin, P. N.: Observations and modeling of dense water cascading from the northwestern Laptev Sea shelf, *Journal of Geophysical Research: Oceans*, 112, <https://doi.org/10.1029/2006JC003882>, 2007.
- Jungclaus, J., Bittner, M., Wieners, K.-H., Wachsmann, F., Schupfner, M., Legutke, S., Giorgetta, M., Reick, C., Gayler, V., Haak, H., de Vrese, P., Raddatz, T., Esch, M., Mauritsen, T., von Storch, J.-S., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahn, T., Kloster, S., Kinne, S., Lasslop, G., Kornbluh, L., Marotzke, J., Matei, D., Meraner, K., 675 Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Peters-von Gehlen, K., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., and Roeckner, E.: [dataset] MPI-M MPIESM1.2-HR model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.741>, 2019.
- Landrum, L. L., DuVivier, A. K., Holland, M. M., Krumhardt, K., and Sylvester, Z.: Defining Antarctic polynyas in satellite observations and climate model output to support ecological climate change research, *EGUsphere*, 2024, 1–40, <https://doi.org/10.5194/egusphere-2024-3490>, 2024.
- 680

- Lavoie, J., Tremblay, B., and Rosenblum, E.: Pacific waters pathways and vertical mixing in the CESM1-LE: Implication for mixed layer depth evolution and sea ice mass balance in the Canada basin, *Journal of Geophysical Research: Oceans*, 127, e2021JC017729, <https://doi.org/10.1029/2021JC017729>, 2022.
- 685 Lee, W.-L. and Liang, H.-C.: [dataset] AS-RCEC TaiESM1.0 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.9684>, 2019.
- Lee, Y. J., Maslowski, W., Cassano, J. J., Clement Kinney, J., Craig, A. P., Kamal, S., Osinski, R., Seefeldt, M. W., Stroeve, J., and Wang, H.: Causes and evolution of winter polynyas north of Greenland, *The Cryosphere*, 17, 233–253, <https://doi.org/10.5194/tc-17-233-2023>, 2023.
- 690 Lovato, T. and Peano, D.: [dataset] CMCC CMCC-CM2-SR5 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1362>, 2020.
- Lovato, T., Peano, D., and Butenschön, M.: [dataset] CMCC CMCC-ESM2 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.13164>, 2021.
- Luneva, M. V., Ivanov, V. V., Tuzov, F., Aksenov, Y., Harle, J. D., Kelly, S., and Holt, J. T.: Hotspots of dense water cascading in the Arctic Ocean: Implications for the Pacific water pathways, *Journal of Geophysical Research: Oceans*, 125, e2020JC016044, <https://doi.org/10.1029/2020JC016044>, 2020.
- 695 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.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., et al.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing CO₂, *Journal of Advances in Modeling Earth Systems*, 11, 998–1038, <https://doi.org/10.1029/2018MS001400>, 2019.
- 700 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.
- 705 Moore, G.: Impact of model resolution on the representation of the wind field along Nares Strait, *Scientific Reports*, 11, 13 271, <https://doi.org/10.1038/s41598-021-92813-9>, 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, 9809, <https://doi.org/10.1038/s41598-023-36179-0>, 2023.
- 710 Morales Maqueda, M. A., Willmott, A. J., and Biggs, N. R. T.: Polynya Dynamics: a Review of Observations and Modeling, *Reviews of Geophysics*, 42, RG1004, <https://doi.org/https://doi.org/10.1029/2002RG000116>, 2004.
- Muilwijk, M., Nummelin, A., Heuzé, C., Polyakov, I. V., Zanowski, H., and Smedsrud, L. H.: Divergence in climate model projections of future Arctic Atlantification, *Journal of Climate*, 36, 1727–1748, <https://doi.org/10.1175/JCLI-D-22-0349.1>, 2023.
- Muilwijk, M., Hattermann, T., Martin, T., and Granskog, M. A.: Future sea ice weakening amplifies wind-driven trends in surface stress and Arctic Ocean spin-up, *Nature Communications*, 15, 6889, <https://doi.org/10.1038/s41467-024-50874-0>, 2024.
- 715 Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., et al.: A higher-resolution version of the max planck institute earth system model (MPI-ESM1. 2-HR), *Journal of Advances in Modeling Earth Systems*, 10, 1383–1413, <https://doi.org/10.1029/2017MS001217>Digital Object Identifier (DOI), 2018.

- Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stoll, J., Folini, D. S., Tegen, I., Wieners, K.-H., Mauritsen, T., Stemmler, I.,
720 Barthel, S., Bey, I., Daskalakis, N., Heinold, B., Kokkola, H., Partridge, D., Rast, S., Schmidt, H., Schutgens, N., Stanelle, T., Stier,
P., Watson-Parris, D., and Lohmann, U.: [dataset] HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP,
<https://doi.org/10.22033/ESGF/CMIP6.1622>, 2019.
- Notz, D. and SIMIP Community: Arctic sea ice in CMIP6, *Geophysical Research Letters*, 47, e2019GL086749,
<https://doi.org/10.1029/2019GL086749>, 2020.
- 725 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, 13, <https://doi.org/10.1186/S40562-016-0045-4>, 2016.
- Peralta-Ferriz, C. and Woodgate, R. A.: The dominant role of the East Siberian Sea in driving the oceanic flow through the Bering
Strait—Conclusions from GRACE ocean mass satellite data and in situ mooring observations between 2002 and 2016, *Geophysical
Research Letters*, 44, 11–472, <https://doi.org/10.1002/2017GL075179>, 2017.
- 730 Preußner, 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.
- Preußner, A., Ohshima, K. I., Iwamoto, K., Willmes, S., and Heinemann, G.: Retrieval of Wintertime Sea Ice Production in Arctic Polynyas
Using Thermal Infrared and Passive Microwave Remote Sensing Data, *Journal of Geophysical Research: Oceans*, 124, 5503–5528,
735 <https://doi.org/doi.org/10.1029/2019JC014976>, 2019.
- Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.:
The Arctic has warmed nearly four times faster than the globe since 1979, *Communications Earth and Environment*, 3, 168,
<https://doi.org/10.1038/s43247-022-00498-3>, 2022.
- Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., and Andrews, T.: [dataset] MOHC HadGEM3-GC31-LL model output prepared for
740 CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.419>, 2018.
- Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., and Andrews, T.: [dataset] MOHC HadGEM3-GC31-MM model output prepared for
CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.420>, 2019.
- Rong, X.: [dataset] CAMS CAMS_CSM1.0 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1399>, 2019.
- Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in: *Medical image computing
745 and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings,
part III 18*, pp. 234–241, Springer, https://doi.org/10.1007/978-3-319-24574-4_28, 2015.
- Roquet, F., Madec, G., McDougall, T. J., and Barker, P. M.: Accurate polynomial expressions for the density and specific volume of seawater
using the TEOS-10 standard, *Ocean Modelling*, 90, 29–43, <https://doi.org/10.1016/j.ocemod.2015.04.002>, 2015.
- Rousset, C., Vancoppenolle, M., Madec, G., Fichfet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Lévy, C., Masson, S.,
750 et al.: The Louvain-La-Neuve sea ice model LIM3. 6: global and regional capabilities, *Geoscientific Model Development*, 8, 2991–3005,
<https://doi.org/10.5194/gmd-8-2991-2015>, 2015.
- Seferian, R.: [dataset] CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP,
<https://doi.org/10.22033/ESGF/CMIP6.1391>, 2018.
- Seland, O., Bentsen, M., Olivie, D. J. L., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A.,
755 Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O. A.,

- Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: [dataset] NCC NorESM2-LM model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.502>, 2019.
- 760 Seland, Ø., Bentsen, M., Seland Graff, L., Olivíé, D., Toniazzo, T., Gjermundsen, A., Debernard, J. B., Gupta, A. K., He, Y., Kirkevåg, A., et al.: The Norwegian Earth system model, NorESM2–evaluation of the CMIP6 DECK and historical simulations, *Geoscientific Model Development Discussions*, 13, 6165–6200, <https://doi.org/10.5194/gmd-13-6165-2020>, 2020.
- Semmler, T., Danilov, S., Rackow, T., Sidorenko, D., Barbi, D., Hegewald, J., Sein, D., Wang, Q., and Jung, T.: [dataset] AWI AWI-CM1.1MR model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.359>, 2018.
- 765 Semmler, T., Danilov, S., Gierz, P., Goessling, H. F., Hegewald, J., Hinrichs, C., Koldunov, N., Khosravi, N., Mu, L., Rackow, T., et al.: Simulations for CMIP6 with the AWI climate model AWI-CM-1-1, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002009, <https://doi.org/10.1029/2019MS002009>, 2020.
- Sepulchre, P., Caubel, A., Ladant, J.-B., Bopp, L., Boucher, O., Braconnot, P., Brockmann, P., Cozic, A., Donnadieu, Y., Estella-Perez, V., et al.: IPSL-CM5A2. An Earth System Model designed for multi-millennial climate simulations, *Geoscientific Model Development*, 13, 3011–3053, <https://doi.org/10.5194/gmd-13-3011-2020>, 2020.
- Smith Jr, W. O. and Barber, D.: *Polynyas: Windows to the world*, vol. 74, Elsevier, 2007.
- 770 Song, Z., Qiao, F., Bao, Y., Shu, Q., Song, Y., and Yang, X.: [dataset] FIO-QLNM FIO-ESM2.0 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.9047>, 2019.
- 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.
- Sun, S. and Solomon, A.: Suitability of the CICE sea ice model for seasonal prediction and positive impact of CryoSat-2 ice thickness initialization, *The Cryosphere*, 18, 3033–3048, <https://doi.org/10.5194/tc-18-3033-2024>, 2024.
- 775 Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., et al.: The Canadian earth system model version 5 (CanESM5. 0.3), *Geoscientific Model Development*, 12, 4823–4873, <https://doi.org/10.5194/gmd-12-4823-2019>, 2019a.
- Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., 780 Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B., and Sigmond, M.: [dataset] CCCma CanESM5–CanOE model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.10205>, 2019b.
- Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B., and Sigmond, M.: [dataset] 785 CCCma CanESM5 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1303>, 2019c.
- Tamura, T. and Ohshima, K. I.: Mapping of sea ice production in the Arctic coastal polynyas, *Journal of Geophysical Research: Oceans*, 116, C07 030, <https://doi.org/doi.org/10.1029/2010JC006586>, 2011.
- Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., Walton, J., and Jones, C.: [dataset] MOHC UKESM1.0-LL model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1569>, 2019.
- 790 Tatebe, H. and Watanabe, M.: [dataset] MIROC MIROC6 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.881>, 2018.
- Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., et al.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, *Geoscientific Model Development*, 12, 2727–2765, <https://doi.org/10.5194/gmd-12-2727-2019>, 2019.

- 795 Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., Knutti, R., Lowe, J., O'Neill, B., Sanderson, B., et al.: Climate model projections from the scenario model intercomparison project (ScenarioMIP) of CMIP6, *Earth System Dynamics*, 12, 253–293, <https://doi.org/10.5194/esd-12-253-2021>, 2021.
- Tian, T., Yang, S., Høyer, J. L., Nielsen-Englyst, P., and Singha, S.: Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis, *Communications Earth & Environment*, 5, 111, <https://doi.org/10.1038/s43247-024-01276-z>, 2024.
- 800 Uotila, P., Iovino, D., Vancoppenolle, M., Lensu, M., and Rousset, C.: Comparing sea ice, hydrography and circulation between NEMO3.6 LIM3 and LIM2, *Geoscientific Model Development*, 10, 1009–1031, <https://doi.org/10.5194/gmd-10-1009-2017>, 2017.
- Vancoppenolle, M., Rousset, C., Blockley, E., Aksenov, Y., Feltham, D., Fichefet, T., Garric, G., Guémas, V., Iovino, D., Keeley, S., Madec, G., Massonnet, F., Ridley, J., Schroeder, D., and Tietsche, S.: SI3, the NEMO Sea Ice Engine, <https://doi.org/10.5281/zenodo.7534900>, 2023.
- 805 Vivier, F., Lourenço, A., Skogseth, R., Goszczko, I., Michel, E., Rousset, C., Bouruet-Aubertot, P., Cuypers, Y., Lansard, B., and Waelbroeck, C.: Dense Water Production in Storfjorden, Svalbard, From a 1-Year Time Series of Observations and a Simple Model: Are Polynyas in a Warming Arctic Exporting Heat to the Deep Ocean?, *Journal of Geophysical Research: Oceans*, 129, e2024JC020878, <https://doi.org/10.1029/2024JC020878>, 2024.
- Voltaire, A.: [dataset] CNRM-CERFACS CNRM-CM6-1 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1375>, 2018.
- 810 Voltaire, A.: [dataset] CNRM-CERFACS CNRM-CM6-1-HR model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.1385>, 2019.
- Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., Legutke, S., Schupfner, M., Wachsmann, F., Gayler, V., Haak, H., de Vrese, P., Raddatz, T., Mauritsen, T., von Storch, J.-S., Behrens, J., Brovkin, V., Claussen, M., Crueger, T., Fast, I., Fiedler, S., Hagemann, S., Hohenegger, C., Jahns, T., Kloster, S., Kinne, S., Lasslop, G., Kornbluh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Müller, W., Nabel, J., Notz, D., Peters-von Gehlen, K., Pincus, R., Pohlmann, H., Pongratz, J., Rast, S., Schmidt, H., Schnur, R., Schulzweida, U., Six, K., Stevens, B., Voigt, A., and Roeckner, E.: [dataset] MPI-M MPIESM1.2-LR model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.742>, 2019.
- 820 Wong, C. H. M., Heuzé, C., Ickes, L., and Zhou, L.: The spatio-temporal variability, trends, and drivers of winter Arctic polynyas, *Journal of Climate*, accepted, <https://doi.org/10.31223/X5VJ0P>, 2026.
- Woodgate, R. A., Aagaard, K., Muench, R. D., Gunn, J., Björk, G., Rudels, B., Roach, A., and Schauer, U.: The Arctic Ocean boundary current along the Eurasian slope and the adjacent Lomonosov Ridge: Water mass properties, transports and transformations from moored instruments, *Deep Sea Research Part I: Oceanographic Research Papers*, 48, 1757–1792, [https://doi.org/10.1016/S0967-0637\(00\)00091-1](https://doi.org/10.1016/S0967-0637(00)00091-1), 2001.
- 825 World Meteorological Organization: Sea Ice Nomenclature (WMO No. 259), World Meteorological Organization (WMO), Geneva, Switzerland, <https://library.wmo.int/idurl/4/41953>, includes term definition: “Polynya: Any non-linear shaped opening enclosed in ice. Polynyas may contain brash ice and/or be covered with new ice, nilas or young ice.”, 2014.
- Yang, Y.-M., Wang, B., Cao, J., Ma, L., and Li, J.: Improved historical simulation by enhancing moist physical parameterizations in the climate system model NESM3.0, *Climate Dynamics*, 54, 3819–3840, <https://doi.org/10.1007/s00382-020-05209-2>, 2020.

- 830 Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., et al.: The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2. 0: Description and basic evaluation of the physical component, *Journal of the Meteorological Society of Japan*. Ser. II, 97, 931–965, <https://doi.org/10.2151/jmsj.2019-051>, 2019a.
- Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yoshimura, H., Shindo, E., Mizuta, R., Ishii, M., Obata, A., and Adachi, Y.: [dataset] MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP, 835 <https://doi.org/10.22033/ESGF/CMIP6.621>, 2019b.
- Zapponini, M. and Goessling, H. F.: Atmospheric destabilization leads to Arctic Ocean winter surface wind intensification, *Communications Earth & Environment*, 5, 262, <https://doi.org/10.1038/s43247-024-01428-1>, 2024.
- Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of higher climate sensitivity in CMIP6 models, *Geophysical Research Letters*, 47, e2019GL085 782, <https://doi.org/10.1029/2019GL085782>, 2020.
- 840 Zhou, L., Heuzé, C., and Mohrmann, M.: Early winter triggering of the Maud Rise polynya, *Geophysical Research Letters*, 49, e2021GL096 246, <https://doi.org/10.1029/2021GL096246>, 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, <https://doi.org/10.1029/2022JC019148>, 2023.
- Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., Wang, Y., Dobrohotoff, P., Sribnovsky, J., Stevens, L., Vohralik, P., 845 Mackallah, C., Sullivan, A., O’Farrell, S., and Druken, K.: [dataset] CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP, <https://doi.org/10.22033/ESGF/CMIP6.2288>, 2019.