No emergence of deep convection in the Arctic Ocean across CMIP6 models

Céline Heuzé[1] and Hailong Liu[2]

[1]University of Gothenburg
[2]Shanghai Jiao Tong University

September 25, 2023

Abstract

As sea ice disappears, the emergence of open ocean deep convection in the Arctic has been suggested. Here, using 36 state-of-the-art climate models and up to 50 ensemble members per model, we show that Arctic deep convection is rare even under the strongest warming scenario. Only 5 models have somewhat permanent convection by 2100, while 11 have had convection by the middle of the run. For all, the deepest mixed layers are in the Eurasian basin, by St Anna Trough. When the models convect, that region undergoes a salinification and increasing wind speeds; it is freshening otherwise. We discuss the causality and potential reasons for the opposite trends. Given the model's different parameterisations, and given that the ensemble members that convect the deepest, most often, are those with the strongest sensitivity, we conclude that differences in deep convection are most likely linked to the model formulation.

No emergence of deep convection in the Arctic Ocean across CMIP6 models

Céline Heuzé[1], and Hailong Liu[2]

[1]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden
[2]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China

Key Points:

- Oceanic deep convection does not emerge and persist in the Arctic in the majority of CMIP6 models, despite a cessation in the Nordic Seas.
- Arctic deep convection occurs only when both surface salinity and winds are increasing, year round, yet most models are freshening.
- The models with the strongest sensitivity, especially with an oceanic polar amplification, have the deepest Arctic mixed layers, most often.

Corresponding author: Céline Heuzé, `celine.heuze@gu.se`

Abstract

As sea ice disappears, the emergence of open ocean deep convection in the Arctic has been suggested. Here, using 36 state-of-the-art climate models and up to 50 ensemble members per model, we show that Arctic deep convection is rare even under the strongest warming scenario. Only 5 models have somewhat permanent convection by 2100, while 11 have had convection by the middle of the run. For all, the deepest mixed layers are in the Eurasian basin, by St Anna Trough. When the models convect, that region undergoes a salinification and increasing wind speeds; it is freshening otherwise. We discuss the causality and potential reasons for the opposite trends. Given the model's different parameterisations, and given that the ensemble members that convect the deepest, most often, are those with the strongest sensitivity, we conclude that differences in deep convection are most likely linked to the model formulation.

Plain Language Summary

Both observations and modelling simulations suggest that deep vertical mixing (or deep convection) in winter may become the new normal in the Arctic as sea ice disappears. These simulations are often done using only one model, so here we used all models available that participated in the Climate Model Intercomparison Project phase 6, for the strongest warming scenario. We show that after removing those that are already inaccurate in the present, and even with a restrictive threshold, most models have no deep convection in the Arctic, or extremely rarely. Only 5 still had deep convection by the time the run finishes in 2100. We investigated the possible links between deep convection and surface salinity, surface temperature, sea ice concentration and surface wind speeds, and found that the salinity was most important. Deep convection regions and periods are associated with a saltier, windier surface, while the rest of the Arctic and/or rest of the run freshens. Causality is unclear; we need higher resolution than monthly output. Similar behaviours within model families, a strong link to the model sensitivity, and cited work make us conclude that ultimately, the differences are probably caused primarily by the different model designs.

1 Introduction

The Arctic Ocean is changing. The resulting reduction in Arctic sea ice extent and thickness (Mallett et al., 2021; Meier & Stroeve, 2022) could enhance vertical mixing: More brine may be rejected year-round as younger, saltier ice desalinates (Peterson, 2018) or as sea ice reforms in winter over the now seasonal-ice areas (Onarheim et al., 2018), while the ice-freed regions may become more susceptible to wind stirring (Timmermans & Marshall, 2020). In the Eurasian Arctic, the process known as Atlantification (Polyakov et al., 2017), whereby warm water of Atlantic origin penetrates further into the Arctic, may be further weakening the stratification and enhancing sea ice melt. These led Polyakov et al. (2017) to hypothesise that the Arctic may start exhibiting deep convection in winter, a result found by Lique et al. (2018) in the 4x CO_2 scenario using the model HiGEM.

However, this hypothesis is so far not confirmed. Peralta-Ferriz and Woodgate (2015) found that a deepening of mixed layers in the Arctic is unlikely, since stratification greatly dominates over the wind effect. The latest observations in the Eurasian Arctic (Schulz et al., under review) yielded mixed layers no deeper than 130 m, even in winter. Besides, in a recent Arctic study using models that participated in the Climate Model Intercomparison Project phase 6 or CMIP6 (Eyring et al., 2016), Muilwijk et al. (2023) showed that there was no agreement among models regarding future stratification and the effect of Atlantification under the strongest warming scenario (SSP5-8.5, O'Neill et al. (2016)). This suggests that HiGEM's deep convection in the Arctic may be a model artefact rather than the future of the Arctic.

We here determine whether deep convection emerges in the Arctic in the future scenario SSP5-8.5 using all CMIP6 models and all their ensemble members for which the mixed layer depth output was available, as described in section 2. In section 3, we detail the spatial and temporal patterns of future Arctic mixed layers and discuss possible reasons for these, focusing on the model biases and their trends in surface properties, sea ice and winds. We conclude in section 4.

2 Data and Methods

2.1 CMIP6 data

To investigate deep convection and its potential drivers in the future Arctic, we use all CMIP6 models and all their ensemble members that had monthly mixed layer depth (“mlotst”), ocean surface salinity (“sos”) and temperature (“tos”), sea ice concentration (“siconc”) and surface wind speed (“sfcWind”) available on any of the Earth System Grid Federation (ESGF) nodes for the future scenario SSP5-8.5, for January 2015 to December 2100. The models and their ensemble members are listed in supp. Tables S1 and S2.

For AWI-CM-1-1-MR and GISS-E2-1-G, we used the sea ice thickness (“sivol”) because sea ice concentration was not available. For CAMS-CSM1-0, we generated the surface salinity from the full-depth salinity (“so”) as the former was not available; we used the salinity of the shallowest level, 5 m depth. Similarly, we used the models’ bathymetry (“deptho”) when it was available, but had to generate it from the full-depth salinity for 11 models, as the last level with salinity data. Finally, for CESM2-WACCM, GFDL-CM4 and MRI-ESM2-0, several grid types were available and we chose for simplicity the regularised grid (“gr”). For the other models, we took the one grid type available; see supp. Tables S1 and S2.

2.2 Methods

The thresholds and choices of this subsection are discussed in supplementary text S1. In agreement with Lique et al. (2018), we consider that there is deep convection in the Arctic if the mixed layer depth (MLD) exceeds 500 m. We here do not quantify overturned volumes. We only perform a binary detection of deep convection, so we select the overall maximum MLD, in space or time depending on the analysis. The ensemble members that exhibit deep convection in the Arctic over the observed part of SSP5-8.5 (2015-2023) are shown on the first figure and subsequently removed from the study, as they are already inaccurate at the beginning of the run. Consequently, of the originally 36 models, 27 remain for most of the analysis. We do not take biases in Nordic Seas MLD into account for model selection, as all CMIP6 models have spurious deep convection there (Heuzé, 2021).

The sea ice edge (supp. Fig S1) is detected as the contour of 15% concentration or 10 cm thickness, averaged over 2040-2060 and 2080-2100. We define the Eurasian basin as the region north of 80°N, longitudes 20°W - 140°E, deeper than 1000 m, and the Nordic Seas as the region of latitudes 66 - 80°N, longitudes 30°W - 20°, deeper than 1000 m - see contours on Fig 1. We perform across-model correlations by comparing, for each model and each ensemble member, their maximum MLD and number of year where the MLD exceeds 500 m in the Eurasian basin, to:

1. Average surface temperature and salinity in the Eurasian basin over the first 20 years of SSP5-8.5;
2. Tropical and Mid-latitude (60°S - 60°N) and Arctic (north of 75°N) warming, as difference between the 2080-2100 and 2015-2035 average ocean surface temperature, which we use as a proxy for the model sensitivity. We also define the oceanic

polar amplification as the difference between the Arctic and tropical/mid-latitude warming.

Besides, we compute correlations and trends for each grid cell, after interpolating all the parameters onto each model's mlotst grid. The correlations and trends are computed for each ensemble member separately, and averaged afterwards. We consider that there is ensemble agreement if more than 50% of the ensemble members have a significant correlation/trend (determined using a t-test at 95% significance) of the same sign; we then present the median value and its ensemble-spread. Finally, we present the composite trends after grouping the models based on their Arctic deep convection behaviour, as described in the result section.

3 Results and Discussion

3.1 Arctic deep convection is rare, restricted both in space and time

The maximum mixed layer depth reached in the Arctic over 2015-2100 varies strongly across models (Fig. 1, note the logarithmic scale). The value does not exceed 100 m for some, such as CAMS-CSM1-0, while others such as CMCC-CM2-SR5 exceed 2000 m in the majority of the Eurasian basin. All models with deep MLD agree that the deepest values are in the Eurasian basin, most commonly by St Anna Trough. Nine models have MLD exceeding the 500 m threshold for deep convection already over 2015 to present for all their ensemble members (red squares on Fig. 1). Interestingly, these are not only by St Anna Trough but also north of Svalbard, suggesting that for these models the Nordic Sea deep convection area extends too far north, most likely following the sea ice edge (shown on supp. Fig. S1). All models presented on Fig. 1 have spuriously deep MLD in the Nordic Seas.

After removing the models and ensemble members that are unrealistic in the present-day regarding their Arctic MLD (asterisks in supp Tables S1 and S2), 27 models remain. The temporal evolution of their MLD reveals four groups of models (Fig. 2 and supp Table S3):

1. The first 11 models have no Arctic deep convection during the entire run. The Arctic MLD time series (blue lines, Fig. 2) are mostly flat, with no year where the MLD exceeds 500 m, regardless of the ensemble member. The maximum MLD across these models and their ensemble members is often of the order of 100 m, i.e. like currently observed in the Arctic (Schulz et al., under review).
2. Six models have deep convection in the Arctic on rare occasions, in the middle of the run. The maximum number of years with deep convection for this group is 17 out of 86 (supp Table S3), but is most often 4 or fewer.
3. Five models also have deep convection in the Arctic in the middle of the run, more often. It starts by 2030, peaks in the first half of the run, and then declines slowly, with the ensemble average back under the 500 m threshold by the end of the run.
4. The last 5 models start convecting in the second half of the run, by 2070, and appear "stably" convecting at the end of the run. Unfortunately, only one ensemble member of one of the models (ACCESS-CM2) is available beyond 2100, so we cannot tell whether deep convection in these models would also decline later.

The Nordic Seas MLD (red lines, Fig. 2) falls below 1000 m for all models, and even below 500 m for two thirds of them. The consistent cessation of Nordic Seas deep convection is not related to the models' behaviour in the Arctic. Therefore, unlike suggested by Lique et al. (2018), deep convection does not migrate to the Arctic in response to its cessation further south.

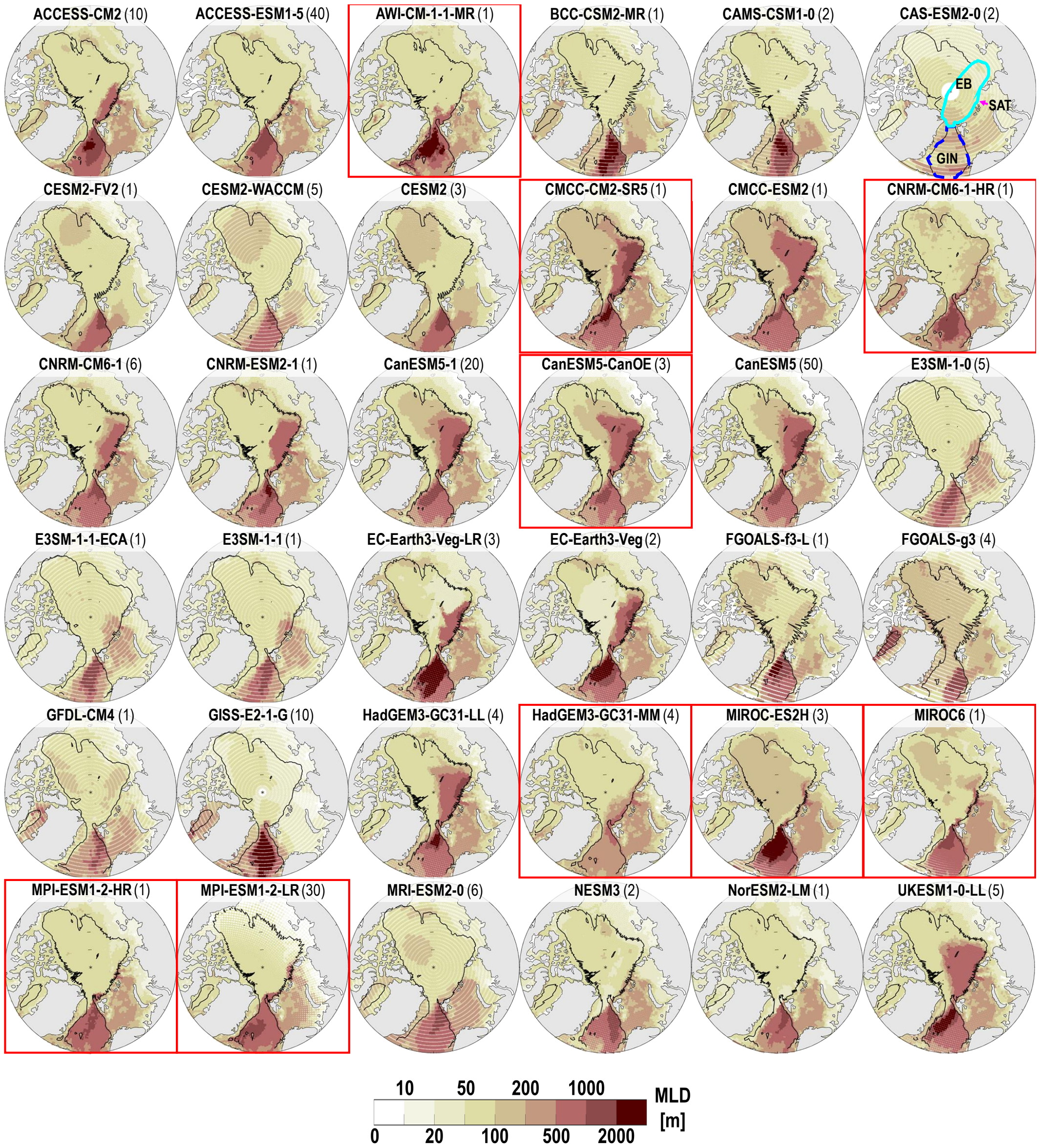

Figure 1. Maximum MLD over January 2015 - December 2100 (shading, logarithmic scale) for SSP5-8.5 at each grid cell. For each model, parentheses indicate the number of ensemble members available (see supp. Tables S1 and S2); when this number is larger than one, the figure shows the ensemble median. Black contours are the 1000 m isobath. Red squares indicate the models for which all ensemble members have deep convection in the Arctic already over 2015-2023, and which are therefore not considered for further analysis. Locations discussed in the manuscript are indicated on the top-right panel: Cyan contour "EB" is the Eurasian Basin, indigo dashed contour "GIN" is the Nordic Seas, and magenta arrow "SAT" is the St Anna Trough.

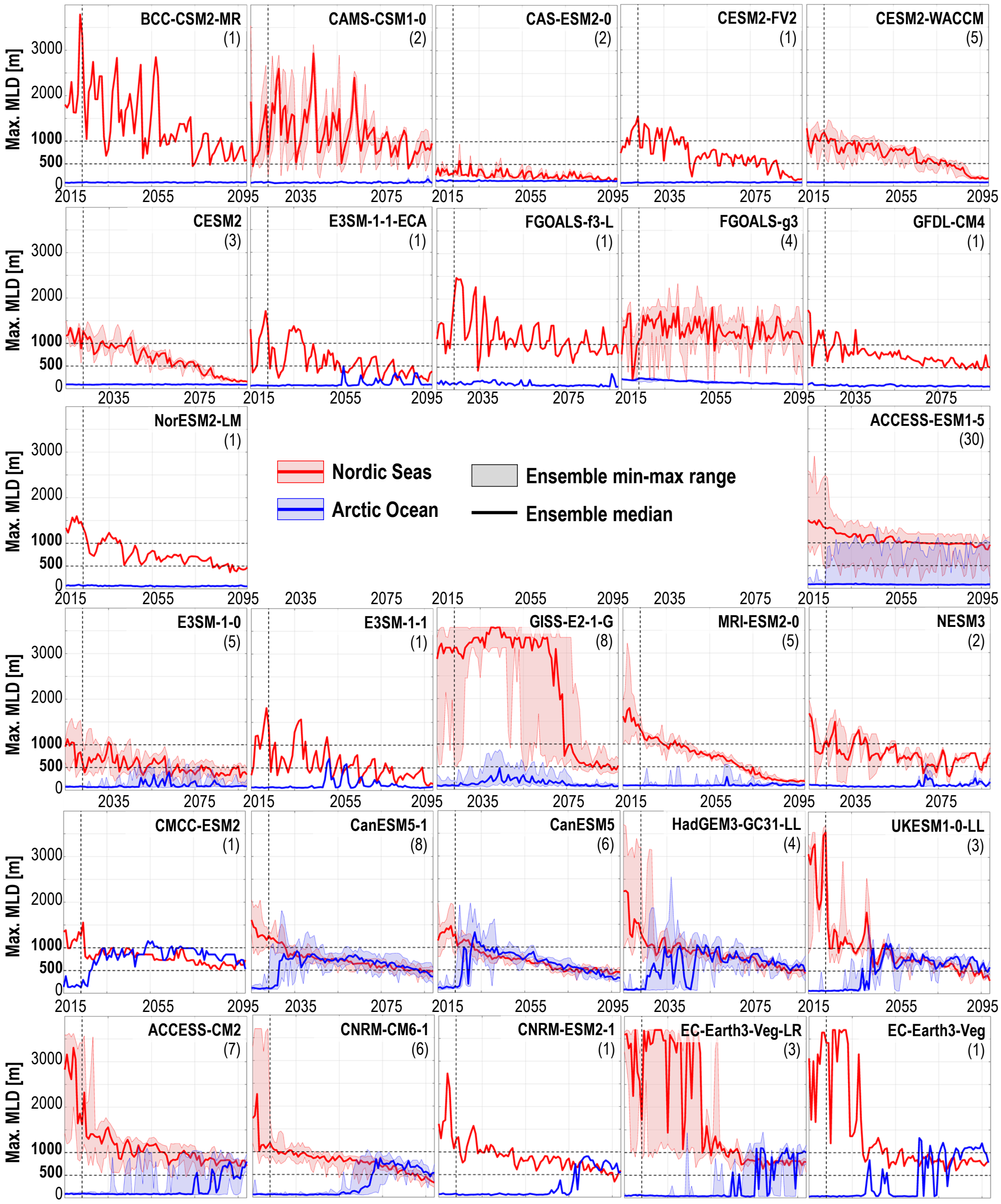

Figure 2. For the 27 models that do not have deep convection in the Arctic over 2015-2023, time series of their yearly maximum MLD in the Nordic Seas (red) and in the Arctic Ocean (blue). For each model, parentheses indicate the number of ensemble members remaining; when this number is larger than one, the figure shows the range across these ensemble members (shading) and the ensemble median (thick line). Horizontal black lines indicate the 500 and 1000 m MLD thresholds, indicative of deep convection. Vertical black line is the year 2023. Models are ordered based on their Arctic behaviour: First two rows and NorESM2-LM, no Arctic deep convection; ACCESS-ESM1-5 and fourth row, rare convection, by the middle of the run; fifth row, convection peaks by the middle of the run and then declines; bottom row, convection starts late in the run.

What causes these different behaviours then? Ensemble members usually have a consistent behaviour (the shading on Fig. 2 usually agrees with the thick line), and models of the same family tend to belong to the same group. The only exception are the two ACCESS models (Fig. 2), but there are large differences in their designs and implemented schemes of relevance for polar regions and deep convection in particular, as discussed in Mohrmann et al. (2021). In the next section, we investigate in more details what could be causing the different Arctic deep convection behaviours, starting with model designs.

3.2 Potential causes: High sensitivity, salinification, stronger winds

The models that have deep convection in the Arctic usually have it at the same location (Fig. 1), but the temporal evolution of the mixed layer yields four groups of models (Fig. 2). Models of the same family tend to belong to the same group; these usually have similar biases and sensitivities, so we first investigate these.

The mean temperature in the Eurasian Basin at the beginning of the SSP5-8.5 run and subsequent MLD are positively correlated across models (Fig. 3, left): the warmer at the beginning of the run, the deeper the mixed layers (correlation 0.40), the more often (0.33). The same results are obtained when considering the ensemble members separately, as on Fig. 3, and the ensemble mean (not shown). There is no such relationship with salinity. The relationship with temperature persists throughout the run, so that the models that warm the most (Fig. 3, centre), and quite strikingly, those for which the ocean surface warms more in the Arctic than in the midlatitudes (Fig. 3, right) are the ones with the deepest MLD, the most often. The models belonging to the E3SM project (cyan contour, Fig. 3) are the exception: They warm strongly but convect rarely or not at all, which could be because their design is very different from that of the other CMIP models (Golaz et al., 2019).

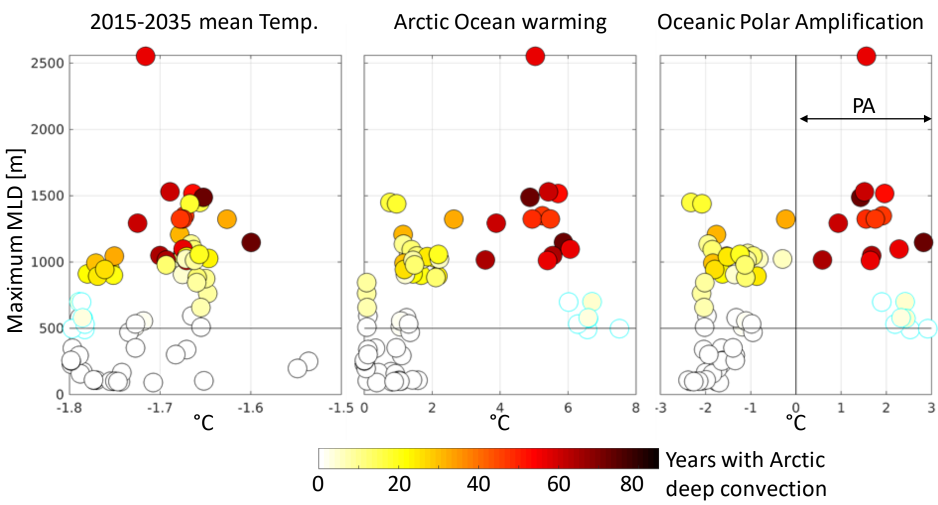

Figure 3. For each model and each ensemble member that were not removed from the study, scatter plot between the maximum MLD over 2015-2100, over the Eurasian Basin, and: Left, the mean ocean surface temperature over the Eurasian basin, over the first 20 years of the run; Centre, the Arctic warming, i.e. the mean ocean surface over the last 20 years of the run minus that of the first 20 years; Right, that Arctic warming minus that of the tropical to mid-latitudes, which we call "Oceanic Polar Amplification". See Methods for latitude definitions. Colors indicate the number of years where the Arctic MLD exceeds the 500 m threshold, as per supp. Table S3. Cyan contours highlight the E3SM project models (see text).

Although the salinity at the beginning of the run has no relationship with the maximum MLD reached, salinity variations are strongly correlated to MLD variations (Table 1) and the correlation differs depending on the Arctic deep convection behaviour of the models. Most models that have no deep convection in the Arctic or rarely (first two blocks) have a positive correlation between their March MLD and the surface salinity one month before, and a negative correlation with the ocean surface temperature: These models have shallower MLD when they are fresher and warmer (as expected in a changing Arctic, e.g. Peralta-Ferriz and Woodgate (2015)). In the two groups of convecting models, especially so in the convective region (lines “DC”, Table 1), the correlation is positive with salinity and temperature: deeper MLD are associated with saltier and warmer surface waters the month prior. Of the four drivers investigated, the MLD is most strongly correlated with salinity or temperature for the vast majority of models (bold fonts on Table 1); usually it is strongest with the salinity for the non or rarely convecting models, and with the temperature for the convecting ones (as expected from the previous paragraph). Correlations are similar when considering possible drivers the summer before (not shown, see also supp. Text S1). Unsurprisingly, the possible drivers are not independent (supp. Table S4). The correlations between temperature and salinity, and between salinity and either sea ice concentration or wind speed are of different signs depending on the convecting behaviour, suggesting that different processes and/or water masses are involved (see next subsection).

From the correlation analysis, we suspect that the models with no or rare deep convection may be freshening, while those with convection may become saltier. A trend analysis confirms this hypothesis:

1. The models with no deep convection in the Arctic become fresher throughout the Arctic, throughout the run (Fig. 4, first two lines). Their trends are rather weak compared to the other model groups. Their winter sea ice does not retreat far into the Arctic, even by the end of the run (supp. Fig. S1).
2. The models with deep convection at the middle of the run, be it rarely (Fig. 4, lines 3 and 4) or peaking and declining (lines 5 and 6) exhibit similar trends. Their ocean surface becomes saltier and winds stronger at the location where mixed layers deepen in the first half of the run; they freshen in the rest of the Arctic. In the second half, when mixed layers are shallow again, the ocean surface freshens everywhere. The main difference between these two groups is in their sea ice trends, with the models whose convection peaks and slowly declines having no winter sea ice by the end of the run (supp. Fig. S1).
3. The models that convect at the end of the run (Fig. 4, last two lines) have the opposite salinity trends: first a freshening, then a salinification in the region where mixed layers deepen, along with stronger winds.

Note that the trend patterns are similar in summer (supp. Figure S2), indicating that the changes occur year round. The conclusion is that Arctic deep convection is associated with both a saltier ocean surface and stronger winds. Stronger winds alone are not enough (see e.g. the no deep convection group, second half of the run).

3.3 On causality

Deep convection in the Nordic Seas ceases for all CMIP6 models, but rarely emerges as a stable feature in the Arctic. In the Arctic, we find that deep convection is associated with a saltier ocean surface and stronger wind speeds, and is most intense and durable in the models with the strongest sensitivity, especially so if the Arctic warms more than the mid-latitudes. But are the trends we observe the causes or consequences of deep convection?

Table 1. Correlation coefficient and its standard deviation for each model between the March MLD and the previous February ocean surface salinity (S), ocean surface temperature (T), sea ice concentration (SIC) and surface wind speed (Wind), for the Eurasian basin (EB) and, for the models with Arctic deep convection, where the maximum MLD of Fig. 1 exceeds 500 m (DC). Models are ordered based on their Arctic deep convection behaviour, as per Fig. 2. Only correlations significant at 95% are shown; for models with more than one ensemble member, median of the correlations of the dominating sign. For models with more than one ensemble member, standard deviation is the across-ensemble spread; spatial spread otherwise. Bold fonts highlight the maximum correlation for each model. Correlation between parameters is shown in supp. Table S4.

Model	Region	MLD vs S	MLD vs T	MLD vs SIC	MLD vs Wind
BCC-CSM2-MR	EB	**0.47** ± 0.15	-0.41 ± 0.17	-0.28 ± 0.35	0.23 ± 0.24
CAMS-CSM1-0	EB	**0.54** ± 0.02	-0.49 ± 0.00	-0.35 ± 0.03	0.29 ± 0.04
CAS-ESM2-0	EB	**0.59** ± 0.04	-0.49 ± 0.06	0.51 ± 0.06	-0.24 ± 0.01
CESM2-FV2	EB	-0.29 ± 0.37	0.22 ± 0.37	-0.23 ± 0.16	**0.31** ± 0.08
CESM2-WACCM	EB	**0.34** ± 0.06	-0.31 ± 0.05	0.25 ± 0.06	0.30 ± 0.02
CESM2	EB	**0.32** ± 0.08	-0.30 ± 0.04	0.30 ± 0.05	0.32 ± 0.02
E3SM-1-1-ECA	EB	0.43 ± 0.14	-0.45 ± 0.47	**0.60** ± 0.41	-0.31 ± 0.31
FGOALS-f3-L	EB	0.51 ± 0.21	-0.51 ± 0.10	**0.59** ± 0.13	-0.26 ± 0.16
FGOALS-g3	EB	**0.89** ± 0.01	-0.28 ± 0.02	0.46 ± 0.04	0.28 ± 0.04
GFDL-CM4	EB	**0.39** ± 0.31	-0.25 ± 0.39	-0.31 ± 0.29	-0.27 ± 0.19
NorESM2-LM	EB	-0.28 ± 0.36	**0.36** ± 0.31	-0.26 ± 0.22	0.26 ± 0.16
ACCESS-ESM1-5	EB	-	-	-	-
	DC	-	-	-	-
E3SM-1-0	EB	0.38 ± 0.01	-0.50 ± 0.12	**0.51** ± 0.05	-0.26 ± 0.04
	DC	-	-	-	-
E3SM-1-1	EB	0.39 ± 0.14	-0.41 ± 0.27	**0.55** ± 0.28	-0.28 ± 0.27
	DC	**0.51** ± 0.01	0.34 ± 0.02	-0.29 ± 0.01	0.43 ± 0.01
GISS-E2-1-G	EB	**0.61** ± 0.03	-0.59 ± 0.04	0.35 ± 0.08	0.27 ± 0.04
	DC	0.48 ± 0.14	**0.67** ± 0.10	-0.37 ± 0.08	0.32 ± 0.09
MRI-ESM2-0	EB	**0.62** ± 0.03	-0.40 ± 0.04	0.29 ± 0.05	0.31 ± 0.04
	DC	-	-	-	-
NESM3	EB	**-0.31** ± 0.01	-0.24 ± 0.01	-0.24 ± 0.01	-0.23 ± 0.02
	DC	-	-	-	-
CMCC-ESM2	EB	**0.59** ± 0.45	0.45 ± 0.37	-0.31 ± 0.43	0.28 ± 0.25
	DC	**0.65** ± 0.11	0.49 ± 0.14	-0.40 ± 0.12	0.29 ± 0.08
CanESM5-1	EB	**0.50** ± 0.03	0.39 ± 0.11	-0.46 ± 0.08	0.36 ± 0.05
	DC	**0.52** ± 0.04	0.39 ± 0.12	-0.48 ± 0.08	0.36 ± 0.05
CanESM5	EB	-	-	-	-
	DC	-	-	-	-
HadGEM3-GC31-LL	EB	-	-	-	-
	DC	-	-	-	-
UKESM1-0-LL	EB	0.56 ± 0.02	**0.64** ± 0.01	-0.51 ± 0.02	0.43 ± 0.05
	DC	0.57 ± 0.02	**0.65** ± 0.02	-0.54 ± 0.02	0.43 ± 0.06
ACCESS-CM2	EB	0.40 ± 0.05	**0.45** ± 0.19	-0.28 ± 0.08	0.28 ± 0.03
	DC	0.44 ± 0.05	**0.76** ± 0.11	-0.51 ± 0.08	0.34 ± 0.04
CNRM-CM6-1	EB	0.40 ± 0.07	**0.73** ± 0.05	-0.62 ± 0.05	0.38 ± 0.05
	DC	0.46 ± 0.05	**0.77** ± 0.08	-0.68 ± 0.05	0.47 ± 0.04
CNRM-ESM2-1	EB	0.44 ± 0.15	**0.77** ± 0.48	-0.53 ± 0.31	0.43 ± 0.13
	DC	0.46 ± 0.09	**0.86** ± 0.10	-0.57 ± 0.16	0.47 ± 0.12
EC-Earth3-Veg-LR	EB	0.44 ± 0.04	**0.85** ± 0.05	-0.55 ± 0.12	0.35 ± 0.05
	DC	0.45 ± 0.03	**0.90** ± 0.05	-0.66 ± 0.04	0.39 ± 0.03
EC-Earth3-Veg	EB	0.47 ± 0.13	**0.85** ± 0.43	-0.49 ± 0.17	0.32 ± 0.09
	DC	0.52 ± 0.10	**0.89** ± 0.07	-0.53 ± 0.12	0.34 ± 0.10

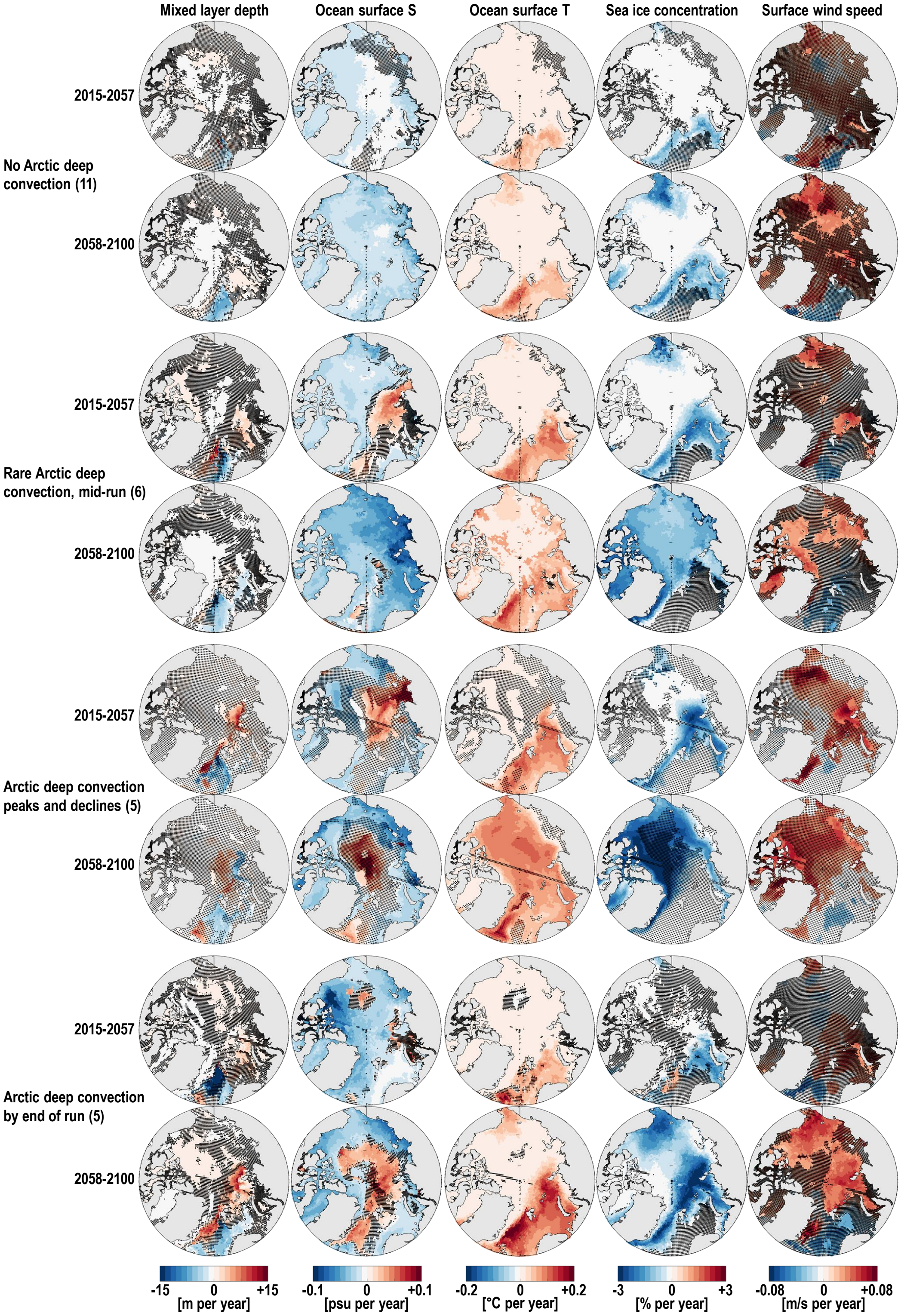

Figure 4. Composite trends based on the models' Arctic deep convection behaviours of Fig. 2, for the first half of the 21st century run (top rows) and the second half (bottom rows), in March MLD (first column), and February ocean surface salinity (S, second), ocean surface temperature (T, third), sea ice concentration (fourth) and surface wind speed (last). Behaviours are described to the left of the figure, and number of models for each behaviour is given in parentheses. For each panel, stippling indicates non-significant trends and/or model disagreement regarding the trend's sign. Straight stippled lines across the North Pole are an artefact of the necessary interpolation.

The Arctic Ocean surface warming and sea ice loss trends, year-round, are to be expected in a warming world (IPCC, 2019). Besides, they are consistent across-models, regardless of their convecting behaviour in the Arctic. Similarly, the models' sensitivities have been attributed to their representation of cloud cover and cloud albedo (Zelinka et al., 2020), which locally can be affected by the heat and moisture fluxes from deep convection (Monroe et al., 2021), but in coarse models is more likely due to each model's cloud parameterisation (Zelinka et al., 2020). Finally, changes in both "normal" winds (Screen et al., 2018) and cyclones (Rinke et al., 2017) are most commonly attributed to the large scale atmospheric circulation, although local changes at the boundary layer because of sea ice loss may accelerate winds (DuVivier et al., accepted).

As for the salinity, Lique et al. (2018) attributed the freshening - salinification dipoles to changes in the large scale oceanic circulation, whereas Davis et al. (2016) attributed the increased salinity to increased vertical mixing. That is, for the former it drives Arctic deep convection; for the latter, it is a consequence of it. A local change in salinity can be also caused by enhanced sea ice formation, but we find a negative correlation between winter salinity and sea ice concentration among the convecting models (supp. Table S4), which makes this causality unlikely given that the sea ice maps do not seem to exhibit polynyas (supp. Fig. S1). Salinification taking place year-round, at the same location as winds are increasing, would instead suggest an enhanced sea ice drift, but the detailed sea ice mass budget of Keen et al. (2021) shows the opposite.

Alternatively, the impact of deeper MLD on the Arctic surface salinity will depend on each model's Atlantic layer. Heuzé et al. (2023) showed strong biases in that layer, with most models having it too deep. Khosravi et al. (2022) further showed that there is no consistency across CMIP6 models regarding the changes of that Atlantic layer during SSP5-8.5, which Muilwijk et al. (2023) linked to the models' lack of consensus regarding future changes in stratification in the Eurasian basin and in Atlantification. As stratification and deep convection are intimately linked, this is another feedback that makes the causality uncertain. One can unfortunately not study the onset of deep convection in more details without higher temporal resolution output. Most likely, the causality will be different for each model, based on their choices of parameterisation in the atmosphere (Zelinka et al., 2020), sea ice (Keen et al., 2021) and ocean (Heuzé et al., 2023), and even their definition of the mixed layer (Griffies et al., 2016). But individual model studies require access to each model's code, and are way beyond the scope of this paper.

4 Conclusions

We used all CMIP6 models and all their ensemble members for which the mixed layer depth and its potential drivers the surface salinity, temperature, sea ice concentration, and wind speed, were available for the strongest warming scenario SSP5-8.5. After removing the ensemble members that had spurious Arctic deep convection (defined, as in Lique et al. (2018), as MLD deeper than 500 m) over 2015-2023, we were left with 27 models, of which 11 had no deep convection in the Arctic over 2015 - 2100; 6 that had it extremely rarely (usually 4 years or fewer), by the middle of the run; 5 for which deep convection peaked in the first half of the run and then declined and disappeared; and 5 in which deep convection emerged in the Arctic in the second half of the run and still convected in 2100. All models exhibit a cessation of deep convection in the Nordic Seas, showing that deep convection in the Arctic is not simply a northward migration of the Nordic Seas ventilation. The Arctic MLD was most strongly correlated with the surface salinity, and the sign of this correlation depended on whether the model convected or not. Similarly, when and where the models are not convecting, their surface salinity freshens; at the location where they do, when they do, it becomes saltier and surface winds are increasing. Neither the exact mechanism triggering deep convection nor the direction of the causality between deep convection and that compound salinification and wind event can be investigated in more details with CMIP6 monthly output. The fact that models of the same family have the same convection behaviour; that the depth and frequency of the maximum MLD is strongly correlated to early-run biases and sensitivity; and that the other processes involved have been linked to individual model parameterisations (Zelinka et al., 2020; Keen et al., 2021; Muilwijk et al., 2023) suggest that the trigger for Arctic deep convection is model-specific, and its determination requires in-depth sensitivity studies for each model. Such in-depth investigation could also lead to model improvement. CMIP6 models consistently exaggerate deep convection both in the North Atlantic and in the Southern Ocean (Heuzé, 2021). Understanding why they have no such consensus in the Arctic could hold the key to a more realistic representation of mixing, globally.

5 Open Research

All CMIP6 data are freely available via the Earth Grid System Federation. For this paper, we primarily used the German Climate Computing Centre (DKRZ) node `https://esgf-data.dkrz.de/search/cmip6-dkrz/`.

Acknowledgments

CH is funded by the Swedish Research Council (dnr 2018-03859).

References

Davis, P., Lique, C., Johnson, H., & Guthrie, J. (2016). Competing Effects of Elevated Vertical Mixing and Increased Freshwater Input on the Stratification and Sea Ice Cover in a Changing Arctic Ocean . *Journal of Physical Oceanography*, *46*, 1531-1553. doi: 10.1175/JPO-D-15-0174.1

DuVivier, A., Holland, M., Vavrus, S., Landrum, L., Shields, C., & Thaker, R. (accepted). Arctic sea ice loss drives increasing Arctic wind speeds with combined impact on surface roughness and boundary layer stability. *Journal of Geophysical Research: Atmospheres*. doi: 10.21203/rs.3.rs-2210756/v1

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

Golaz, J., Caldwell, P., Van Roekel, L., Petersen, M., Tang, Q., Wolfe, J., ... Baldwin, S. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*, 2089-2129. doi: 10.1029/2018MS001603

Griffies, S., Danabasoglu, G., Durack, P., Adcroft, A., Balaji, V., Böning, C., ... Fox-Kemper, B. (2016). OMIP contribution to CMIP6: Experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project. *Geoscientific Model Development*, *9*, 3231-3296. doi: 10.5194/gmd-9-3231-2016

Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Science*. doi: 10.5194/os-17-59-2021

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

IPCC. (2019). *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (H. Pörtner et al., Eds.). Cambridge University Press.

Keen, A., Blockley, E., Bailey, D., Boldingh Debernard, J., Bushuk, M., Delhaye, S., ... Ponsoni, L. (2021). An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *The Cryosphere*, *15*, 951-982. doi: 10.5194/tc-15-951-2021

Khosravi, N., Wang, Q., Koldunov, N., Hinrichs, C., Semmler, T., Danilov, S., & Jung, T. (2022). The Arctic Ocean in CMIP6 models: Biases and projected changes in temperature and salinity. *Earth's Future*, *10*, e2021EF002282. doi: 10.1029/2021EF002282

Lique, C., Johnson, H., & Plancherel, Y. (2018). Emergence of deep convection in the Arctic Ocean under a warming climate. *Climate Dynamics*, *50*, 3833-3847. doi: 10.1007/s00382-017-3849-9

Mallett, R., Stroeve, J., Tsamados, M., Landy, J., Willatt, R., Nandan, V., & Liston, G. E. (2021). Faster decline and higher variability in the sea ice thickness of the marginal Arctic seas when accounting for dynamic snow cover. *The Cryosphere*, *15*, 2429–2450. doi: 10.5194/tc-15-2429-2021

Meier, W., & Stroeve, J. (2022). An updated assessment of the changing Arctic sea ice cover. *Oceanography*, *35*, 10-19.

Mohrmann, M., Heuzé, C., & Swart, S. (2021). Southern Ocean polynyas in CMIP6 models. *The Cryosphere*, *15*, 4281-4313. doi: 10.5194/tc-15-4281-2021

Monroe, E., Taylor, P., & Boisvert, L. (2021). Arctic cloud response to a perturbation in sea ice concentration: The North Water polynya. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD034409. doi: 10.1029/2020JD034409

Muilwijk, M., Nummelin, A., Heuzé, C., Polyakov, I., Zanowski, H., & Smedsrud, L. (2023). Divergence in climate model projections of future Arctic Atlantification. *Journal of Climate*, *36*, 1727-1748. doi: 10.1175/JCLI-D-22-0349.1

Onarheim, I., Eldevik, T., Smedsrud, L., & Stroeve, J. (2018). Seasonal and regional manifestation of Arctic sea ice loss. *Journal of Climate*, *31*, 4917-4932. doi: 10.1175/JCLI-D-17-0427.1

O'Neill, B., Tebaldi, C., Van Vuuren, D., Eyring, V., Friedlingstein, P., Hurtt, G., … Meehl, G. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*, 3461–3482. doi: 10.5194/gmd-9-3461-2016

Peralta-Ferriz, C., & Woodgate, R. (2015). Seasonal and interannual variability of pan-Arctic surface mixed layer properties from 1979 to 2012 from hydrographic data, and the dominance of stratification for multiyear mixed layer depth shoaling. *Progress in Oceanography*, *134*, 19-53. doi: 10.1016/j.pocean.2014.12.005

Peterson, A. (2018). Observations of brine plumes below melting Arctic sea ice. *Ocean Science*, *14*, 127–138. doi: 10.5194/os-14-127-2018

Polyakov, I., Pnyushkov, A., Alkire, M., Ashik, I., Baumann, T., Carmack, E., … Krishfield, R. (2017). Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science*, *356*, 285-291. doi: 10.1126/science.aai8204

Rinke, A., Maturilli, M., Graham, R., Matthes, H., Handorf, D., Cohen, L., … Moore, J. (2017). Extreme cyclone events in the Arctic: Wintertime variability and trends. *Environmental Research Letters*, *12*, 094006. doi: 10.1088/1748-9326/aa7def

Schulz, K., Koenig, Z., Muilwijk, M., Bauch, D., Hoppe, C., Droste, E., … Granskog, M. (under review). The Eurasian Arctic Ocean along the MOSAiC drift (2019-2020): An interdisciplinary perspective on properties and processes. *Elementa*. doi: 10.31223/X5TT2W

Screen, J., Deser, C., Smith, D., Zhang, X., Blackport, R., Kushner, P., … Sun, L. (2018). Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models. *Nature Geosci*, *11*, 155–163. doi: 10.1038/s41561-018-0059-y

Timmermans, M., & Marshall, J. (2020). Understanding Arctic Ocean circulation: A review of ocean dynamics in a changing climate. *Journal of Geophysical Research: Oceans*, *125*, e2018JC014378. doi: 10.1029/2018JC014378

Zelinka, M., Myers, T., McCoy, D., Po-Chedley, S., Caldwell, P., Ceppi, P., …

Taylor, K. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*, e2019GL085782. doi: 10.1029/2019GL085782

No emergence of deep convection in the Arctic Ocean across CMIP6 models

Céline Heuzé[1], and Hailong Liu[2]

[1]Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden
[2]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China

Key Points:

- Oceanic deep convection does not emerge and persist in the Arctic in the majority of CMIP6 models, despite a cessation in the Nordic Seas.
- Arctic deep convection occurs only when both surface salinity and winds are increasing, year round, yet most models are freshening.
- The models with the strongest sensitivity, especially with an oceanic polar amplification, have the deepest Arctic mixed layers, most often.

Corresponding author: Céline Heuzé, `celine.heuze@gu.se`

Abstract

As sea ice disappears, the emergence of open ocean deep convection in the Arctic has been suggested. Here, using 36 state-of-the-art climate models and up to 50 ensemble members per model, we show that Arctic deep convection is rare even under the strongest warming scenario. Only 5 models have somewhat permanent convection by 2100, while 11 have had convection by the middle of the run. For all, the deepest mixed layers are in the Eurasian basin, by St Anna Trough. When the models convect, that region undergoes a salinification and increasing wind speeds; it is freshening otherwise. We discuss the causality and potential reasons for the opposite trends. Given the model's different parameterisations, and given that the ensemble members that convect the deepest, most often, are those with the strongest sensitivity, we conclude that differences in deep convection are most likely linked to the model formulation.

Plain Language Summary

Both observations and modelling simulations suggest that deep vertical mixing (or deep convection) in winter may become the new normal in the Arctic as sea ice disappears. These simulations are often done using only one model, so here we used all models available that participated in the Climate Model Intercomparison Project phase 6, for the strongest warming scenario. We show that after removing those that are already inaccurate in the present, and even with a restrictive threshold, most models have no deep convection in the Arctic, or extremely rarely. Only 5 still had deep convection by the time the run finishes in 2100. We investigated the possible links between deep convection and surface salinity, surface temperature, sea ice concentration and surface wind speeds, and found that the salinity was most important. Deep convection regions and periods are associated with a saltier, windier surface, while the rest of the Arctic and/or rest of the run freshens. Causality is unclear; we need higher resolution than monthly output. Similar behaviours within model families, a strong link to the model sensitivity, and cited work make us conclude that ultimately, the differences are probably caused primarily by the different model designs.

1 Introduction

The Arctic Ocean is changing. The resulting reduction in Arctic sea ice extent and thickness (Mallett et al., 2021; Meier & Stroeve, 2022) could enhance vertical mixing: More brine may be rejected year-round as younger, saltier ice desalinates (Peterson, 2018) or as sea ice reforms in winter over the now seasonal-ice areas (Onarheim et al., 2018), while the ice-freed regions may become more susceptible to wind stirring (Timmermans & Marshall, 2020). In the Eurasian Arctic, the process known as Atlantification (Polyakov et al., 2017), whereby warm water of Atlantic origin penetrates further into the Arctic, may be further weakening the stratification and enhancing sea ice melt. These led Polyakov et al. (2017) to hypothesise that the Arctic may start exhibiting deep convection in winter, a result found by Lique et al. (2018) in the 4x CO_2 scenario using the model HiGEM.

However, this hypothesis is so far not confirmed. Peralta-Ferriz and Woodgate (2015) found that a deepening of mixed layers in the Arctic is unlikely, since stratification greatly dominates over the wind effect. The latest observations in the Eurasian Arctic (Schulz et al., under review) yielded mixed layers no deeper than 130 m, even in winter. Besides, in a recent Arctic study using models that participated in the Climate Model Intercomparison Project phase 6 or CMIP6 (Eyring et al., 2016), Muilwijk et al. (2023) showed that there was no agreement among models regarding future stratification and the effect of Atlantification under the strongest warming scenario (SSP5-8.5, O'Neill et al. (2016)). This suggests that HiGEM's deep convection in the Arctic may be a model artefact rather than the future of the Arctic.

We here determine whether deep convection emerges in the Arctic in the future scenario SSP5-8.5 using all CMIP6 models and all their ensemble members for which the mixed layer depth output was available, as described in section 2. In section 3, we detail the spatial and temporal patterns of future Arctic mixed layers and discuss possible reasons for these, focusing on the model biases and their trends in surface properties, sea ice and winds. We conclude in section 4.

2 Data and Methods

2.1 CMIP6 data

To investigate deep convection and its potential drivers in the future Arctic, we use all CMIP6 models and all their ensemble members that had monthly mixed layer depth ("mlotst"), ocean surface salinity ("sos") and temperature ("tos"), sea ice concentration ("siconc") and surface wind speed ("sfcWind") available on any of the Earth System Grid Federation (ESGF) nodes for the future scenario SSP5-8.5, for January 2015 to December 2100. The models and their ensemble members are listed in supp. Tables S1 and S2.

For AWI-CM-1-1-MR and GISS-E2-1-G, we used the sea ice thickness ("sivol") because sea ice concentration was not available. For CAMS-CSM1-0, we generated the surface salinity from the full-depth salinity ("so") as the former was not available; we used the salinity of the shallowest level, 5 m depth. Similarly, we used the models' bathymetry ("deptho") when it was available, but had to generate it from the full-depth salinity for 11 models, as the last level with salinity data. Finally, for CESM2-WACCM, GFDL-CM4 and MRI-ESM2-0, several grid types were available and we chose for simplicity the regularised grid ("gr"). For the other models, we took the one grid type available; see supp. Tables S1 and S2.

2.2 Methods

The thresholds and choices of this subsection are discussed in supplementary text S1. In agreement with Lique et al. (2018), we consider that there is deep convection in the Arctic if the mixed layer depth (MLD) exceeds 500 m. We here do not quantify overturned volumes. We only perform a binary detection of deep convection, so we select the overall maximum MLD, in space or time depending on the analysis. The ensemble members that exhibit deep convection in the Arctic over the observed part of SSP5-8.5 (2015-2023) are shown on the first figure and subsequently removed from the study, as they are already inaccurate at the beginning of the run. Consequently, of the originally 36 models, 27 remain for most of the analysis. We do not take biases in Nordic Seas MLD into account for model selection, as all CMIP6 models have spurious deep convection there (Heuzé, 2021).

The sea ice edge (supp. Fig S1) is detected as the contour of 15% concentration or 10 cm thickness, averaged over 2040-2060 and 2080-2100. We define the Eurasian basin as the region north of 80°N, longitudes 20°W - 140°E, deeper than 1000 m, and the Nordic Seas as the region of latitudes 66 - 80°N, longitudes 30°W - 20°, deeper than 1000 m - see contours on Fig 1. We perform across-model correlations by comparing, for each model and each ensemble member, their maximum MLD and number of year where the MLD exceeds 500 m in the Eurasian basin, to:

1. Average surface temperature and salinity in the Eurasian basin over the first 20 years of SSP5-8.5;
2. Tropical and Mid-latitude (60°S - 60°N) and Arctic (north of 75°N) warming, as difference between the 2080-2100 and 2015-2035 average ocean surface temperature, which we use as a proxy for the model sensitivity. We also define the oceanic

polar amplification as the difference between the Arctic and tropical/mid-latitude warming.

Besides, we compute correlations and trends for each grid cell, after interpolating all the parameters onto each model's mlotst grid. The correlations and trends are computed for each ensemble member separately, and averaged afterwards. We consider that there is ensemble agreement if more than 50% of the ensemble members have a significant correlation/trend (determined using a t-test at 95% significance) of the same sign; we then present the median value and its ensemble-spread. Finally, we present the composite trends after grouping the models based on their Arctic deep convection behaviour, as described in the result section.

3 Results and Discussion

3.1 Arctic deep convection is rare, restricted both in space and time

The maximum mixed layer depth reached in the Arctic over 2015-2100 varies strongly across models (Fig. 1, note the logarithmic scale). The value does not exceed 100 m for some, such as CAMS-CSM1-0, while others such as CMCC-CM2-SR5 exceed 2000 m in the majority of the Eurasian basin. All models with deep MLD agree that the deepest values are in the Eurasian basin, most commonly by St Anna Trough. Nine models have MLD exceeding the 500 m threshold for deep convection already over 2015 to present for all their ensemble members (red squares on Fig. 1). Interestingly, these are not only by St Anna Trough but also north of Svalbard, suggesting that for these models the Nordic Sea deep convection area extends too far north, most likely following the sea ice edge (shown on supp. Fig. S1). All models presented on Fig. 1 have spuriously deep MLD in the Nordic Seas.

After removing the models and ensemble members that are unrealistic in the present-day regarding their Arctic MLD (asterisks in supp Tables S1 and S2), 27 models remain. The temporal evolution of their MLD reveals four groups of models (Fig. 2 and supp Table S3):

1. The first 11 models have no Arctic deep convection during the entire run. The Arctic MLD time series (blue lines, Fig. 2) are mostly flat, with no year where the MLD exceeds 500 m, regardless of the ensemble member. The maximum MLD across these models and their ensemble members is often of the order of 100 m, i.e. like currently observed in the Arctic (Schulz et al., under review).
2. Six models have deep convection in the Arctic on rare occasions, in the middle of the run. The maximum number of years with deep convection for this group is 17 out of 86 (supp Table S3), but is most often 4 or fewer.
3. Five models also have deep convection in the Arctic in the middle of the run, more often. It starts by 2030, peaks in the first half of the run, and then declines slowly, with the ensemble average back under the 500 m threshold by the end of the run.
4. The last 5 models start convecting in the second half of the run, by 2070, and appear "stably" convecting at the end of the run. Unfortunately, only one ensemble member of one of the models (ACCESS-CM2) is available beyond 2100, so we cannot tell whether deep convection in these models would also decline later.

The Nordic Seas MLD (red lines, Fig. 2) falls below 1000 m for all models, and even below 500 m for two thirds of them. The consistent cessation of Nordic Seas deep convection is not related to the models' behaviour in the Arctic. Therefore, unlike suggested by Lique et al. (2018), deep convection does not migrate to the Arctic in response to its cessation further south.

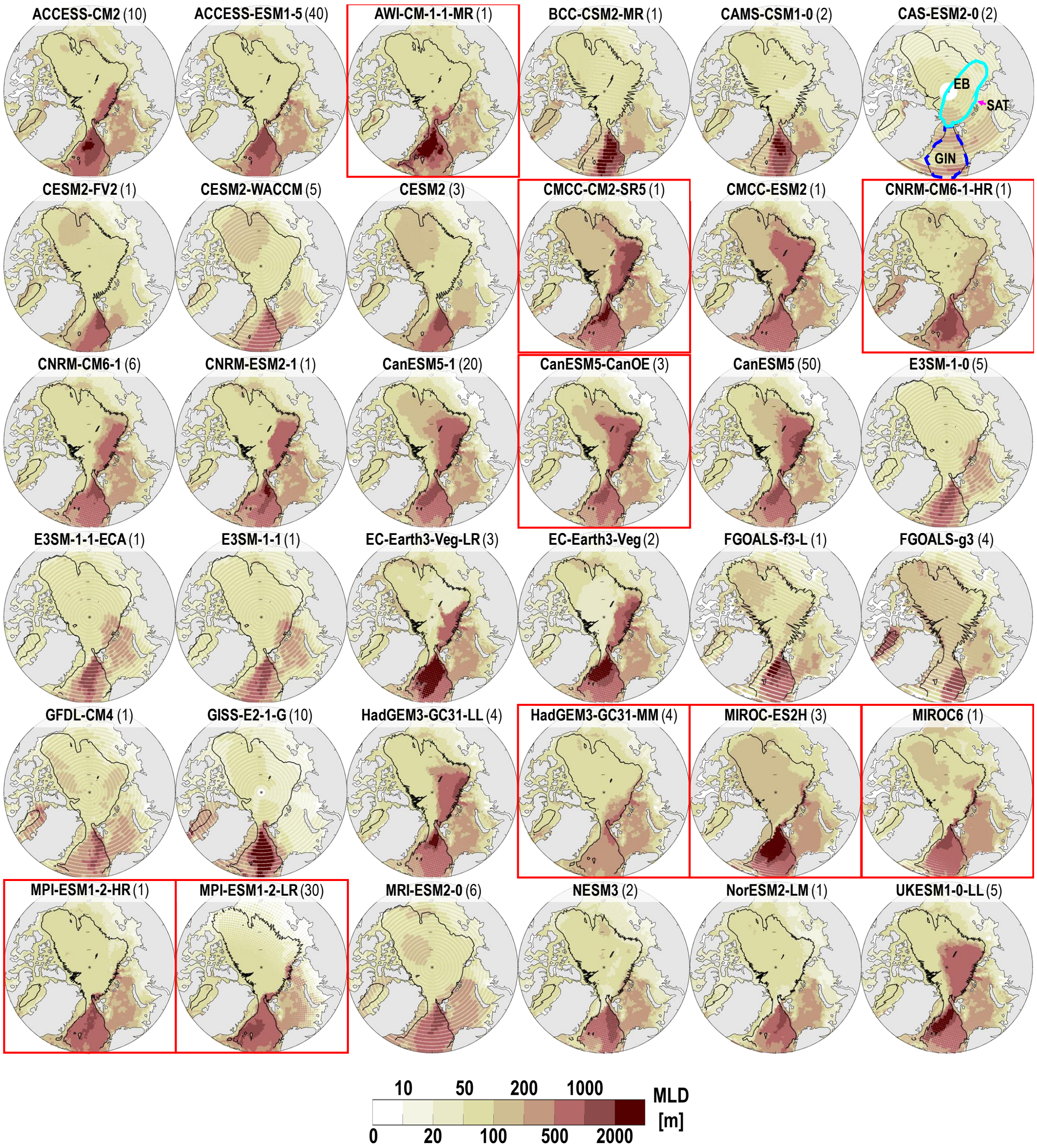

Figure 1. Maximum MLD over January 2015 - December 2100 (shading, logarithmic scale) for SSP5-8.5 at each grid cell. For each model, parentheses indicate the number of ensemble members available (see supp. Tables S1 and S2); when this number is larger than one, the figure shows the ensemble median. Black contours are the 1000 m isobath. Red squares indicate the models for which all ensemble members have deep convection in the Arctic already over 2015-2023, and which are therefore not considered for further analysis. Locations discussed in the manuscript are indicated on the top-right panel: Cyan contour "EB" is the Eurasian Basin, indigo dashed contour "GIN" is the Nordic Seas, and magenta arrow "SAT" is the St Anna Trough.

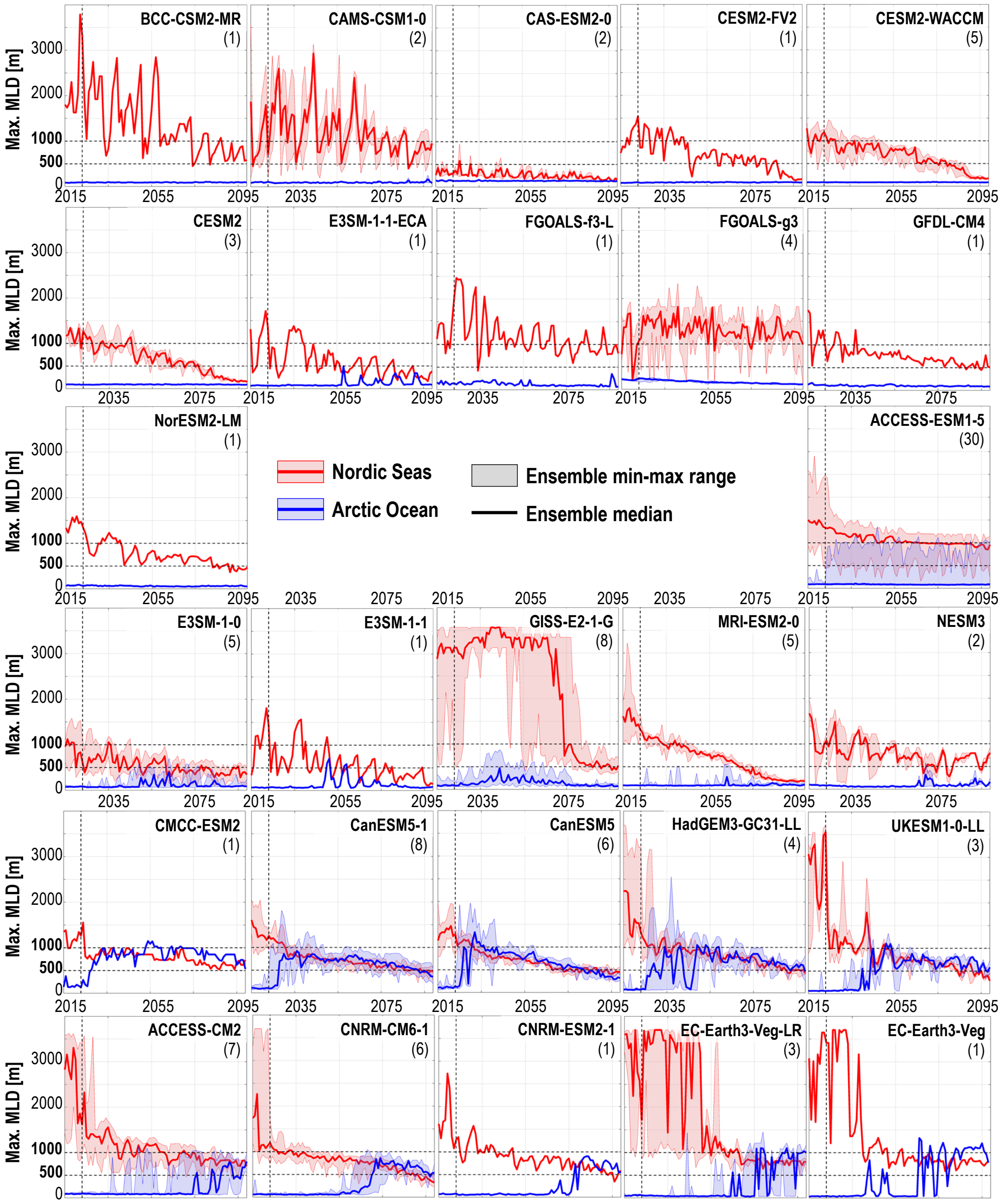

Figure 2. For the 27 models that do not have deep convection in the Arctic over 2015-2023, time series of their yearly maximum MLD in the Nordic Seas (red) and in the Arctic Ocean (blue). For each model, parentheses indicate the number of ensemble members remaining; when this number is larger than one, the figure shows the range across these ensemble members (shading) and the ensemble median (thick line). Horizontal black lines indicate the 500 and 1000 m MLD thresholds, indicative of deep convection. Vertical black line is the year 2023. Models are ordered based on their Arctic behaviour: First two rows and NorESM2-LM, no Arctic deep convection; ACCESS-ESM1-5 and fourth row, rare convection, by the middle of the run; fifth row, convection peaks by the middle of the run and then declines; bottom row, convection starts late in the run.

What causes these different behaviours then? Ensemble members usually have a consistent behaviour (the shading on Fig. 2 usually agrees with the thick line), and models of the same family tend to belong to the same group. The only exception are the two ACCESS models (Fig. 2), but there are large differences in their designs and implemented schemes of relevance for polar regions and deep convection in particular, as discussed in Mohrmann et al. (2021). In the next section, we investigate in more details what could be causing the different Arctic deep convection behaviours, starting with model designs.

3.2 Potential causes: High sensitivity, salinification, stronger winds

The models that have deep convection in the Arctic usually have it at the same location (Fig. 1), but the temporal evolution of the mixed layer yields four groups of models (Fig. 2). Models of the same family tend to belong to the same group; these usually have similar biases and sensitivities, so we first investigate these.

The mean temperature in the Eurasian Basin at the beginning of the SSP5-8.5 run and subsequent MLD are positively correlated across models (Fig. 3, left): the warmer at the beginning of the run, the deeper the mixed layers (correlation 0.40), the more often (0.33). The same results are obtained when considering the ensemble members separately, as on Fig. 3, and the ensemble mean (not shown). There is no such relationship with salinity. The relationship with temperature persists throughout the run, so that the models that warm the most (Fig. 3, centre), and quite strikingly, those for which the ocean surface warms more in the Arctic than in the midlatitudes (Fig. 3, right) are the ones with the deepest MLD, the most often. The models belonging to the E3SM project (cyan contour, Fig. 3) are the exception: They warm strongly but convect rarely or not at all, which could be because their design is very different from that of the other CMIP models (Golaz et al., 2019).

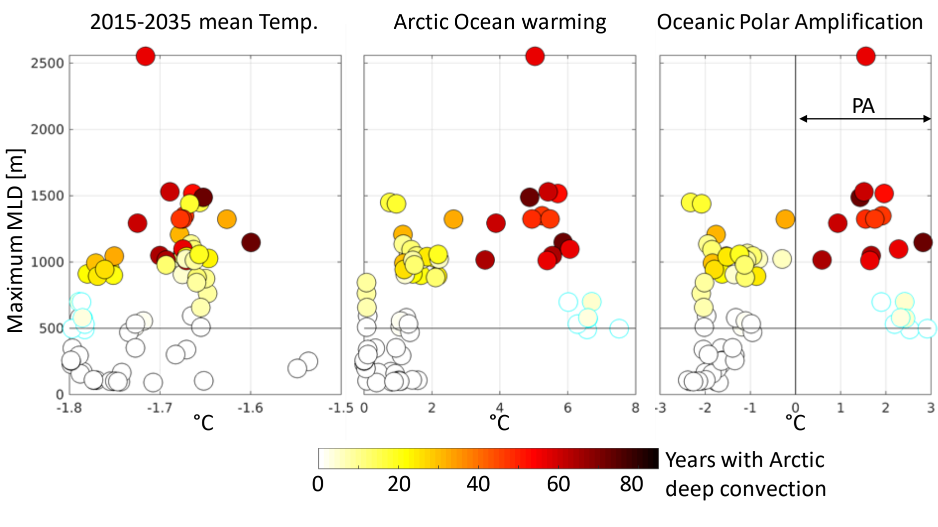

Figure 3. For each model and each ensemble member that were not removed from the study, scatter plot between the maximum MLD over 2015-2100, over the Eurasian Basin, and: Left, the mean ocean surface temperature over the Eurasian basin, over the first 20 years of the run; Centre, the Arctic warming, i.e. the mean ocean surface over the last 20 years of the run minus that of the first 20 years; Right, that Arctic warming minus that of the tropical to mid-latitudes, which we call "Oceanic Polar Amplification". See Methods for latitude definitions. Colors indicate the number of years where the Arctic MLD exceeds the 500 m threshold, as per supp. Table S3. Cyan contours highlight the E3SM project models (see text).

Although the salinity at the beginning of the run has no relationship with the maximum MLD reached, salinity variations are strongly correlated to MLD variations (Table 1) and the correlation differs depending on the Arctic deep convection behaviour of the models. Most models that have no deep convection in the Arctic or rarely (first two blocks) have a positive correlation between their March MLD and the surface salinity one month before, and a negative correlation with the ocean surface temperature: These models have shallower MLD when they are fresher and warmer (as expected in a changing Arctic, e.g. Peralta-Ferriz and Woodgate (2015)). In the two groups of convecting models, especially so in the convective region (lines "DC", Table 1), the correlation is positive with salinity and temperature: deeper MLD are associated with saltier and warmer surface waters the month prior. Of the four drivers investigated, the MLD is most strongly correlated with salinity or temperature for the vast majority of models (bold fonts on Table 1); usually it is strongest with the salinity for the non or rarely convecting models, and with the temperature for the convecting ones (as expected from the previous paragraph). Correlations are similar when considering possible drivers the summer before (not shown, see also supp. Text S1). Unsurprisingly, the possible drivers are not independent (supp. Table S4). The correlations between temperature and salinity, and between salinity and either sea ice concentration or wind speed are of different signs depending on the convecting behaviour, suggesting that different processes and/or water masses are involved (see next subsection).

From the correlation analysis, we suspect that the models with no or rare deep convection may be freshening, while those with convection may become saltier. A trend analysis confirms this hypothesis:

1. The models with no deep convection in the Arctic become fresher throughout the Arctic, throughout the run (Fig. 4, first two lines). Their trends are rather weak compared to the other model groups. Their winter sea ice does not retreat far into the Arctic, even by the end of the run (supp. Fig. S1).
2. The models with deep convection at the middle of the run, be it rarely (Fig. 4, lines 3 and 4) or peaking and declining (lines 5 and 6) exhibit similar trends. Their ocean surface becomes saltier and winds stronger at the location where mixed layers deepen in the first half of the run; they freshen in the rest of the Arctic. In the second half, when mixed layers are shallow again, the ocean surface freshens everywhere. The main difference between these two groups is in their sea ice trends, with the models whose convection peaks and slowly declines having no winter sea ice by the end of the run (supp. Fig. S1).
3. The models that convect at the end of the run (Fig. 4, last two lines) have the opposite salinity trends: first a freshening, then a salinification in the region where mixed layers deepen, along with stronger winds.

Note that the trend patterns are similar in summer (supp. Figure S2), indicating that the changes occur year round. The conclusion is that Arctic deep convection is associated with both a saltier ocean surface and stronger winds. Stronger winds alone are not enough (see e.g. the no deep convection group, second half of the run).

3.3 On causality

Deep convection in the Nordic Seas ceases for all CMIP6 models, but rarely emerges as a stable feature in the Arctic. In the Arctic, we find that deep convection is associated with a saltier ocean surface and stronger wind speeds, and is most intense and durable in the models with the strongest sensitivity, especially so if the Arctic warms more than the mid-latitudes. But are the trends we observe the causes or consequences of deep convection?

Table 1. Correlation coefficient and its standard deviation for each model between the March MLD and the previous February ocean surface salinity (S), ocean surface temperature (T), sea ice concentration (SIC) and surface wind speed (Wind), for the Eurasian basin (EB) and, for the models with Arctic deep convection, where the maximum MLD of Fig. 1 exceeds 500 m (DC). Models are ordered based on their Arctic deep convection behaviour, as per Fig. 2. Only correlations significant at 95% are shown; for models with more than one ensemble member, median of the correlations of the dominating sign. For models with more than one ensemble member, standard deviation is the across-ensemble spread; spatial spread otherwise. Bold fonts highlight the maximum correlation for each model. Correlation between parameters is shown in supp. Table S4.

Model	Region	MLD vs S	MLD vs T	MLD vs SIC	MLD vs Wind
BCC-CSM2-MR	EB	**0.47** ± 0.15	-0.41 ± 0.17	-0.28 ± 0.35	0.23 ± 0.24
CAMS-CSM1-0	EB	**0.54** ± 0.02	-0.49 ± 0.00	-0.35 ± 0.03	0.29 ± 0.04
CAS-ESM2-0	EB	**0.59** ± 0.04	-0.49 ± 0.06	0.51 ± 0.06	-0.24 ± 0.01
CESM2-FV2	EB	-0.29 ± 0.37	0.22 ± 0.37	-0.23 ± 0.16	**0.31** ± 0.08
CESM2-WACCM	EB	**0.34** ± 0.06	-0.31 ± 0.05	0.25 ± 0.06	0.30 ± 0.02
CESM2	EB	**0.32** ± 0.08	-0.30 ± 0.04	0.30 ± 0.05	0.32 ± 0.02
E3SM-1-1-ECA	EB	0.43 ± 0.14	-0.45 ± 0.47	**0.60** ± 0.41	-0.31 ± 0.31
FGOALS-f3-L	EB	0.51 ± 0.21	-0.51 ± 0.10	**0.59** ± 0.13	-0.26 ± 0.16
FGOALS-g3	EB	**0.89** ± 0.01	-0.28 ± 0.02	0.46 ± 0.04	0.28 ± 0.04
GFDL-CM4	EB	**0.39** ± 0.31	-0.25 ± 0.39	-0.31 ± 0.29	-0.27 ± 0.19
NorESM2-LM	EB	-0.28 ± 0.36	**0.36** ± 0.31	-0.26 ± 0.22	0.26 ± 0.16
ACCESS-ESM1-5	EB	-	-	-	-
	DC	-	-	-	-
E3SM-1-0	EB	0.38 ± 0.01	-0.50 ± 0.12	**0.51** ± 0.05	-0.26 ± 0.04
	DC	-	-	-	-
E3SM-1-1	EB	0.39 ± 0.14	-0.41 ± 0.27	**0.55** ± 0.28	-0.28 ± 0.27
	DC	**0.51** ± 0.01	0.34 ± 0.02	-0.29 ± 0.01	0.43 ± 0.01
GISS-E2-1-G	EB	**0.61** ± 0.03	-0.59 ± 0.04	0.35 ± 0.08	0.27 ± 0.04
	DC	0.48 ± 0.14	**0.67** ± 0.10	-0.37 ± 0.08	0.32 ± 0.09
MRI-ESM2-0	EB	**0.62** ± 0.03	-0.40 ± 0.04	0.29 ± 0.05	0.31 ± 0.04
	DC	-	-	-	-
NESM3	EB	**-0.31** ± 0.01	-0.24 ± 0.01	-0.24 ± 0.01	-0.23 ± 0.02
	DC	-	-	-	-
CMCC-ESM2	EB	**0.59** ± 0.45	0.45 ± 0.37	-0.31 ± 0.43	0.28 ± 0.25
	DC	**0.65** ± 0.11	0.49 ± 0.14	-0.40 ± 0.12	0.29 ± 0.08
CanESM5-1	EB	**0.50** ± 0.03	0.39 ± 0.11	-0.46 ± 0.08	0.36 ± 0.05
	DC	**0.52** ± 0.04	0.39 ± 0.12	-0.48 ± 0.08	0.36 ± 0.05
CanESM5	EB	-	-	-	-
	DC	-	-	-	-
HadGEM3-GC31-LL	EB	-	-	-	-
	DC	-	-	-	-
UKESM1-0-LL	EB	0.56 ± 0.02	**0.64** ± 0.01	-0.51 ± 0.02	0.43 ± 0.05
	DC	0.57 ± 0.02	**0.65** ± 0.02	-0.54 ± 0.02	0.43 ± 0.06
ACCESS-CM2	EB	0.40 ± 0.05	**0.45** ± 0.19	-0.28 ± 0.08	0.28 ± 0.03
	DC	0.44 ± 0.05	**0.76** ± 0.11	-0.51 ± 0.08	0.34 ± 0.04
CNRM-CM6-1	EB	0.40 ± 0.07	**0.73** ± 0.05	-0.62 ± 0.05	0.38 ± 0.05
	DC	0.46 ± 0.05	**0.77** ± 0.08	-0.68 ± 0.05	0.47 ± 0.04
CNRM-ESM2-1	EB	0.44 ± 0.15	**0.77** ± 0.48	-0.53 ± 0.31	0.43 ± 0.13
	DC	0.46 ± 0.09	**0.86** ± 0.10	-0.57 ± 0.16	0.47 ± 0.12
EC-Earth3-Veg-LR	EB	0.44 ± 0.04	**0.85** ± 0.05	-0.55 ± 0.12	0.35 ± 0.05
	DC	0.45 ± 0.03	**0.90** ± 0.05	-0.66 ± 0.04	0.39 ± 0.03
EC-Earth3-Veg	EB	0.47 ± 0.13	**0.85** ± 0.43	-0.49 ± 0.17	0.32 ± 0.09
	DC	0.52 ± 0.10	**0.89** ± 0.07	-0.53 ± 0.12	0.34 ± 0.10

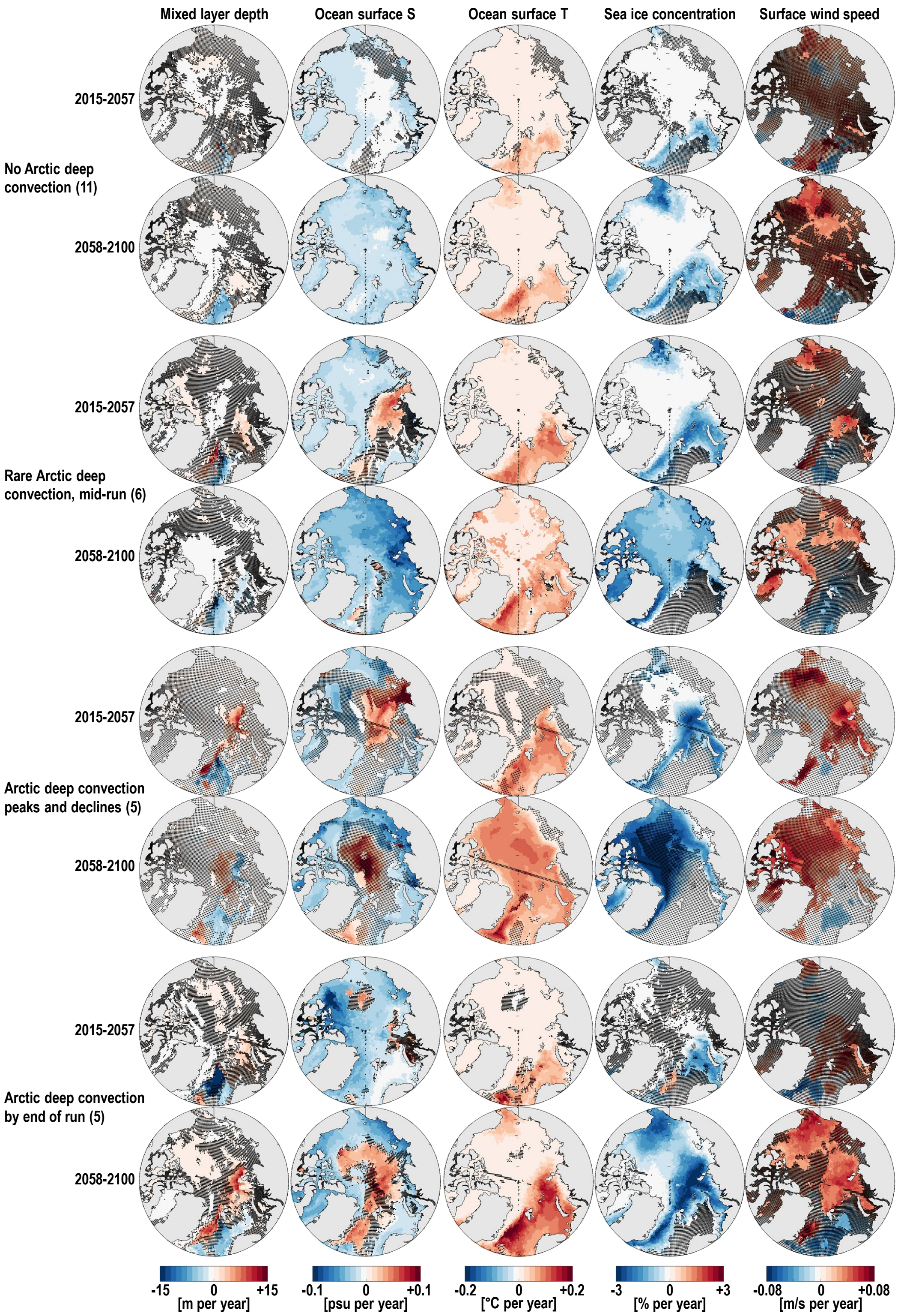

Figure 4. Composite trends based on the models' Arctic deep convection behaviours of Fig. 2, for the first half of the 21st century run (top rows) and the second half (bottom rows), in March MLD (first column), and February ocean surface salinity (S, second), ocean surface temperature (T, third), sea ice concentration (fourth) and surface wind speed (last). Behaviours are described to the left of the figure, and number of models for each behaviour is given in parentheses. For each panel, stippling indicates non-significant trends and/or model disagreement regarding the trend's sign. Straight stippled lines across the North Pole are an artefact of the necessary interpolation.

The Arctic Ocean surface warming and sea ice loss trends, year-round, are to be expected in a warming world (IPCC, 2019). Besides, they are consistent across-models, regardless of their convecting behaviour in the Arctic. Similarly, the models' sensitivities have been attributed to their representation of cloud cover and cloud albedo (Zelinka et al., 2020), which locally can be affected by the heat and moisture fluxes from deep convection (Monroe et al., 2021), but in coarse models is more likely due to each model's cloud parameterisation (Zelinka et al., 2020). Finally, changes in both "normal" winds (Screen et al., 2018) and cyclones (Rinke et al., 2017) are most commonly attributed to the large scale atmospheric circulation, although local changes at the boundary layer because of sea ice loss may accelerate winds (DuVivier et al., accepted).

As for the salinity, Lique et al. (2018) attributed the freshening - salinification dipoles to changes in the large scale oceanic circulation, whereas Davis et al. (2016) attributed the increased salinity to increased vertical mixing. That is, for the former it drives Arctic deep convection; for the latter, it is a consequence of it. A local change in salinity can be also caused by enhanced sea ice formation, but we find a negative correlation between winter salinity and sea ice concentration among the convecting models (supp. Table S4), which makes this causality unlikely given that the sea ice maps do not seem to exhibit polynyas (supp. Fig. S1). Salinification taking place year-round, at the same location as winds are increasing, would instead suggest an enhanced sea ice drift, but the detailed sea ice mass budget of Keen et al. (2021) shows the opposite.

Alternatively, the impact of deeper MLD on the Arctic surface salinity will depend on each model's Atlantic layer. Heuzé et al. (2023) showed strong biases in that layer, with most models having it too deep. Khosravi et al. (2022) further showed that there is no consistency across CMIP6 models regarding the changes of that Atlantic layer during SSP5-8.5, which Muilwijk et al. (2023) linked to the models' lack of consensus regarding future changes in stratification in the Eurasian basin and in Atlantification. As stratification and deep convection are intimately linked, this is another feedback that makes the causality uncertain. One can unfortunately not study the onset of deep convection in more details without higher temporal resolution output. Most likely, the causality will be different for each model, based on their choices of parameterisation in the atmosphere (Zelinka et al., 2020), sea ice (Keen et al., 2021) and ocean (Heuzé et al., 2023), and even their definition of the mixed layer (Griffies et al., 2016). But individual model studies require access to each model's code, and are way beyond the scope of this paper.

4 Conclusions

We used all CMIP6 models and all their ensemble members for which the mixed layer depth and its potential drivers the surface salinity, temperature, sea ice concentration, and wind speed, were available for the strongest warming scenario SSP5-8.5. After removing the ensemble members that had spurious Arctic deep convection (defined, as in Lique et al. (2018), as MLD deeper than 500 m) over 2015-2023, we were left with 27 models, of which 11 had no deep convection in the Arctic over 2015 - 2100; 6 that had it extremely rarely (usually 4 years or fewer), by the middle of the run; 5 for which deep convection peaked in the first half of the run and then declined and disappeared; and 5 in which deep convection emerged in the Arctic in the second half of the run and still convected in 2100. All models exhibit a cessation of deep convection in the Nordic Seas, showing that deep convection in the Arctic is not simply a northward migration of the Nordic Seas ventilation. The Arctic MLD was most strongly correlated with the surface salinity, and the sign of this correlation depended on whether the model convected or not. Similarly, when and where the models are not convecting, their surface salinity freshens; at the location where they do, when they do, it becomes saltier and surface winds are increasing. Neither the exact mechanism triggering deep convection nor the direction of the causality between deep convection and that compound salinification and wind event can be investigated in more details with CMIP6 monthly output. The fact that models of the same family have the same convection behaviour; that the depth and frequency of the maximum MLD is strongly correlated to early-run biases and sensitivity; and that the other processes involved have been linked to individual model parameterisations (Zelinka et al., 2020; Keen et al., 2021; Muilwijk et al., 2023) suggest that the trigger for Arctic deep convection is model-specific, and its determination requires in-depth sensitivity studies for each model. Such in-depth investigation could also lead to model improvement. CMIP6 models consistently exaggerate deep convection both in the North Atlantic and in the Southern Ocean (Heuzé, 2021). Understanding why they have no such consensus in the Arctic could hold the key to a more realistic representation of mixing, globally.

5 Open Research

All CMIP6 data are freely available via the Earth Grid System Federation. For this paper, we primarily used the German Climate Computing Centre (DKRZ) node `https://esgf-data.dkrz.de/search/cmip6-dkrz/`.

Acknowledgments

CH is funded by the Swedish Research Council (dnr 2018-03859).

References

Davis, P., Lique, C., Johnson, H., & Guthrie, J. (2016). Competing Effects of Elevated Vertical Mixing and Increased Freshwater Input on the Stratification and Sea Ice Cover in a Changing Arctic Ocean . *Journal of Physical Oceanography*, *46*, 1531-1553. doi: 10.1175/JPO-D-15-0174.1

DuVivier, A., Holland, M., Vavrus, S., Landrum, L., Shields, C., & Thaker, R. (accepted). Arctic sea ice loss drives increasing Arctic wind speeds with combined impact on surface roughness and boundary layer stability. *Journal of Geophysical Research: Atmospheres*. doi: 10.21203/rs.3.rs-2210756/v1

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

Golaz, J., Caldwell, P., Van Roekel, L., Petersen, M., Tang, Q., Wolfe, J., ... Baldwin, S. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*, 2089-2129. doi: 10.1029/2018MS001603

Griffies, S., Danabasoglu, G., Durack, P., Adcroft, A., Balaji, V., Böning, C., ... Fox-Kemper, B. (2016). OMIP contribution to CMIP6: Experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project. *Geoscientific Model Development*, *9*, 3231-3296. doi: 10.5194/gmd-9-3231-2016

Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models. *Ocean Science*. doi: 10.5194/os-17-59-2021

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

IPCC. (2019). *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (H. Pörtner et al., Eds.). Cambridge University Press.

Keen, A., Blockley, E., Bailey, D., Boldingh Debernard, J., Bushuk, M., Delhaye, S., ... Ponsoni, L. (2021). An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *The Cryosphere*, *15*, 951-982. doi: 10.5194/tc-15-951-2021

Khosravi, N., Wang, Q., Koldunov, N., Hinrichs, C., Semmler, T., Danilov, S., & Jung, T. (2022). The Arctic Ocean in CMIP6 models: Biases and projected changes in temperature and salinity. *Earth's Future*, *10*, e2021EF002282. doi: 10.1029/2021EF002282

Lique, C., Johnson, H., & Plancherel, Y. (2018). Emergence of deep convection in the Arctic Ocean under a warming climate. *Climate Dynamics*, *50*, 3833-3847. doi: 10.1007/s00382-017-3849-9

Mallett, R., Stroeve, J., Tsamados, M., Landy, J., Willatt, R., Nandan, V., & Liston, G. E. (2021). Faster decline and higher variability in the sea ice thickness of the marginal Arctic seas when accounting for dynamic snow cover. *The Cryosphere*, *15*, 2429–2450. doi: 10.5194/tc-15-2429-2021

Meier, W., & Stroeve, J. (2022). An updated assessment of the changing Arctic sea ice cover. *Oceanography*, *35*, 10-19.

Mohrmann, M., Heuzé, C., & Swart, S. (2021). Southern Ocean polynyas in CMIP6 models. *The Cryosphere*, *15*, 4281-4313. doi: 10.5194/tc-15-4281-2021

Monroe, E., Taylor, P., & Boisvert, L. (2021). Arctic cloud response to a perturbation in sea ice concentration: The North Water polynya. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD034409. doi: 10.1029/2020JD034409

Muilwijk, M., Nummelin, A., Heuzé, C., Polyakov, I., Zanowski, H., & Smedsrud, L. (2023). Divergence in climate model projections of future Arctic Atlantification. *Journal of Climate*, *36*, 1727-1748. doi: 10.1175/JCLI-D-22-0349.1

Onarheim, I., Eldevik, T., Smedsrud, L., & Stroeve, J. (2018). Seasonal and regional manifestation of Arctic sea ice loss. *Journal of Climate*, *31*, 4917-4932. doi: 10.1175/JCLI-D-17-0427.1

O'Neill, B., Tebaldi, C., Van Vuuren, D., Eyring, V., Friedlingstein, P., Hurtt, G., … Meehl, G. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*, 3461–3482. doi: 10.5194/gmd-9-3461-2016

Peralta-Ferriz, C., & Woodgate, R. (2015). Seasonal and interannual variability of pan-Arctic surface mixed layer properties from 1979 to 2012 from hydrographic data, and the dominance of stratification for multiyear mixed layer depth shoaling. *Progress in Oceanography*, *134*, 19-53. doi: 10.1016/j.pocean.2014.12.005

Peterson, A. (2018). Observations of brine plumes below melting Arctic sea ice. *Ocean Science*, *14*, 127–138. doi: 10.5194/os-14-127-2018

Polyakov, I., Pnyushkov, A., Alkire, M., Ashik, I., Baumann, T., Carmack, E., … Krishfield, R. (2017). Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science*, *356*, 285-291. doi: 10.1126/science.aai8204

Rinke, A., Maturilli, M., Graham, R., Matthes, H., Handorf, D., Cohen, L., … Moore, J. (2017). Extreme cyclone events in the Arctic: Wintertime variability and trends. *Environmental Research Letters*, *12*, 094006. doi: 10.1088/1748-9326/aa7def

Schulz, K., Koenig, Z., Muilwijk, M., Bauch, D., Hoppe, C., Droste, E., … Granskog, M. (under review). The Eurasian Arctic Ocean along the MOSAiC drift (2019-2020): An interdisciplinary perspective on properties and processes. *Elementa*. doi: 10.31223/X5TT2W

Screen, J., Deser, C., Smith, D., Zhang, X., Blackport, R., Kushner, P., … Sun, L. (2018). Consistency and discrepancy in the atmospheric response to Arctic sea-ice loss across climate models. *Nature Geosci*, *11*, 155–163. doi: 10.1038/s41561-018-0059-y

Timmermans, M., & Marshall, J. (2020). Understanding Arctic Ocean circulation: A review of ocean dynamics in a changing climate. *Journal of Geophysical Research: Oceans*, *125*, e2018JC014378. doi: 10.1029/2018JC014378

Zelinka, M., Myers, T., McCoy, D., Po-Chedley, S., Caldwell, P., Ceppi, P., …

Taylor, K. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*, e2019GL085782. doi: 10.1029/2019GL085782

Supporting Information for "No emergence of deep convection in the Arctic Ocean across CMIP6 models"

Céline Heuzé[1], and Hailong Liu[2]

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

[2]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China

Contents of this file

1. Text S1
2. Figures S1 to S2
3. Tables S1 to S4

Introduction

This document provides one supplementary text, two supplementary figures and four supplementary tables, referred to in the main text as Figures S1-s2 and Tables S1-S4, as well as the corresponding references.

Text S1. Choices of thresholds and robustness Globally, a deeper threshold than 500 m is commonly used to determine deep convection. Using a threshold of 1000 m instead of that of 500 m (chosen by consistency with Lique, Johnson, and Plancherel (2018)) does not significantly change the results. It only increased the number of models that are not convecting in the Arctic from 11 to 18, as can be seen on the main text Fig. 2. Time series of the yearly maximum mixed layer depth have been produced with the actual maximum regardless of time; April, month of deepest mixed layers for the few convecting models (not shown); and March, the month traditionally used and where the non-convecting models have their deepest mixed layers. The differences between the results were not significant. We therefore perform correlations and show results using the March mixed layers, for consistency with other studies.

The sensitivies and oceanic polar amplification values are not significantly affected by the choice of latitude bands. That is, the model order remains the same, only the individual models' magnitudes are changed. We chose to present the results that exclude 60°N to 75°N, as these are dominated by the biased Nordic Seas.

Finally, we tested different lags for the correlations and trends ranging from 1 to 6 months, as we aim to determine which process could be driving changes in the mixed layer depth. We show primarily those with one month lag (i.e. February salinity, temperature, sea ice and winds), that had the strongest signal-to-noise ratio. We also show as supplementary material the 6 months (September) values, although the reader must bear in mind that most models lose their September sea ice early (see supp. Fig. S1), therefore the statistics involving sea ice become less robust. We also tested different thresholds for the ensemble and model agreement. Because of the low number of ensemble members for most models (especially so after removing the ones that have Arctic deep convection over 2015-2023), 66% was too restrictive: When rounded, it often meant "all ensemble members/models". Our choice of 50% threshold does not change the results significantly. In particular, it does not change the sign or relative magnitude of the correlations/trends. The trends are presented over exactly half of the run, but we tested with different number of years, with the entire run, and visually verified them for every model and ensemble member. We chose to present them separately over the two halves of the run as this matches the models' Arctic deep convection behaviour, as will be discussed.

References

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C., ... Held et al., I. (2019). The GFDL global ocean and sea ice model OM4. 0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*, 3167-3211. doi: 10.1029/2019MS001726

Cao, J., Wang, B., Yang, Y.-M., Ma, L., Li, J., Sun, B., ... Wu, L. (2018). The NUIST Earth System Model (NESM) version 3: description and preliminary evaluation. *Geoscientific Model Development*, *11*. doi: 10.5194/gmd-11-2975-2018

Cherchi, A., Fogli, P., Lovato, T., Peano, D., Iovino, D., Gualdi, S., ... Navarra, A. (2019). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems*, *11*, 185-209. doi: 10.1029/2018MS001369

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., & Emmons et al., L. K. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*. doi: 10.1029/2019MS001916

Dong, X., Jin, J., Liu, H., Zhang, H., Zhang, M., Lin, P., ... Lin, Z. (2021). CAS-ESM2.0 model datasets for the CMIP6 ocean model intercomparison project phase 1 (OMIP1). *Advances in Atmospheric Sciences*, *38*, 307-316. doi: 10.1007/s00376-020-0150-3

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., ... Bernadello et al., R. (2021). The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6. *Geosci. Model Dev. Discuss.*. doi: 10.5194/gmd-2020-446

Golaz, J., Caldwell, P., Van Roekel, L., Petersen, M., Tang, Q., Wolfe, J., ... Baldwin, S. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*, 2089-2129. doi: 10.1029/2018MS001603

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., ... Kawamiya, M. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, *13*. doi: 10.5194/gmd-13-2197-2020

He, B., Bao, Q., Wang, X., Zhou, L., Wu, X., Liu, Y., ... Li, J. (2019). CAS FGOALS-f3-L model datasets for CMIP6 historical atmospheric model intercomparison project simulation. *Advances in Atmospheric Sciences*, *36*, 771-778. doi: 10.1007/s00376

-019-9027-8

Kelley, M., Schmidt, G., Nazarenko, L., Bauer, S., Ruedy, R., Russell, G., ... Canuto, V. (2020). GISS-E2. 1: Configurations and climatology. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002025. doi: 10.1029/2019MS002025

Kuhlbrodt, T., Jones, C., Sellar, A., Storkey, D., Blockley, E., Stringer, M., ... Calvert, D. (2018). The low-resolution version of HadGEM3 GC3. 1: Development and evaluation for global climate. *Journal of Advances in Modeling Earth Systems*, *10*. doi: 10.1029/2018MS001370

Li, L., Yu, Y., Tang, Y., Lin, P., Xie, J., Song, M., ... Pu, Y. (2020). The flexible global ocean-atmosphere-land system model grid-point version 3 (FGOALS-g3): description and evaluation. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002012. doi: 10.1029/2019MS002012

Lique, C., Johnson, H., & Plancherel, Y. (2018). Emergence of deep convection in the Arctic Ocean under a warming climate. *Climate Dynamics*, *50*, 3833-3847. doi: 10.1007/s00382-017-3849-9

Lovato, T., Peano, D., ButenschÖn, M., Materia, S., Iovino, D., Scoccimarro, E., ... Masina, S. (2022). CMIP6 simulations with the CMCC Earth system model (CMCC-ESM2). *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002814. doi: 10.1029/2021MS002814

Mackallah, C., Chamberlain, M., Law, R., Dix, M., Ziehn, T., Bi, D., ... Evans, B. e. (2022). ACCESS datasets for CMIP6: methodology and idealised experiments. *Journal of Southern Hemisphere Earth Systems Science*, *72*, 93-116. doi:

10.1071/ES21031

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... Fast, I. (2019). 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*. doi: 10.1029/2018MS001400

Müller, W., Jungclaus, J., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., ... Ilyina, T. (2018). A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1. 2-HR). *Journal of Advances in Modeling Earth Systems*, *10*. doi: 10.1029/2017MS001217

Rong, X. Y., Li, J., & Chen, H. M. (2019). Introduction of CAMS-CSM model and its participation in CMIP6. *Climate Change Res*, *6*. doi: 10.12006/j.issn.1673-1719.2019.186

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., ... Sénési, S. (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2019MS001791

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L., ... Schwinger, J. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*, 6165-6200. doi: 10.5194/gmd-13-6165-2020

Sellar, A., Jones, C., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., ... Woodward, S. (2019). UKESM1: Description and evaluation of the UK Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*, 4513-4558. doi: 10.1029/2019MS001739

Semmler, T., Danilov, S., Gierz, P., Goessling, H., Hegewald, J., Hinrichs, C., ... Sein, D. (2020). Simulations for CMIP6 with the AWI climate model AWI-CM-1-1. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002009. doi: 10.1029/2019MS002009

Sigmond, M., Anstey, J., Arora, V., Digby, R., Gillett, N., Kharin, V., ... Virgin, J. (2023). Improvements in the Canadian Earth System Model (CanESM) through systematic model analysis: CanESM5. 0 and CanESM5.1. *Geoscientific Model Development Discussions*, 1-54. doi: 10.5194/gmd-2023-52

Swart, N., Cole, J., Kharin, V., Lazare, M., Scinocca, J., Gillett, N., ... Jiao, Y. (2019). The Canadian Earth System Model version 5 (CanESM5. 0.3). *Geoscientific Model Development*, *12*. doi: 10.5194/gmd-2019-177

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., ... Kimoto, M. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, *12*. doi: 10.5194/gmd-12-2727-2019

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., & Colin et al., J. (2019). Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, *11*. doi: 10.1029/2019MS001683

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., ... Liu, X. (2019). The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6. *Geoscientific Model Development*, *12*. doi: 10.5194/gmd-12-1573-2019

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., … Yabu, S. (2019). 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*. doi: 10.2151/jmsj.2019-051

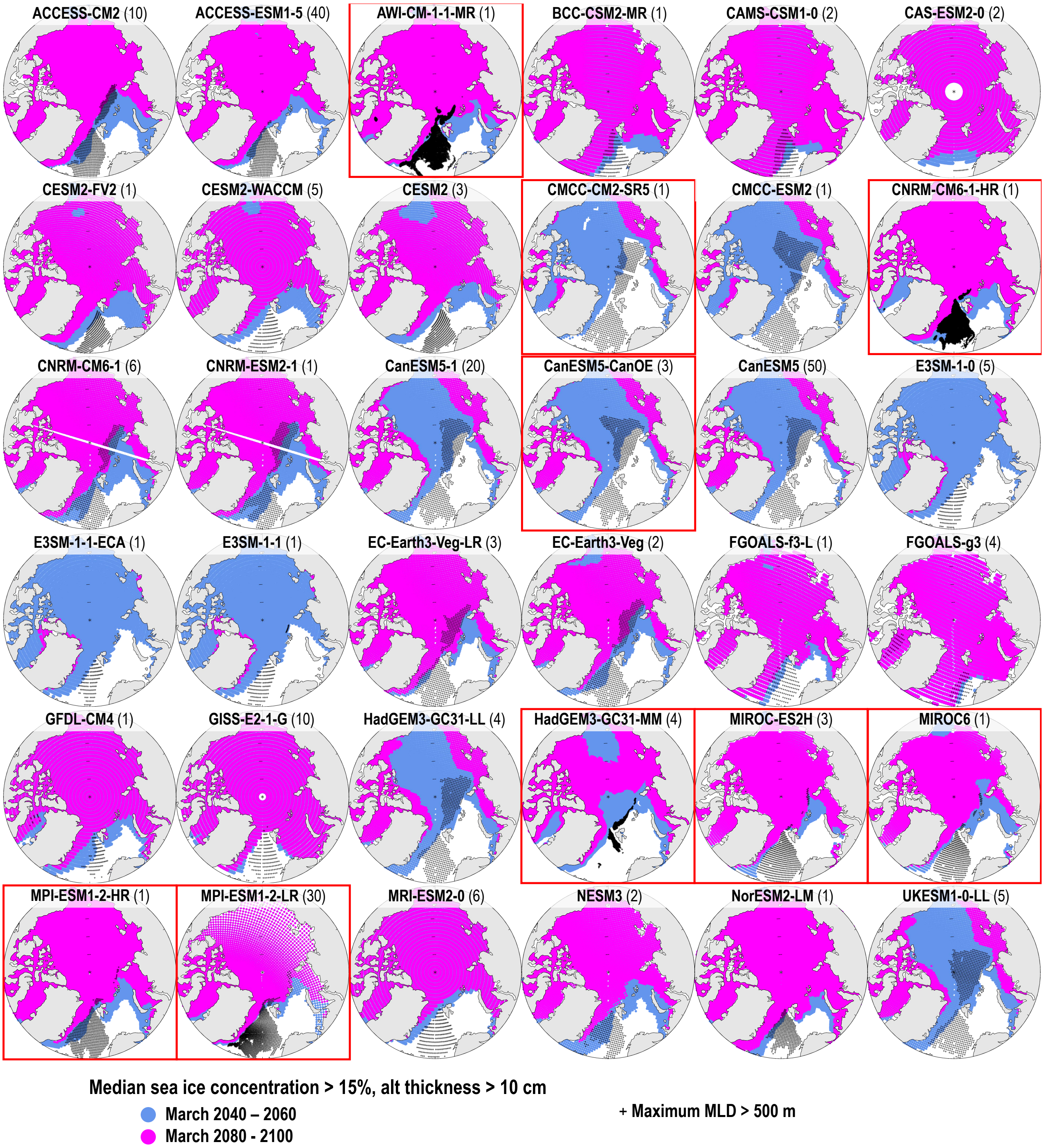

Figure S1. Median 2040-2060 (blue) and 2080-2100 (magenta overlaying the blue) March sea ice extent for all 36 models. Threshold is 15% sea ice concentration or 10 cm sea ice thickness. Black + indicate where the 2015-2100 maximum mixed layer depth (MLD) of Fig. 1 exceeds 500 m. Red boxes as per Fig. 1.

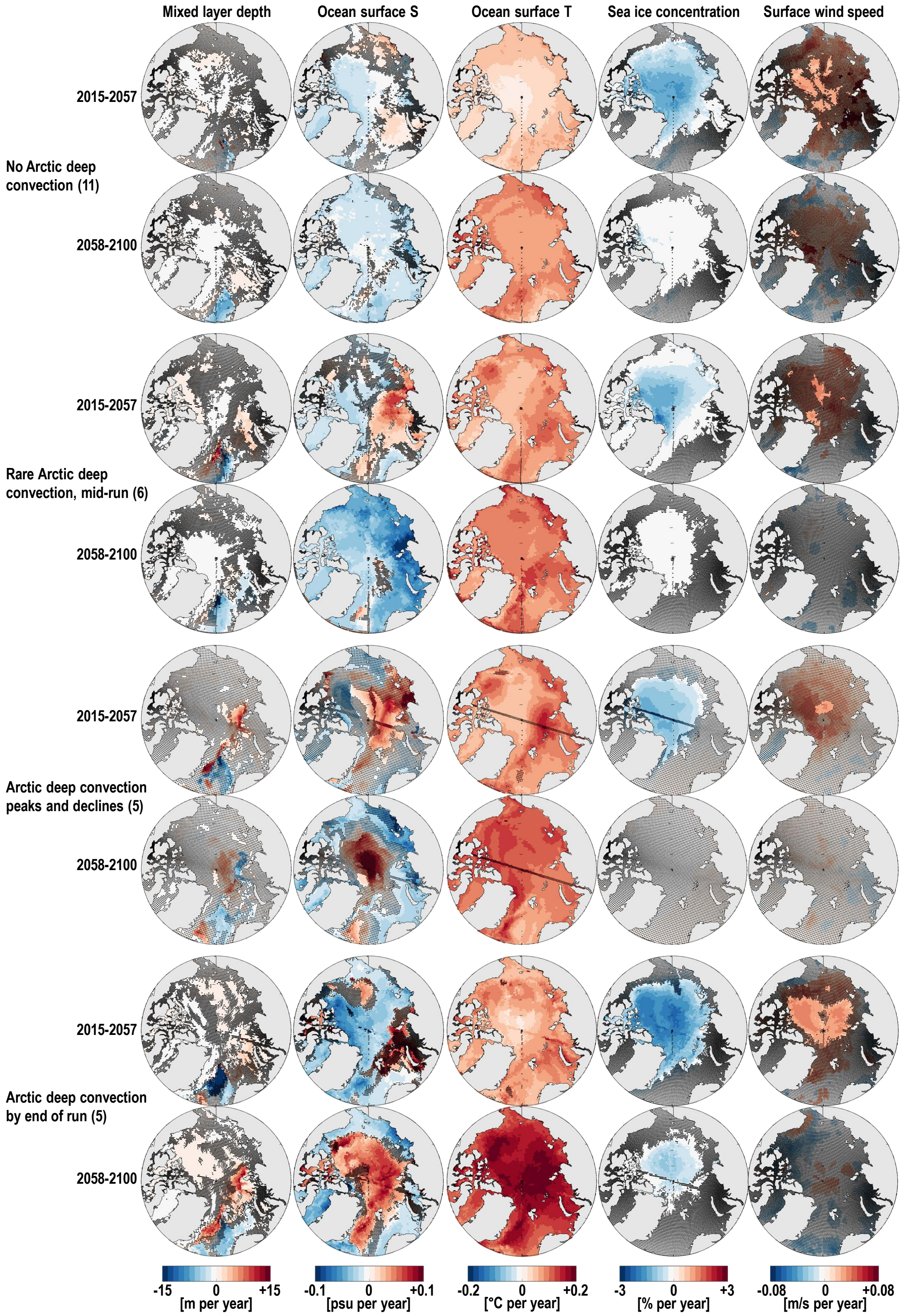

Figure S2. Composite trends based on the models' Arctic deep convection behaviours, for the first half of the 21st century run (top rows) and the second half (bottom rows), in March mixed layer depth (MLD, first column), and September ocean surface salinity, ocean surface temperature, sea ice concentration and surface wind speed. Behaviours are described to the left of the figure, and number of models for each behaviour is given in parentheses. For each panel, stippling indicates non-significant trends and/or model disagreement regarding the trend's sign.

Table S1. Models, ensemble members and corresponding grid type used in this study, along with the model's reference [a]

Model	Ensemble members	Grid	Reference
ACCESS-CM2	r1i1p1f1; r2i1p1f1; r3i1p1f1*; r4i1p1f1; r5i1p1f1; r6i1p1f1; r7i1p1f1; r8i1p1f1; r9i1p1f1*; r10i1p1f1*	gn	Mackallah et al. (2022)
ACCESS-ESM1-5	r1i1p1f1; r2i1p1f1; r3i1p1f1*; r4i1p1f1*; r5i1p1f1; r6i1p1f1; r7i1p1f1*; r8i1p1f1; r9i1p1f1; r10i1p1f1; r11i1p1f1*; r12i1p1f1; r13i1p1f1; r14i1p1f1; r15i1p1f1*;r16i1p1f1; r17i1p1f1; r18i1p1f1*; r19i1p1f1; r20i1p1f1; r21i1p1f1; r22i1p1f1*; r23i1p1f1*; r24i1p1f1; r25i1p1f1; r26i1p1f1; r27i1p1f1; r28i1p1f1; r29i1p1f1; r30i1p1f1; r31i1p1f1*; r32i1p1f1; r33i1p1f1; r34i1p1f1; r35i1p1f1; r36i1p1f1; r37i1p1f1; r38i1p1f1*; r39i1p1f1; r40i1p1f1	gn	Mackallah et al. (2022)
AWI-CM-1-1-MR*	r1i1p1f1*	gn	Semmler et al. (2020)
BCC-CSM2-MR	r1i1p1f1	gn	Wu et al. (2019)
CAMS-CSM1-0	r1i1p1f1; r2i1p1f1	gn	Rong, Li, and Chen (2019)
CAS-ESM2-0	r1i1p1f1; r3i1p1f1	gn	Dong et al. (2021)
CESM2-FV2	r1i2p2f1	gn	Danabasoglu et al. (2020)
CESM2-WACCM	r1i1p1f1; r2i1p1f1; r3i1p1f1; r4i1p1f1; r5i1p1f1	gr	Danabasoglu et al. (2020)
CESM2	r4i1p1f1; r10i1p1f1; r11i1p1f1	gn	Danabasoglu et al. (2020)
CMCC-CM2-SR5*	r1i1p1f1*	gn	Cherchi et al. (2019)
CMCC-ESM2	r1i1p1f1	gn	Lovato et al. (2022)
CNRM-CM6-1-HR*	r1i1p1f2*	gn	Voldoire et al. (2019)
CNRM-CM6-1	r1i1p1f2; r2i1p1f2; r3i1p1f2; r4i1p1f2; r5i1p1f2; r6i1p1f2	gn	Voldoire et al. (2019)
CNRM-ESM2-1	r1i1p1f2	gn	Séférian et al. (2019)
CanESM5-1	r1i1p1f1; r1i1p2f1; r2i1p1f1; r2i1p2f1; r3i1p1f1*; r3i1p2f1*; r4i1p1f1*; r4i1p2f1*; r5i1p1f1*; r5i1p2f1; r6i1p1f1*; r6i1p2f1*; r7i1p1f1; r7i1p2f1; r8i1p1f1*;r8i1p2f1; r9i1p1f1*; r9i1p2f1*; r10i1p1f1*; r10i1p2f1*	gn	Sigmond et al. (2023)
CanESM5-CanOE*	r1i1p2f1*; r2i1p2f1*; r3i1p2f1*	gn	Swart et al. (2019)

[a] Ensemble members marked with an asterisk (*) have deep convection in the Arctic over 2015 - 2023 already and were excluded from the correlation and trend analyses. Models marked with an asterisk (*) have deep convection in the Arctic over 2015 - 2023 already for all their ensemble members.

Table S2. Table S1 continued

Model	Ensemble members	Grid	Reference
CanESM5	r1i1p1f1*; r1i1p2f1*; r2i1p1f1*; r2i1p2f1*; r3i1p1f1*; r3i1p2f1*; r4i1p1f1*; r4i1p2f1; r5i1p1f1*; r5i1p2f1*; r6i1p1f1*; r6i1p2f1*; r7i1p1f1*; r7i1p2f1*; r8i1p1f1*;r8i1p2f1; r9i1p1f1*; r9i1p2f1*; r10i1p1f1*; r10i1p2f1*; r11i1p1f1*; r11i1p2f1*; r12i1p1f1*; r12i1p2f1*; r13i1p1f1*;r13i1p2f1*; r14i1p1f1*; r14i1p2f1; r15i1p1f1*; r15i1p2f1*; r16i1p1f1*; r16i1p2f1; r17i1p1f1*; r17i1p2f1; r18i1p1f1*; r18i1p2f1*; r19i1p1f1*; r19i1p2f1*; r20i1p1f1*; r20i1p2f1*; r21i1p1f1*; r21i1p2f1*; r22i1p1f1*; r22i1p2f1*; r23i1p1f1*; r23i1p2f1*; r24i1p1f1*; r24i1p2f1*; r25i1p1f1*; r25i1p2f1*	gn	Swart et al. (2019)
E3SM-1-0	r1i1p1f1; r2i1p1f1; r3i1p1f1; r4i1p1f1; r5i1p1f1	gr	Golaz et al. (2019)
E3SM-1-1-ECA	r1i1p1f1	gr	Golaz et al. (2019)
E3SM-1-1	r1i1p1f1	gr	Golaz et al. (2019)
EC-Earth3-Veg-LR	r1i1p1f1; r2i1p1f1; r3i1p1f1	gn	Döscher et al. (2021)
EC-Earth3-Veg	r1i1p1f1*; r2i1p1f1	gn	Döscher et al. (2021)
FGOALS-f3-L	r1i1p1f1	gn	He et al. (2019)
FGOALS-g3	r1i1p1f1; r2i1p1f1; r3i1p1f1; r4i1p1f1	gn	Li et al. (2020)
GFDL-CM4	r1i1p1f1	gr	Adcroft et al. (2019)
GISS-E2-1-G	r1i1p1f2; r1i1p3f1; r1i1p5f1*; r2i1p1f2; r2i1p3f1; r2i1p5f1; r3i1p1f2*; r3i1p3f1; r4i1p1f2; r4i1p3f1	gn	Kelley et al. (2020)
HadGEM3-GC31-LL	r1i1p1f3; r2i1p1f3; r3i1p1f3; r4i1p1f3	gn	Kuhlbrodt et al. (2018)
HadGEM3-GC31-MM*	r1i1p1f3*; r2i1p1f3*; r3i1p1f3*; r4i1p1f3*	gn	Kuhlbrodt et al. (2018)
MIROC-ES2H*	r1i1p4f2*; r2i1p4f2*; r3i1p4f2	gn	Hajima et al. (2020)
MIROC6*	r1i1p1f1*	gn	Tatebe et al. (2019)
MPI-ESM1-2-HR*	r2i1p1f1*	gn	Müller et al. (2018)
MPI-ESM1-2-LR*	r1i1p1f1*; r2i1p1f1*; r3i1p1f1*; r4i1p1f1*; r5i1p1f1*; r6i1p1f1*; r7i1p1f1*; r8i1p1f1*; r9i1p1f1*; r10i1p1f1*; r11i1p1f1*; r12i1p1f1*; r13i1p1f1*; r14i1p1f1*; r15i1p1f1*; r16i1p1f1*; r17i1p1f1*; r18i1p1f1*; r19i1p1f1*; r20i1p1f1*; r21i1p1f1*; r22i1p1f1*; r23i1p1f1*; r24i1p1f1*; r25i1p1f1*; r26i1p1f1*; r27i1p1f1*; r28i1p1f1*; r29i1p1f1*; r30i1p1f1*	gn	Mauritsen et al. (2019)
MRI-ESM2-0	r1i1p1f1; r1i2p1f1; r2i1p1f1; r3i1p1f1; r4i1p1f1; r5i1p1f1*	gr	Yukimoto et al. (2019)
NESM3	r1i1p1f1; r2i1p1f1	gn	Cao et al. (2018)
NorESM2-LM	r1i1p1f1	gn	Seland et al. (2020)
UKESM1-0-LL	r1i1p1f2; r2i1p1f2*; r3i1p1f2; r4i1p1f2; r8i1p1f2*	gn	Sellar et al. (2019)

Table S3. Table version of Figure 2: Models; number of ensemble members used[a]; across-ensemble minimum - median - maximum of the yearly maximum mixed layer depth (in meters); across-ensemble minimum - mode - maximum number of years with deep convection in the Arctic; type of Arctic deep convection behaviour (for the composites)[b]

Model	Nb. members	Across-ensemble max MLD min - median - max	Across-ensemble years min - mode - max	Type
ACCESS-CM2	7	883 - 1041 - 1134	10 - 23 - 25	Late
ACCESS-ESM1-5	30	661 - 1019 - 1356	2 - 4 - 17	MidR
BCC-CSM2-MR	1	104	0	NoDC
CAMS-CSM1-0	2	198 - 224 - 251	0 - 0 - 0	NoDC
CAS-ESM2-0	2	157 - 165 - 173	0 - 0 - 0	NoDC
CESM2-FV2	1	97	0	NoDC
CESM2-WACCM	5	99 - 101 - 102	0 - 0 - 0	NoDC
CESM2	3	108 - 108 - 111	0 - 0 - 0	NoDC
CMCC-ESM2	1	1147	73	MidD
CNRM-CM6-1	6	894 - 946 - 1045	22 - 22 - 33	Late
CNRM-ESM2-1	1	901	21	Late
CanESM5-1	8	1011 - 1193 - 1808	43 - 65 - 69	MidD
CanESM5	6	1262 - 1497 - 1969	49 - 55 - 60	MidD
E3SM-1-0	5	492 - 575 - 698	0 - 1 - 4	MidR
E3SM-1-1-ECA	1	497	0	NoDC
E3SM-1-1	1	701	4	MidR
EC-Earth3-Veg-LR	3	1207 - 1439 - 1450	16 - 16 - 31	Late
EC-Earth3-Veg	1	1322	33	Late
FGOALS-f3-L	1	350	0	NoDC
FGOALS-g3	4	229 - 255 - 258	0 - 0 - 0	NoDC
GFDL-CM4	1	166	0	NoDC
GISS-E2-1-G	8	302 - 625 - 875	0 - 0 - 10	MidR
HadGEM3-GC31-LL	4	1325 - 1418 - 2552	47 - 47 - 72	MidD
MRI-ESM2-0	5	352 - 511 - 554	0 - 0 - 2	MidR
NESM3	2	295 - 437 - 579	0 - 0 - 1	MidR
NorESM2-LM	1	92	0	NoDC
UKESM1-0-LL	3	1097 - 1325 - 1517	47 - 47 - 56	MidD

[a] The ensemble members are those listed in supp. Tables S1 and S2 without an asterisk, i.e. that do not have deep convection in the Arctic already over 2015-2023.

[b] Types are: “NoDC”, no deep convection in the Arctic over 2015 - 2100 in SSP5-8.5; “MidR”, rare Arctic deep convection, mid-run; “MidD”, Arctic deep convection peaks mid-run and declines; “Late”, Arctic deep convection late in the run.

Table S4. Correlation coefficient[a] for each model between the February ocean surface salinity (S), temperature (T), sea ice concentration (SIC) and surface wind speed (Wind), for the Eurasian basin (EB) and, for the models with Arctic deep convection, where the maximum MLD of Fig. 1 exceeds 500 m (DC). Bold fonts highlight the maximum correlation for each model.

Model	Region	S vs T	S vs SIC	S vs Wind	T vs SIC	T vs Wind	SIC vs Wind
BCC-CSM2-MR	EB	**-0.84**	0.26	0.25	-0.42	-0.23	-0.24
CAMS-CSM1-0	EB	**-0.98**	-0.26	0.24	-0.24	0.26	-0.39
CAS-ESM2-0	EB	-0.76	0.86	-0.34	**-0.78**	0.29	-0.32
CESM2-FV2	EB	**-0.98**	0.55	-0.40	-0.69	0.39	-0.36
CESM2-WACCM	EB	-0.78	0.60	-0.33	**-0.89**	0.36	-0.36
CESM2	EB	-0.77	0.76	-0.45	**-0.86**	0.37	-0.46
E3SM-1-1-ECA	EB	0.57	-0.41	0.35	**-0.75**	0.48	-0.58
FGOALS-f3-L	EB	0.23	0.25	0.27	**-0.66**	0.29	-0.29
FGOALS-g3	EB	-0.27	**0.54**	0.28	-0.32	-0.26	-0.24
GFDL-CM4	EB	**-0.90**	0.46	-0.29	-0.58	0.35	-0.31
NorESM2-LM	EB	**-0.89**	0.30	0.27	-0.41	0.27	-0.39
ACCESS-ESM1-5	EB	-	-	-	-	-	-
	DC	-	-	-	-	-	-
E3SM-1-0	EB	-0.25	0.27	0.25	**-0.80**	0.43	-0.55
	DC	-	-	-	-	-	-
E3SM-1-1	EB	-0.48	0.40	-0.22	**-0.84**	0.44	-0.54
	DC	-	-0.27	0.42	**-0.86**	0.56	-0.60
GISS-E2-1-G	EB	**-0.99**	0.64	0.24	-0.65	0.23	-0.30
	DC	0.72	-0.52	0.49	**-0.75**	0.48	-0.44
MRI-ESM2-0	EB	-0.70	0.61	-0.25	**-0.90**	0.31	-0.32
	DC	-	-	-	-	-	-
NESM3	EB	0.43	-0.48	0.32	**-0.84**	0.30	-0.35
	DC	-	-	-	-	-	-
CMCC-ESM2	EB	0.64	-0.62	0.40	**-0.80**	0.47	-0.58
	DC	0.70	-0.66	0.41	**-0.76**	0.47	-0.57
CanESM5-1	EB	0.32	-0.40	0.39	**-0.84**	0.54	-0.63
	DC	0.34	-0.42	0.39	**-0.83**	0.57	-0.67
CanESM5	EB	-	-	-	-	-	-
	DC	-	-	-	-	-	-
HadGEM3-GC31-LL	EB	-	-	-	-	-	-
	DC	-	-	-	-	-	-
UKESM1-0-LL	EB	0.70	-0.74	0.48	**-0.84**	0.52	-0.58
	DC	0.72	-0.75	0.50	**-0.84**	0.55	-0.61
ACCESS-CM2	EB	-0.31	-0.35	-0.23	**-0.83**	0.40	-0.44
	DC	0.55	-0.53	0.30	**-0.82**	0.41	-0.45
CNRM-CM6-1	EB	0.28	-0.30	0.31	**-0.82**	0.40	-0.50
	DC	0.39	-0.40	0.33	**-0.86**	0.50	-0.59
CNRM-ESM2-1	EB	0.38	-0.39	0.34	**-0.79**	0.50	-0.54
	DC	0.44	-0.44	0.37	**-0.84**	0.54	-0.57
EC-Earth3-Veg-LR	EB	0.41	-0.54	0.33	**-0.77**	0.43	-0.45
	DC	0.49	-0.62	0.41	**-0.81**	0.44	-0.50
EC-Earth3-Veg	EB	0.41	-0.44	0.32	**-0.76**	0.34	-0.37
	DC	0.56	-0.54	0.34	**-0.77**	0.40	-0.48

[a] Only correlations significant at 95% are shown; for models with more than one ensemble member, median of the correlations of the dominating sign.