~~Sea~~ Evaluation of sea ice variability in three ocean reanalyses and two regional models in the Barents and Kara Seas ~~: models versus reanalyses~~in 1991–2020

Cecilia Äijälä[1], Lucía Gutiérrez-Loza[2], Wolfgang Dorn[3], Siv K. Lauvset[2], Wieslaw Maslowski[4], and Petteri Uotila[1]

[1]University of Helsinki, Institute for Atmospheric and Earth System Research / Physics, Helsinki, Finland
[2]NORCE Research, Bjerknes Centre for Climate Research, Bergen, Norway
[3]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[4]Naval Postgraduate School, Monterey, CA, United States

Correspondence: Cecilia Äijälä (cecilia.aijala@helsinki.fi)

Abstract.

Satellite measurements indicate that the Barents and Kara Seas have experienced some of the most significant ~~declines in sea ice concentration and an increase~~ sea ice declines and increases in sea surface temperature in the Arctic. Understanding the processes driving these changes is of utmost importance to assess the implications of climate change in the Arctic and globally. Given the observational limitations in these high-latitude areas, coupled models and reanalysis products play a key role. However, it is essential to ~~better understand the strengths and limitations of such data products~~determine how reliably these products represent the observed sea ice decline and mixed-layer warming. Models are often evaluated on a pan-Arctic scale, but ~~results~~ the model skill might differ when considering sub-regions. This study examines two ~~regional coupled ocean-atmosphere models (HIRHAM-NAOSIM~~ coupled regional climate models (HIRHAM–NAOSIM 2.2 and RASM), two ~~coupled model intercomparison project, phase 6 (CMIP6) models (CNRM-ESM2_1, NorESM2-MM), two~~ regional ocean reanalyses (RARE 1.15.2 and TOPAZ 4B), one global ocean reanalysis (ORAS5)~~and OSI-SAF~~, and OSI SAF satellite observations over the period 1991–2020. We focus on spring (~~March-April-May~~March–May) and autumn (~~August-September-October~~August–October) sea ice area and sea ice concentration, the edge of the sea ice, and the temperature of the mixed layer. We also examine correlations between sea ice and mixed-layer temperature anomalies and five climate indices to assess how the datasets respond to large-scale climate variability. The results show a statistically significant decrease in sea ice area and sea ice concentration in the satellite measurements and in all reanalyses. However, ~~only one of the coupled models,~~ HIRHAM–NAOSIM ~~, exhibits a statistically significant sea ice~~ shows only a near-significant decline in the Barents Sea, ~~albeit weaker than in the reanalyses, and no coupled model~~ RASM shows no long-term change, and none of the two models shows a significant trend in the Kara Sea. Additionally, the reanalyses show a more pronounced warming of the mixed layer than the coupled models, particularly in autumn and in the Kara Sea. ORAS5 and TOPAZ agree most closely with the observed sea ice, partly because assimilated observations constrain their historical states, whereas RARE reproduces the decreasing trends but overestimates sea ice and shows anomalous mixed-layer temperature behaviour in the Barents Sea. The reanalyses also represent the relationships with

the climate indices more consistently than the coupled models, particularly for the Atlantic Multidecadal Oscillation and the Northern Hemisphere ocean heat content. Among the coupled models, HIRHAM–NAOSIM represents changes in the Barents Sea better than RASM, but in the Kara Sea, where none of the models captures the decreasing trends, RASM has a more realistic mean sea-ice state. Although the coupled models reproduce interannual variability, the results indicate that in particular the effects of large-scale changes of the ocean circulation are not realistically reproduced in the coupled models.

1 Introduction

The Arctic sea ice is an important part of the Arctic Ocean and the global climate system, as it influences ocean circulation and atmospheric processes by regulating the exchanges of moisture, momentum, and heat between the ocean and atmosphere. The reduction of sea ice has been linked to feedback mechanisms (Goosse et al., 2018), leading to Arctic Amplification (Stroeve and Notz, 2018). The two ~~ocean basins~~ marginal seas that have experienced the largest sea ice decline and surface warming in the Arctic are the Barents and Kara Seas (Årthun et al., 2012; Lind et al., 2018; Onarheim et al., 2018) (Fig. 1). The observed warming trends in this region, along with increased salinity and altered vertical fluxes that lead to reduced stratification, have been associated with intensified Atlantic Water (AW) inflow to the Arctic ~~,~~and with local and regional feedback processes, affecting both physical and ecological systems (Asbjørnsen et al., 2020; Ingvaldsen et al., 2021). These observed changes characterise the region as a warming hotspot (Lind et al., 2018) and as a mediator of climate change between the North Atlantic and the central Arctic basins (Smedsrud et al., 2013). Furthermore, the ~~Barents-Kara Seas~~ Barents and Kara Sea warming has also been identified as a key Arctic storyline, i.e., a physically self-consistent unfolding of past or plausible future events or pathways ~~(Shepherd et al., 2018), explaining a large proportion of the inter-model variability in pan-Arctic surface summer climate response to global warming among the Coupled Model Intercomparison Project, phase 6 (CMIP6) models (Levine et al., 2024)~~(Shepherd et al., 2018; Levine et al., 2024).

This paper evaluates several reanalysis products and regional ocean models and focuses on assessing what biases and trends in sea ice and mixed layer temperature in these products and models tell us about their strengths and weaknesses. This allows us to define concrete guidance to users on the appropriate use of such products and models for different purposes. Understanding the performance of reanalysis and models in these regions is crucial due to their role as hotspots in the current Arctic climate change.

The Barents and Kara Seas are shallow marginal seas of the Arctic Ocean located north of Norway and Siberia on the Siberian shelf. The ~~average depth of the~~ Barents Sea is ~~229 m, while the Kara Sea is shallower, especially in the east, with an average depth of 127 m, and 40 % of its area is less than 50 m deep. The Barents Sea is bounded by Svalbard and Franz Josef Land in the north and Novaya Zemlya in the east. The western border is towards~~ connected to the Greenland Sea ~~, where warm AW flows in~~ and receives warm Atlantic Water through the Barents Sea Opening (BSO)~~and later reaches the central Arctic Ocean through the Barents Sea. On the contrary~~. In contrast, the Kara Sea is rather isolated from the adjacent seas, with Novaya Zemlya in the west as a barrier to the Barents Sea, with inflow through the Barents Sea Exit (BSX) to the north ~~,~~and the shallow and narrow Kara ~~Strait~~ Gate to the south. The ~~Kara Sea is bounded to the east by Siberia and the Severnaya~~

~~Zemlya archipelago, with connections to the central Arctic Ocean to the north and the Laptev Sea in the northeast. The~~ mean annual freshwater inflow to the Kara Sea relative to its area is the largest in the world due to the inflow from the Ob, Yenisei, and other large rivers, affecting its stratification and circulation patterns. These seas play distinct roles in the Arctic system — for instance, the role of the Barents Sea as a key gateway between the North Atlantic and the Arctic Ocean (Skagseth et al., 2008; Smedsrud et al., 2013)~~and the~~ . This has a disproportionately large influence ~~of~~ on the Kara Sea landfast ice and on the halocline stability of the Eurasian Basin and beyond~~(Liu et al., 2025)~~, as noted by Liu et al. (2025).

Several studies have looked into the sea-ice variability in the Barents and/or Kara Seas and the processes behind it (e.g., Duan et al., 2019a . Warm Atlantic waters flow into the Barents Sea through the BSO, are progressively modified and lose heat to the atmosphere. Part of this modified water subsequently exits the Barents Sea through the BSX toward the Kara Sea, while Barents Sea waters also enter the Kara Sea through the Kara Gate. Shabanov et al. (2024) found that the strongest trends in Kara Sea ice conditions occur near pathways of Atlantic-water inflow, highlighting the role of Atlantification in the observed sea-ice decline. Several studies have also linked variability in Atlantic heat transport through the BSO to interannual variability in Barents Sea sea ice (Årthun et al., 2012; Wang et al., 2019, e.g.,). While this oceanic inflow can enhance sea-ice melt and contribute to Atlantification, atmospheric forcing also plays an important role in sea-ice variability, with air temperature, winds, and radiative fluxes exerting strong control, and can also influence sea ice indirectly by modulating ocean heat transport into the Barents Sea (Duan et al., 2019a; Wang et al., 2019). Recent work provides further evidence that the ocean can be a dominant driver of interannual sea-ice variability in parts of the region: Dörr et al. (2024) found that ocean heat transport dominates in the central and northeastern Barents Sea, whereas atmospheric drivers are more important in the northern Barents and Kara Sea. Overall, the atmosphere, ocean, and sea ice form a tightly coupled system, and internal feedbacks make it difficult to attribute observed variability to any single driver (Duan et al., 2019a; Chatterjee et al., 2024).

Although the Barents and Kara Seas are often studied as a single Barents–Kara region ~~(e.g., Li et al. (2022) and Dörr et al. (2024) , among others), these basins~~ (e.g. Li et al., 2022; Dörr et al., 2024), these marginal seas differ substantially in terms of circulation, stratification, and seasonal sea ice cycle. The Barents Sea is ice-covered in the northern parts in the winter and can become nearly ice-free across the ~~basin~~ region in summer, whereas the colder Kara Sea is typically fully ice-covered in winter and becomes largely ice-free in summer (Duan et al., 2019a, b). Observed sea ice decline has also been more pronounced in the Barents Sea than in the Kara Sea (Kumar et al., 2021). These differences motivate us to consider the two regions separately.

Advances in the observing and modelling capacities in recent decades have provided tools to widen our understanding of the polar regions. ~~Nowadays, observations~~Observations, reanalysis products, and models are ~~often~~ complementary tools used to better understand the Arctic climate system, the processes behind recent changes, and possible future storylines. Nevertheless, existing data products and models present both advantages and disadvantages. Ocean reanalyses combine models and observations via data assimilation; these products give an estimate of the ocean and sea ice (Balmaseda et al., 2015) and are important for studying processes that are only partly or indirectly observed, as well as for initialising and forcing models, even though they often struggle in data-sparse regions (Storto et al., 2019). Global Earth system models ~~, like the Coupled Model Intercomparison Project phase 6 (CMIP6) models (Eyring et al., 2016),~~ provide a representation of the whole Earth's climate and possible pathways for the future. However, the resolution of these models is usually too coarse to simulate all interactions

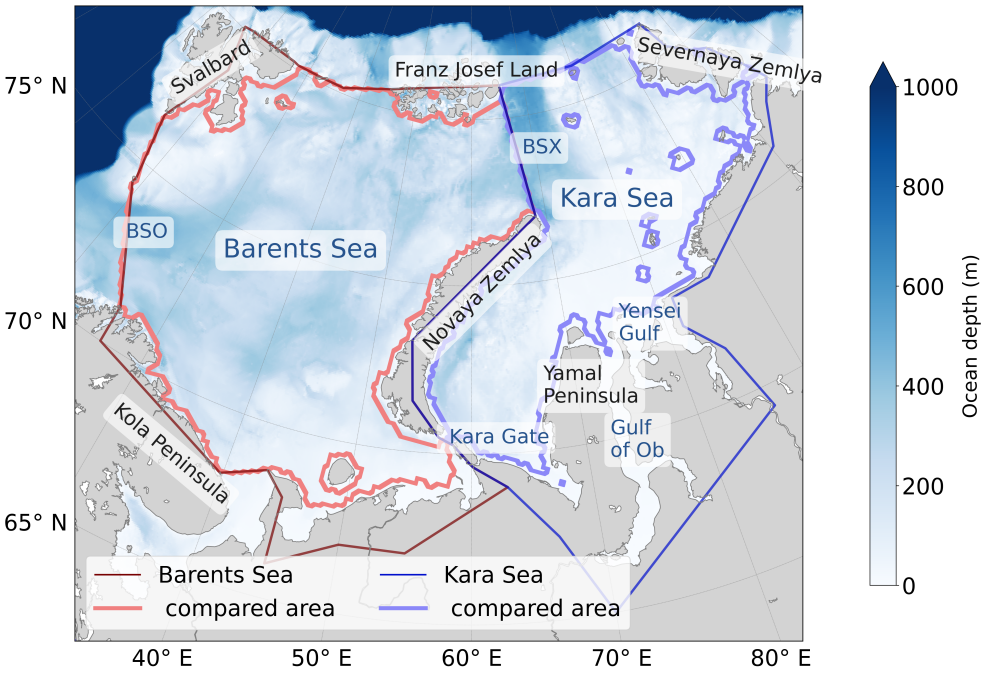


Figure 1. Barents and Kara Seas Topography and International Hydrographic Organization domains (International Hydrographic Organization and Sieger, 2012) (darker lines), and regions with data in all models (lighter colours), which are used for area-averaged comparisons. These are 1255 thousand km^2 for the Barents Sea and 682 thousand km^2 for the Kara Sea. BSO stands for Barents Sea opening ~~,~~and BSX for Barents Sea exit. Topographic data from GEBCO Bathymetric Compilation Group (2026)

and feedback between the components. Docquier et al. (2020) showed that a horizontal resolution of 1°, which is common for the global models, is not sufficient to correctly simulate the ocean currents entering the Arctic through the Barents Sea. Furthermore, they highlight the need for a high horizontal resolution to correctly represent the Arctic sea ice in the models. ~~Regional coupled~~ Coupled regional ocean-atmosphere models offer the ability to better resolve regional processes and associated feedbacks than their global counterparts due to their finer resolution. These models also offer the ability to respond to real-world changes in the large-scale atmospheric circulation by quasi-observed lateral boundary conditions from reanalysis data.

Table 1. Datasets used in this study.

Data source	Data type
HIRHAM–NAOSIM Regional Model 1/12° (∼ 9 km) 50 / 31 1991–2020 NAOSIM RASM OSI SAF	Regional Model Satellite Observation
NorESM ORAS	Global Model Reanalysis
CNRM TOPAZ	Global Model Regional Reanalysis
RARE	Regional Reanalysis
TOPAZ HIRHAM–NAOSIM	Regional Reanalysis Model
ORAS Global Reanalysis RASM	0.25° Regional Model 75 1

Datasets used in this study.

2 Reanalysis and regional model data

In this study, we investigate the capability of various global and regional reanalyses and coupled models to represent the sea ice and its recent changes in the Barents and Kara Seas. Understanding the performance of reanalysis and models in these regions is crucial due to their role as hotspots in the current Arctic climate change.

3 Data

We use data evaluate sea ice variability and mixed layer temperature from two coupled regional atmosphere-ocean models, HIRHAM–NAOSIM (Dorn et al., 2019) and RASM (Cassano et al., 2017); two global CMIP6 (Eyring et al., 2016) Earth system models, CNRM-ESM2-1 (Séférian et al., 2019) and NorESM-MM (Seland et al., 2020); data from two regional and one global (Cassano et al., 2017; Seefeldt et al., 2024), and two regional ocean reanalyses, TOPAZ version 4b (Arctic Ocean Physical Reanalysis (), (Xie et al., 2017; E.U. Copernicus Marine Service Information (CMEMS)) and RARE version 1.15.2 (Carton and Chepurin, 2023)and the global ORAS5 (Zuo et al., 2019). Additionally, we use OSI-SAF . We use OSI SAF satellite sea ice concentration (SIC) data OSI-450-a (EUMETSAT, 2022) as a comparison point for the sea ice model and reanalysis data. Additionally, we use the global reanalysis ORAS5 (Zuo et al., 2019), a global benchmark that is known to perform well (Zuo et al., 2019). ORAS5 is also the source for the ocean boundary conditions in HIRHAM–NAOSIM. ORAS5 is evaluated together with the other reanalyses for the sea ice variables, but is used as the comparison point for the mixed layer (ML) temperature evaluations. The datasets are described below and summarised in Table 1.

2.1 Coupled Regional Models

HIRHAM–NAOSIM is a coupled regional atmosphere–ocean–sea ice model applied over a circum-Arctic domain using rotated latitude–longitude grids with a horizontal resolution of 1/4° (∼ 27 km) and 40 vertical layers in the atmosphere component HIRHAM and 1/12° (∼ 9 km) and 50 vertical layers in the ocean–sea ice component NAOSIM. Detailed information on

the model components and their coupling is given by Dorn et al. (2019), and information on the current model version 2.2, particularly regarding recently introduced reparametrizations, is given by Aue et al. (2023) in their supplementary material. The simulation used for the present study was ~~running~~ run for the period 1979–2021, with the first part of the simulation disregarded to let the model find its own steady state. It is driven by ERA5 reanalysis data (Hersbach et al., 2020) at HIRHAM's lateral boundaries and at HIRHAM's lower and NAOSIM's upper boundaries, which are not covered by the domain of the other model component. For NAOSIM's open lateral boundaries in the North Atlantic, ORAS5 reanalysis data ~~(Zuo et al., 2019)~~ were used.

~~RASM Cassano et al. (2017)~~ The Regional Arctic System Model (RASM) is a fully coupled ~~atmosphere–land–ocean–sea~~ atmosphere-land-ocean-sea ice model. It uses ~~WRF (~~the Weather Research and Forecasting ~~)~~ (WRF; Powers et al., 2017) version 3.71 for atmosphere, a regionally configured version of ~~POP (~~Parallel Ocean Program ~~)~~ (POP; Smith et al., 2010) version 2.1 for the ocean and ~~CICE (~~the Los Alamos National Laboratory Sea Ice Model ~~)~~ CICE version 6 (Hunke et al., 2018) for sea ice. Additionally, it uses the ~~VIC (~~Variable Infiltration Capacity (VIC) land hydrology model and the streamflow routing ~~(RVIC)~~ (RVIC; Hamman et al., 2017) model for transporting the freshwater flux from the land to the ocean. The ocean and sea ice have a resolution of 1/12° (∼ 9 km) ~~in~~ on a rotated grid with 45 vertical levels for the ocean ~~,~~ and 5 thickness categories for sea ice.

2.2 ~~CMIP6 Earth System Models~~

~~We use two CMIP6 (Eyring et al., 2016) Earth system models, CNRM-ESM2-1 run r1i1p1f2 (hereafter CNRM) and NorESM2-MM run r1i1p1f1 (hereafter NorESM) (see Table 1). These models were chosen following the analysis described in (Levine et al., 2024) , where the future simulation of CNRM represents a warm humid Arctic Ocean and cold dry continents, while NorESM represents a cold dry Arctic Ocean and warm humid continents, as described in (Mottram et. al (submitted)).~~

~~The CNRM atmosphere component is based on ARPEGE-Climat_v6.3, and run with 150 km resolution while the ocean model is NEMO (Nucleus for European Models of the Ocean) version 3.6 (?) which is coupled to GELATO (Global Experimental Leads and ice for ATmosphere and Ocean) sea ice model (Salas Mélia, 2002) version 6 on a grid with a nominal resolution of 1° and 75 vertical ocean layers (Séférian et al., 2019).~~

~~NorESM uses CAM6-Nor for the atmosphere, BLOM for the ocean and CICE 5.1.2 for sea ice. Both the atmosphere and the ocean are run on a 1° grid (Seland et al., 2020), and the data were provided re-gridded to 70 vertical ocean layers.~~

2.2 Ocean Reanalysis

The Arctic Ocean physical reanalysis TOPAZ 4B (hereafter TOPAZ) ~~reanalysis (Xie et al. (2017),)~~ is a regional ocean reanalysis for the polar region north of 50° (Xie et al., 2017) N. It uses ~~HYCOM (~~the Hybrid Coordinate Ocean Model ~~)~~ ~~(Chassignet et al., 2003)~~ (HYCOM; Chassignet et al., 2003) version 2.2 and a ~~single category version of an early~~ single-category version of CICE ~~(Xie et al., 2023)~~ 3 for sea ice (Xie and Bertino, 2025). ERA5 ~~(Hersbach et al., 2020) data~~ data (Hersbach et al., 2020) is used as atmospheric forcing. The model uses the deterministic ensemble Kalman filter (DEnKF) to assimilate multiple ocean and sea ice observations simultaneously (Xie et al., 2023), such as SIC and drift from ~~OSI-SAF~~OSI SAF, gridded sea surface temperature (SST) from OSTIA, and in situ temperature and salinity profiles from the CORA database. Additionally, sea ice

thickness ~~from~~ for the winter months from CS2SMOS ~~,~~ and sea surface salinity based on the SMOS satellite are assimilated after 2010. For river runoff, the reanalysis uses a climatology based on ECMWF reanalysis and the TRIP hydrological model (Xie and Bertino, 2025). The data ~~is~~ are provided at a standard polar stereographic grid with 12.5 km resolution and 40 depth levels.

The Regional Arctic Reanalysis RARE 1.15.2 (hereafter RARE) is a regional reanalysis with a southern boundary at 44.7° N Carton and Chepurin (2023). It uses the GFDL MOM5/SIS1 numerical model and ERA5 as surface forcing. The model assimilates World Ocean Database (WOD) hydrographic profiles as well as satellite and in situ SST profiles (Carton and Chepurin, 2023). Sea ice thickness estimates of ~~PIOMAS~~ the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS; Zhang and Rothrock, 2003) 2.1 ~~(Lindsay and Zhang, 2006)~~ are used to constrain sea ice using a net surface heat flux relaxation term with a 30-day time scale~~, so that too thick areas~~ . Areas with thicker sea ice compared to PIOMAS receive more surface heat flux, and ~~too thin receives~~ thinner areas receive less (Carton and Chepurin, 2023). The continental discharge is based on a dataset combined from the World Meteorological Organisation Global Runoff Data Centre estimates and discharge estimates from R-ArcticNET ~~()~~ and the Arctic Great Rivers Observatory~~()~~; gaps in the dataset are filled with climatological monthly estimates. The dataset is provided on a regular 0.1° Mercator grid with 75 depth levels.

~~ORAS5 (~~Ocean Reanalysis System 5 ~~,~~ (ORAS5) (hereafter ORAS) ~~(Zuo et al., 2019)~~ is a global eddy-permitting ocean-sea ice ensemble reanalysis system by ECMWF. The resolution is 0.25°, and it has 75 vertical levels. ORAS is produced with the OCEAN5 system (Zuo et al., 2019), which uses NEMO V3.4 coupled to the LIM2 sea-ice model and the NEMOVAR ocean assimilation system. Data assimilated ~~includes~~ include in-situ T/S profiles from EN4 (Gouretski and Reseghetti, 2010) and SST and SIC through HadISST2 + OSTIA (Donlon et al., 2012), which is based on satellite observations. The atmospheric forcing comes from ERA-Interim from 1979 to 2014 and re-processed observations.

3 Methods

In this study, we primarily analyse monthly and seasonal means and trends of ~~ML temperature and~~ sea ice characteristics and ocean ML temperature, focusing on the spring (March, April, May~~, hereafter MAM~~) and autumn (August, September, October~~, hereafter ASO~~) seasons. These seasons are chosen because April and September ~~correspond to~~ are usually the months of maximum and minimum sea ice in the Barents Sea, respectively.

The study period is 1991–2020, however, ~~some of the datasets cover shorter periods~~ ORAS cover only 1991–2018 (see Table 1)~~, for which the available period is analysed~~. Monthly multiyear climatologies for each dataset are used to calculate monthly anomalies within the datasets. For these, the full period of each model has been used ~~. The difference in the period lengths is deemed~~ as we deemed using a shorter period length for ORAS preferable to a shorter common ~~24~~ 28 year time-period~~for~~ , leading to more robust results. ~~Analyses have also been done for~~ Sensitivity analyses were performed with the shorter common time period ~~(not shown),~~ without major changes to the conclusions (not shown).

Given that the regional models and reanalysis products used in this study have different grids, a bilinear interpolation was used to regrid all datasets. The equal-area 12.5 km EASE2 grid was used for this purpose. Additionally, only grid cells with data in all models were used for calculations of area averages (see Fig. 1).

3.1 Sea ice variables

For sea ice, the analysed variables are ~~sea ice concentration (SIC)~~SIC, sea ice area (SIA), defined as the sum of grid-cell area multiplied by the SIC for all cells with SIC >15 %~~)~~, and Integrated Ice Edge Error (IIEE), defined as the area where the model or reanalysis and ~~OSI-SAF~~ OSI SAF satellite observations disagree on the ice concentration being above or below 15 %~~)~~. The IIEE is calculated following Goessling et al. (2016):

$$\mathrm{IIEE} = O + U \tag{1}$$

with

$$O = \int_A \max(c_\mathrm{m} - c_\mathrm{o}, 0)\, dA \tag{2}$$

and

$$U = \int_A \max(c_\mathrm{o} - c_\mathrm{m}, 0)\, dA\,, \tag{3}$$

where O and U stand for over- and underestimation, A is the area of either the Barents or Kara Seas (see Fig. 1), $c = 1$ where the sea-ice concentration is greater than 15% and $c = 0$ elsewhere. The subscript m denotes the model or reanalysis data, and the subscript o denotes the satellite observations.

3.2 Ocean variables

For the ocean, we focus on the mean temperature of the ~~ocean mixed layer (ML)~~ML. The depth of the ML is calculated for each month using the density anomaly (σ) criterion, as the depth at which the density difference to the surface (5 m) is equal to 0.03 kg m^{-3}, following the method ~~by~~ of Sallée et al. (2006, 2013). Density at 5 m depth is used as the surface ~~,~~because it is the uppermost depth available in all models, and the depth of the ML is where the density differs from the density at 5 m by σ. The monthly mean temperature for each grid cell is calculated ~~for this depth~~ from the surface to the depth of the ML. The monthly mean temperature for each region is then calculated by taking the area-weighted mean of each grid cell.

3.3 ~~Statistics calculations~~Climate indexes

In addition to the sea ice and ML temperature datasets, we use time series of five climate indices to analyse potential drivers of sea ice variability and inter-product differences by correlating the indices with SIC and ML temperature. The Indices are: North Atlantic Oscillation (NAO; NOAA CPC, b), Atlantic Multidecadal Oscillation (AMO; van Oldenborgh et al., 2009), Pacific Decadal Oscillation (PDO; Newman et al., 2016), East Atlantic/ Western Russia teleconnection (EA/WR; NOAA CPC, a), and

the Northern Hemisphere Ocean Heat Content (0-700m) (NH OHC; NOAA NODC). We used the AMO and PDO versions based on NOAA Extended Reconstructed SST V5 (ERSST v5; Huang et al., 2017).

3.4 Statistical analysis

The trends are calculated using a linear least-squares regression. The statistical significance of the trends is assessed using the p-value for the null hypothesis that the regression slope equals zero. Trends are considered significant at the $p < $ ~~0.1~~ 0.05 level. The significance of the differences between satellite-observed ~~sea-ice concentration~~ SIA or SIC and models/reanalyses, presented in ~~section 4.3is~~ Sect. 4.3, as well as differences between ORAS and other dataset ML temperatures, are calculated using a Student's t-test. The Pearson's correlation coefficient (PCC), r, is used to quantify the correlation between the sea-ice area and ML temperature in ~~section 4.4,~~ Sect. 4.4, as well as the correlation between climate indexes and sea ice and ML temperatures with:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} , \tag{4}$$

where $\bar{x}$ is the mean of vector x_i, and $\bar{y}$ is the mean of vector y_i. The Fishers z-test is used, with a threshold of $p < 0.05$, to determine if differences between the correlations are significant:

$$Z = \frac{z_1 - z_2}{SE} , \tag{5}$$

where z_1 and z_2 are calculated from r using the Fisher's transformation

$$z_n = \frac{1}{2} \ln\left(\frac{1 + r_n}{1 - r_n}\right) \quad \text{with} \quad n \in \{1, 2\} . \tag{6}$$

The standard error SE in (5) is given by

$$SE = \sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}} , \tag{7}$$

where N_1 and N_2 represent the sample sizes of datasets 1 and 2, respectively.

4 Results and discussion

4.1 Sea ~~Ice Area~~ice area

The mean and trend of the SIA for the seasons with maximum (spring) and minimum (autumn) sea ice are shown in Table 2. Figure 2 shows the difference between the mean SIA of each month and the monthly area climatology of each dataset.

4.1.1 Barents Sea sea ice area

~~The mean and trend of the sea ice area for the seasons MAM and ASO are shown in table 2.~~ In the Barents Sea, the satellite observations from ~~OSI-SAF~~ OSI SAF indicate a mean ~~sea ice area of 400~~ SIA of 412 000 km^2 in spring ~~(i.e., MAM) and 28~~ and

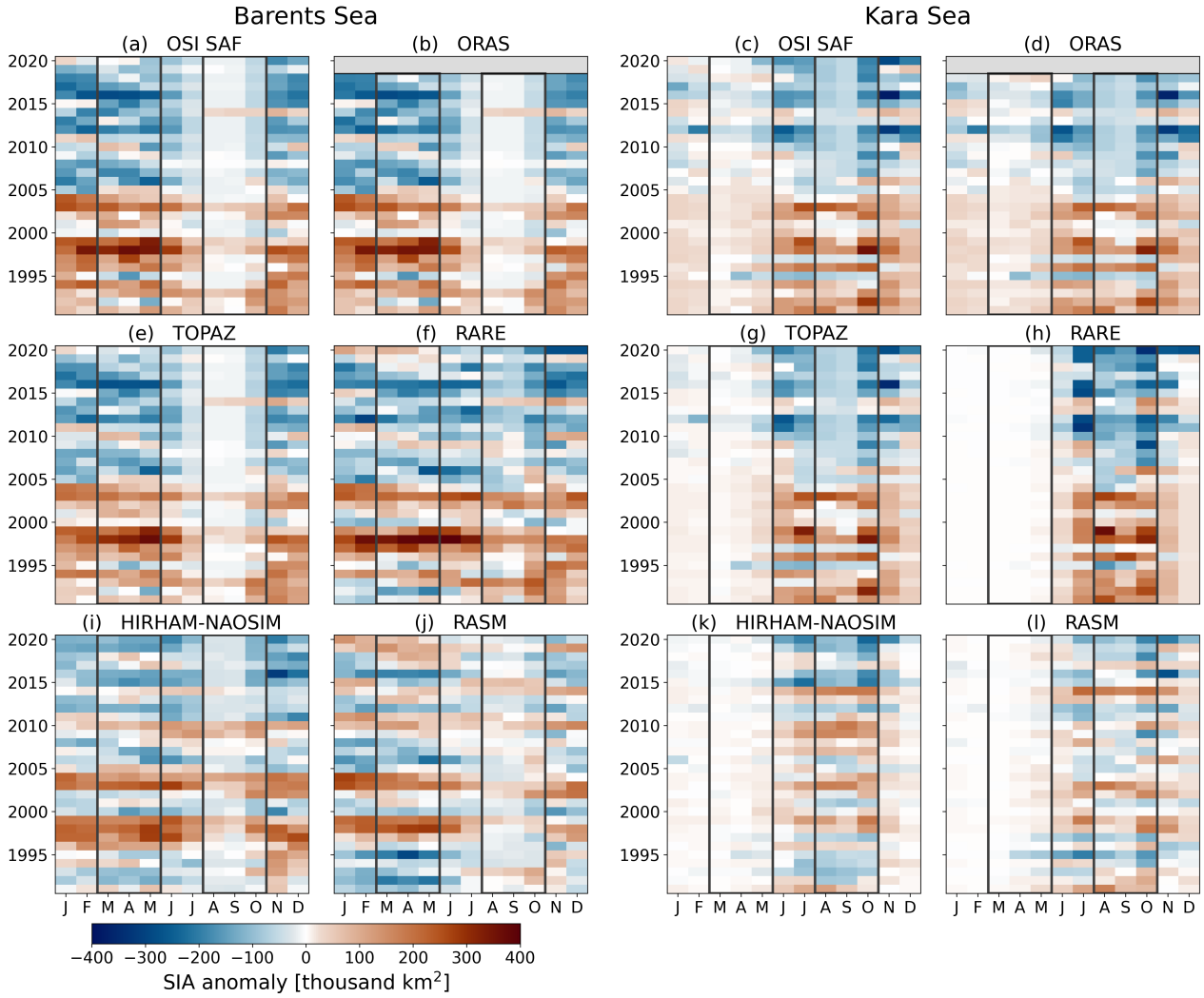


Figure 2. ~~Sea ice area~~ SIA anomaly (SIAa) in each ~~model~~ dataset for the Barents (left) and Kara Sea (right). Spring ~~(MAM)~~ and autumn ~~(ASO)~~ seasons are marked with dark grey boxes. The anomaly is calculated by taking the difference from the monthly multiyear mean for each month.

29 000 km^2 in ~~the autumn(i.e., ASO)~~autumn. According to the observations, the ~~sea ice area~~ SIA shows statistically significant decadal trends in both seasons during the period 1991–2020 (Table 2). All reanalyses also show a statistically significant SIA trend in both seasons~~in the sea ice area, while of the coupled models only HIRHAM–NAOSIM shows a statistically significant trend.~~

~~Figure 2shows the difference between the mean sea ice area of each monthand the monthly area climatology of each model. In the Barents Sea, values above the mean can be seen mostly~~ , while the coupled regional models do not. The monthly SIA

Table 2. Spring and autumn mean SIA [thousand km^2], decadal SIA trend [thousand km^2/decade], and p-value for Barents and Kara Seas. The values are calculated using the time periods shown in Table 1. Mean values are bold when the yearly seasonal means differ significantly from OSI SAF (t-test, $p < 0.05$). Trends and p-values are bold when the given $p < 0.05$. (least-squares regression)

	Barents Sea				
	Spring				
Model	Mean	Trend	p-value		Mean
~~HIRHAM–NAOSIM 611~~ **~~44.4~~ ~~0.059~~** ~~66~~ **~~20.6~~ ~~0.074~~** OSI SAF	~~586~~ *412*	~~-1.0~~ **-99.2**	~~0.550~~ **0.000**	~~172~~	~~-13.7~~ 29
ORAS	~~0.550~~ 404	~~144~~ **-113.6**	~~-20.0~~ **0.000**	~~0.19 NorESM~~	~~655~~ 25
TOPAZ	~~-46.2~~ 438	~~0.121 CNRM~~ **-93.5**	~~556~~ **0.000**	~~-44.7~~	~~0.339~~ 30
RARE	~~616~~ **636**	**~~-55.0~~-56.3**	**~~0.039~~0.036**	~~121~~	**~~-71.2~~124**
HIRHAM–NAOSIM	**629**	~~605~~ **~~-8.0~~ ~~0.013~~** ~~108~~ -45.3	**~~-80.6~~** 0.057	**~~0.000~~** ~~ORAS~~	~~392~~ **67**
RASM	**~~-72.5~~550**	**~~0.000~~** ~~OSI-SAF~~ 20.2	~~400~~ 0.431	**~~-96.2~~**	**~~0.000~~62**

~~MAM and ASO sea ice area mean thousand km^2, decadal trend thousand km^2/ 10 yr, and p-value for Barents and Kara Seas. Calculated for the periods in Table 1. Trends and p-values are bold when $p < 0.1$.~~

anomalies (SIAa) for OSI SAF (Fig. 2a) show a similar pattern for each month, indicating that a trend is present throughout the year. In general, the SIA is larger than the multiyear mean before 2006, ~~and values below the~~ while smaller SIA values in comparison to the multiyear mean can be seen ~~mostly after 2005 (Fig. 2n).~~ after 2005. The drop below the climatological mean coincides with winter 2005–2006, when the second-lowest April sea ice extent ~~to date of~~ in the Barents Sea to date was recorded (Fig. 2~~n~~a). In 2005 and 2006~~the maximum sea ice area~~ , the maximum SIA of the Arctic had two consecutive ~~record low~~ record-low years, with ~~the negative sea ice anomalies most dominant~~ persistent negative anomalies in the eastern Arctic basin (Comiso, 2006; Li et al., 2022). Regarding winter extreme ~~sea ice areas~~SIA, the minimum April extent in 2016 is shown as dark blue between February and May~~, while~~ . In contrast, distinctly high winter values ~~occur~~ occurred in 1998 and 2003.

From the ocean reanalyses, ORAS ~~slightly~~ underestimates the mean area in both spring and autumn, while TOPAZ ~~slightly overestimates it, but~~ overestimates it. However, the differences are small and not statistically significant (Table 2). In the spring, ORAS slightly overestimates the trend and TOPAZ slightly underestimates the trend, while the opposite is true in the autumn. ORAS and TOPAZ ~~show similar temporal patterns as OSI-SAF, which is not surprising as they assimilate OSI-SAF~~ exhibit temporal variability consistent with OSI SAF, including similar interannual and year-to-year variations (Fig. 2~~j, m, n~~a, b, e and A1).

The ~~third ocean reanalysis , RARE ,~~ ocean reanalysis RARE shows a temporal pattern rather similar to ~~OSI-SAF~~ OSI SAF (Fig. 2~~i, n). Their~~ a, f), with, for example, consistent high and low ~~sea ice area yearscoincide, for example. The most notable difference is that RARE shows interannual declines in August and September, unlike OSI-SAF, ORAS and TOPAZ. This is because the mean RARE sea ice area in both spring and autumn is clearly overestimated , resulting in the largest,~~ SIA years. The RARE SIA is, however, overestimated compared to OSI SAF, especially in the autumn (Table 2). RARE

also has the largest statistically significant decadal trend in autumn~~due to greater ice loss (ASO; Table 2)~~. This greater absolute ice loss occurs from a more extensive mean ice cover. The trend in spring, however, is smaller than that of the other ocean reanalyses and ~~OSI-SAF~~OSI SAF. Thus, RARE reproduces the direction and broad temporal variability of the observed change, but performs less well in representing the mean SIA and the seasonal magnitude of the trend than the other reanalyses. The closer agreement of ORAS and TOPAZ with OSI SAF is expected as both products assimilate satellite-derived sea-ice observations, which constrain their sea-ice state toward observations, whereas RARE does not directly assimilate satellite SIC. This difference likely contributes to the better representation of mean SIA in ORAS and TOPAZ, although sea-ice performance also depends on the assimilation method, atmospheric forcing, and underlying ocean–sea-ice model (Uotila et al., 2019; Storto et al., 2019; Zhang et al., 2021; Xiong et al., 2025). Thus, ORAS and TOPAZ are more suitable for reproducing the observed historical SIA, but their closer agreement with OSI SAF does not necessarily indicate a better representation of the physical processes governing sea-ice variability.

~~Of regional coupled~~ From the coupled regional models, the mean HIRHAM–NAOSIM spring ~~sea ice area~~ SIA in the Barents Sea is ~~close~~ overestimated compared to OSI SAF and closer to that of RARE, and both spring and autumn mean areas are ~~clearly~~ significantly higher than those of ~~OSI-SAF~~ OSI SAF, though the SIA in autumn is only about half that of RARE (Table 2). The ~~interannual trend in spring in~~ linear trends for HIRHAM–NAOSIM ~~is a lot smaller than that of OSI-SAF, even smaller than RARE, while the autumn~~ have p-values between 0.05 and 0.1 and are thus not statistically significant. In spring, the trend is smaller than for any reanalysis, while in autumn the trend is close to that of ~~OSI-SAF~~OSI SAF. The negative anomalies are in the majority during the last ten years, signifying an interannual decline (Fig. 2~~a~~i). In contrast to the three ocean reanalyses and ~~OSI-SAF~~OSI SAF, HIRHAM–NAOSIM has an oscillating ~~sea ice area~~ SIA with three above-average sea ice periods separated by periods with below-average values (Fig. 2a, ~~i, j, m).~~

~~Of coupled models, the~~ b, e, f, i). RASM spring mean ~~sea ice area is the smallest, and the closest to OSI-SAF, while NorESM has the highest and the most overestimated mean area (Table 2). In autumn , the CNRM mean value is the lowest among the coupled models, with similar values to TOPAZ, ORAS, and OSI-SAF. RASM has an autumn mean close to~~ SIA is smaller than that of RARE and HIRHAM–NAOSIM ~~, and NorESM has the second-highest autumn mean after RARE~~but still statistically different from OSI SAF, while the autumn mean is close to that of HIRHAM–NAOSIM 2). RASM shows ~~interannual variations but no clear decadal trends~~ no clear trend, only interannual variations (Fig. 2~~e~~).

~~Regarding interannual variations , CNRM shows oscillations as alternating negative and positive anomalies (Fig. 2f), while in NorESM values below average dominate around 2010. There is, however, no statistically significant~~ j). The interannual variations in the SIAa for both coupled regional models are similar to each other and closely follow OSI SAF and the reanalyses as well, even though they lack the trend (Fig. 2 ~~b).~~

4.1.2 ~~Kara Sea sea ice area~~

~~As~~ and A1a). This suggests that the coupled models reproduce much of the interannual variability in the Barents Sea, but not the magnitude of the observed long-term decline. Of the two coupled models, HIRHAM–NAOSIM performs better in terms of the SIA trend, even though the trend is not statistically significant at $p < 0.05$, whereas RASM has a smaller mean SIA bias. For

coupled atmosphere-ocean models, it can not be expected that the results follow observations exactly, as they have their own atmospheric models, reducing the similarity with the ocean reanalyses, which are (directly) forced by the ERA atmospheric reanalysis. The coupled models represent a possible pathway of the climate system rather than a direct reproduction of observed conditions. Although coupled regional models can be driven by quasi-observed data at the lateral boundaries, the ~~OSI-SAF, TOPAZ and ORAS sea ice areas in the Kara Sea are relatively close to each other (Table 2 and Fig. 2l, o, p). TOPAZ slightly overestimates the mean area compared to OSI-SAF in both spring and autumn, while ORAS slightly underestimates it , as was the case in the Barents Sea. Of coupled models ,~~ simulation inside the model domain evolves freely. Consequently, small differences in the initial state can lead to different simulation results.

4.1.2 Kara Sea sea ice area

In the Kara Sea, the spring SIA does not statistically differ between OSI SAF, ORAS, and HIRHAM–NAOSIM~~has the lowest spring sea ice area, being the same as OSI-SAF. In other coupled models (RASM, NorESM and CNRM), spring sea ice areas are identical, indicating an entirely ice-covered Kara Sea~~ , while TOPAZ, RASM, and RARE overestimate the mean, but the differences are small (Table 2). In autumn, the ~~CNRM mean sea ice area is unrealistically the smallest of all products analysed in this study, while other coupled models (~~mean SIA of ORAS and TOPAZ does not statistically differ from OSI SAF, while RASM, HIRHAM–NAOSIM, ~~RASM and NorESM)~~ and RARE overestimate the mean ~~autumn sea ice area compared to OSI-SAF.~~

~~In~~ SIA, with the RARE mean being largest. The differences between the datasets are small in spring because the Kara Sea is typically almost completely ice-covered during this season, limiting the variability of an area-based metric and making SIA relatively insensitive to concentration differences within the compact ice cover (Duan et al., 2019a, b; Kumar et al., 2021) . Autumn therefore provides a more discriminating test of model performance in the Kara Sea, ~~only OSI-SAF~~ because differences in melt, oceanic and atmospheric heat input, and the timing of freeze-up are directly reflected in the position of the ice edge and total ice area.

OSI SAF and the reanalyses show a statistically significant decreasing trend in both seasons (Table 2). In spring, the ~~sea ice area declines less than in autumn, when the TOPAZ decline is a bit less than that of ORAS and OSI-SAF (Table 2)~~ SIA trend is small, with the TOPAZ trend smaller than OSI SAF and ORAS, and the RARE trend is ~~quite small, even though it~~ even smaller. Although the RARE trend is statistically significant, its magnitude is small relative to its mean spring SIA (Table 2) and (Fig. 2c, d, g, h). In the autumn, however, the RARE trend is ~~clearly~~ larger than in TOPAZ, ORAS, and ~~OSI-SAF~~OSI SAF, of which TOPAZ has the largest trend and ORAS the smallest. ~~Moreover, RARE overestimates the mean sea ice area in spring and Autumn in the Kara Sea~~Similar to the Barents Sea, the large RARE trend occurs together with an overestimation of the mean autumn SIA, indicating that RARE captures the direction of the observed change but not the mean state of the SIA.

Departures from the climatological mean ~~sea ice area in OSI-SAF, TOPAZ and ORAS~~ SIA in the datasets are low in winter and spring (January–April), when the Kara Sea is practically completely ice-covered~~. The RARE monthly~~ , even though some change can be seen in OSI SAF and ORAS. The interannual sea ice anomaly ~~pattern resembles OSI-SAF, TOPAZ and ORAS,~~

even though its changes in winter appear smaller and in summer larger patterns in the reanalyses resemble OSI SAF (Fig. 2k, l, o, p).

As c, d, g, h), similarly to what was seen in the Barents Sea, . In HIRHAM–NAOSIM, SIA in the Kara Sea is below average in 1992–1995 and decline again below average in 2015, with a clear drop in summer and autumn (Fig. 2k and A1b), even though no statistical linear trend can be seen. RASM does not display an evident any interannual sea ice decline, and its trends are not statistically significant (Table 2 and Fig. 2g). Somewhat analogously, CNRM sea ice area in the Kara Sea oscillates inter-annually as in l). Again, the interannual variation of the coupled regional model resembles that of OSI SAF and the reanalyses, but the difference to OSI SAF is larger than for the Barents Sea (Fig . 2f and h). The HIRHAM–NAOSIM sea ice area A1b). Thus, while the coupled models reproduce some of the observed interannual variability, neither reproduces the long-term SIA decline in the Kara Seaoccurred in 2015, with a clear drop in summer and autumn (Fig. 2c).

4.2 Sea Ice Edgeice edge error

Integrated Ice Edge Error (IIEE) of the (a) Barents and (b) Kara Seas, as well as (c, d) Overestimation - Underestimation of the sea ice edge, and (e,f) time series of the yearly mean Overestimation-Underestimation. For the Kara Sea, only values during May–December are shown, as there is no ice edge during winter and spring in most years.

Table 3. IIEE mean and ice edge Overestimation - Underestimation (O-U) mean, trend, and p-value for Barents and Kara Sea. Means and trends are expressed in thousand km^2. Trends and p-values are bold when $p < 0.05$, indicating a significant trend. The trends are calculated from the time series shown in Figure A2e, f. For the Kara Sea, only the May–December period is analyzed due to the region being fully ice covered in the winter.

	Barents Jan–Dec				
	IIEE	O–U			IIEE
Model	Mean	Mean	Trend	p-value	Mean
HIRHAM–NAOSIM 237 220 **37.1** **0.007** 123 92 **63.7** **0.000** RASM ORAS	142 31	103 -20	**100.0**5.4	0.000	87 29
NorESM TOPAZ	251 47	212 -23	**93.4**12.5	**0.001**0.000	132 30
CNRM RARE	175 160	54 147	**68.5**19.9	**0.098**0.004	142 80
RARE HIRHAM–NAOSIM	155 243	142 224	**19.0**38.6	0.006	75 64 **11.7** **0.018** TOPAZ 4
ORAS RASM	30 146	-19 106	**5.4**102.5	0.000	25 93

IIEE mean and ice edge Overestimation - Underestimation mean, trend and p-value for Barents and Kara Sea. Means and trends are expressed in thousand km^2. Trends and p-values are bold when $p < 0.1$, indicating a significant trend. The trends are calculated from the time series shown in Figure ?? (e,f).

Given the good sea ice area agreement between ORAS , TOPAZ and OSI-SAFmean SIA agreement of ORAS and TOPAZ with OSI SAF discussed in Sect. 4.1.1, it is not surprising that their integrated ice edge errors (IIEE), relative to OSI-SAFOSI SAF, are much smaller than those of the other datasets . Also, their (Table 3). Furthermore, the monthly mean IIEEs vary

relatively little (Table 3 and Fig. **??**for these products show only small variations (Fig. A2), both in the Barents Sea and in the Kara Sea. In the Barents Sea, ORAS and TOPAZ IIEEs are quite similar from July to December, while TOPAZ has a higher IIEE from January to June (Fig. **??**a) due to the due to a greater underestimation of the ice edge location than ORAS (Fig. **??**A2a, c). In the Kara Sea, the largest difference between ORAS and TOPAZ occurs in October and November when TOPAZ overestimates the edge while ORAS underestimates it. The ice edge of TOPAZ tends to recede too fast in early summer and refreeze too fast in early winter, which is most likely the reason for this difference (Laurent Bertino and Jiping Xie, 2025). The trend of the annual mean O-U is small for ORAS and TOPAZ and decreases over time, indicating increasing accuracy in their ice edge simulation (Fig. **??**Table 3 and Fig. A2e).

After ORAS and TOPAZ, RASM has the third smallest IIEE in the Barents Sea (Table 3) with clear seasonal IIEE and O-U cycles having maxima in June and minima in September (Fig. **??**a,c), which coincides with the minimum sea icearea month. The RASM O-U time series shows a distinct annual trend, starting with slightly negative values and ending with values around 0.2 (Fig. **??**e). This is in line with its non-significant trend in sea ice area (Table 2)This is consistent with the good agreement in SIA shown in Sect. 4.1. Thus, both products perform well not only in representing the total amount of sea ice, but also its spatial extent.

The RARE mean IIEE in the Barents Sea is only slightly larger than that of RASM (Table 3), with a similar seasonal cycle but somewhat larger monthly IIEE values due to an overestimation of the sea ice edge IIEE of RASM and RARE are similar to each other, with the maximum IIEE in the summer (June or July) (Fig. **??**a, c). In contrast, the RARE O-U trend is clearly smaller than that of RASM (Table 3) due to the smaller annual O-U trends after 2010 (Fig. **??**e).

The ESMs have difficulties in correctly identifying the ice edge due to their coarse resolution (Table 1). Therefore, the IIEE is expected to be higher than for the other higher resolution models. The CNRM monthly A2a, b). RARE has a larger mean IIEE in the Barents Sea is close to RARE and RASM from July to November but clearly higher in other months (Fig. **??**a), as can be expected due to the coarser resolution. As for all the products, the CNRM monthly minimum IIEE occurs in September, but the monthly maximum is one month earlier, in May. Unlike most other products analysed, CNRM underestimates the ice edge from September to November and overestimates it in other months comparably moderately (Fig. **??**c), resulting in a relatively small positive mean O-U and RASM a larger one in the Kara Sea (Table 3). The CNRM O-U interannual trend variability is particularly high before 2006 with big swings between negative and positive values (Fig. **??**c) with strongly negative values in 1993 and 2003, coinciding with the years of low sea ice area (Fig. 2f).

The coupled models HIRHAM–NAOSIM and NorESM have the highest IIEEs and O-Us in the Barents Sea is larger in RARE than in RASM for both regions. The main difference between RARE and RASM is in the O-U trend (Table 3). At least for NorESM this could be expected due to its coarse resolution. However, the monthly HIRHAM–NAOSIM IIEE is the highest among all products from October to December, with the maximum in December, although its monthly IIEE is close to those of RASM, RARE and CNRM from July to September (Fig. **??**a). From January to September, the NorESM monthly IIEE is the highest, and rather close to the HIRHAM–NAOSIM monthly IIEE from April to June. The O-U trend of HIRHAM–NAOSIM is comparably small and relatively close to TOPAZ, ORAS, and RARE. This is likely associated with its declining sea ice area and Fig. A2e, f). The trend is over 5 times larger in RASM than in RARE, because the ice edge in RASM does not change

much over the study period (~~Fig. 2a). In contrast, the NorESM O-U trend is quite high, close to RASM, which has the highest O-U trend, with the sea ice edge error increasing over time (Fig. ??c).~~ Sec. 4.3), which results in the model underestimating the ice edge at the start of the study period and significantly overestimating it at the end. Although RARE and RASM have similar mean IIEE values, their behaviour differs substantially. RARE captures the observed decrease in SIA (Section 4.1), while the larger O–U trend in RASM reflects its lack of a comparable long-term change.

~~In the Kara Sea, RARE has a slightly smaller mean IIEE than RASM in every month (Table 3 and Fig. ??b). The RARE O-U trend is small compared to the RASM, which shows the largest O-U trend among the analysed products due to its unrealistic decadal sea ice area trend (Fig. 2e). The~~ HIRHAM–NAOSIM ~~mean IIEE is the third-highest mean IIEE among the analysed products, with monthly values either third or second in the Kara Sea (Table 3 and Fig. ??b). It has~~ has the highest mean IIEE and the highest mean O-U ~~and an exceptionally high monthly O-U in October (Fig. ??d). The HIRHAM–NAOSIM O-U trend is almost as large as in RASM, with increasing sea ice edge errors toward the end of the time series~~ (Table 3~~and Fig. ??f~~). The ~~NorESM mean IIEE is the second-highest~~ error is largest in December in the Barents Sea and in October in the Kara Sea ~~, with the highest monthly values among the products in August and September (Table 3 and Fig. ??~~(Fig. A2a, b). ~~From the coupled models, it is the only one that has no statistically significant O-U trend (Table 3). The CNRM coupled model has the highest mean IIEE~~ In the Barents Sea, the trend is smaller than the RASM trend, but in the Kara Sea ~~. In contrast to the other coupled models, the CNRM mean O-U is negative, indicating more underestimation than overestimation in the ice edge (Table 3), due to underestimations from August to November~~ the trend is similar. The variability from year to year in HIRHAM–NAOSIM is larger than in the other models (Fig. ~~??d), as it was seen in the ASO mean area as well (Table 2~~A2e, f). The ~~CNRM O-U shows large interannual variability, with a particularly low O-U value in 2003, as was the case in the Barents Sea as well (Fig. ??f)~~IIEE results show that a realistic SIA does not necessarily imply a realistic ice edge. For example, RARE reproduces the observed SIA trends better than RASM, while HIRHAM–NAOSIM has a smaller SIA bias than RARE but a larger ice edge error. Evaluating both metrics therefore provides a more complete assessment of model performance.

4.3 Sea ice concentration

The SIC difference between the analysed products and ~~OSI-SAF~~ OSI SAF is shown in Fig. 3 ~~, both for the season with the maximum sea ice area (spring) and for the season with the minimum sea ice area (autumn). The~~ for the spring and autumn. The mean SIC and the decadal trends for these two seasons are shown in Fig. ~~??~~ 4 and Fig. ~~??~~5. The evolution of the SIC over the study period 1991–2020, along with the differences between the reanalysis and model products against the observed ~~OSI-SAF concentration~~OSI SAF concentration, is discussed below.

4.3.1 Spring sea ice concentration

In spring, the ~~OSI-SAF~~ OSI SAF 1991–2020 mean sea ice edge (SIC =15%) is located at, or slightly north of, the 75° N latitude until about 50° E longitude, where it turns south-west toward the southern part of the Kola ~~peninsula~~ Peninsula (Fig. 3n and Fig. ~~??n~~4a). The boundary between the ~~inner~~ pack ice and the marginal ice zone (MIZ) (SIC = 80%, later called the inner MIZ boundary) goes from south-east Svalbard ~~,~~ to the middle of the Barents Sea Exit (BSX) at the ~~Barents–Kara~~

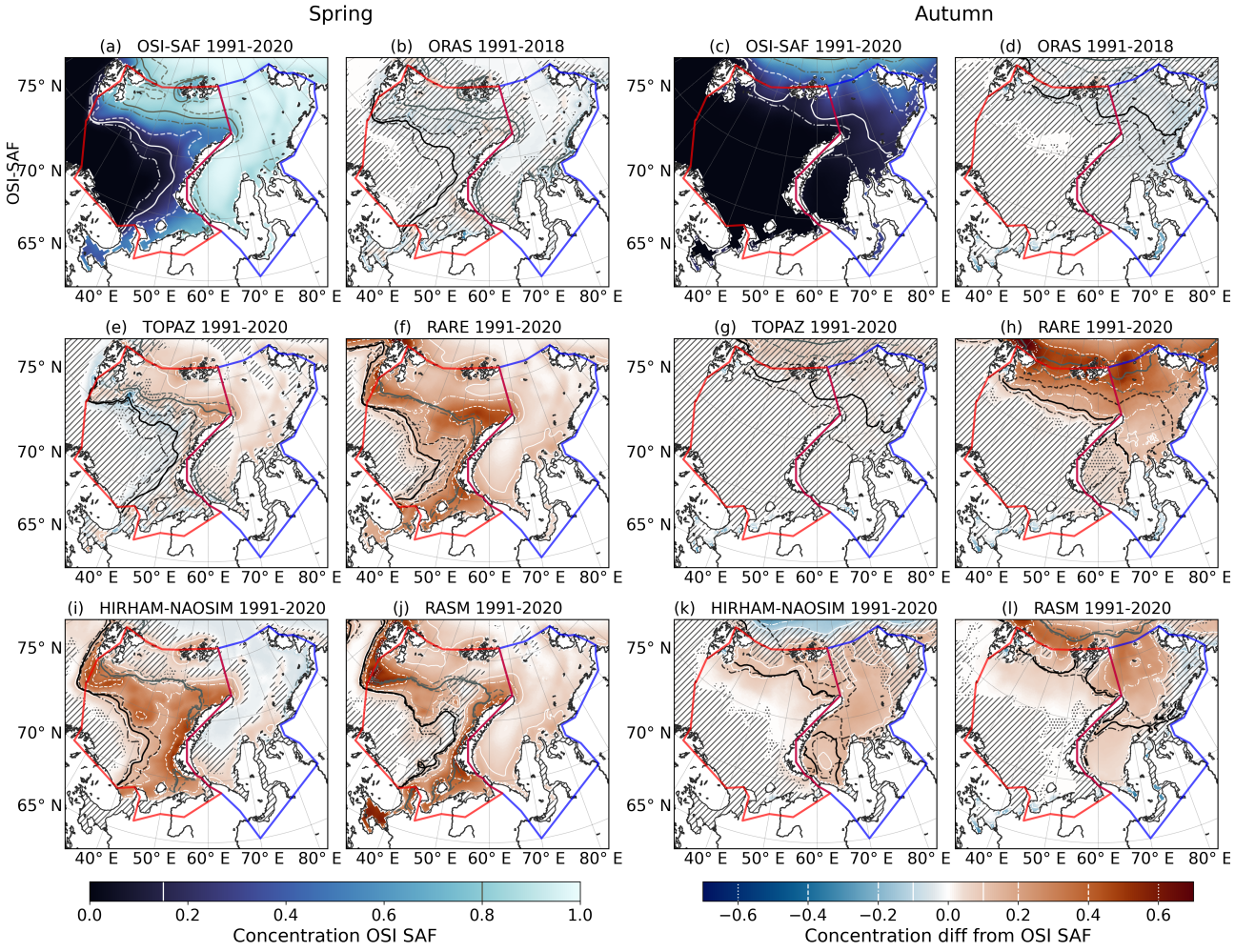


Figure 3. Mean difference between model ~~/~~or reanalysis and ~~OSI-SAF sea ice concentration~~ OSI SAF SIC for ~~MAM~~ spring (left) and ~~ASO~~ autumn (right), and mean ~~OSI-SAF~~ OSI SAF concentration (~~a~~ a and ~~b~~c). The difference plots (~~a~~ m~~b~~, ~~c~~d–~~l~~) have been masked where the difference is not statistically significant (t-test, ~~$p < 0.1$~~"////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$). The black (white) solid contour lines show the mean ice edge (SIC = 15%), and the grey contour shows the ~~edge~~ inner boundary of the marginal ice zone (MIZ) (SIC = 80%) for the whole time period. To represent the sea ice edge at the start and the end of the time-series, the dash-dotted lines show the 1991–2000, and the dashed lines the mean from ~~2011 to the end of the dataset~~ 2011-2020, (2011-2018 for ORAS) for the sea ice edge and the ~~marginal ice zone~~inner MIZ boundary.

Barents and Kara Sea border between Novaya Zemlya and Franz Josef Land (see Fig. 1)~~, and then~~ . Then it turns south towards northern Novaya Zemlya at around 70° E. Another small area with SIC < 80% exists west of Franz Josef Land. In the south,

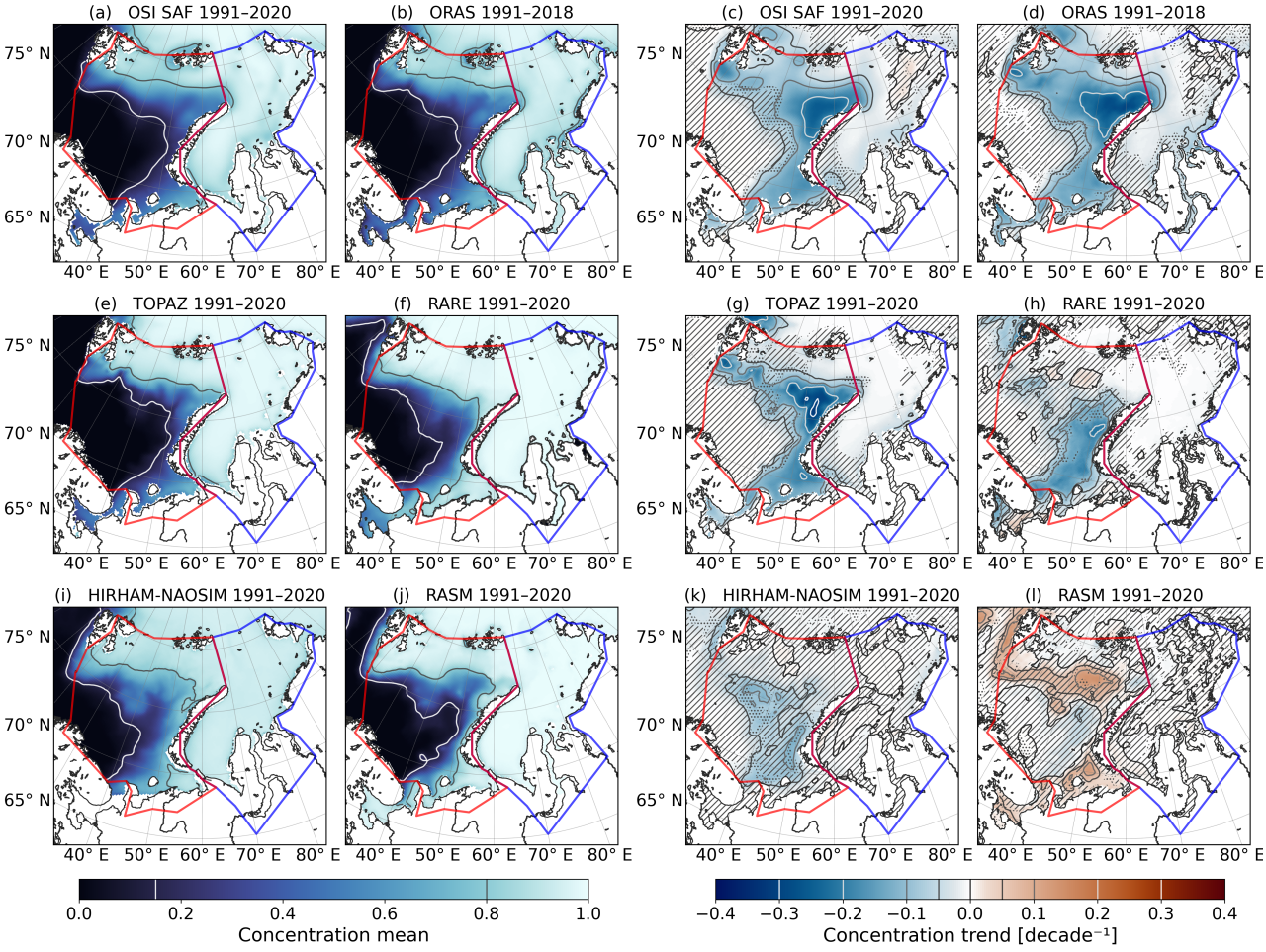


Figure 4. Average ~~sea ice concentration~~ SIC mean (left) and trend (right) of all the datasets in spring ~~(MAM)~~ over the whole period, presented as the fraction of sea ice cover per grid cell. The mean isolines (15% and 80% ice concentration) are used to distinguish the ~~marginal ice zone~~MIZ. In the trend subplots, the not statistically significant areas have been masked ("///~~masks areas~~/" where ~~the trend is not significant (~~$p > 0.1$ and "...." where $0.05 < p < 0.1$).

the ~~boundary between the inner pack and the marginal ice zone~~ inner MIZ boundary is located on average ~~on the Kara Sea side of the Kara Strait~~at the Kara Gate.

During the first decade, in 1991–2000, the ~~OSI-SAF~~ OSI SAF sea ice edge is somewhat ~~south-west of the one described above, and the boundary between the inner pack and the marginal ice zone goes from southern Svalbard to northern Novaya Zemlya , and close to the Kara Strait in the south~~ southwest and the inner MIZ boundary between Svalbard and Novaya Zemlya is south of the one described above (Fig. 3~~a~~a). In 2011–2020, the ~~OSI-SAF~~ largest retreat of the OSI SAF sea ice edge ~~location~~

~~south of Svalbard has not changed much compared to earlier decades. Meanwhile, further east, the sea ice edge has clearly retreated, and touches~~ occurs east of 45° E and is northward and eastward. The ice edge reaches Novaya Zemlya at around 75° N ~~, and then goes~~ and then extends from southern Novaya Zemlya to ~~southern Kola peninsula~~ the southern Kola Peninsula (Fig. 3~~n). In this last decade~~a). During this period, the SIC in the Barents Sea is generally below 80%, except in some small areas ~~south of Franz Josef Land and east of Svalbard, while in the Kara Sea, the pack ice~~ in the north-east and north-west; the inner MIZ boundary has retreated at the BSX and the Kara ~~Strait~~ Gate compared to earlier decades. Moreover, there is now a small ~~marginal ice zone~~ MIZ area north of the Yenisey Gulf.

The ~~OSI-SAF~~ OSI SAF SIC shows statistically significant negative trends for the period 1991–2020 across most of the Barents Sea ice-covered region in ~~MAM~~ spring (Fig. ~~??~~p4c). The negative trend is largest north-west of Novaya Zemlya, close to the region where the sea ice edge retreats at the end of the time series, where the negative trend is larger than 0.2 decade^{-1}. In the Kara Sea, the negative ~~OSI-SAF~~ OSI SAF SIC trends are smaller than in the Barents Sea, below $\pm$0.05 decade^{-1} and not always statistically significant. ~~The small trend is mostly negative, except in the north-east, where a small area has a small positive trend.~~ An exception to this is at the BSX, Kara ~~Strait~~ Gate, and north of the Yenisei Gulf, where the trends are slightly larger~~.~~

~~In MAM,~~ , with a negative trend above 0.1 decade^{-1} at the BSX.

ORAS is the ~~ORAS SIC is the closest to OSI-SAF~~dataset with a SIC closest to OSI SAF, with no large statistically significant differences in the study area (Fig. 3~~j~~b). TOPAZ overestimates SIC in the northern Barents Sea and in the Kara Sea. This ~~, together with the IIEE underestimation (Sect 4.2),~~ results in a narrower ~~marginal ice zone than OSI-SAF, and~~ MIZ than in OSI SAF and in a correspondingly extensive pack ice area in TOPAZ in ~~MAM~~ spring (Fig. 3~~m~~e). The statistically significant overestimation tends to be larger in the Northern Barents Sea than in the Kara Sea, but is mostly less than 0.1 decade^{-1} except ~~in the northern Barents Sea between Svalbard and Franz Josef land~~ at the northern edge of the Barents Sea and north of Novaya Zemlya at the BSX (Fig. 3~~m~~e). TOPAZ SIC trends form a similar spatial pattern to ~~OSI-SAF~~ OSI SAF (compare Fig. ~~??o and p~~4c and g), but with a stronger maximum negative trend west of Novaya Zemlya. In the northernmost part of the Barents Sea, however, the TOPAZ SIC trend is not significant and ~~smaller than OSI-SAF and ORAS .~~

~~As with TOPAZ, RARE also~~ is smaller than OSI SAF and ORAS and the area with a negative trend above 0.1 decade^{-1} does not extend into the Kara Sea. Overall, ORAS and TOPAZ reproduce the observed spring SIC and its spatial trend pattern well, although TOPAZ overestimates SIC and has a narrower MIZ indicating a sharper transition between compact ice and open water than in OSI SAF. Although both ORAS and TOPAZ assimilate satellite-derived sea-ice information, they employ different ocean and sea-ice models, vertical coordinate systems, and assimilation approaches. The opposite SIC biases therefore suggest that differences in model physics, sea-ice thermodynamics, ice-ocean coupling, and the implementation of assimilation are at least as important as the assimilated observations themselves. This may be related to the too-rapid seasonal retreat and early refreezing discussed in Sect. 4.2.

RARE overestimates SIC relative to ~~OSI-SAF in MAM but with a clearly larger margin~~ OSI SAF in spring over most of the sea ice covered area (Fig. 3~~i~~f). The ~~RARE~~ ice edge and the ~~boundary between the inner pack and the marginal ice zone~~ inner MIZ boundary are too extensive, ~~and the concentration difference to OSI-SAF is statistically significant in almost the whole~~

ice-covered regionbeing further south and east than the OSI SAF boundary. The RARE SIC is more than 0.1 over OSI-SAF above OSI SAF in the southern Kara Sea, in the Kara Sea at the BSX, and in most of the ice-covered Barents Sea, and the difference is over . The difference is above 0.4 south of Svalbard and west of Novaya Zemlya. The RARE ice edge in MAM retreated eastward from 70–75° N spring retreated eastward and northward between 1991–2000 and 2011–2020 (Fig. 3if),
460 while the boundary between the inner pack and the marginal ice zone did not change almost at allinner MIZ boundary changed wery little. The RARE SIC west of Novaya Zemlya, extending toward the Kola peninsula, and the negative trend there has a negative trend in the Barents Sea south of 75° N that is mostly below 0.2 decade^{-1}., but no meaningful significant trend north of that or in the Kara Sea. (Fig. 4h). Thus, RARE captures part of the observed spring SIC decline in the southern Barents Sea, but overestimates SIC and shows little change farther north and in the Kara Sea. This is consistent with its SIA and ice-edge
465 overestimations in Sect. 4.1 and 4.2 and may partly reflect that RARE constrains sea ice indirectly through a PIOMAS-based surface heat-flux relaxation rather than directly assimilating satellite SIC (Carton and Chepurin, 2023).

All coupled models (RASM, HIRHAM–NAOSIM, CNRM and NorESM)overestimate the MAM SIC in comparison to OSI-SAF in the Barents Sea. Furthermore, their marginal ice zones do not extend to the Kara Sea Similar to RARE, the RASM SIC is also significantly overestimated for most of the region, but with an area where the difference is not statistically significant
470 in the middle to northern part of the Barents Sea between 40° E and 60° E (Fig. 3a,b,e,fj). RASM has is the model with the largest concentration differencescompared to OSI-SAF, in particular , with even above 0.6 decade^{-1} differences south of Svalbard and west on the west coast of Novaya Zemlya, with statistically significant differences across most of the ice-covered area (Fig. 3e). Moreover, the RASM . The sea ice edge and the extent of the marginal ice zone indicate limited decadal changes , as seen in previous sections.

475 The spring NorESM SIC is significantly higher than that of OSI-SAF, with the highest differences above 0.4 west of Novaya Zemlya only retreats in a small area west of Novaya Zemlya, and the changes to the inner MIZ boundary are small. The SIC is statistically significant in small areas close to the inner MIZ boundary in the Barents Sea, where the trend is positive, with some small regions with a positive trend above 0.1 decade^{-1} (Fig. 3b). Its sea iceedge is too far west, especially south of 75°N in the Barents Sea, which explains the highest IIEE of all products (Table 3). The MAM SIC difference between OSI-SAF and
480 CNRM is smaller than for NorESM, and broadly less statistically significant 4j, l). The absence of widespread negative SIC trends despite the observed decline indicates that RASM fails to reproduce the long-term reduction of sea ice. This is consistent with the lack of a significant SIA trend discussed in Sect. 4.1.

The HIRHAM–NAOSIM spring SIC difference from OSI SAF (Fig. 3g) . Both products also have small areas with small statistically significant underestimations in the north-east corner of the domain and east of Novaya Zemlya.

485 The HIRHAM–NAOSIM sea ice edge in MAM i) differs from the other datasets in that it is underestimated in the north and western half of the Kara Sea and overestimated in the southeastern Kara Sea and in most of the ice-covered Barents Sea. The ice edge is located too far south-west compared to OSI-SAF southwest compared to OSI SAF (Fig. 3a). The HIRHAM–NAOSIM boundary between the inner pack and the marginal ice zone are rather close to that of OSI-SAF, except it turns south at 50, i), while the inner MIZ boundary does not extend east of 60° E instead of continuing east, towards the Kara Sea. Also, the
490 HIRHAM–NAOSIM pack ice boundary does not retreat between 1991–2000 and 2011–2020 N. The ice edge retreats, but not

as much as ~~that of OSI-SAF. The HIRHAM–NAOSIM SIC is typically larger than OSI-SAF, especially in the marginal ice zone, except in the northern and eastern~~ for OSI SAF, and the negative SIC trend is only statistically significant in small parts of the ~~study area, and in the central Kara Sea , where HIRHAM–NAOSIM slightly underestimates sea ice concentration.~~

~~Regarding the coupled model SIC trends in MAM, NorESM and RASM show small but significantly positive trends across some quite limited regions in the~~ MIZ in the Barents Sea ~~, surrounded by large areas of non-significant trends~~ and at the north-eastern coast in the Kara Sea (Fig. ~~??~~d,g). ~~Additionally, NorESM shows some limited areas in the Barents Sea with weak but significantly negative trends. Also~~3i and 4k). Of the coupled models, HIRHAM–NAOSIM ~~and CNRM have some limited areas with significantly negative SIC trends~~ reproduces the observed spring SIC decline best, but only in the Barents Sea ~~(Fig. ??c,h), the negative trend is stronger in CNRM than in HIRHAM–NAOSIM but the statistically significant area is much smaller. In HIRHAM–NAOSIM, the significant trend area is located in the central and southern Barents Sea, similar to the trend in RARE.~~ and with significant decreasing trends only present in parts of the MIZ. This indicating that the model underestimates the spatial extent of the ongoing sea ice retreat

4.3.2 Autumn sea ice concentration

In autumn, the ~~OSI-SAF~~ OSI SAF SIC is above 15% in the northern part of the study area (Fig. 3~~p~~c). The mean sea ice edge goes from north of Svalbard to Franz Josef Land, ~~and~~ from Franz Josef Land to the middle of the BSX, then continues ~~east~~eastward, and finally turns ~~south~~ southward at around 80° E (Fig. 3~~p~~). ~~OSI-SAF~~ c and 5a). OSI SAF shows significant negative trends in the whole ice-covered area (Fig. ~~??~~p5a). In 1991–2000, the sea ice edge goes from Svalbard to the northern tip of Novaya Zemlya ~~,~~and then from the north-east of Novaya Zemlya to the Yamal Peninsula. In comparison, the 2011–2020 mean sea ice edge goes north of the Barents Sea and along the northern boundary of the Kara Sea (Fig. 3~~p~~c).

Autumn SIC ~~between OSI-SAF and TOPAZ , and OSI-SAF~~ of TOPAZ and ORAS do not differ significantly from OSI SAF (Fig. 3~~l, o~~d, g), and the ORAS and TOPAZ decadal trends are almost identical to those of ~~OSI-SAF~~ OSI SAF (Fig. ~~??l, o~~5d, g). Compared to ~~OSI-SAF~~OSI SAF, RARE overestimates the autumn sea ice edge and concentration, with the largest differences in the northern part of the study region (Fig. 3~~k~~). h). For autumn, RARE is the dataset with the largest concentration overestimation. The RARE mean sea ice edge goes from southern Svalbard to the northern tip of Novaya Zemlya, then continuing to the Yamal Peninsula (Fig. ~~s 3k or ??~~3h or 5i). ~~The RARE region with a significant negative SIC trend resembles the corresponding OSI-SAF region. Still, it extends over a larger area in the south compared to OSI-SAF (See Fig. ??k and Fig. ??p), which is not surprising as the sea ice has a larger extent, and there is thus more ice to melt~~As for OSI SAF, the RARE negative trend in SIC extends over the sea ice covered area (Fig. 5h), but he maximum trend is stronger and located farther south than in OSI SAF. This can partly be explained by RARE retaining sea ice in areas where OSI SAF, ORAS, and TOPAZ become ice-free, leaving more ice available to decline.

~~RASM and NorESM overestimate the average 1991–2020 SIC in autumn across the whole study region, except at some coastal locations~~ The HIRHAM–NAOSIM sea ice edge in the Barents Sea goes from southern Svalbard to northern Novaya Zemlya, similar to RARE, but in the Kara Sea the edge is even further south (Fig. 3~~d,g). The NorESM mean sea ice edge goes further south than RASM, while the RASM mean sea ice edge goes slightly further south than that of OSI-SAF. In both models,~~

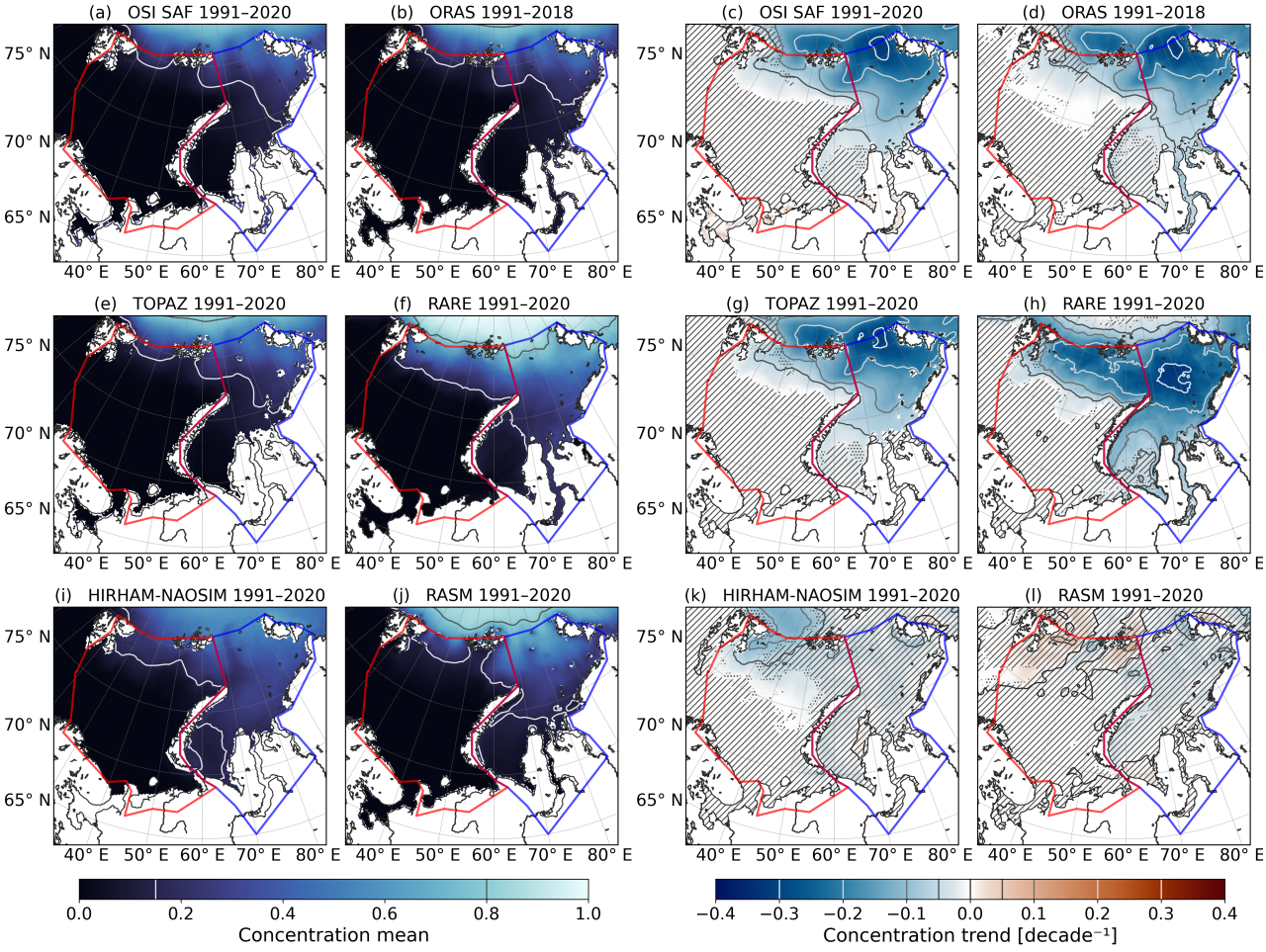


Figure 5. Average ~~sea ice concentration~~ SIC mean (left) and trend (right) of all the datasets in autumn ~~(ASO)~~ over the whole period, presented as the fraction of sea ice cover per grid cell. The mean isolines (15% and 80% ice concentration) are used to distinguish the ~~marginal ice zone~~MIZ. In the trend subplots, the not statistically significant areas have been masked ("///~~masks areas~~/" where ~~the trend is not significant~~ (~~$p > 0.1$~~ and "..." where $0.05 < p < 0.1$).

525 ~~the ice edge only retreats a bit north in the Kara Sea from 1991–2000 to 2011–2020, while no clear change can be seen in the Barents Sea~~k). Most of the ice covered area has an overestimated SIC, except in the northernmost part where the differences compared to OSI SAF are not statistically significant. The overestimations are, however, clearly less than for RARE. The negative SIC trend in HIRHAM–NAOSIM is statistically significant and above 0.05 decade^{-1} only in a small region east of Svalbard (Fig. 5k).

Table 4. Spring and autumn mean ML temperature [°C], decadal trends [°C/decade] , and p-values for Barents and Kara Seas. Calculated for the maximum period of each dataset. Mean values are bold when the seasonal means differ significantly from ORAS5 (t-test, p < 0.05, 1991-2018). Trends and p-values are bold when the given p < 0.05 (least-squares regression).

	Barents Sea			
	Spring			
Model	Mean	Trend	p	
HIRHAM–NAOSIM -0.4 **0.23** **0.001** 3.7 **0.56** **0.003** -1.7 0.00 0.323 0.1 0.23 ORAS	0.207 RASM *0.1*	-0.3 **0.36**	-0.03 **0.000**	0.721
TOPAZ	0.15 0.1	0.458 CNRM **0.37**	-0.1 **0.000**	0.12
RARE	0.7 **0.6**	**0.24**	**0.044** **0.042**	
TOPAZ HIRHAM–NAOSIM	0.1 **-0.5**	**0.37** **0.22**	**0.000** **0.001**	4.4
ORAS RASM	0.1 **-0.3**	**0.37** -0.03	**0.000** 0.722	4.2

MAM and ASO mean ML temperature °C, decadal trends °C/ 10 yr, and p-values for Barents and Kara Seas. Calculated for the maximum period of each dataset. Trends and p-values are bold when $p < 0.1$.

The autumn HIRHAM–NAOSIM SIC is slightly overestimated across the entire sea ice-covered area. In the Barents Sea, RASM mean autumn sea ice edge is further north than the HIRHAM–NAOSIM mean and RARE sea ice edge extends too far south from Svalbard to northern Novaya Zemlya, and there is only a small ice-free area in the southern Kara Sea , and the SIC is overestimated less than in RARE but more than in HIRHAM–NAOSIM, with overestimation above 0.2 decade^{-1} in the northern parts (Fig. 3c).

In CNRM, the mean SIC does not differ significantly from OSI-SAF in the Barents Sea in ASO. In contrast, in the Kara Sea, the SIC is clearly underestimated l). However, the SIC is not statistically significantly different from OSI SAF in the eastern Kara Sea at the coast. RASM does not have any meaningful significant trend in the SIC in autumn (Fig. 3h). The CNRM mean sea ice edge goes from Svalbard to the BSX, and continues eastward towards the south of Severnaya Zemlya.

The coupled model autumn sea ice area and concentration trends are not significant, with few exceptions (see Table 2 and Fig. **??**c,d,g,h). In 5l). Unlike in spring, where HIRHAM–NAOSIM reproduces part of the observed decline in the Barents Sea MIZ, neither coupled model reproduces the widespread autumn SIC decline. HIRHAM–NAOSIM performs somewhat better by showing a small area of significant decline east of Svalbard, HIRHAM–NAOSIM has a statistically significant negative SIC trend(Fig. **??**c), and the NorESM negative SIC trend is statistically significant in an area east of Novaya Zemlya and in a small area at the east coast of Svalbard (Fig. **??**d)whereas RASM shows no meaningful autumn trend. Together with their overestimated SIC, this indicates that the seasonal sea ice retreat is too weak in both models, particularly in the Kara Sea.

4.4 Ocean temperature of the mixed layer

ML temperature anomaly in each model for the Barents (left) and Kara Sea (right). Spring (MAM) and autumn (ASO) seasons are marked with dark grey boxes. The anomaly is calculated by taking the difference from the monthly multiyear mean for each month. Note the different colour scales in the figures.

The mean temperature and trends of the ML are shown in Table 4 for the seasons MAM and ASOspring and autumn. In the Barents Sea, the spring (MAM) mean temperature varies from −0.6 0.5 °C (NorESM) to 0.7 HIRHAM–NAOSIM) to 0.6 °C (RARE) between the models. RARE is considerably warmer than ORAS and TOPAZ, while HIRHAM–NAOSIM and RASM are colder. The mean ML temperature in the regions stays mostly below 5 °C, except in NorESM and RARE, where some warmer waters can be seen (Fig. ??c ,iA3d). For the other models, the difference in the regional mean seems to depend more on how widespread the warmer waters are than the difference in temperature of the warmer waters (compare Fig. A3a, c to g, h). In the Kara Sea, the MAM spring variation is smaller , and the ML is close to freezing (Table 4). In autumn(ASO), the ML temperature in , the Barents Sea is between 2.1 (NorESM) and 5.5 °C (CNRM), and between −0.4 (NorESM) and 2.3 °C (CNRM)in the Kara Seamean ML temperature does not significantly differ between the reanalyses in either the Barents or the Kara Sea, while both models are significantly colder (Table 4). The mean ML temperature stays mostly below 10 °C in the Barents Sea, except in the south-west in RASM , CNRM and RARE (Fig. ??e,f,i) , A4) and mostly below 5°C in the Kara Sea, except in the south-west in RASM, CNRM, RARERARE, and TOPAZ (Fig. ??e, f, i,). A4c, d, h). The reason why the regional mean for HIRHAM–NAOSIM and RASM is similar, even if RASM has these warmer coastal waters, is that many of these are outside of the compared region (Fig. 1). The colder mean mixed-layer temperatures in the coupled models relative to the reanalyses, particularly in autumn, are consistent with their tendency to maintain too much sea ice (Sections 4.1 and 4.3). A more extensive ice cover reduces the direct exposure of the ocean to atmospheric forcing and can limit seasonal ocean warming, suggesting that at least part of the temperature bias may be linked to the sea-ice bias rather than to ocean processes alone.

The ML temperature trend is positive and statistically significant in the Barents Sea in both MAM and ASO spring and autumn for the reanalyses and for HIRHAM–NAOSIM, being stronger in the autumn (Table 4). In MAM spring, the warming trend is highest IN HIRHAM-NAOSIM, ORASand TOPAZ in the ice free areas , HIRHAM–NAOSIM, ORAS, and TOPAZ is highest in the ice-free areas or close to the ice edge (Fig. ??c, l, o)while in RARE the highest A3b, e, i), while in the autumn ORAS and TOPAZ have a significant warming trend over the whole Barents Sea. The warming trend is weaker and the region of significant warming is smaller in HIRHAM–NAOSIM and does not extend as far north and east. In RARE, the highest spring warming coincides with the areas of strong deeper areas (Fig. 1), where Atlantic water enters the Barents Sea (Skagseth et al., 2008). This partly coincides with the area of the strongest decreasing SIC trend in the marginal ice zone MIZ east of Novaya Zemlya (Fig. ??k and ??4j and A3k). In RASM, NorESM and CNRM, the Barents Sea trend is not The trend in spring at these locations is stronger than in ORAS, and the two areas are separated by a nonsignificant cooling trend. In autumn, RARE has two areas with a cooling trend, of which the eastern one at the coast of Novaya Zemlya is statistically significant. These are located close to the same areas where there was the strong warming trend in the spring. The ML temperature in this location is also on average warmer than in the other reanalyses (Fig. A3d,f). Of these, the eastern one is located west of Novaya Zemlya, in the same location as the strong warming in spring and is statistically significant. In autumn, RARE also shows a

Table 5. Pearson's correlation coefficients between SIA and ML temperature of Barents and Kara Sea, denoted as r, and between the SIAa (Fig. 2) and ML temperature anomaly (MLTa) (Fig. A5), denoted as r_{diff}. For r_{diff}, bold numbers are statistically significantly different from ORAS r_{diff} using Fisher's z-test ($p < 0.05$), while italic shows where Barents and Kara Sea differ from each other.

Model																							
HIRHAM–NAOSIM	-0.903	-0.681	-0.648	**-0.832**	-0.912	-0.766	**-0.837**	**-0.855**	RASM	-0.917	-0.631	-0.553	**-0.819**	-0.936	-0.676	-0.642	**-0.838**	NorESM					
CNRM	-0.838	-0.702	-0.676	**-0.820**	TOPAZ																		
RARE																							
TOPAZ	-0.885	-0.690	**-0.568**	**-0.877**	HIRHAM–NAOSIM																		
ORAS	-0.858	-0.672	**-0.546**	**-0.836**	RASM																		

Pearson's correlation coefficients between sea ice area and ML temperature of Barents and Kara Sea, denoted as r, and between the sea ice area anomaly (Fig. 2) and ML temperature anomaly (Fig. **??**), denoted as r_{diff}. r_{diff} is calculated both for the monthly data of the whole year and for the summer (APR–SEP) and winter (OCT-MAR) half years separately. Bold numbers are statistically significantly different from full year r_{diff} using Fisher's z-test ($p < 0.05$)

statistically significant deepening of the ML west of Novaya Zemlya (not shown). The contrasting spring warming and autumn cooling in this region may indicate differences in how vertical mixing and the seasonal storage and release of ocean heat are represented in RARE compared to the other reanalyses. The present analysis cannot isolate the respective contributions of the ocean model, assimilation scheme, and atmospheric forcing to this behaviour.

In the Kara Sea, the trend is small for all ~~models~~ datasets in spring, when the region is ice-covered (Table 4). ~~In Autumn~~The limited spring warming is therefore consistent with the weak spring SIA and SIC trends, because the compact ice cover (SIC > 80%) restricts atmosphere–ocean heat exchange and keeps the ML temperature close to freezing. In autumn, the trend is statistically significant in the reanalyses and close to 1 °C/decade (Table 4). The ~~ASO~~ autumn ML temperature trend is ~~clearly~~ above 1 °C/decade in large areas of the Kara Sea in all reanalyses and is stronger in the Kara Sea than in the Barents Sea~~in all the reanalyses. The Kara Sea average ML temperature is smallest in ORAS (Table 4), where the trend is below 1 °C/decade north of 75° N, but strongest of all reanalyses at the cost, being over 2 °C/decade in some areas~~. HIRHAM–NAOSIM shows no statistically significant trend in the Kara Sea where the sea ice is overestimated (Section 4.3.2 and Fig. 5i). RASM has no widespread statistically significant trends in either region in spring or autumn, only small areas of warming in the Kara Sea and cooling in the Barents Sea (Fig. **??**~~l). In RARE and TOPAZ,~~ A4i, j). Thus, HIRHAM–NAOSIM reproduces the Barents Sea warming better than RASM, but neither coupled model captures the pronounced autumn warming in the Kara Sea. For HIRHAM–NAOSIM, the ~~above 1 °C/decade warming is more widespread and covers most of the area, except the north, being somewhat broader and stronger in RARE (Fig. **??**k, o)~~weak Kara Sea warming coincides with its extensive sea-ice cover, which may limit seasonal atmosphere–ocean heat exchange. RASM, by contrast, shows weak ML temperature trends despite a less extensive sea-ice bias, suggesting that its lack of warming cannot be attributed to excessive ice cover alone.

Figure ?? shows the ML temperature anomaly (difference between monthly mean and monthly climatology) and Table 5 shows the Pearson's correlation coefficient (PCC) calculated between sea ice area PCC calculated between SIA and ML temperature and between the anomalies (shown in Fig. 2 and ??A5). The temperature average ML temperature anomalies (MLTa) (Fig. A5) follows similar patterns to that of the sea ice area SIA (Fig 2), discussed in Section Sect. 4.1.1 and 4.1.2. The reanalyses show a decline with 2005–2006 showing a switch from mainly below average and above average temperatures in both regions, though in the Kara Sea, the year-to-year variations are mainly only seen in summer and autumn (for all models). RASM has only year-to-year variation, in both regions, with no trend. Some oscillation is present in HIRHAM–NAOSIM in the Barents Sea and in CNRM in both regions, as seen in the sea ice area (section 4.1.1). The correlation coefficient The PCC (r) (Table 5) between the sea ice and temperature is negative for all datasets , and and ranges between -0.83 0.85 and -0.94, and the correlation coefficient . The PCC between the anomalies (r_{diff}), where the effect of the annual variation is removed, is also negative and ranges between -0.63 and $-0.86$for all datasets . The correlation coefficient . These high correlations are in line with previous studies (Kumar et al., 2021; Li et al., 2022). The consistently strong negative correlations indicate that all datasets capture the expected tight coupling between mixed-layer temperature and sea ice, whereby warmer upper-ocean conditions favour sea-ice loss and reduced ice cover enhances ocean heat uptake. RARE seems to have a higher correlation between the anomalies is higher in than the other datasets and is significantly higher than ORAS in both the Barents Sea and the Kara Seathan in . In the Barents Seafor all datasets except CNRM and TOPAZ. When calculating , only the RARE anomaly PCC (r_{diff}) differs significantly from ORAS, while in the Kara Sea the difference compared to ORAS is statistically significant for RARE and HIRHAM–NAOSIM. The correlation coefficients between the anomalies are slightly lower in the Kara Sea than in the Barents Sea for ORAS and TOPAZ and slightly higher for the other datasets. However, the differences are statistically significant (r_{diff}) for the light summer (Apr–Sep) months and dark winter months (Oct–Mar) separately, the correlation coefficient for winter is always higher than for the full year for all datasetsfor both regions. This difference is statistically significant using Fisher's z-test ($p < 0.05$) for all datasets except for RARE in the Barents Sea and CNRM in the Kara Sea, and only CNRM in the Kara Sea has a summer value slightly higher than the winter value. In the Barents Sea, the correlation coefficient for summer is always smaller than that for the whole year, but the difference is only statistically significant for TOPAZ and ORAS. In the Kara Sea, only RASM shows a lower correlation coefficient for summer than for the whole year, z-test, p < 0.05) only for RARE and most differences for the whole year are statistically not significant HIRHAM–NAOSIM. The stronger coupling in RARE may indicate that sea-ice variability is more tightly linked to local ML temperature variability than in the other products. One possible explanation is that RARE constrains sea ice indirectly through surface heat flux relaxation based on PIOMAS (Sec. 2.2)Carton and Chepurin (2023). This interpretation is consistent with the product-specific spatial temperature behaviour described above, although the present results do not identify the mechanism responsible for the stronger correlation.

4.5 Climate index correlation

Coupled models and reanalyses were evaluated against OSI-SAF satellite measurements regarding their performance in representing the sea ice

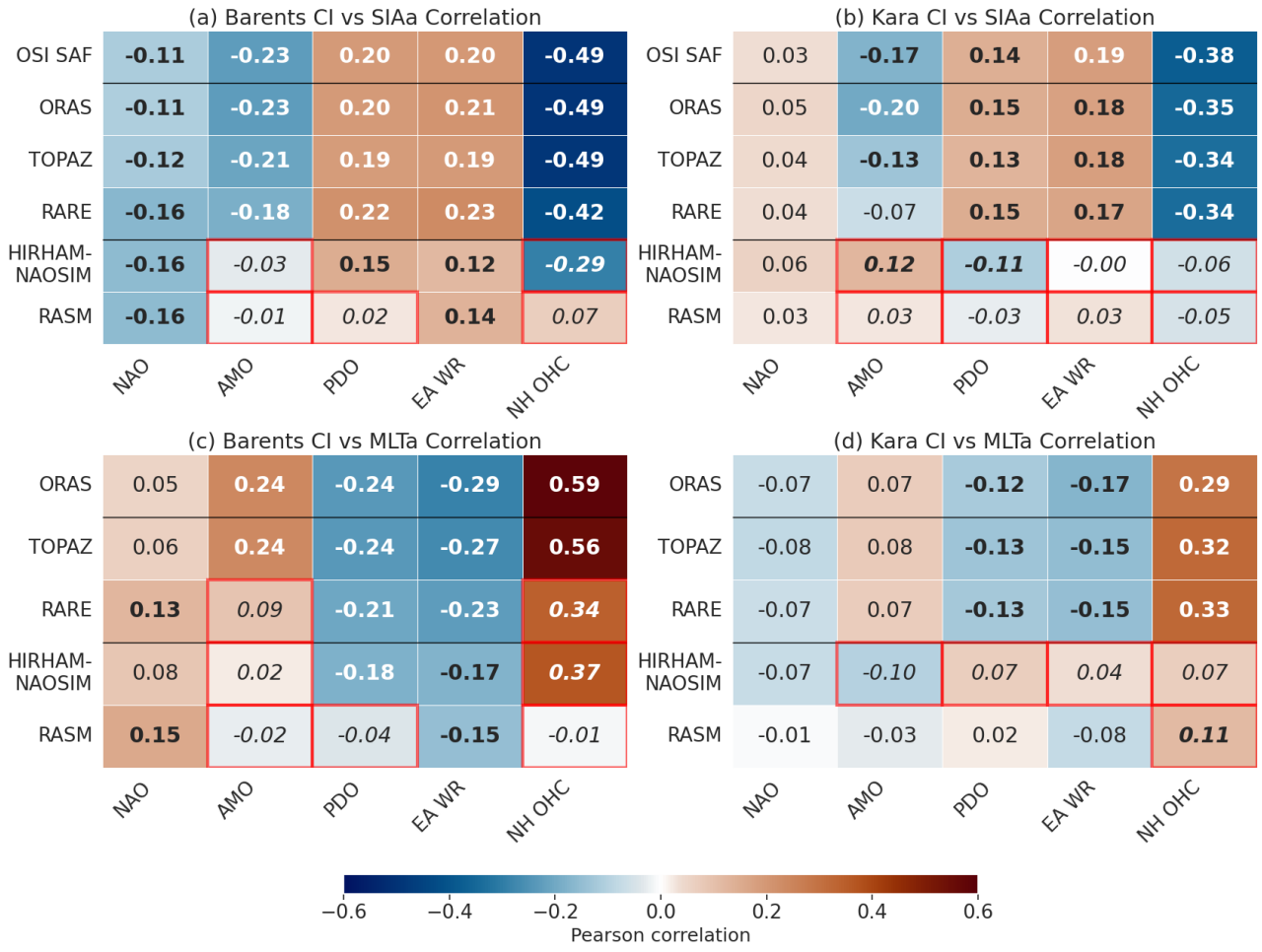


Figure 6. Correlation between climate indices and SIAa (a,b) and average MLTa (c,d) for the Barents and Kara Seas. Correlation values in bold indicate that the correlation is statistically significant ($p < 0.05$), and the red squares indicate where the correlations are statistically different (Fisher's z-test, $p < 0.05$) from OSI SAF for SIA and ORAS for MLT. NAO: North Atlantic Oscillation; AMO: Atlantic Multidecadal Oscillation; PDO: Pacific Decadal Oscillation; EA/WR: East Atlantic/ Western Russia teleconnection; NH OHC: Northern Hemisphere Ocean Heat Content (0–700m).

PCC of five climate indices (Sec. 3.3) to SIAa and average MLTa were calculated for both regions to explore the teleconnections in the data (Fig. 6). Maps for the correlation of the climate indices per gridcell to the SIC anomaly (SICa) and MLTa, from which the SIAa and average MLTa have been calculated, can be found in the appendix (Fig. A6, A7, A8, A9 and A10). Overall, none of the climate-index correlations differ significantly among OSI SAF, ORAS, and TOPAZ for SIAa, or between ORAS and TOPAZ for MLTa. This agreement is expected given the small differences among these datasets found in the preceding sections.

For NAO, there are no significant differences between the datasets (Fig. 6). The positive NAO phase is characterised by a stronger North Atlantic pressure gradient and enhanced westerly circulation, whereas the negative phase is associated with a weaker pressure gradient and weaker or more meridional circulation (NOAA CPC, b). A positive NAO leads to warm and humid westerly winds and enhanced heat transport into the Barents Sea region, leading to decreasing sea ice and a statistically significant correlation in SIAa in the Barents ~~and Kara Seas. The satellite measurements show asignificant decreasing trend in the sea ice area in the Barents and Kara Seas in both MAM and ASO. The trend is larger in the Barents Sea in MAM than in ASO, and larger in the~~ Sea (Fig. 6a). The effect is, however, weaker further away from the Atlantic in the Kara ~~Sea in ASO than in MAM. This is because~~ Sea, and there the SIAa correlation is not statistically significant (Fig. 6b). The correlation map (Fig. A6a, b, e, f, i, j) shows that the negative correlation with sea ice is significant mostly in the northern Barents Sea, and a non-significant negative correlation extends into the northeastern Kara Sea. The rest of the Kara Sea is ~~still mostly fully ice-covered in MAM, while in ASO the two regions, but especially~~ positively correlated but only statistically significant in small areas.

A positive AMO is characterised by warmer-than-average North Atlantic sea-surface temperatures and has been linked to subsequent temperature variability in the Barents Sea ~~, are already mostly ice-free at the end of the study period. The decreasing trends seen here agree with previous studies (e.g., Duan et al., 2019b; Kumar et al., 2021; Liu et al., 2022; Shabanov et al., 2024)~~ .

~~The reanalyses examined show significant decreasing trends in sea ice area and concentration as well. From the coupled atmosphere-ocean models, however, only HIRHAM–NAOSIM shows a statistically significantdecreasing trend in the Barents Sea ice area, but on a lower level of significance than the satellite data and the reanalyses. Also, the regions with significant trends in sea ice concentration are smaller. In~~ (Seip and Wang, 2023), increasing the heat transport to the Barents Sea. Consistent with this relationship, a positive AMO is associated with a negative SIAa and positive MLTa correlation for OSI SAF, ORAS, and TOPAZ, which is about double the correlation with NAO (Fig. 6a, c). In the Kara Sea, we see a weaker, though still significant, correlation in the SIAa, but not in the MLTa as the effect of the Atlantic is again weaker in the Kara Sea ~~, none of the coupled models show a statistically significant change.~~

~~Mean sea ice area is overestimated in most of the coupled models and in the RARE reanalysis due to both an overestimation of the concentration and a shifted location of the ice edge. CNRM is an exception,as it underestimates the sea ice, especially in the Kara Seain ASO. TOPAZ and ORAS mean sea ice area and sea ice edge are closest to the satellite measurements, which is not surprising, as they assimilate satellite data. However, TOPAZ tends to overestimate the sea ice concentration in spring, while ORAS slightly underestimates~~ (Fig. 6b, d). This difference between SIAa and average MLTa may arise because sea-ice variability integrates the cumulative effects of atmospheric and oceanic forcing, whereas the ML is close to freezing and insulated from the atmosphere during the winter and spring months by the ice cover, dampening the correlation. From the correlation map (Fig. A7a,b,e) one can see that the negative correlation to SICa is highest for OSI SAF, ORAS, and TOPAZ in the northern Barents Sea. A weaker negative correlation appears in the Kara Sea. A notable feature is an area of positive statistically significant correlation to SICa in the centre of the Barents Sea in OSI SAF, ORAS, and TOPAZ. This could be because currents and winds collects ice here from the surrounding. For MLTa (Fig. A7d,g), ORAS and TOPAZ have a positive

correlation for most of the area - strongest in the western Barents Sea and weakest in the Kara Seaice concentration. Even though RARE does not assimilate satellite data,it is able to produce a negative trend, but its mean sea ice area is higher , where only small areas are statistically significant. In RARE (Fig. A7f,h), the sea ice correlation is similar but weaker than in the other reanalyses. This indicates that data assimilation has a strong impact on the skill to represent the sea ice correctly, as shown in previous comparisons of Arctic reanalysis products and prediction systems (e.g., Zhang et al., 2021; Xiong et al., 2025), and it is consistent with documented inter-reanalysis differences arising from different assimilation strategies (see Uotila et al., 2019) . Thus,while TOPAZ and ORAS provide a useful reference for the coupled ocean–sea ice state and its variability,their high skill in sea ice primarily reflects the impact of assimilation,and does not necessarily indicate superior physical process representation compared to the free-running coupled models. Even though reanalyses are important tools and sources of data for research, they can not be used for everything (Storto et al., 2019). For MLTa, however, there are some areas with strong negative correlation. These regions coincide with the anomalous ML behaviour identified for RARE in Sect. 4.4, particularly west of Novaya Zemlya (Fig. A3d,f and A4d,f). The negative correlations therefore appear to be associated with the anomalous ML behaviour in RARE, although the underlying cause of this product-specific relationship cannot be determined from the present analysis. For HIRHAM–NAOSIM (Fig. A7i,k), the SICa correlation looks similar to the reanalyses for the western Barents Sea, but is positive for the eastern Barents Sea and for the Kara Sea. The same pattern, although with opposite correlation, can be seen for the MLTa as well. RASM correlations are generally weaker (Fig. A7j,l), with Barents Sea SICa divided into negative correlation in the north and positive in the south, and with Kara Sea showing no large areas of statistically significant correlation.

For coupled atmosphere-ocean models,it can not be expected that the results follow observations exactly, as they represent a possible pathway of the climate system rather than a direct reproduction of observed conditions. Although coupled regional models can be driven by quasi-observed data at the lateral boundaries, the simulation inside the model domain evolves freely. Consequently, small differences in the initial state can lead to different simulation results. Their goal is to simulate realistic climate dynamics and to reproduce statistical properties of the climate. From the coupled models, RASM has the smallest IIEE, with results close to those of RARE. However, as mentioned before, there is no trend in the sea ice area, while the ice edge overestimation-underestimation shows the largest trend, because the ice edge does not move during the study period. On the other hand, For PDO, a warm (cold) central northern Pacific is related to a negative (positive) PDO phase. When the central northern Pacific is warm (negative PDO), warm-air advection to the Barents and Kara Seas increases, causing a negative SIAa and positive MLTa. This results in a positive (negative) correlation for the SIAa (MLTa). The PDO correlation is opposite to the AMO correlations and of approximately the same size (Fig. 6). The correlation patterns are similar to those for AMO too (Fig. A8), but with more spread for the SICa correlations in the reanalyses. RARE PDO-MLTa correlation (Fig. A8h) shows an area of positive correlation, but not as large and strong as for the AMO. HIRHAM–NAOSIM shows a large IIEE with the largest overestimation-underestimation in the Barents Sea. Trend-wise it is the best performing model PDO correlations (Fig. A8i,k) show a similar east-west divide as for AMO, but the divide is more eastward, resulting in only PDO showing a significantly different correlation (Fig. 6). RASM is significantly different from OSI-SAF and ORAS for both regions (Fig. A8j,l). The correlation for both SICa and MLTa to PDO in RASM is generally non-significant and spatially inconsistent.

The positive (negative) EA-WR produces negative (positive) temperature anomalies over eastern Europe, including the Ural Mountains. As a result, a positive (negative) phase pushes more (less) continental air into the Barents and Kara seas, leading to a generally positive EA-WR - SIA correlation and negative EA-WR - MLTa correlation (Fig. 6), which is seen in all reanalyses and in the Barents Sea ~~, while it does not perform better than the other modelsin~~ also for the coupled models. In the Kara Sea~~. This might have to do with the differences between the basins, the dominant physical processes, and how well they are represented~~ , however, neither coupled regional model has a statistically significant correlation either for SIAa or MLTa. From the map (Fig. A9) we see that the correlations are weaker in the models ~~.~~ than in the reanalyses. HIRHAM–NAOSIM has even an opposite, non-significant, correlation in the northeastern Barents Sea and in most of the Kara Seas, while RASM has a non-significant opposite correlation in the northern parts of the study area.

~~Several studies have looked into the sea-ice variability in the Barents and/or Kara Seasand the processes behind it (e.g., Duan et al., 2019a . Warm Atlantic waters flow into the Barents Sea through the BSO, where the water cools and mixes, releasing heat into the atmosphere. Part of this modified water continues into the Kara Sea through the BSX and Kara Gate. Shabanov et al. (2024) stated that the warm water inflow is~~ The NH OHC is the climate index that is most affected by the global warming trend and which shows the strongest correlation to both SIAa and MLTa for OSI SAF, ORAS, TOPAZ, and RARE (Fig. 6). It is a good indicator for overall climate warming, as 90% of the ~~most important process for the sea ice decline in the Kara Sea . While this oceanic inflow can enhance sea-ice melt and contribute to Atlantification, several studies (e.g., Liu et al., 2022) showed that atmospheric forcing can dominate interannual to decadal sea-ice variability, with air temperature, winds, and radiative fluxes exerting strong control. Recent work also indicates that the relative importance of these drivers varies spatially (Dörr et al., 2024). Overall, the atmosphere, ocean, and sea ice form a tightly coupled system, and internal feedbacks make it difficult to attribute observed variability to any single driver (Duan et al., 2019a; Chatterjee et al., 2024).~~

~~The temperature of the ML tells a similar story as the sea ice. A statistically significant increasing trend of the mean ML temperature can be seen~~ global warming heat is absorbed into the ocean (IPCC, 2021). For OSI SAF, ORAS, and TOPAZ, the correlation with SICa is negative and strongest in the northern Barents Sea, whereas the MLTa correlation is positive and strongest in the western Barents Sea (Fig. A10a, b, d, e, g). For RARE, the SICa correlation is slightly weaker than for the other reanalyses, but always negative, while in MLTa the biases in the Barents Sea, that was mentioned before in Sect. 4.4 and for AMO and PDO, can again be seen (Fig. A10f, h). The correlations in HIRHAM–NAOSIM in the Barents Sea ~~in both MAM and ASO for the reanalyses and HIRHAM–NAOSIM, while in~~ are in the same direction as in the reanalyses for both SICa and MLTa, but they are weaker and significantly different from those for OSI SAF, ORAS, and TOPAZ (Fig. A10i, k). In the Kara Sea, ~~a clear statistically significant trend can only be seen in the reanalyses in ASO. The correlation between the sea ice area anomaly and the ML temperature anomaly is high in all datasets, which is in line with previous studies (Kumar et al., 2021; Li et al., 2022). There is no clear difference between the different classes of datasets (models vs. reanalyses) , even though there is some variation between the datasets. RARE seems to have a somewhat higher correlation between the anomalies than TOPAZ and ORAS, but more research would be needed to see if this difference is significant. The mean ML temperatures between the reanalyses are also quite similar, although the magnitude and spatial pattern of the autumn trend differ between them. Taken together, the consistently high correlations indicate that oceanic conditions play an~~

~~important role in shaping the sea-ice variability , even if atmospheric processes are not explicitly examined here. This makes it particularly important to investigate the ocean component~~ however, the HIRHAM–NAOSIM does not show a statistically significant correlation. For RASM the only region with a wider statistically significant correlation is the southeastern Kara Sea, where there is a weak correlation to MLTa (Fig. A10l).

For the coupled regional models, HIRHAM–NAOSIM does better in the Barents Sea than RASM for both SIAa and average MLTa correlations. In HIRHAM–NAOSIM, only the AMO and NH OHC correlations differ significantly from OSI SAF (or ORAS in the case of MLTa), while for RASM, AMO, PDO, and NH OHC correlations differ significantly. In the Kara Sea, HIRHAM–NAOSIM SICa (MLTa) correlations with the climate indices differ significantly from OSI SAF (ORAS) for all climate indices except NAO, while RASM SICa climate-index correlations differ significantly for the same indices but do not show the opposite AMO and PDO correlations found in HIRHAM–NAOSIM. For MLTa, only the NH OHC correlation differs significantly in RASM. The differences from observations and reanalyses may partly arise because the coupled regional models simulate interactive atmosphere–ocean–sea-ice processes, whereas the ocean reanalyses are driven by prescribed atmospheric forcing and constrained by assimilated observations. However, the differences between HIRHAM–NAOSIM and RASM also indicate model spread arising from their different components and configurations. The significantly different AMO and NH OHC correlations in both coupled models suggest that the influence of North Atlantic variability is not fully represented. This is consistent with the findings of Yamagami et al. (2022), who showed that Gulf Stream warming contributes to Barents–Kara sea-ice decline and that an underestimation of this warming can lead to an underestimation of sea-ice loss.

A plausible interpretation is that biases in the representation of Atlantic-origin warming affect the ability of the coupled models to ~~understand why they fail to~~ reproduce the observed decline in sea ice and increase in mixed-layer temperature. Enhanced warming in the ~~Kara Sea, and for most of them also in the Barents Sea.~~ North Atlantic can increase heat transport through the Nordic Seas and Barents Sea, contributing to upper-ocean warming and sea-ice retreat in the Barents–Kara region. The weaker mixed-layer warming trends and weaker sea-ice decline seen in the coupled models are therefore consistent with an under-representation of this mechanism. Although the present analysis cannot determine whether the deficiency originates from the simulation of remote North Atlantic processes, the propagation of heat anomalies into the Arctic, or local ocean–atmosphere–sea-ice feedbacks, the results suggest that Atlantic warming signals are not transmitted to the Barents and Kara Seas.

~~It was shown by Yamagami et al. (2022) that the CMIP6 GCMs have difficulties in reproducing~~

5 Conclusions

Coupled models and reanalyses were evaluated against OSI SAF satellite measurements regarding their performance in representing the sea ice in ~~regions such as the Barents-Kara sea, and that the CMIP6 multi-model mean underestimates the Barents-Kara SIC negative trend in winter. They show that underestimation of the Gulf Stream surface warming is a likely reason for the too small Barents-Kara negative trend . The present results are consistent with this finding, as CNRM and NorESM underestimate the magnitude of the observed sea ice decline. From the coupled regional models, HIRHAM–NAOSIM does not cover the Gulf~~

~~Stream region between 30–80° W and 30–50° N as NAOSIM 's southern border is at around 50° N, while RASM only partly covers the area in question. This means that the possible influence of remote North Atlantic warming , which has been shown to contribute to Barents~~the Barents and Kara Seas. The satellite measurements show a significant decreasing trend in the SIA in both seas in both spring and autumn. All three reanalyses reproduce this decline in both seasons, whereas neither coupled regional model reproduces a statistically significant SIA decline in the Kara Sea, and only HIRHAM–~~Kara sea ice decline (Yamagami et al., 2022), may only partly be represented in these regional model simulations, and would instead be included within their lateral forcing.~~ NAOSIM shows a near-significant decline in the Barents Sea.

~~The two CMIP6 models chosen for this study represent high and low Barents-Kara sea warmingaccording to the storyline analysis by Levine et al. (2024). Thus, they are models far from the multi-model mean and not necessarily the best CMIP6 model in terms of closeness to observations. NorESM is the model for low BK warming, and CNRM is the one for high BK warming. It is, however, surprising that neither of the two models shows a significant trend during the historical period, though they clearly differ in the future scenarios. The largest differences between the models during the period of 1991–2014 are that CNRM clearly underestimates the autumn sea ice extent and concentration~~ORAS and TOPAZ show the closest agreement with OSI SAF for sea-ice area, sea-ice concentration, and ice-edge location. RARE, HIRHAM–NAOSIM, and RASM generally overestimate sea ice extent and concentration, resulting in larger ice-edge errors. Among the coupled models, HIRHAM–NAOSIM reproduces aspects of the observed decline better, whereas RASM shows no long-term change in sea ice conditions. The mixed-layer temperature results are consistent with the sea-ice analysis. The reanalyses show significant warming trends, particularly in autumn, whereas the coupled models exhibit substantially weaker warming. In the Kara Sea, significant autumn warming is only reproduced by the reanalyses. Climate index correlations further show that ORAS and TOPAZ reproduce the observed variables most consistently. RARE's problems seem to stem mostly from the anomalous behaviour of the ML temperature discussed in Sect. 4.4 and 4.5. HIRHAM–NAOSIM generally performs better than RASM in the Barents Sea, while both models struggle considerably with the Kara Sea, RASM slightly less. Overall, the reanalyses, particularly ORAS and TOPAZ, provide the most realistic representation of recent sea-ice and mixed-layer temperature variability in the Barents and Kara Seas. The coupled regional models reproduce some aspects of interannual variability but underestimate the observed long-term changes, especially in the Kara Sea~~, while NorESM clearly overestimates it~~ . ~~NorESM also have a clearly higher sea ice thickness than CNRM (not shown), which might be the reason for its lower than average warming in the future projections~~

Overall, ORAS and TOPAZ perform almost equally well in representing the observed sea ice conditions, although ORAS performs slightly better in the spring SIC. Both products are therefore suitable for studies requiring a realistic representation of recent observed climate, while it is still important to take into account their data assimilation if SIC and SST. RARE is also able to produce the trend seen in the satellite observations, but overestimates the sea ice and shows anomalous behaviour in the Barents Sea MLT (Sec. 4.4 and 4.5). We suggest further research and possibly modelling studies into the reasons behind the seen anomalies, whether they result from the ocean model, from the assimilation scheme or from the forcing.

~~Even though the Barents and Kara Seas are a relatively small (11%) part of the Arctic Ocean, they play a larger role than their size indicates for Arctic sea ice and the global climate (Peng et al., 2025). They have contributed to 30–40% of the Arctic~~

~~winter sea iceloss in recent years (Onarheim et al., 2018; Liu et al., 2022), and the processes of ocean–sea-ice–atmosphere interaction in the region affect the ocean and sea ice further downstream. Models currently struggle to represent this region. One reason is resolution, as at least 0.25° resolution is needed to correctly simulate the ocean currents~~ For applications where models and not reanalyses are needed, which of the models we would recommend depends on the use case and the region. These applications include the ones analysing air-sea fluxes, which reanalyses are not able to realistically simulate. RASM has a better mean state of the sea ice, but for any study looking into the changes happening in the Arctic, the total lack of a trend in the Barents Sea ~~(Docquier et al., 2020), which is not the case in the CMIP6 models~~is a problem. Therefore, we recommend HIRHAM–NAOSIM over RASM in such cases for the Barents Sea. However, ~~as seen in this study, even higher resolution models struggle. Docquier et al. (2020) also found that the effect of increasing the resolution on ocean heat transport and March sea ice area in~~ for the Kara Sea, where both models lack the trend, the better mean sea ice representation of RASM might make it the better choice. For HIRHAM–NAOSIM, further research could focus on looking more closely into the behaviour of the model over the Kara Sea and also the eastern part of the Barents Sea~~differs between models, suggesting that more work is needed than simply increasing the resolution. Further model development is therefore required to better understand the drivers of~~ . Our research only looked into the first ensemble member of the HIRHAM–NAOSIM model. An interesting question would be whether the lack of a trend, too wide autumn sea ice, and the opposite correlations in the climate indices persist over all the ensemble members. If they do, then a deeper study into the model to find the processes behind the Kara Sea performance is suggested. For RASM, the biggest problems of the model come down to the ~~declining sea ice trends in~~ total lack of trend in the sea ice and ML temperature. More research into the reasons of this could considerably improve the model. For example, the Atlantic Ocean lateral boundary temperature trends might result in sea ice and ML temperature trends in RASM. Future studies could also account for delayed responses to large-scale climate variability through lead-lag analysis and use methods such as multiple regression or machine learning to distinguish the combined and potentially nonlinear effects of the different drivers. Even though the Barents and Kara Seas ~~and to improve~~ cover only approximately 11% of the Arctic Ocean, they have contributed 30–40% of the recent Arctic winter sea ice loss and influence the ocean, sea ice, and atmosphere farther downstream. Improving the representation of ~~current trends in models in order to better predict future changes~~these regions in coupled models is therefore important not only for regional studies, but also for understanding Arctic and global climate variability and change.

Code and data availability. RARE 1.15.2 (Carton and Chepurin, 2023) data was acquired through the UMD Ocean Climate Lab https://www2.atmos.umd.edu/~carton/index_files2/rare1.15.2_download.htm. TOPAZ4b data, was acquired from the CMEMS website https://doi.org/10.48670/moi-00007. ORAS5 (Zuo et al., 2019) was aquired from the Universität Hamburg website https://www.cen.uni-hamburg.de/icdc/data/ocean/easy-init-ocean/ecmwf-oras5.html. OSI SAF data (EUMETSAT, 2022) is available at EUMETSAT https://doi.org/10.15770/EUM_SAF_OSI_0013. The NorESM2-MM and CNRM-ESM2-1 simulations are part of the World Climate Research Programme (WCRP)'s CMIP6 archived simulations, which can be found on the Earth System Grid Federation (ESGF, https://pcmdi.llnl.gov/CMIP6/, Lawrence Livermore National Laboratory, 2025). HIRHAM–NAOSIM data are available at the tape archive of the German Climate Computing Center (DKRZ) via https://hdl.handle.net/21.14106/a6d312c42d4501e75bd9de9186b323206ff5a65b (Dorn, 2024, dataset DKRZ_

LTA_049_ds00009). RASM data was acquired from Naval Postgraduate School, Monterey, CA, United States, through WM. Climate index data was aquired from Royal Netherlands Meteorological Institute (KNMI) Code for producing plots of this manuscript as well as a subset of the data needed for this can be found at https://doi.org/10.5281/zenodo.22141085 (Äijälä, 2026)

Appendix A: Additional figures

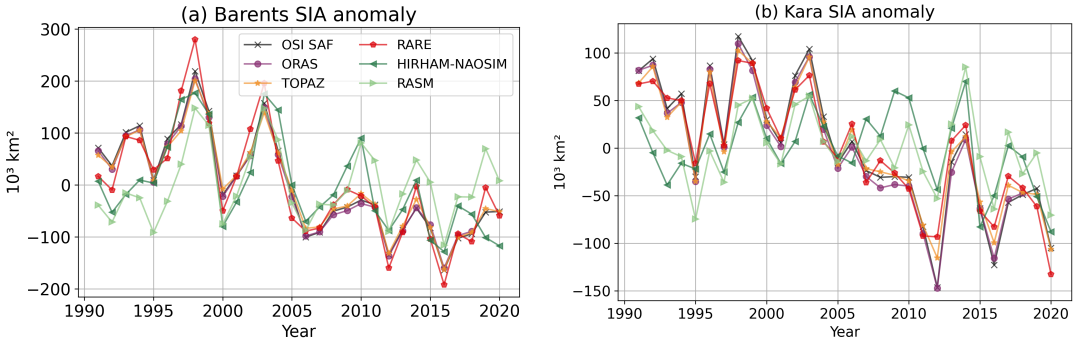


Figure A1. ~~Average of~~ Yearly mean ~~ML temperature~~ of the monthly SIA anomalies for Barents (left) and ~~trend~~ Kara Sea (right) ~~of all the datasets~~ in ~~spring (MAM) over the whole period~~each model shown in Fig 2.

A1

Author contributions. **C. Ä.:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization **L. G.-L.:** Conceptualization, Writing - Original Draft, Writing - Review & Editing **W. D.:** Conceptualization, Resources, Writing - Review & Editing, Project administration, Funding acquisition **S. K. L.:** Conceptualization, Writing - Review & Editing, Supervision, Project administration, Funding acquisition **W. M.:** Conceptualization, Resources, Writing - Review & Editing **P. U.:** Conceptualization, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

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

Acknowledgements. This study was undertaken as part of the European Union's Horizon 2020 research and innovation program under grant agreement no. 101003590 (PolarRES project; https://polarres.eu). The authors thank the European Commission for facilitating this research. The work was also supported by the Research Council of Finland (contract 364876). WD also acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project 268020496 TRR 172, within the Transregional Collaborative Research Center "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms (AC)^3". The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources. This work also used resources of the Deutsches Klimarechenzentrum (DKRZ) under project ID aa0049. This study has been conducted using E.U. Copernicus Marine Service Information (https://doi.org/10.48670/moi-00007) ChatGPT and Microsoft Copilot assisted with initial data visualization code and with light language editing of an earlier draft. All code was reviewed and further refined by CÄ.

References

Asbjørnsen, H., Årthun, M., Skagseth, , and Eldevik, T.: Mechanisms Underlying Recent Arctic Atlantification, Geophysical Research Letters, 47, e2020GL088 036, https://doi.org/10.1029/2020GL088036, 2020.

Aue, L., Röntgen, L., Dorn, W., Uotila, P., Vihma, T., Spreen, G., and Rinke, A.: Impact of three intense winter cyclones on the sea ice cover in the Barents Sea: A case study with a coupled regional climate model, Frontiers in Earth Science, 11, https://doi.org/10.3389/feart.2023.1112467, 2023.

Balmaseda, M., Hernandez, F., Storto, A., Palmer, M., Alves, O., Shi, L., Smith, G., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y.-S., Chepurin, G., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M., Masina, S., Masuda, S., Meyssignac, B., Mogensen, K., Parent, L., Peterson, K., Tang, Y., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J.-F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., Wilmer-Becker, K., and Gaillard, F.: The Ocean Reanalyses Intercomparison Project (ORA-IP), Journal of Operational Oceanography, 8, 80–97, https://doi.org/10.1080/1755876X.2015.1022329, 2015.

Carton, J. A. and Chepurin, G. A.: RARE: The Regional Arctic Reanalysis, Journal of Climate, 36, 2333–2348, https://doi.org/10.1175/JCLI-D-22-0340.1, 2023.

Cassano, J. J., DuVivier, A., Roberts, A., Abel, M. R., Seefeldt, M., Brunke, M., Craig, A., Fisel, B., Gutowski, W., Hamman, J., Higgins, M., Maslowski, W., Nijssen, B., Osinski, R., and Zeng, X.: Development of the Regional Arctic System Model (RASM): Near-Surface Atmospheric Climate Sensitivity, Journal of Climate, 30, 5729–5753, https://doi.org/10.1175/JCLI-D-15-0775.1, 2017.

Chassignet, E. P., Smith, L. T., Halliwell, G. R., and Bleck, R.: North Atlantic Simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the Vertical Coordinate Choice, Reference Pressure, and Thermobaricity, Journal of Physical Oceanography, 33, 2504–2526, https://doi.org/10.1175/1520-0485(2003)033<2504:NASWTH>2.0.CO;2, 2003.

Chatterjee, S., Semmler, T., Screen, J., Furevik, T., Selivanova, J., Raj, R. P., Murukesh, N., and Ravichandran, A. M.: Atmosphere–Ocean–Sea Ice Feedbacks Sustain Recent Barents Sea Ice Loss despite Cooler Atlantic Water Inflow, Journal of Climate, 37, 6519–6532, https://doi.org/10.1175/JCLI-D-24-0020.1, 2024.

Comiso, J. C.: Abrupt decline in the Arctic winter sea ice cover, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL027341, 2006.

Docquier, D., Fuentes-Franco, R., Koenigk, T., and Fichefet, T.: Sea Ice—Ocean Interactions in the Barents Sea Modeled at Different Resolutions, Frontiers in Earth Science, 8, https://doi.org/10.3389/feart.2020.00172, 2020.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sensing of Environment, 116, 140–158, https://doi.org/10.1016/j.rse.2010.10.017, 2012.

Dorn, W.: HIRHAM–NAOSIM simulations for the period 1979–2021 with ice–ocean initialization from ORAS5 (P3), DOKU at DKRZ [data set], https://hdl.handle.net/21.14106/a6d312c42d4501e75bd9de9186b323206ff5a65b, 2024.

Dorn, W., Rinke, A., Köberle, C., Dethloff, K., and Gerdes, R.: Evaluation of the Sea-Ice Simulation in the Upgraded Version of the Coupled Regional Atmosphere-Ocean- Sea Ice Model HIRHAM–NAOSIM 2.0, Atmosphere, 10, 431, https://doi.org/10.3390/atmos10080431, 2019.

Duan, C., Dong, S., and Wang, Z.: New insights into the quantitative description of seasonal sea ice variability process in the semi-closed Kara Sea, Regional Studies in Marine Science, 32, https://doi.org/10.1016/j.rsma.2019.100856, 2019a.

Duan, C., Dong, S., Xie, Z., and Wang, Z.: Temporal variability and trends of sea ice in the Kara Sea and their relationship with atmospheric factors, Polar Science, 20, 136–147, https://doi.org/10.1016/j.polar.2019.03.002, 2019b.

Dörr, J., Årthun, M., Docquier, D., Li, C., and Eldevik, T.: Causal Links Between Sea-Ice Variability in the Barents-Kara Seas and Oceanic and Atmospheric Drivers, Geophysical Research Letters, 51, e2024GL108 195, https://doi.org/10.1029/2024GL108195, 2024.

E.U. Copernicus Marine Service Information (CMEMS): TOPAZ4B, Marine Data Store (MDS), https://doi.org/10.48670/moi-00007, accessed on 01 Nov 2024.

EUMETSAT: OSI SAF Global sea ice concentration climate data record 1978-2020, https://doi.org/10.15770/EUM_SAF_OSI_0013, 2022.

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

GEBCO Bathymetric Compilation Group: The GEBCO_2026 Grid - a continuous terrain model for oceans and land at 15 arc-second intervals, https://doi.org/10.5285/4f68d5c7-45eb-f999-e063-7086abc036fa, 2026.

Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E., and Jung, T.: Predictability of the Arctic sea ice edge, Geophysical Research Letters, 43, 1642–1650, https://doi.org/10.1002/2015GL067232, 2016.

Goosse, H., Kay, J. E., Armour, K. C., Bodas-Salcedo, A., Chepfer, H., Docquier, D., Jonko, A., Kushner, P. J., Lecomte, O., Massonnet, F., Park, H.-S., Pithan, F., Svensson, G., and Vancoppenolle, M.: Quantifying climate feedbacks in polar regions, Nature Communications, 9, 1919, https://doi.org/10.1038/s41467-018-04173-0, 2018.

Gouretski, V. and Reseghetti, F.: On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database, Deep Sea Research Part I: Oceanographic Research Papers, 57, 812–833, https://doi.org/10.1016/j.dsr.2010.03.011, 2010.

Hamman, J., Nijssen, B., Roberts, A., Craig, A., Maslowski, W., and Osinski, R.: The coastal streamflow flux in the Regional Arctic System Model, Journal of Geophysical Research: Oceans, 122, 1683–1701, https://doi.org/10.1002/2016JC012323, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JC012323, 2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, Journal of Climate, 30, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1, 2017.

Hunke, E., Allard, R., Bailey, D., Blain, P., Craig, T., Damsgaard, A., Dupont, F., DuVivier, A., Grumbine, R., Holland, M., Jeffery, N., Lemieux, J.-F., Roberts, A., Turner, M., and Winton, M.: CICE Consortium/CICE Version 6.0.0, https://doi.org/10.5281/zenodo.1893041, language: eng, 2018.

Ingvaldsen, R. B., Assmann, K. M., Primicerio, R., Fossheim, M., Polyakov, I. V., and Dolgov, A. V.: Physical manifestations and ecological implications of Arctic Atlantification, Nature Reviews Earth & Environment, 2, 874–889, https://doi.org/10.1038/s43017-021-00228-x, 2021.

International Hydrographic Organization, I. H. O. and Sieger, R.: Limits of oceans and seas in digitized, machine readable form, https://doi.org/10.1594/PANGAEA.777975, 2012.

IPCC: Climate Change 2021: The Physical Science Basis, https://doi.org/10.1017/9781009157896, 2021.

Kumar, A., Yadav, J., and Mohan, R.: Spatio-temporal change and variability of Barents-Kara sea ice, in the Arctic: Ocean and atmospheric implications, Science of The Total Environment, 753, 142 046, https://doi.org/10.1016/j.scitotenv.2020.142046, 2021.

Laurent Bertino and Jiping Xie: SYNTHESIS QUALITY OVERVIEW DOCUMENT (SQO) Associated to extended quality information document (QUID): CMEMS-ARC-QUID-002-003 QUID Version: 1.5 Associated to Product ID: ARCTIC_MULTIYEAR_PHY_ 002_003, 2025.

Levine, X. J., Williams, R. S., Marshall, G., Orr, A., Seland Graff, L., Handorf, D., Karpechko, A., Köhler, R., Wijngaard, R. R., Johnston, N., Lee, H., Nieradzik, L., and Mooney, P. A.: Storylines of summer Arctic climate change constrained by Barents–Kara seas and Arctic tropospheric warming for climate risk assessment, Earth System Dynamics, 15, 1161–1177, https://doi.org/10.5194/esd-15-1161-2024, 2024.

Li, M., Ke, C., Cheng, B., Shen, X., He, Y., and Sha, D.: The Roles of Sea Ice Export, Atmospheric and Oceanic Factors in the Seasonal and Regional Variability of Arctic Sea Ice during 1979–2020, Remote Sensing, 14, https://doi.org/10.3390/rs14040904, 2022.

Lind, S., Ingvaldsen, R. B., and Furevik, T.: Arctic warming hotspot in the northern Barents Sea linked to declining sea-ice import, Nature Climate Change, 8, 634–639, https://doi.org/10.1038/s41558-018-0205-y, 2018.

Lindsay, R. W. and Zhang, J.: Assimilation of ice concentration in an ice–ocean model, J. Atmos. Oceanic Technol., 23, 742–749, 2006.

Liu, Y., Losch, M., Tremblay, L. B., and Janout, M.: Landfast ice in the Kara Sea stabilizes the Arctic halocline and may slow down Atlantification of the Eurasian Basin, Communications Earth and Environment, 6, https://doi.org/10.1038/s43247-025-02360-8, 2025.

Liu, Z., Risi, C., Codron, F., Jian, Z., Wei, Z., He, X., Poulsen, C. J., Wang, Y., Chen, D., Ma, W., Cheng, Y., and Bowen, G. J.: Atmospheric forcing dominates winter Barents-Kara sea ice variability on interannual to decadal time scales, Proceedings of the National Academy of Sciences, 119, e2120770 119, https://doi.org/10.1073/pnas.2120770119, 2022.

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Lorenzo, E. D., Mantua, N. J., Miller, A. J., Minobe, S., Nakamura, H., Schneider, N., Vimont, D. J., Phillips, A. S., Scott, J. D., and Smith, C. A.: The Pacific Decadal Oscillation, Revisited, Journal of Climate, 29, 4399–4427, https://doi.org/10.1175/JCLI-D-15-0508.1, 2016.

NOAA CPC: East Atlantic/ Western Russia, https://www.cpc.ncep.noaa.gov/data/teledoc/eawruss.shtml, accessed 2026-06-02, a.

NOAA CPC: North Atlantic Oscillation (NAO) Index, https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml, accessed 2026-06-02, b.

NOAA NODC: monthly NH upper ocean mean temperature, climexp.knmi.nl/getindices.cgi?WMO=NODCData/temp700_nh&STATION=NH_upper_ocean_mean_temperature, accessed 2026-06-02.

Onarheim, I. H., Eldevik, T., Smedsrud, L. H., and Stroeve, J. C.: Seasonal and Regional Manifestation of Arctic Sea Ice Loss, Journal of Climate, 31, 4917–4932, https://doi.org/10.1175/JCLI-D-17-0427.1, 2018.

Peng, Y., Duan, A., Shen, Z., Wei, J., and Cao, S.: Emergent Constraint on CMIP6 Model Projections of Winter Ice-Free Periods in the Barents–Kara Sea, Geophysical Research Letters, 52, e2025GL117 660, https://doi.org/10.1029/2025GL117660, 2025.

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J. L., Gochis, D. J., Ahmadov, R., Peckham, S. E., Grell, G. A., Michalakes, J., Trahan, S., Benjamin, S. G., Alexander, C. R., Dimego, G. J., Wang, W., Schwartz, C. S., Romine, G. S., Liu, Z., Snyder, C., Chen, F., Barlage, M. J., Yu, W., and Duda, M. G.: The Weather Research and Forecasting Model: Overview,

System Efforts, and Future Directions, Bulletin of the American Meteorological Society, 98, 1717–1737, https://doi.org/10.1175/BAMS-D-15-00308.1, 2017.

Royal Netherlands Meteorological Institute (KNMI): Climate Explorer, https://climexp.knmi.nl, accessed 2026-06-02.

Salas Mélia, D.: A global coupled sea ice–ocean model, Ocean Modelling, 4, 137–172, https://doi.org/10.1016/S1463-5003(01)00015-4, 2002.

Sallée, J.-B., Wienders, N., Speer, K., and Morrow, R.: Formation of Subantarctic Mode Water in the southeastern Indian Ocean, Ocean Dynamics, 56, 525–542, https://doi.org/10.1007/s10236-005-0054-x, 2006.

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., and Wang, Z.: Assessment of Southern Ocean mixed-layer depths in CMIP5 models: Historical bias and forcing response, Journal of Geophysical Research: Oceans, 118, 1845–1862, https://doi.org/10.1002/jgrc.20157, 2013.

Seefeldt, M. W., Cassano, J. J., Lee, Y. J., Maslowski, W., Craig, A. P., and Osinski, R.: Evaluation of dynamical downscaling in a fully coupled regional earth system model, Frontiers in Earth Science, 12, https://doi.org/10.3389/feart.2024.1392031, 2024.

Seip, K. L. and Wang, H.: High-Resolution Lead–Lag Relations Between Barents Sea Temperatures, the AMOC and the AMO During 1971–2018, Atmosphere-Ocean, 61, 306–317, https://doi.org/10.1080/07055900.2023.2251426, 2023.

Seland, , Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: 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, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Shabanov, P., Osadchiev, A., Shabanova, N., and Ogorodov, S.: Decline in Ice Coverage and Ice-Free Period Extension in the Kara and Laptev Seas during 1979–2022, Remote Sensing, 16, https://doi.org/10.3390/rs16111875, 2024.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., Senior, C. A., Sobel, A. H., Stainforth, D. A., Tett, S. F. B., Trenberth, K. E., van den Hurk, B. J. J. M., Watkins, N. W., Wilby, R. L., and Zenghelis, D. A.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic Change, 151, 555–571, https://doi.org/10.1007/s10584-018-2317-9, 2018.

Skagseth, , Furevik, T., Ingvaldsen, R., Loeng, H., Mork, K. A., Orvik, K. A., and Ozhigin, V.: Volume and Heat Transports to the Arctic Ocean Via the Norwegian and Barents Seas, in: Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate, edited by Dickson, R. R., Meincke, J., and Rhines, P., pp. 45–64, Springer Netherlands, Dordrecht, ISBN 978-1-4020-6774-7, https://doi.org/10.1007/978-1-4020-6774-7_3, 2008.

Smedsrud, L. H., Esau, I., Ingvaldsen, R. B., Eldevik, T., Haugan, P. M., Li, C., Lien, V. S., Olsen, A., Omar, A. M., Otterå, O. H., Risebrobakken, B., Sandø, A. B., Semenov, V. A., and Sorokina, S. A.: The Role of the Barents Sea in the Arctic Climate System, Reviews of Geophysics, 51, 415–449, https://doi.org/10.1002/rog.20017, 2013.

Smith, A. R., Jones, A. P., Briegleb, A. B. P., Bryan, A. F. O., Danabasoglu, A. G., Dennis, A. J. M., Dukowicz, A. J., Eden, A. C., Fox-Kemper, A. B., Gent, A. P. R., Hecht, A. M., Jayne, A. S., Jochum, A. M., Large, A. W. G., Lindsay, A. K., Maltrud, A. M., Norton, A. N. J., Peacock, A. S. L., Vertenstein, A. M., and Yeager, A. S.: The Parallel Ocean Program (POP) reference manual: Ocean component of the Community Climate System Model (CCSM), https://opensky.ucar.edu/islandora/object/manuscripts%3A825, 2010.

Storto, A., Alvera-Azcárate, A., Balmaseda, M. A., Barth, A., Chevallier, M., Counillon, F., Domingues, C. M., Drevillon, M., Drillet, Y., Forget, G., Garric, G., Haines, K., Hernandez, F., Iovino, D., Jackson, L. C., Lellouche, J.-M., Masina, S., Mayer, M., Oke, P. R., Penny,

S. G., Peterson, K. A., Yang, C., and Zuo, H.: Ocean Reanalyses: Recent Advances and Unsolved Challenges, Frontiers in Marine Science, 6, https://doi.org/10.3389/fmars.2019.00418, 2019.

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

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: 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, 11, 4182–4227, https://doi.org/10.1029/2019MS001791, 2019.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V. S., Marnela, M., Massonnet, F., Mignac, D., Peterson, K. A., Sadikni, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J., and Zhang, Z.: An assessment of ten ocean reanalyses in the polar regions, Climate Dynamics, 52, 1613–1650, https://doi.org/10.1007/s00382-018-4242-z, 2019.

van Oldenborgh, G. J., te Raa, L. A., Dijkstra, H. A., and Philip, S. Y.: Frequency- or amplitude-dependent effects of the Atlantic meridional overturning on the tropical Pacific Ocean, Ocean Science, 5, 293–301, https://doi.org/10.5194/os-5-293-2009, 2009.

Wang, Q., Wang, X., Wekerle, C., Danilov, S., Jung, T., Koldunov, N., Lind, S., Sein, D., Shu, Q., and Sidorenko, D.: Ocean Heat Transport Into the Barents Sea: Distinct Controls on the Upward Trend and Interannual Variability, Geophysical Research Letters, 46, 13 180–13 190, https://doi.org/10.1029/2019GL083837, 2019.

Xie, J. and Bertino, L.: PRODUCT USER MANUAL For Arctic Ocean Physical Reanalysis Product, 2025.

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013, Ocean Science, 13, 123–144, https://doi.org/10.5194/os-13-123-2017, 2017.

Xie, J., Raj, R. P., Bertino, L., Martínez, J., Gabarró, C., and Catany, R.: Assimilation of sea surface salinities from SMOS in an Arctic coupled ocean and sea ice reanalysis, Ocean Science, 19, 269–287, https://doi.org/10.5194/os-19-269-2023, 2023.

Xiong, Z., Yan, Y., Xu, J., Yao, C., and Xu, Y.: Comparative analysis of sea ice conditions in five reanalysis datasets along the Northern Sea Route, Ocean Dynamics, 75, 37, https://doi.org/10.1007/s10236-025-01684-5, 2025.

Yamagami, Y., Watanabe, M., Mori, M., and Ono, J.: Barents-Kara sea-ice decline attributed to surface warming in the Gulf Stream, Nature Communications, 13, 3767, https://doi.org/10.1038/s41467-022-31117-6, 2022.

Zhang, J. and Rothrock, D. A.: Modeling Global Sea Ice with a Thickness and Enthalpy Distribution Model in Generalized Curvilinear Coordinates, Monthly Weather Review, 131, 845–861, https://doi.org/10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2, 2003.

Zhang, Y.-F., Bushuk, M., Winton, M., Hurlin, B., Yang, X., Delworth, T., and Jia, L.: Assimilation of Satellite-Retrieved Sea Ice Concentration and Prospects for September Predictions of Arctic Sea Ice, Journal of Climate, 34, 2107–2126, https://doi.org/10.1175/JCLI-D-20-0469.1, 2021.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Science, 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.

Äijälä, C.: Code for Barents and Kara Seas model and reanalysis comparision analysis, https://doi.org/10.5281/zenodo.22141085, 2026.

Årthun, M., Eldevik, T., Smedsrud, L. H., Skagseth, , and Ingvaldsen, R. B.: Quantifying the Influence of Atlantic Heat on Barents Sea Ice Variability and Retreat in: Journal of Climate Volume 25 Issue 13 (2012), Journal of Climate, 25, https://doi.org/10.1175/JCLI-D-11-00466.1, 2012.

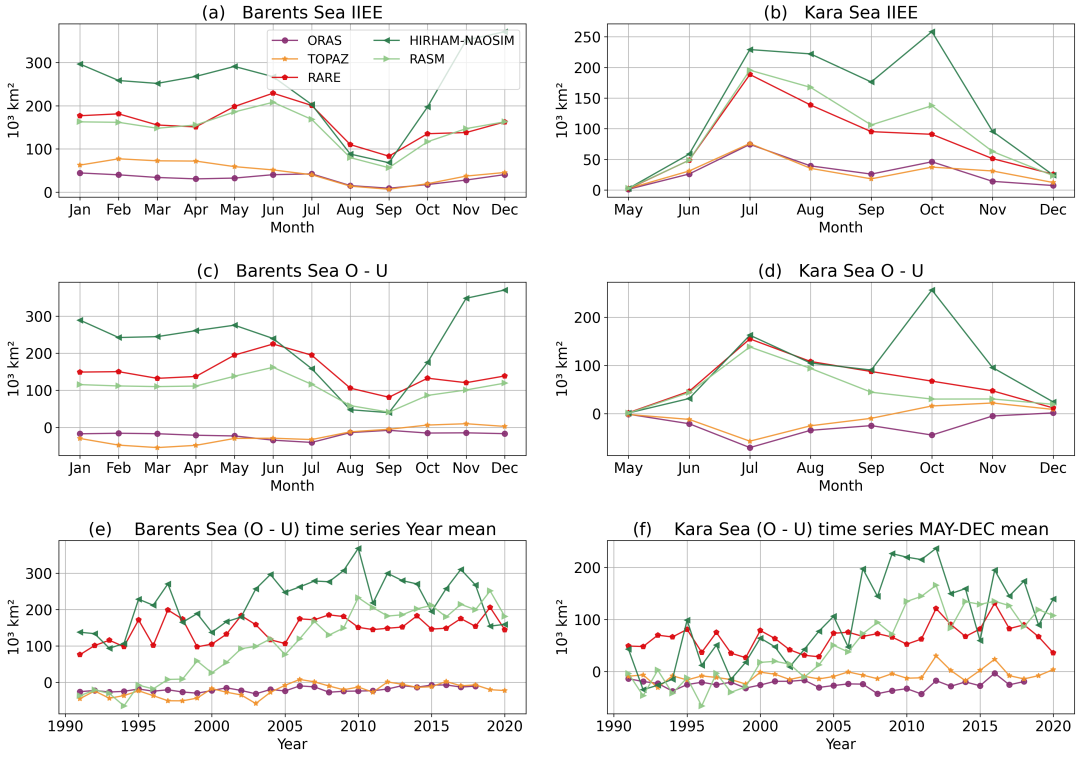


Figure A2. ~~Average~~ Integrated Ice Edge Error (IIEE) of ~~mean ML temperature~~ the (~~left~~a) Barents and ~~trend~~ (~~right~~b) Kara Seas, as well as (c, d) Overestimation - Underestimation of ~~all~~ the ~~datasets in autumn~~ sea ice edge, and (~~ASO~~e, f) ~~over~~ time series of the ~~whole period~~yearly mean Overestimation-Underestimation. For the Kara Sea, only values during May–December are shown, as there is no ice edge during winter and spring in most years.

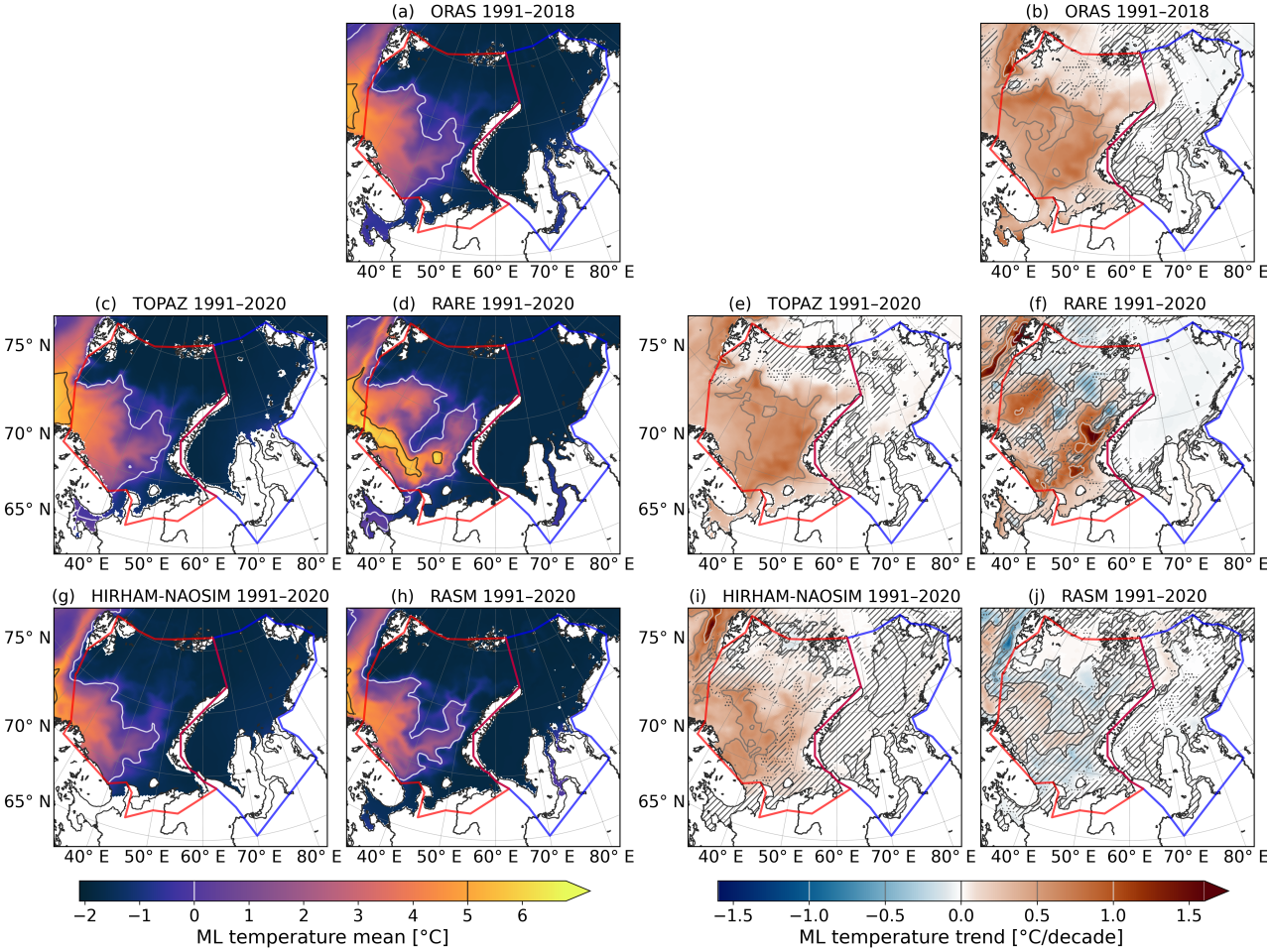


Figure A3. Average of mean ML temperature (left) and trend (right) of all the datasets in spring over the whole period. In the trend subplots, the not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).

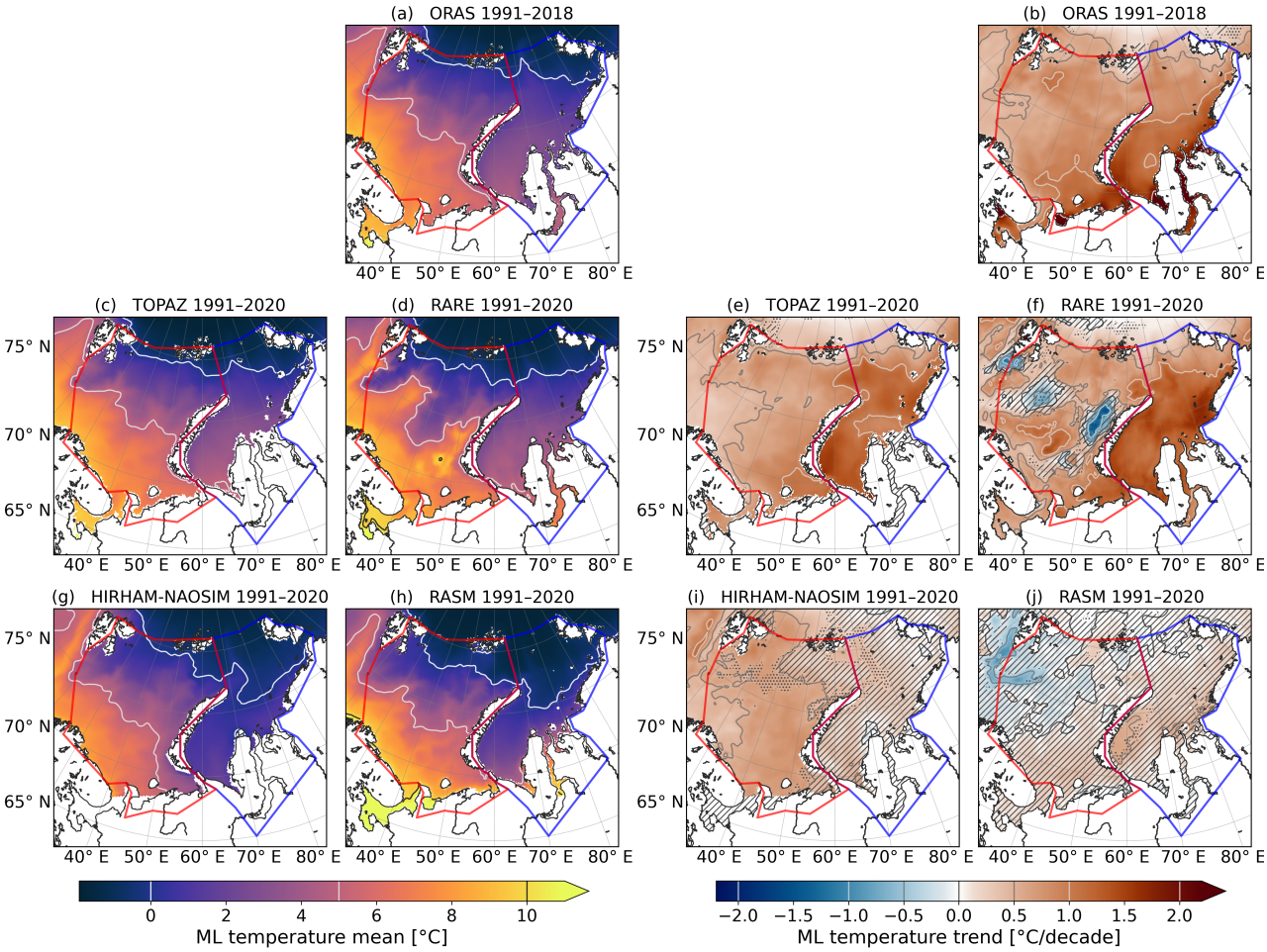


Figure A4. Average of mean ML temperature (left) and trend (right) of all the datasets in autumn over the whole period. In the trend subplots, the not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).

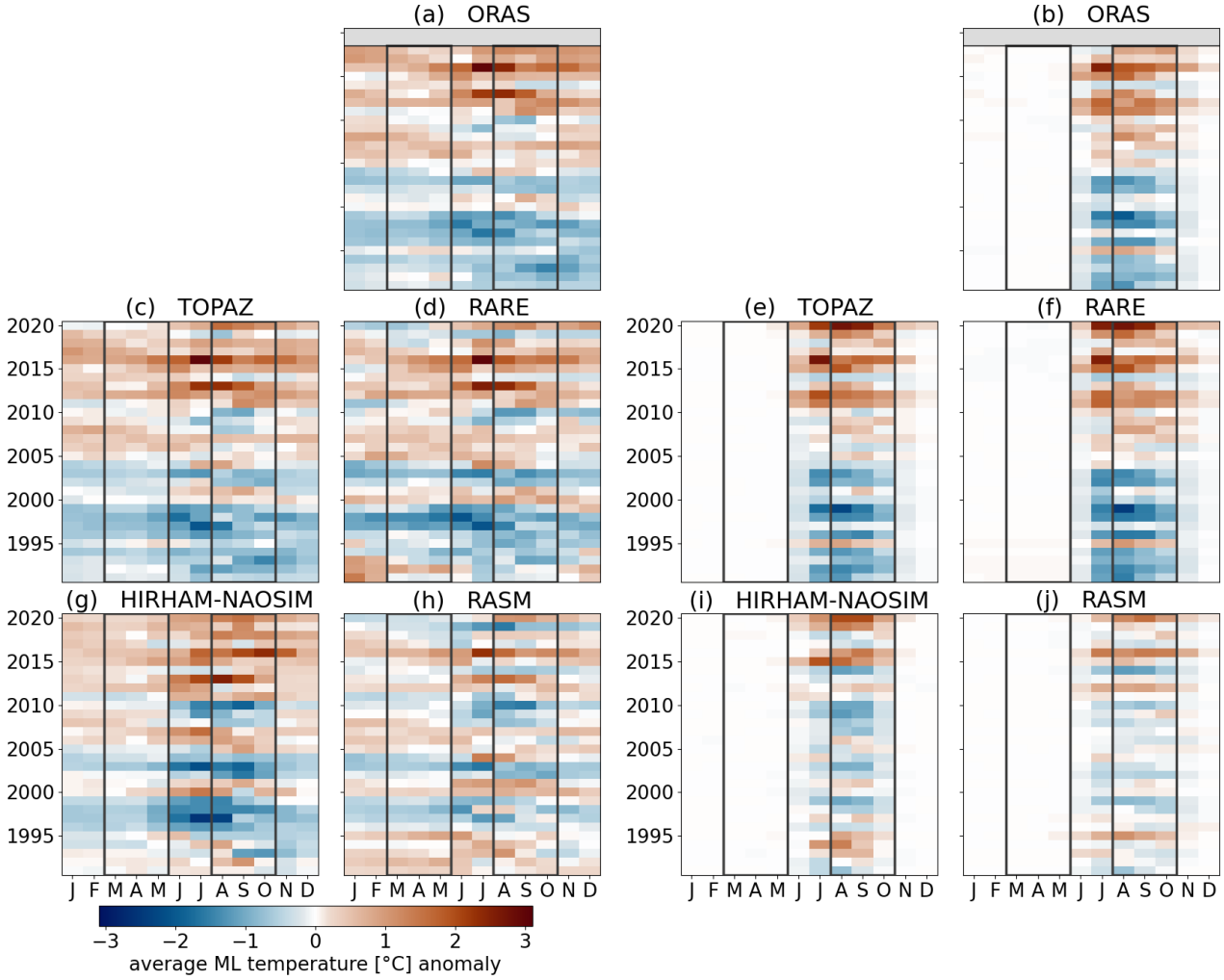


Figure A5. ML temperature anomaly in each model for the Barents (left) and Kara Sea (right). Spring and autumn seasons are marked with dark grey boxes. The anomaly is calculated by taking the difference from the monthly multiyear mean for each month. Note the different colour scales in the figures.

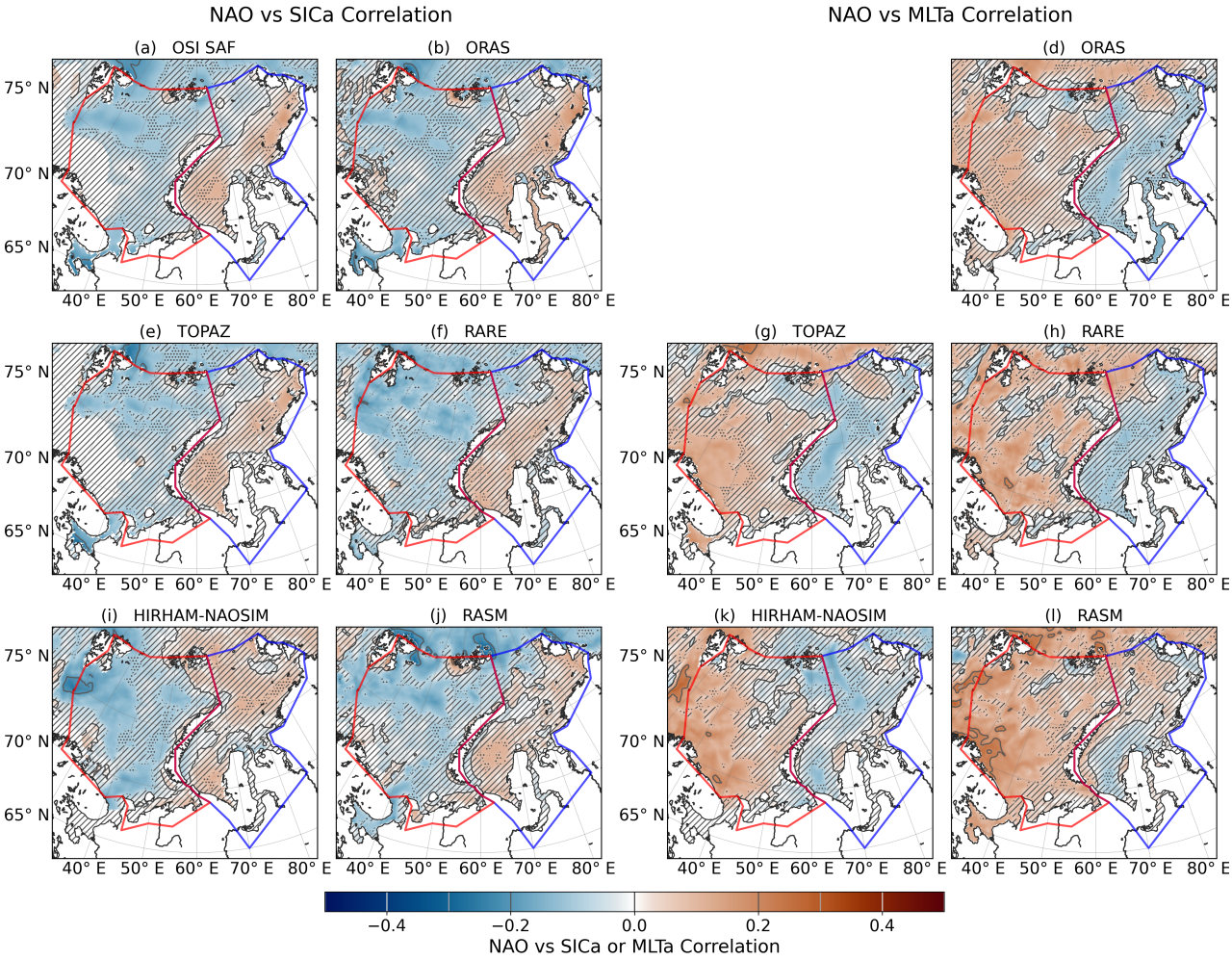


Figure A6. Correlation between climate index NAO and SIC anomaly (left) and MLT anomaly (right). The not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).

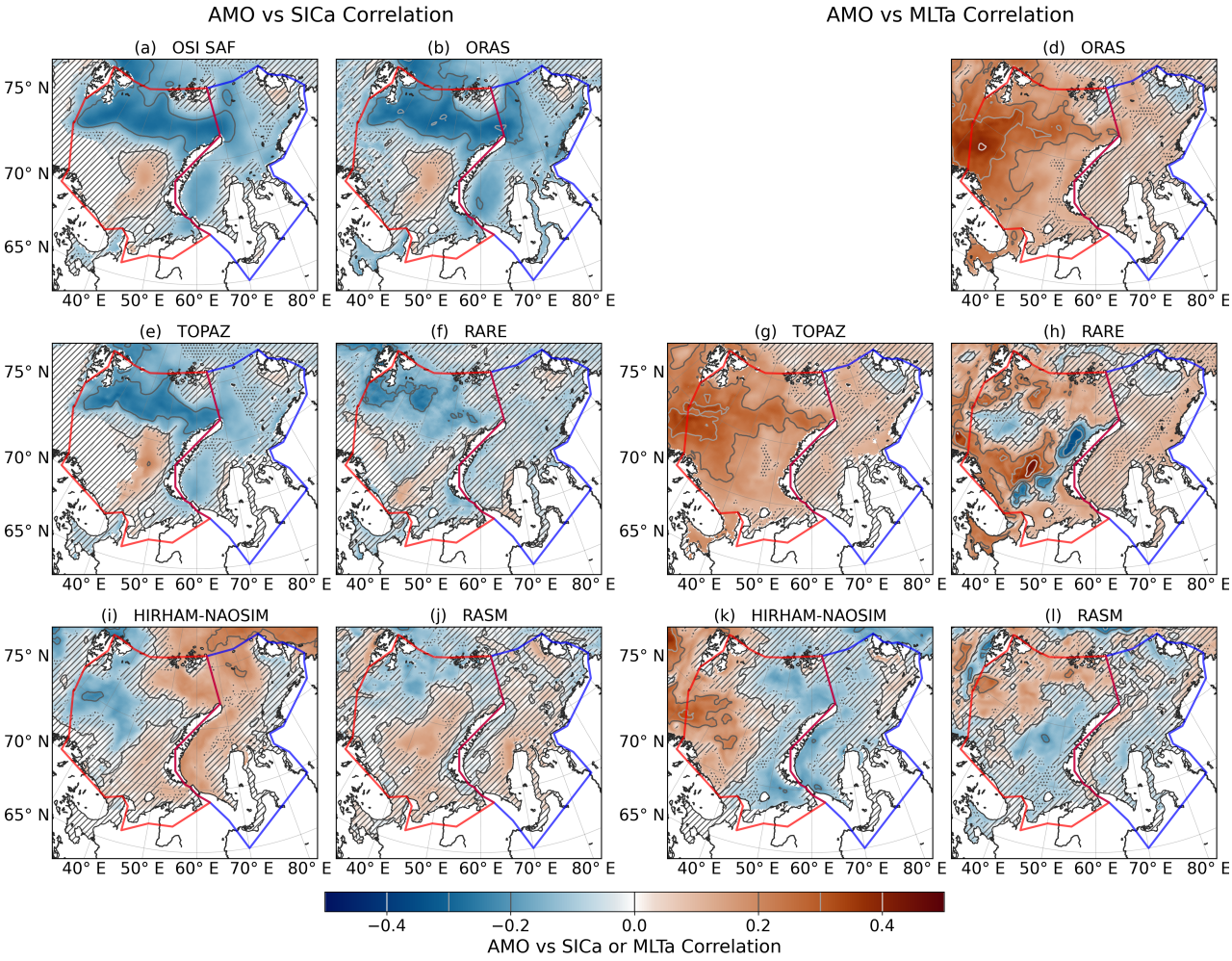


Figure A7. Correlation between climate index AMO and SIC anomaly (left) and MLT anomaly (right). The not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).

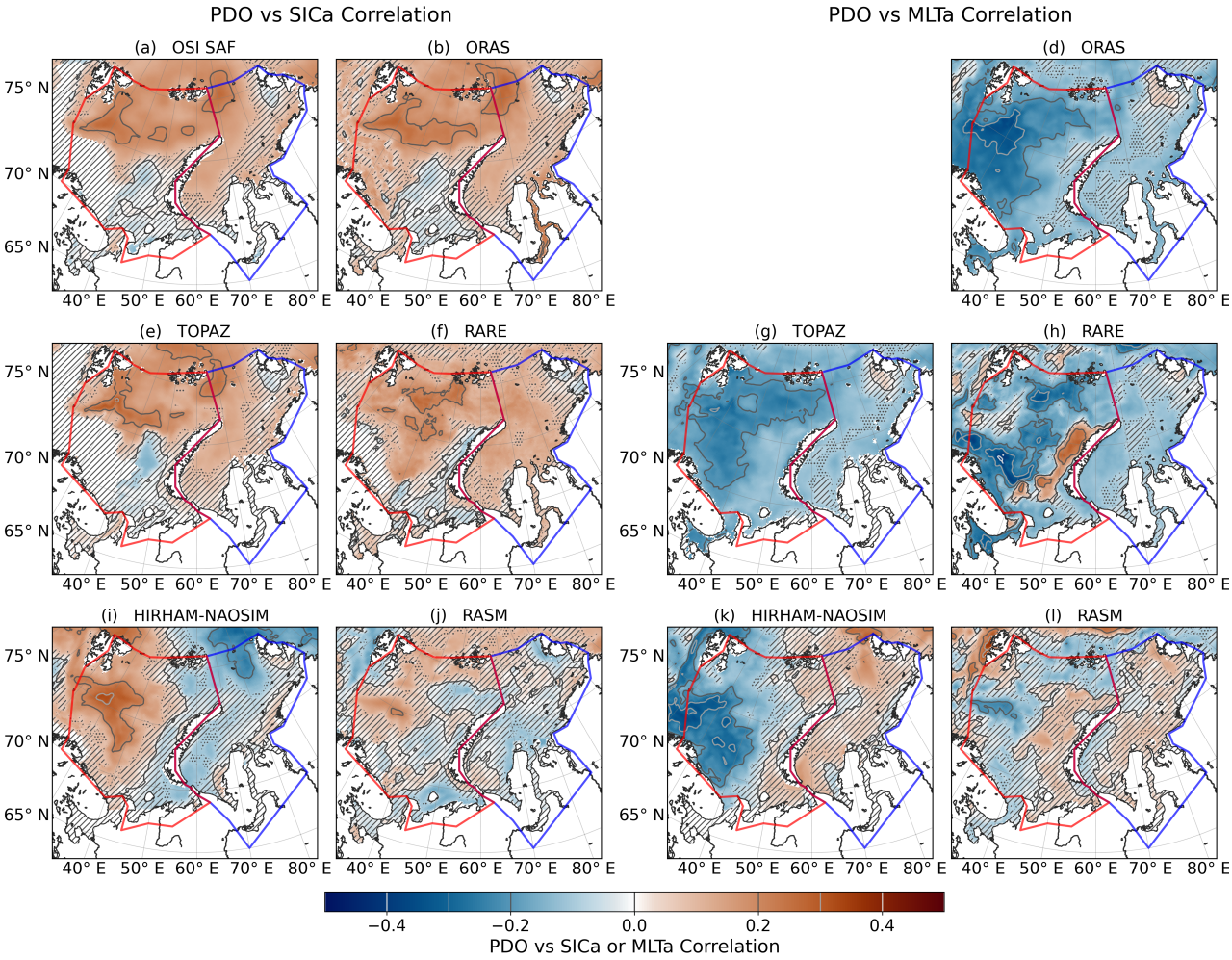


Figure A8. Correlation between climate index PDO and SIC anomaly (left) and MLT anomaly (right). The not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).

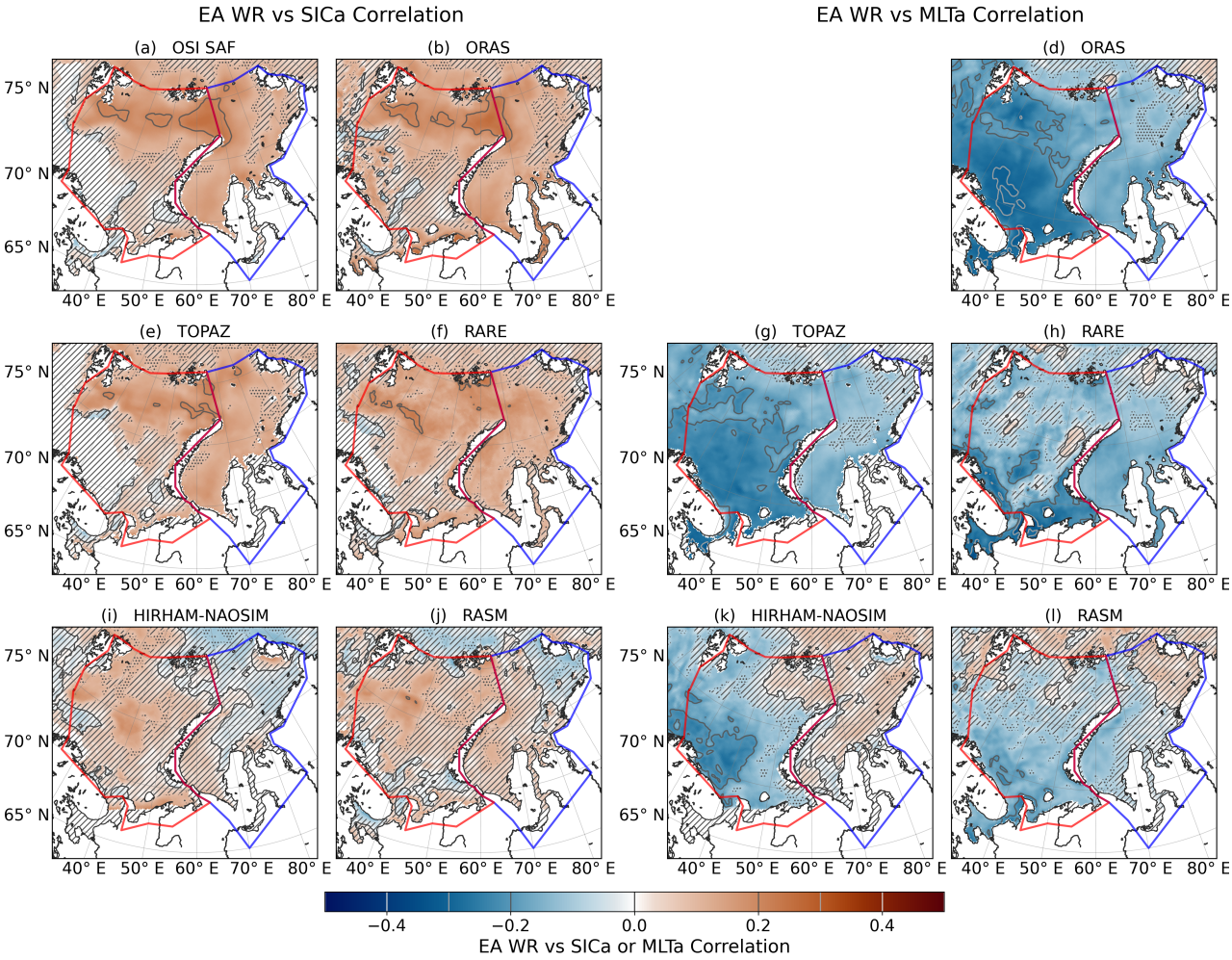


Figure A9. Correlation between climate index EA/WR and SIC anomaly (left) and MLT anomaly (right). The not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).

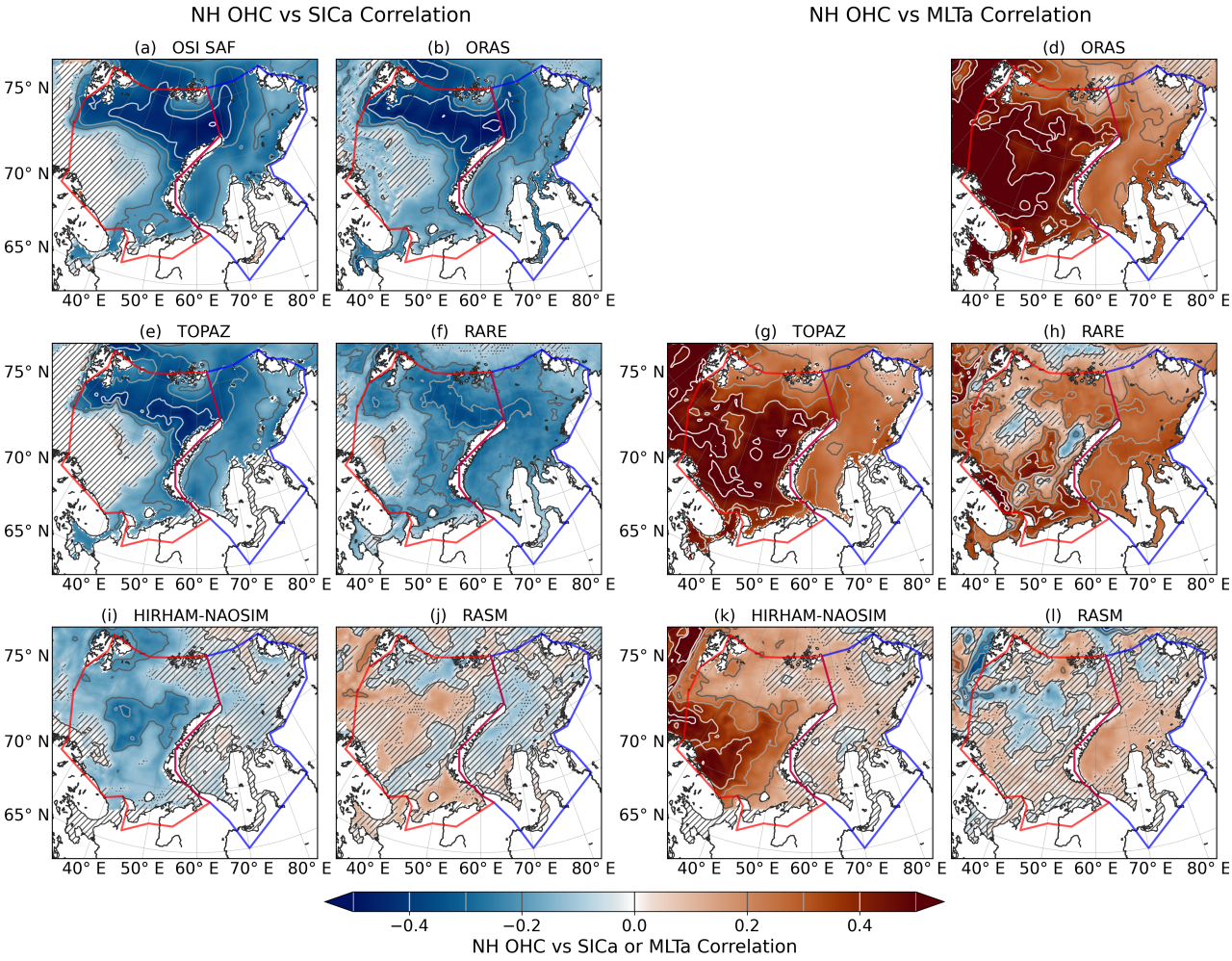


Figure A10. Correlation between climate index NH OHC and SIC anomaly (left) and MLT anomaly (right). The not statistically significant area have been masked ("////" where $p > 0.1$ and "...." where $0.05 < p < 0.1$).