



Impact of initialization methods on the predictive skill in NorCPM: an Arctic–Atlantic case study

Leilane Passos^{1,2,3} · H. R. Langehaug^{2,3} · M. Årthun^{1,2} · T. Eldevik^{1,2} · I. Bethke^{1,2} · M. Kimmritz⁴

Received: 29 July 2021 / Accepted: 20 July 2022
© The Author(s) 2022

Abstract

The skilful prediction of climatic conditions on a forecast horizon of months to decades into the future remains a main scientific challenge of large societal benefit. Here we assess the hindcast skill of the Norwegian Climate Prediction Model (NorCPM) for sea surface temperature (SST) and sea surface salinity (SSS) in the Arctic–Atlantic region focusing on the impact of different initialization methods. We find the skill to be distinctly larger for the Subpolar North Atlantic than for the Norwegian Sea, and generally for all lead years analyzed. For the Subpolar North Atlantic, there is furthermore consistent benefit in increasing the amount of data assimilated, and also in updating the sea ice based on SST with strongly coupled data assimilation. The predictive skill is furthermore significant for at least two model versions up to 8–10 lead years with the exception for SSS at the longer lead years. For the Norwegian Sea, significant predictive skill is more rare; there is relatively higher skill with respect to SSS than for SST. A systematic benefit from more complex data assimilation approach can not be identified for this region. Somewhat surprisingly, skill deteriorates quite consistently for both the Subpolar North Atlantic and the Norwegian Sea when going from CMIP5 to corresponding CMIP6 versions. We find this to relate to change in the regional performance of the underlying physical model that dominates the benefit from initialization.

Keywords Decadal prediction · Subpolar North Atlantic · Norwegian Sea · Prediction system · NorCPM

1 Introduction

Skillful decadal climate predictions are beneficial to society by potentially providing information to stakeholders, contributing to political and economical decision-making, and by guiding the planning of climate adaptation measures (Vera et al. 2010; Kushnir et al. 2019). Dynamical predictions are achieved by initialization of climate models with the observed climate state through data assimilation (DA) or alternative synchronization methods. The initialization can reduce the forecast error by considering the internal variability and the mean forced response in the climate system

(Meehl et al. 2009; Yeager et al. 2012), thus achieving skillful predictions regionally up to a decade (Smith et al. 2019).

Decadal climate predictions are still in their early stages of development, the first initialized coupled ocean-atmospheric efforts started in the beginning of the 2000s (e.g. Smith et al. 2007; Keenlyside et al. 2008). Since then, many procedures and techniques have been developed to deal with inherent initialization problems such as initial “shocks”, model drift, and a growing variety of assimilated data. Developments range from the primary initialization of the surface ocean component to additional initialization of the subsurface ocean and other components of the climate system (Morioka et al. 2018); from updates of single model components (Weakly-Coupled DA) to cross-component updates during DA in particular model components (which we denote here as Strongly-Coupled DA) (Penny et al. 2019) are current aspects under development and investigation.

Comparisons between multiple DA approaches and techniques using the same model can help to understand how different initialization procedures affect the skill of certain physical processes (Polkova et al. 2019b), as for example the Atlantic Meridional Overtuning Circulation. To this

✉ Leilane Passos
leilane.passos@uib.no

¹ University of Bergen, Bergen, Norway

² Bjerknes Centre for Climate Research, Bergen, Norway

³ Nansen Environmental and Remote Sensing Center, Bergen, Norway

⁴ Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

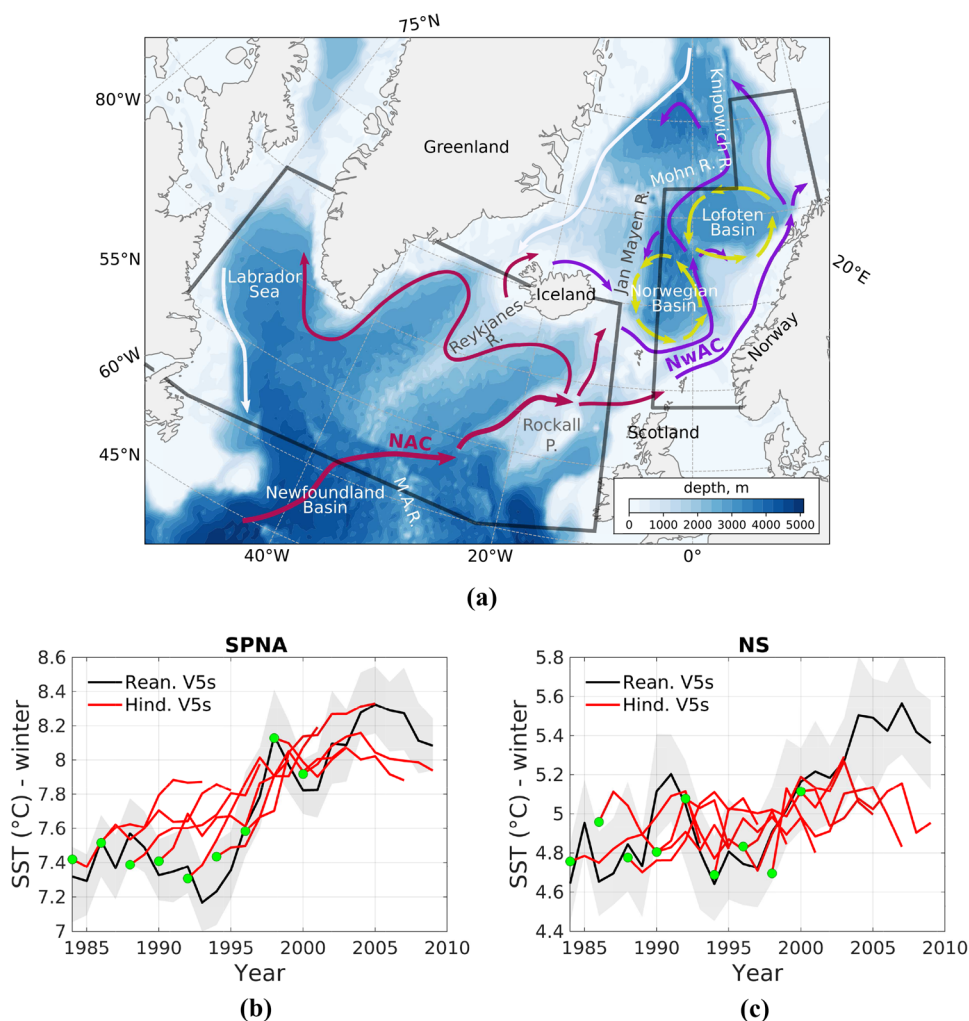
end, the Norwegian Climate Prediction Model (NorCPM), developed by the Bjerknes Center for Climate Research, is a valuable tool. NorCPM consists of a fully-coupled Earth-system model performing ensemble-based sequential DA, based in the Norwegian Earth System model (NorESM). The Norwegian initiative has been through many buildup stages, evolving from surface ocean initialization (Counillon et al. 2016) to assimilation of subsurface observations (Wang et al. 2017) and initializing the sea ice component (Kimmritz et al. 2018). A comparison of these stages and their effects on skill, focusing on a specific region or physical process, can contribute to the NorCPM development by identifying strengths and deficiencies of the different initialization techniques.

On interannual-to-decadal timescales, the Arctic-Atlantic region (Fig. 1a) is characterized by poleward propagation of thermohaline anomalies from the Subpolar North Atlantic (SPNA) towards the Arctic through the Norwegian Sea (NS) (Eldevik et al. 2009; Årthun and Eldevik 2016). Årthun et al. (2017) identified that the propagation is mainly driven by

advection, with a characteristic time scale of 14 years and propagation speed of 2 cm/s. According to the anomalies travel time the prediction horizon varies between 7 and 10 years influencing climate variability over Scandinavia and the state of sea ice in the Arctic (Årthun et al. 2017). Based on observation-based ocean data, Buckley et al. (2019) identified a decorrelation time of winter-time SST and upper ocean heat content of 4–6 years in the SPNA and 2–3 years in the NS. Based on predictions systems, the SPNA is one of the most predictable areas in the world (Yeager et al. 2012; Buckley et al. 2019). Some models have demonstrated significant skill up to 10 years in advance for sea surface temperature and upper ocean heat content (van Oldenborgh et al. 2012; Matei et al. 2012; Yeager et al. 2012; Buckley et al. 2019; Smith et al. 2020).

Despite the NS receiving thermohaline anomalies from the SPNA (Årthun and Eldevik 2016), Langehaug et al. (2017) found skill of only up to 1–3 years in the NS, which is less than what is found for the SPNA region (Matei et al. 2012). Assessing similarities and differences in predictive

Fig. 1 **a** Schematic diagram highlighting the SPNA area (60°W–10°W, 50°N–65°N), according to Robson et al. (2012), and the NS, according to Asbjørnsen et al. (2019), in transparent black. The main currents analyzed in this study are the North Atlantic Current (NAC; pink) and the Norwegian Atlantic Current (NwAC; purple) with its recirculation (yellow). The white arrows indicate cold and fresh surface water from the Arctic. Ridges and Plateau are indicated with R. and P. respectively. **b** Hindcast winter SST in the SPNA in version V5s (red). **c** Hindcast winter SST in the NS in version V5s (red). The black curves show the respective reanalysis used to initialize the model. Grey shading represents the spread of the ensemble members of the reanalysis. The starting time of each hindcast is indicated by a green circle



skill between these two regions for different initialization techniques will improve our understanding of which approaches are most beneficial in enhancing predictive skill and identify sources of uncertainty in the NS.

The present work focuses on investigating the sensitivity of decadal predictive skill of NorCPM to different initialization techniques, and assessing which one leads to higher predictive skill for the SPNA and NS. The variables analyzed are those constrained by the DA: sea surface temperature (SST) and sea surface salinity (SSS). These surface variables are important for marine ecosystems, Arctic sea ice, and atmospheric climate (Årthun et al. 2012, 2018a, b). Here, skill is defined as the ability of the prediction system (NorCPM) to reproduce the same variability of these quantities as in the reanalysis used to initialize NorCPM. In addition, SST variability in NorCPM is compared to that from HadSST2. Five versions of NorCPM are systematically analyzed (Table 1), to primarily investigate how different representations of the ocean initial state and external forcings can affect the predictive skill in the SPNA and the NS for different lead years.

2 Data and methods

2.1 Norwegian climate prediction model (NorCPM)

NorCPM (Counillon et al. 2014, 2016) is based on the Norwegian Earth System Model (NorESM) (Bentsen et al. 2013) adding the Ensemble Kalman Filter (EnKF) as the DA method (Evensen 2003). The fully coupled NorCPM consists of MICOM (ocean), CAM4 (atmosphere), CICE4 (sea ice), CLM4 (land model), and the coupler CPL7; its structure is largely based on the The Community Climate System Model version 4 (CCSM4) (Gent et al. 2011), while its sea ice and land models are based on the Community Earth System Model version 1.0.3 (CESM1) (Vertenstein et al. 2012). The

ocean/sea-ice and atmospheric components have horizontal resolution of 1×1 and 1.9×2.5 , respectively.

The DA in NorCPM is based on anomaly initialization. In this process, monthly anomalies are assimilated, generating a reanalysis field that is used to initialize the decadal hindcast. In anomaly DA, the choice of the climatology reference period is relevant to calculate the mean and the subsequent anomalies. In this study there are two reference periods (Table 1). The DA is applied in the ocean component using sea surface temperature data from HadSST2 (Rayner et al. 2006) and, additionally, subsurface hydrographic profiles from EN4.2.1 (Good et al. 2013) dependent on the particular version. Each reanalysis has 30 ensemble members.

The predictions are initialized from the reanalysis and run freely for 10 years, where each hindcast has between 5 and 20 members depending on the version. In this study, we used results from five different versions: V5, V5w, and V5s, based on NorESM1-ME (Bentsen et al. 2013; Tjiputra et al. 2013) defined here as NorCPM-CMIP5; V6w, and V6s based on NorCPM1 (Bethke et al. 2021) defined here as NorCPM-CMIP6. Details of each version including their respective reference period for calculating anomalies are described in Table 1.

The version V5 assimilates SST through weakly-coupled data assimilation method (WCDA) in the ocean component (described further bellow). Version V5w uses the same assimilation method, but in addition to SST also assimilates hydrographic profiles of temperature and salinity. The implementation of the T-S profiles DA is described in Wang et al. (2017). These authors report an improvement of the system accuracy, generating a reanalyses field suitable to be used at seasonal-to-decadal predictions.

The version V5s (Bethke et al. 2021) assimilates SST and hydrographic profiles, updating the sea ice state through strongly-coupled data assimilation (SCDA) with updates and post-processing of the sea ice state following Kimmritz et al. (2018). The transition from WCDA to SCDA is the pathway

Table 1 NorCPM versions characteristics

| Version (Underlying model) | Assimilated data | CMIP forcing | Assimilation method | Ensemble size | Climatology reference period | Initialization frequency |
|-------------------------------|--|--------------|---------------------|---------------|------------------------------|--------------------------|
| V5 (NorESM1-ME) | SST (HadSST2) | CMIP5 | WCDA | 5 (20) | 1950–2010 | Every 2 years |
| V5w (NorESM1-ME) | SST (HadSST2) Hydrographic profiles (EN4.2.1) | CMIP5 | WCDA | 5 | 1980–2010 | Every 2 years |
| V5s (NorESM1-ME) | SST (HadSST2) Hydrographic profiles (EN4.2.1) | CMIP5 | SCDA | 5 (10) | 1980–2010 | Every 2 years |
| V6w (NorCPM1) | SST (HadSST2) Hydrographic profiles (EN4.2.1) | CMIP6 | WCDA | 5 (10) | 1980–2010 | Every 1 years |
| V6s (NorCPM1) | SST (HadSST2)Hydrographic profiles (EN4.2.1) | CMIP6 | SCDA | 5 (10) | 1950–2010 | Every 1 years |

The total number of members for those versions with more than 5 is in brackets

defined for major operational forecast centers (Penny et al. 2017). Both DA approaches have two stages; the analysis stage where the DA is performed, and the forecast stage when the components of the system interact through the coupler. In NorCPM, the WCDA method only updates the other components during the forecast stage. The information exchange between ocean and the other components (sea-ice, land, atmosphere) is achieved dynamically through cross-component fluxes during the forecast stage. By contrast, in the NorCPM's SCDA implementation, the information exchange between ocean and sea ice is additionally realized by using the cross-domain error covariance, allowing ocean observations to instantaneously impact the sea ice state variables during the analysis stage (there is no update for atmosphere and land in the analysis step). Therefore, the information exchange between ocean and sea ice is achieved statistically during the analysis stage, but also dynamically during the forecast stage. Additionally to the SCDA, the version V5s also has the observation error variance of the T-S profiles (EN4.2.1) inflated in the areas with sea ice concentration higher than 50% (Bethke et al. 2021). This additional inflation was applied to deal with the sparsity of TS profiles underneath the sea ice that makes the observations there unreliable.

The implementations V6w and V6s (Bethke et al. 2021) have the same approach as V5w and V5s, but use external forcings as prescribed by the Coupled Model Intercomparison Project Phase 6 (CMIP6), while CMIP5 forcing is used for V5, V5w and V5s. Besides that, the version NorCPM-CMIP6 has re-tuning, and minor code modifications unrelated to forcing upgrades (like bug fixes). The detailed description of the code changes and their effects is in Bethke et al. (2021). The main difference between CMIP5 and CMIP6 protocols is the set of climate forcings applied such as greenhouse gases (Meinshausen et al. 2017), ozone concentrations, atmospheric aerosols (tropospheric and volcanic) (Thomason et al. 2018), and solar forcing (Matthes et al. 2016). The CMIP6 protocol is more precise than the CMIP5 protocol in the way these external forcings should be implemented, however different models may require adaptations in the forcing implementation due to their specific features and limitations (Lurton et al. 2020; Sellar et al. 2020). In NorCPM, the re-tuning due to the CMIP6 forcings's implementation included an increase of the condensation threshold for low clouds and a decrease of the snow albedo over sea ice adjusting parameters that affect snow metamorphosis (Bethke et al. 2021).

V5, V5w and V5s were initialized each second year, while V6w and V6s were initialized every year. All simulations are initialized on November 1st. Figure 1b, c show an example of the SST winter time series of the reanalysis (black curves) and the hindcast (red curves) in the SPNA and NS of the version V5s. Both areas show a positive SST

trend between 1985 and 2010 for the reanalysis and also for the hindcast (although smaller trends in the hindcast compared to the reanalysis in the NS). The simulation period of the reanalysis and hindcast for each version is shown in Fig. 2. In order to compare all versions, the only overlapping period is 1983–1999, comprising 9 initializations. To compare all hindcast versions with the same number of ensemble members, we use 5 members, which is the total ensemble number of V5w. For those versions with higher number of ensemble members we randomly selected five members. The ensemble members selected represent the variance of the ensemble mean. An example of the skill variance of V5 for all 20 members and for 5 members used in the calculation is in Fig. 3. Additionally, in the implementations with yearly initialization frequency (V6w and V6s), only every second start date was used. In this way, all versions were systematically compared with the same ensemble size and initialization frequency.

2.2 Hindcast skill and uncertainty

In this study, the predictive skill is quantified using the anomaly correlation coefficient (ACC) and root mean square error (RMSE), which are usual ways to analyze decadal climate prediction skill according to Goddard et al. (2013). The skill assessment of the hindcast for SST and SSS for each version is done against the ensemble mean of the respective NorCPM reanalysis (30 members) and HadSST2. The skill is measured using the ACC as a function of the lead year. The first lead year here is the next year after the start date, and so on. SST and SSS are averaged (without weighting) in time every 3 lead years (1–3, 2–4, 4–6, 5–7, 6–8, 7–9, and 8–10), and then averaged in space (not detrended) for the SPNA and the NS regions, as defined in Fig. 1a. After that, the ACC is calculated according to Eq. (1).

$$ACC_{LY} = \frac{\sum_{ini=1}^9 (hind_{ini} - \overline{hind}_{LY})(rean_{ini} - \overline{rean}_{LY})}{\sqrt{\sum_{ini=1}^9 (hind_{ini} - \overline{hind}_{LY})^2} \sqrt{\sum_{ini=1}^9 (rean_{ini} - \overline{rean}_{LY})^2}} \quad (1)$$

where LY is the lead year and ini denote a certain start date for the relevant lead year. $hind$ is the ensemble mean of the hindcast of a start date, and \overline{hind} means the mean of all start dates for the relevant lead year. $rean$ is the ensemble mean of the respective reanalysis of a start date, and \overline{rean} means the mean of all start dates for the correspondent lead year. The ACCs are subject to sampling uncertainty due to the small number of initializations and limited ensemble size. To account for this uncertainty we computed the 25–75% bootstrap confidence interval and we consider two ACCs statistically separated if their confidence intervals do not overlap. The significance level is calculated by the standard two-sided Students t-test (O'Mahony 1986) at 90% and

not accounting for auto and cross/covariances due to the short comparison period and, therefore, the number of initializations. The predictive skill is calculated for the winter period, defined here as January–April average. This study focus on winter because of the persistence of SST anomalies in this season, when the atmosphere–ocean coupling is most vigorous and creates a sea surface temperature anomaly that reaches the base of the deep winter mixed layer and reemerges in the following winter, contributing to the natural decadal climate variability in the North Atlantic (Alexander and Deser 1995; Watanabe and Kimoto 2000). The drift was calculated according to Choudhury et al. (2017) Method Mean Drift Correction. We calculated the drift of the ensemble mean using Eq. (2).

$$Drift_{lead} = \frac{1}{n} \sum_{i=1}^n Hind_{lead,i} - Rean_{lead,i} \quad (2)$$

where *lead* is the lead year and *n* is the number of initializations. Equation (2) gives the drift as a constant value for each lead year. There is drift in the SPNA (Fig. 1b). However, removing this drift from the time series for each lead year does not affect the correlation/ACC, as it is only subtracted a constant value. The same goes for RMSE, removing the drift from the time series before calculating RMSE does not affect the results. This is because we consider anomalies relative to their respective climatologies, when calculating RMSE, according to Goddard et al. (2013). In the other side, the drift would have an impact on the RMSE, if it was calculated the difference between hindcast and reanalysis without subtracting the respective climatology. Consider that the drift removal does not affect the metrics used in this study, all the results presented here are not drift corrected.

3 Predictive skill in different versions of NorCPM

The predictive skill of SST and SSS in different versions of NorCPM in the SPNA (black lines) and in the NS (red lines) is shown in Fig. 4. The skill is higher in the SPNA than in the NS in most versions and lead years. In the SPNA, the version V5w and V5s are the only ones with significant skill at all lead years for SST (Fig. 4 upper panel). For SSS, only V5s has values higher than or close the significance level (Fig. 4 lower panel). Skill differences between implementations and versions are smaller at shorter lead years, and become more pronounced at medium and longer lead years.

In the NS, there is no single implementation that performs higher than the significance level at all lead years neither for SST nor for SSS (red lines in Fig. 4 upper and lower panels). For SST, at shorter lead years, V5 is the only version with skill higher than the significance level, while at medium

lead years V5w is the one with highest skill. At longer lead years, V5s and V6w have the highest skill in the NS (Fig. 4 upper panel). For SSS, the versions with the highest skill are similar to the ones for SST (Fig. 4 lower panel), although the differences between versions at shorter lead years are smaller for SSS than for SST. We note that V6s presents a high anti-correlation at longer lead years for both SST (−0.8) and SSS (−0.75).

The predictive skill in the CMIP6 versions, V6w and V6s, is lower than the respective CMIP5 versions, V5w and V5s. The differences between them are more pronounced at medium and longer lead years. In the next subsections, we assess in more details the effects of the different initialization techniques on the predictive skill in the SPNA and NS.

3.1 Skill effects of assimilating subsurface data using different climatology reference periods

The assimilation of subsurface data in the ocean is important for a realistic representation of mass transport, mixed-layer depth and eddy kinetic energy; however, the subsurface ocean has only been adequately observed in the last decades, and differences in the frequency and quality of observed temperature and salinity data can result in spurious signals in the assimilated field (Yang et al. 2017). In order to deal with this problem, one approach is for the models to calculate ocean transport processes by only initializing SST. In this way, the subsurface field is initialized indirectly (Keenlyside et al. 2008); a similar approach was used in version V5. However, recent studies have shown a skill improvement in the North Atlantic when assimilating subsurface data using, for example, lagged-initialization methods (Tatebe et al. 2012; Kröger et al. 2018). To evaluate the effects of subsurface initialization on the predictive skill in NorCPM we compare versions V5 and V5w in the SPNA and in the NS. Nevertheless, differences between these versions are not limited to the type of data assimilated, since they use different climatology reference periods (Table 1). We will briefly come back to this in the discussion.

The inclusion of subsurface data increases the predictive skill and decreases the error of the hindcast for SST and SSS in the SPNA. In this area, the differences are more pronounced at medium and longer lead years (Fig. 5a, c, Supplementary Figures 15a, c). Unlike the SPNA, SST skill in the NS is higher when using only surface data assimilation at shorter lead years (Fig. 5b). At medium lead years, the addition of subsurface data does not generate statistically significant differences between the two versions. For SSS in the NS, adding subsurface data assimilation improves the predictive skill at medium lead years (Fig. 5d). The RMSE for SST in the NS is slightly higher at shorter lead years in V5w, after that the error is statistically equal for both versions (Supplementary Figure 15b). In the NS, the RMSE

for SSS is the same at short lead years between versions and slightly higher in V5w at medium and longer lead years (Supplementary Figure 15d). We note that using HadSST2 data instead of the reanalysis, we find similar results for both regions (Fig. 5a, b and Supplementary Figure 15a, b).

3.2 Skill effects of weakly vs strongly-coupled DA and ensemble inflation

Sea ice is an important component of the climate system, since it helps regulate the heat transfer between the ocean and the atmosphere with a global effect on climate scales, impacting the slow-evolving thermohaline circulation (Holland et al. 2001; Liu et al. 2019). This component was thus chosen to initiate the implementation of SCDA in NorCPM (see Kimmritz et al. 2018, for DA of sea ice concentration updating the sea ice and the ocean state). In the strongly-coupled method the observations are used to update other model components through cross-covariance error, being an alternative to improve consistency between the analyzed state of each component and eliminate initial shocks (Penny et al. 2017). In this approach, the main challenge is to deal with the differences in spatial and temporal scales between ocean, atmosphere and sea ice. The scales of ocean and sea ice are more alike than those of ocean/sea ice and atmosphere, making the jointly update of ocean-sea ice a natural “starter” for SCDA. Idealized and non-idealized studies have demonstrated advances in the strongly-coupled approach between ocean and atmosphere (Lu et al. 2015a, b; Sluka et al. 2016) and between ocean and sea ice (Kimmritz et al. 2018). Considering that the SCDA approach has recently been developed, it is important to understand whether it has a positive impact on the predictive skill, or whether it transfers biases from one component to another (Penny et al. 2017).

The evaluation of the predictive skills of V5w (WCDA) and V5s (SCDA+inflation) is shown in Fig. 6. In addition to the jointly update of the ocean and sea ice during analysis stage (SCDA), an error inflation is applied as described in Sect. 2. The effect of these implementations on the predictive skill differs depending on the area. In the SPNA, V5s shows slightly higher predictive skill and lower RMSE than V5w for both SST and SSS (Fig. 6a, c, Supplementary Figure 16a, c) at all lead times. In the NS, WCDA (V5w) has higher skill than V5s at shorter and medium lead times for SST (Fig. 6b) and at medium lead times for SSS (Fig. 6d). The RMSE of SST and SSS is statistically similar for both versions in the NS (Supplementary Figure 16b, d). Using HadSST2 data instead of the reanalysis, we find similar results for both regions (Fig. 6a, b, Supplementary Figure 16a, b).

The difference between V6w and V6s is similar to the described above for V5w and V5s. Also in NorCPM-CMIP6 versions, the SCDA with additional inflation has a positive

impact over the skill in the SPNA for almost all lead years; and for the NS, it maintains the skill at first lead years leading to a significant anti-correlation at longer lead years (Supplementary Figures 17, 18). An overall assessment of the spatial peculiarities between V5w and V5s are detailed in Sect. 4.

3.3 Skill differences between NorCPM-CMIP5 and NorCPM-CMIP6

The NorCPM-CMIP6 versions include in its underlying physical model the CMIP6 forcings, re-tuning, and minor code modifications. Thus, the differences discussed here are not regarding to the different CMIP forcings, but it is regarding to different model versions in addition to different CMIP forcings. The underlying model of NorCPM-CMIP5 versions is NorESM1-ME, while the underlying model of NorCPM-CMIP6 versions is called NorCPM1; the differences between the underlying versions are described in Table 2. To compare the predictive skill between NorCPM-CMIP5 and NorCPM-CMIP6, we here compare V5w (CMIP5) and V6w (CMIP6), since both have the same initialization approach (Table 1). For NorCPM-CMIP6 we use the same number of initializations as NorCPM-CMIP5 (every second year), as described in Sect. 2.

In the SPNA, the NorCPM-CMIP5 version has SST skill higher than the significance level for all lead years. The NorCPM-CMIP6 version has the same SST skill/RMSE as NorCPM-CMIP5 at shorter lead years (up to 3–5 years), but lower than the significance level after 5–7 years (Fig. 7a) in addition to a higher RMSE (Supplementary Figure 19a). SSS has statistically the same predictive skill in both versions up to 2–4 years. After 3–5 lead years, the skill of SSS in NorCPM-CMIP6 is lower than in NorCPM-CMIP5 (Fig. 7c) while the RMSE is statistically the same for both versions (Supplementary Figure 19c).

In the NS, the NorCPM-CMIP5 version has higher SST skill than the NorCPM-CMIP6 version at shorter lead years (up to 4–6 years); after that the skill is statistically equal, and then lower at 8–10 lead years (Fig. 7b). At all lead years the RMSE is statistically similar for

Table 2 NorCPM’s underlying models used to verify the effects of CMIP forcings on the North Atlantic skill

| Version (Hindcast) | CMIP forcing | Code modifications |
|---------------------------|--------------|--------------------|
| NorESM1-ME (NorCPM-CMIP5) | CMIP5 | NO |
| NorCPM1_1 (No Hind.) | CMIP6 | NO |
| NorCPM1 (NorCPM-CMIP6) | CMIP6 | YES |

For a better understanding of the code’s modifications refer to Bethke et al. (2021)

Fig. 2 Experiment time period for each version of NorCPM. The black line indicate the period of the reanalysis used and the respective hindcast (grey line). The vertical size of the grey lines indicate the initialization frequency of each experiment, which varies between every year (V6w and V6s) and every second year (V5, V5w, and V5s). The red lines indicate the overlapping period used for comparison

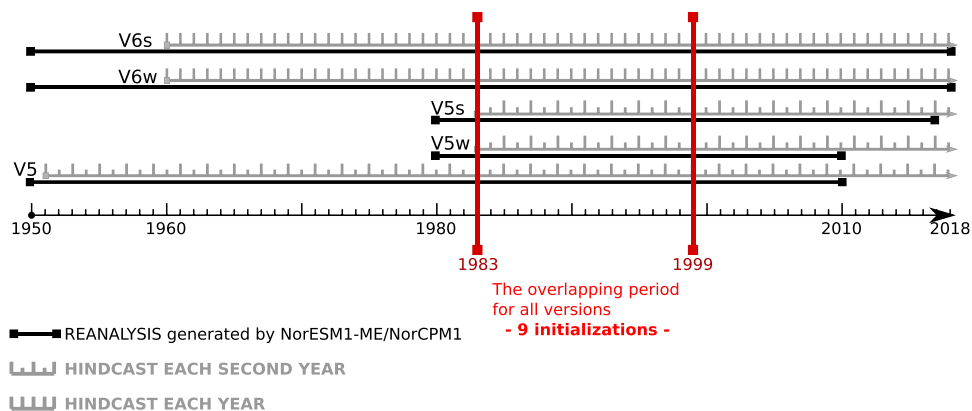
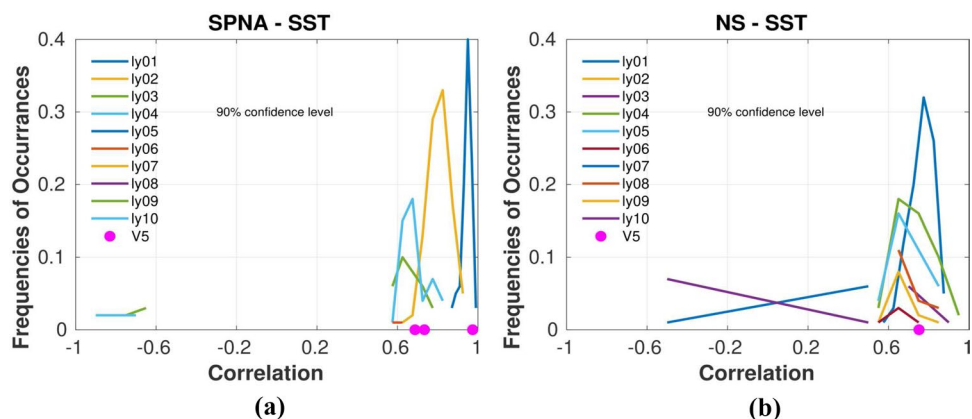


Fig. 3 Skill variance for all lead years of V5's 20 members and its 5 members ensemble mean used to calculate the results. The results are shown for 90% of confidence level. The hindcast is correlated with its own reanalysis



both versions (Supplementary Figure 19b). For SSS in the NS, the predictive skill of NorCPM-CMIP5 version is higher than NorCPM-CMIP6 at almost all lead years with the biggest differences between versions at medium lead years (Fig. 7d), it is also when the RMSE for V6w is higher than in V5w (Supplementary Figure 19d). Using HadSST2 data instead of the reanalysis, we find similar results for both regions (Fig. 7a, b, Supplementary Figure 19a, b).

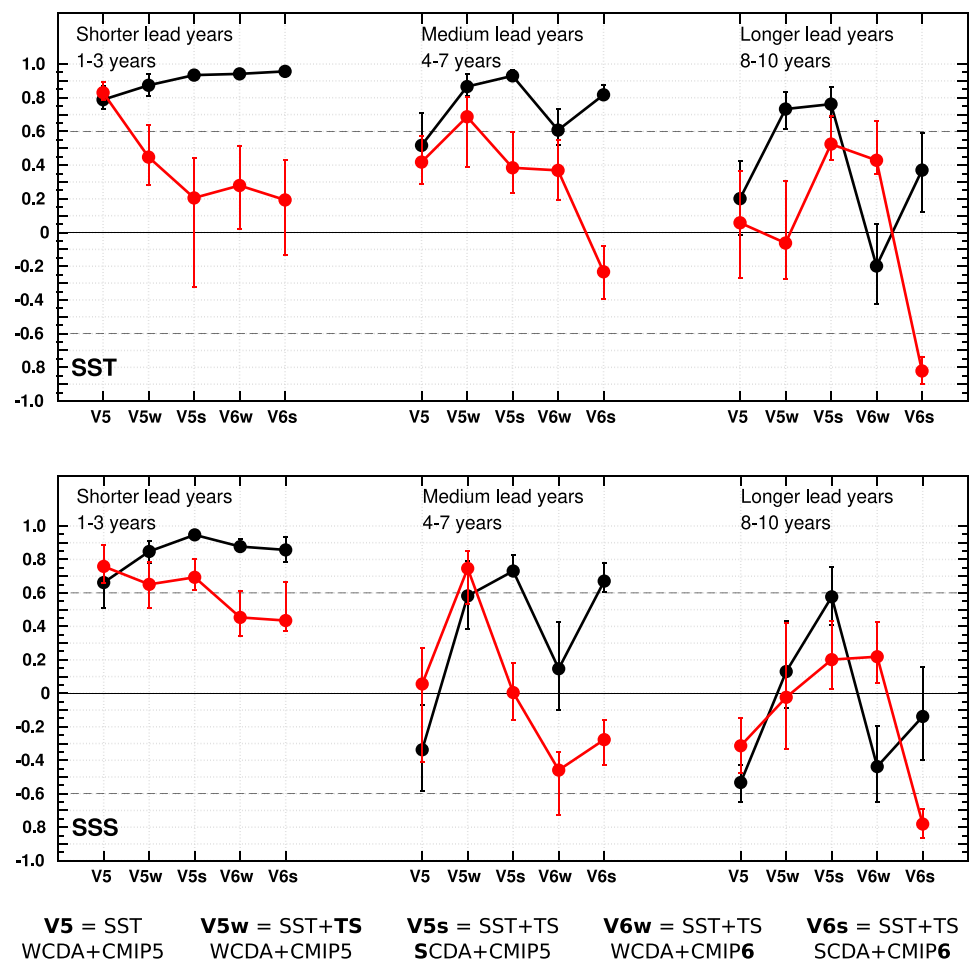
The drop of skill of the most recent V6w compared to its previous version V5w is also verified between V6s and V5s (not shown). The implementation of the new set of CMIP6 forcings caused the underlying model, NorCPM1, to need further adjustments and parametrizations, while NorCPM-CMIP5 (NorESM1-ME) had been adjusted for a longer period. It was challenging to systematically verify and reach the same level of tuning for NorCPM-CMIP6 as NorCPM-CMIP5 in time for the CMIP submission. Recently, it was verified that the land surface types and transient land-use [used in the new version] caused an unrealistic land-cryosphere cooling trend over the historical period in NorCPM1 (Bethke et al. 2021). This issue in the historical period affects the skill in both regions in different ways (Fig. 8).

3.4 CMIP6 forcing effects on historical simulations

Considering the differences between versions described in Sect. 3.3, it is not possible to analyze the effect of the CMIP6 forcings on the predictive skill of the hindcast using NorCPM-CMIP6. Nevertheless, since there is a historical simulation with CMIP6 forcing and no code updates (NorCPM1_1), we compare it with the historical simulations from the underlying models of NorCPM-CMIP5 (NorESM-ME) and NorCPM-CMIP6 (NorCPM1). The details of the historical simulations are described in Table 2.

The effect of the CMIP6 forcings on the SPNA is overall positive (compare red and black curve; Fig. 8a), increasing the models skill above the significance level. The positive effect on the skill due to the CMIP6 forcings is also found in 28 historical simulations by Borchert et al. (2021) in the SPNA. Despite of that, in the NS the effect of the new forcings is the opposite. The historical simulation from NorCPM1_1 (CMIP6) has a lower and not significant correlation compared to that from NorESM-ME (CMIP5) (compare red and black curve; Fig. 8). The lower skill could be related to the trends in SST. The historical simulation with the highest skill in the NS (NorCPM1) is the only one with a warming trend similar to that in HadSST2 in the period

Fig. 4 Anomaly correlation coefficient for winter (Jan–Apr) SST (upper panel) and SSS (lower panel) at shorter, medium and longer lead years in different versions of NorCPM in the SPNA (black) and in the NS (red). The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level. The respective anomaly correlation coefficient calculated using the linearly detrended time series are in grey for the SPNA and in light red for the NS



analyzed (Fig. 9). The result from Fig. 4 highlights that the effect of the new CMIP6 forcings on the skill is not similar for all regions.

4 Overall predictive skill between the Subpolar North Atlantic and the Norwegian Sea

According to the analysis made in Sect. 3, the version V5s (CMIP5, SCDA+error inflation) has the overall highest skill in the SPNA with significant values for almost all lead years for SST and SSS (Fig. 4). In the NS, the version V5w (CMIP5, WCDA) is the only version with significant values at medium lead years for SST and SSS (Fig. 4). Considering this, the comparison between the SPNA and the NS in this section will be evaluated based on V5s and V5w. The spatial evolution of skill with lead year for SST and SSS for both versions is shown in Figs. 10 and 11, respectively.

In the SPNA, the version V5w has a large area with skill higher than 0.6 at 1–3 lead years. Near Newfoundland Basin is the only area where the correlation at this lead year is

null or negative (Fig. 10a). This pattern remains up to 7–9 lead years (Fig. 10k), with the exception of another area with low skill that forms at lead year 2–4 near the Rockall Plateau (Fig. 10b). This is an area where branches of NAC merge and flow northward towards the NS (Danialut et al. 2016); this low skill area remains up to 5–7 lead years (Fig. 10i). At lead years 8–10, significant predictive skill in the SPNA is mainly localized in the Labrador Sea and south of Greenland (Fig. 10l). Comparing version V5w and V5s at 1–3 lead years, V5s shows higher skill in most of the SPNA with predictive skill higher than 0.8 (Fig. 10e). At 1–3 lead years there is a lower skill area near the Newfoundland Basin, and at 2–4 lead years there is also a lower skill area near the Rockall Plateau (Fig. 10e, f). However, in V5s these two areas with poor skill are smaller compared to that in V5w. This pattern remains the same up to 6–8 lead years (Fig. 10n). After that, the two areas with poor skill grow and merge, and at 8–10 lead years significant predictive skill in the SPNA is localized in the eastern part, near to Iceland, and in the Labrador Sea (Fig. 10p). Most parts of the Labrador Sea still have values higher than 0.8, which is not seen in V5w at 8–10 lead years (Fig. 10l). In both

Fig. 5 Anomaly correlation coefficient for winter (Jan–Apr) between NorCPM hindcasts V5 and V5w and the respective reanalysis and HadSST2 for SST from SPNA (a) and NS (b) and SSS from SPNA (c) and NS (d). The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level

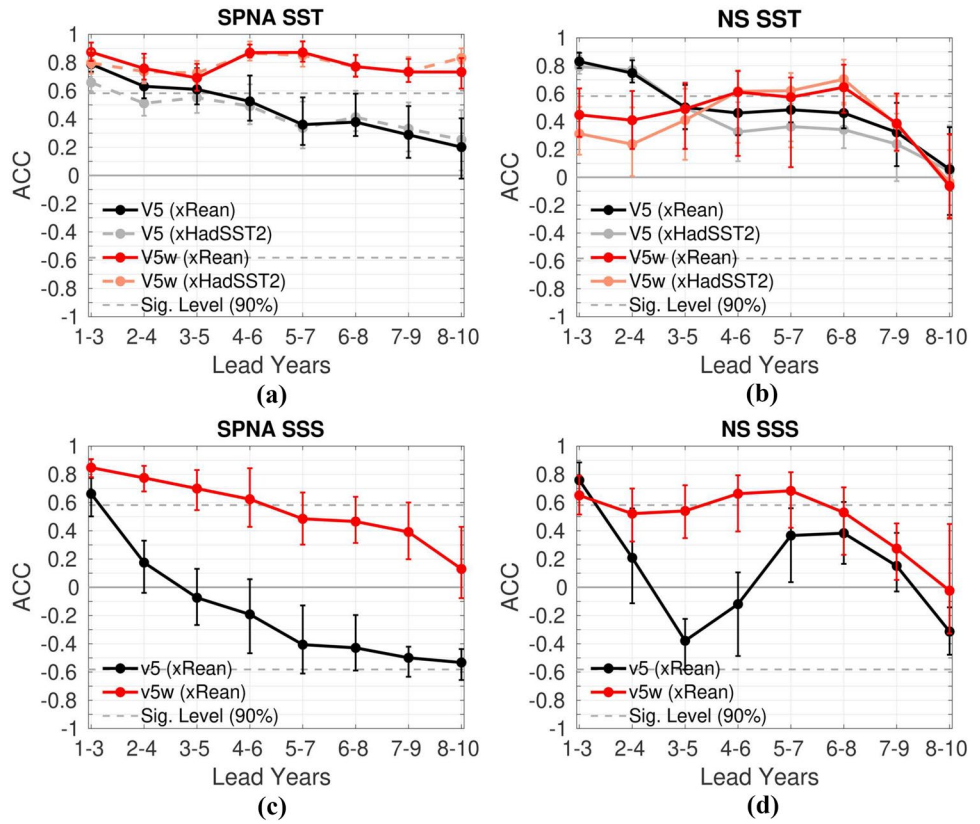


Fig. 6 Anomaly correlation coefficient between NorCPM hindcasts V5w and V5s and the respective reanalysis and HadSST2 for winter (Jan–Apr) SST from SPNA (a) and NS (b) and SSS from SPNA (c) and NS (d). The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level

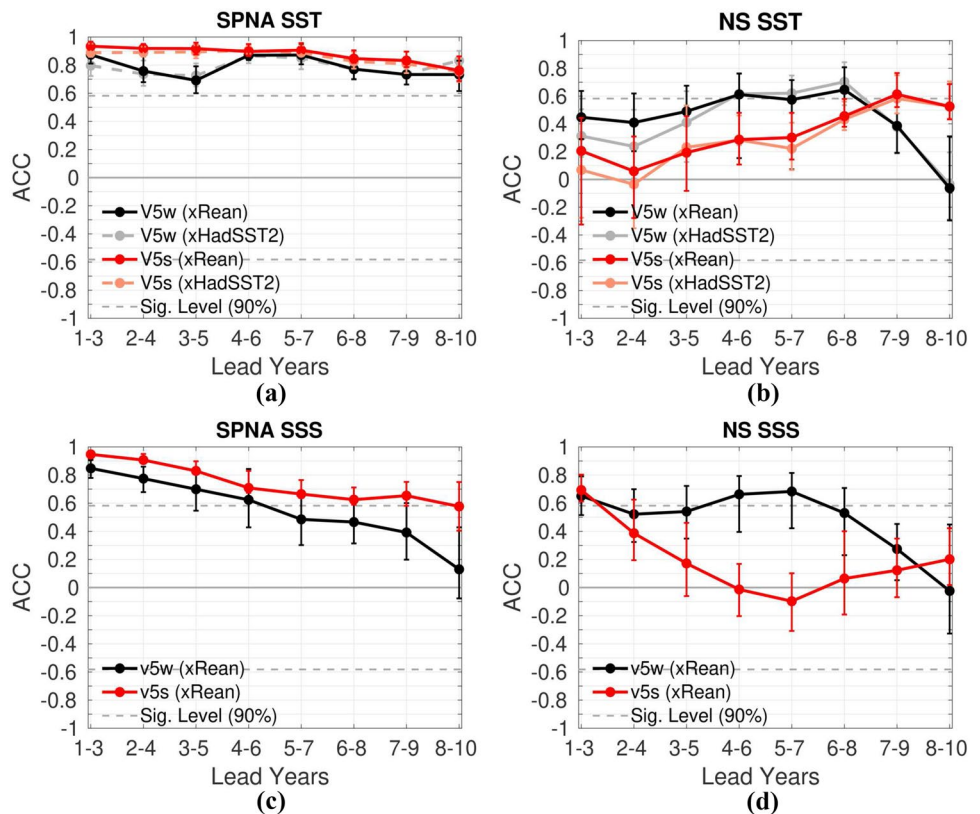


Fig. 7 Anomaly correlation coefficient between NorCPM hindcasts V5w and V6w and the respective reanalysis and HadSST2 for winter (Jan–Apr) SST from SPNA (a) and NS (b) and SSS from SPNA (c) and NS (d). The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level

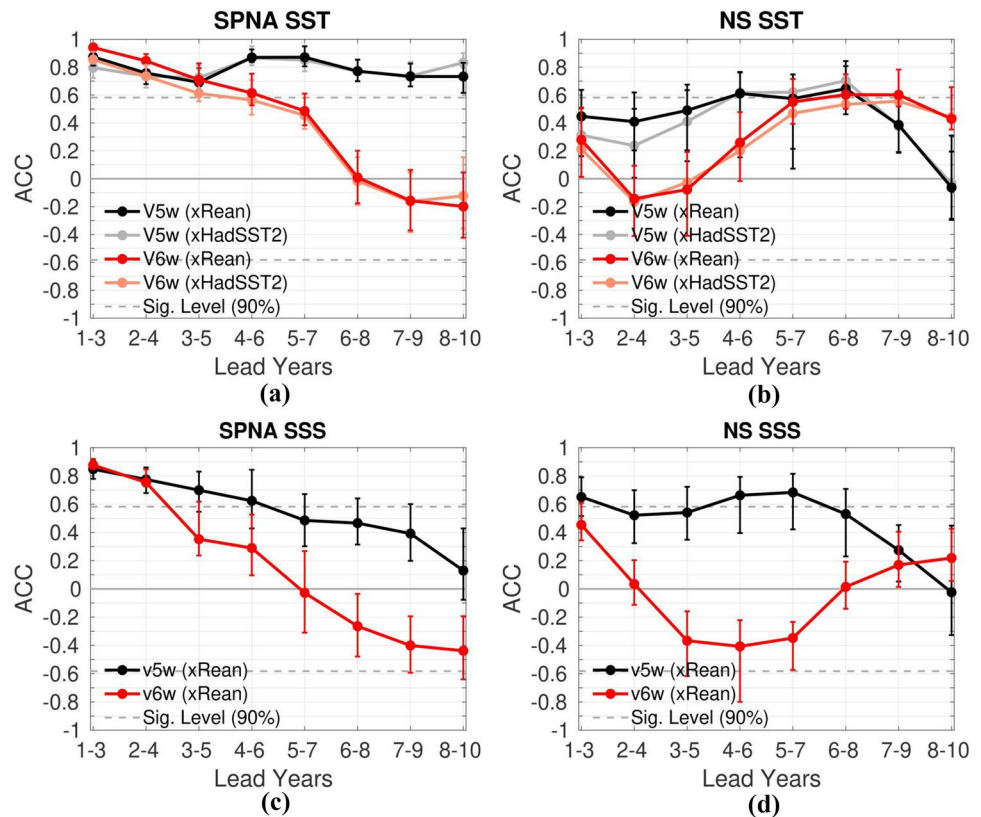
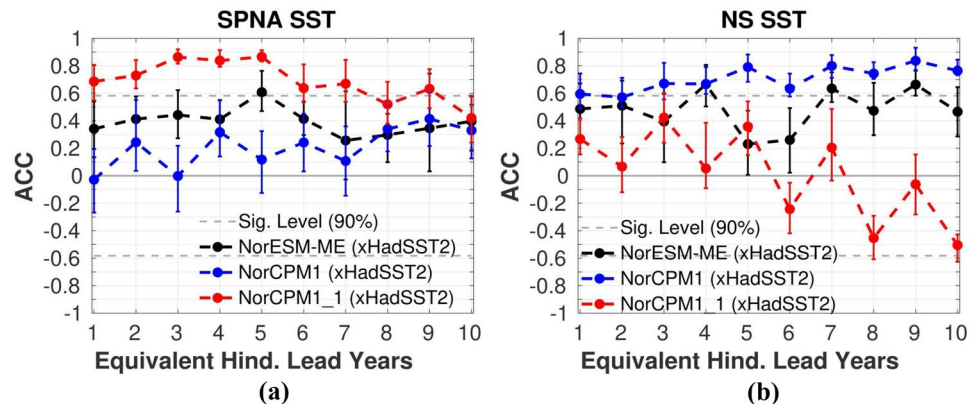


Fig. 8 Anomaly correlation coefficient between historical simulations of NorESM-ME, NorCPM1, end NorCPM1_1 and HadSST2 for winter (Jan–Apr) SST from SPNA (a) and NS (b). The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level



versions, NorCPM struggles to represent SST variability near Newfoundland Basin and Rockall Plateau.

In the NS, the version V5w has predictive skill of 0.6–0.8 at 1–3 lead years in most of the area from the Greenland–Scotland Ridge to the Knipovich Ridge (Fig. 10a). The skill remains in the area, in a narrow region close to Norway, up to 7–9 lead years (Fig. 10k). At 4–6 lead years, skill is also seen in the Barents Sea (Fig. 10d). At 5–7 lead years, an anti-correlation area forms within the Norwegian and Lofoten basins (Fig. 10i). This feature grows at the subsequent lead years becoming significant from 6 to 8 lead years (Fig. 10j). At lead years 8–10, the anti-correlation area

expands and there is no significant skill in the NS (Fig. 10l). In the V5s case, the significant predictive skill in the NS at 1–3 lead years is only seen in a narrow region close to southern Norway (Fig. 10e). At 3–5 lead years, the area with skill extends to the northern Norway (Fig. 10g), and to the Barents Sea at 4–6 lead years (Fig. 10h). However, in this version the anti-correlation forms early, at 2–4 lead years, in addition to being in a larger area comprising the Norwegian and Lofoten basins (Fig. 10f). The anti-correlation is significant from 3–5 to 6–8 lead years (Fig. 10g, h, m, n). After that, the anti-correlation is not significant and the area decreases in size (Fig. 10o, p). Unlike V5w, in V5s at 8–10

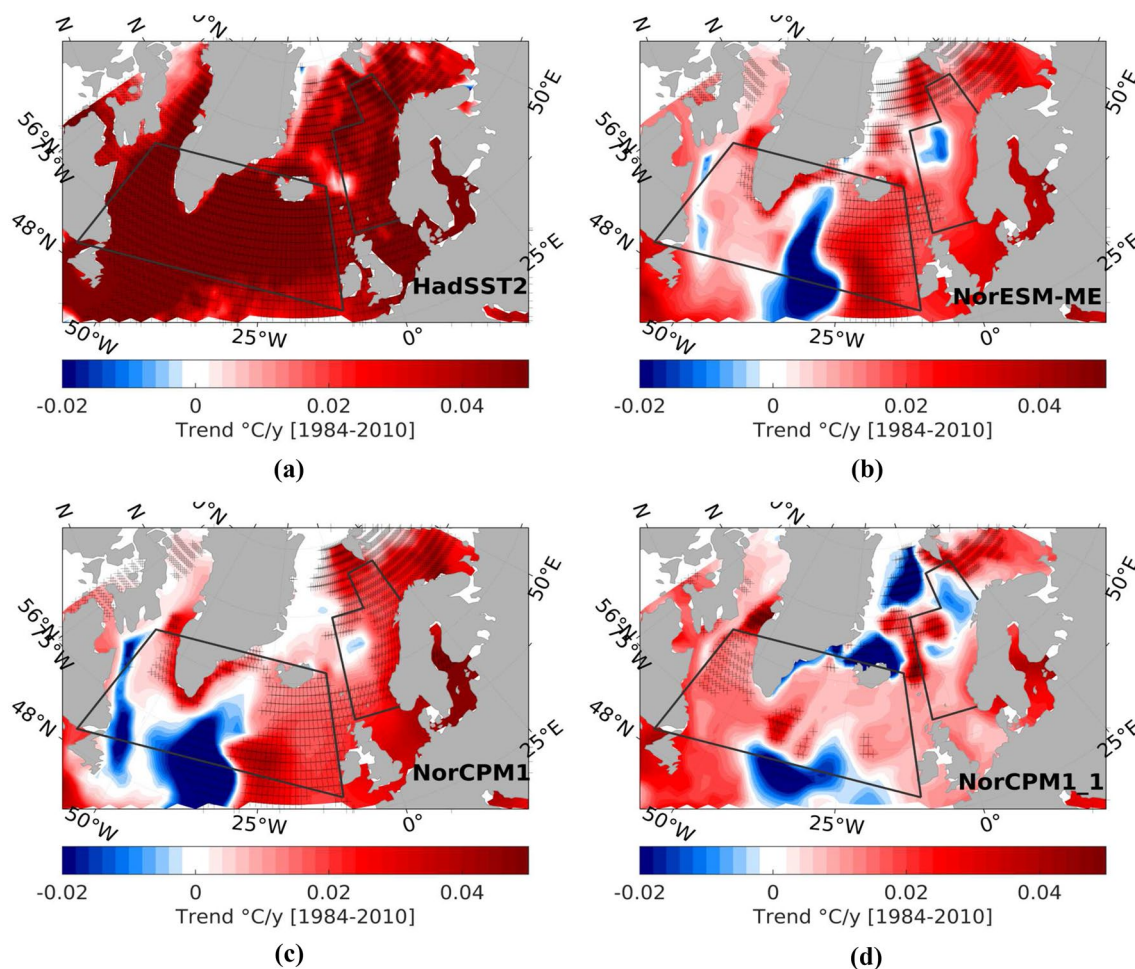


Fig. 9 Linear trend calculated for each grid point between 1984 and 2010 for HadSST2 (a), NorESM-ME (b), NorCPM1 (c), and NorCPM1_1 (d). Significant values ($p < 0.05$) are indicated by the hatched regions

lead years there is a large area in the NS, extending to the SPNA, with predictive skill of 0.6–0.8 (Fig. 10p). These are the only lead years where V5s performs better than V5w in the NS.

To better understand the differences between V5s and V5w, we have investigated the predictive skill of these two version in the very first months after the initialization (Nov–Apr; Fig. 12). In the SPNA, the skill is equal/higher for V5s than for V5w for all lead months for SST/SSS (Fig. 12a, c). However, in the NS, after the third lead month, the SST skill of V5s drops below that of V5w (Fig. 12b); for SSS the skill drops for V5s after the first lead month (Fig. 12c).

To investigate these differences in the NS more in detail, we consider a zonal section at 67.5° N, across the southern part of the Nordic Seas (map in Fig. 12). We calculate the RMSE for sea ice area, SST, and SSS for the first months after initialization (Fig. 13). Overall, V5s has higher RMSE than V5w. Right after the initialization, an error starts to develop in V5s for

both the sea ice area and SST between 20° W and 10° W, and synchronously increases with time (Fig. 13). The analysis of ACC shows a skill decrease in V5s for both sea ice area and SST between 30° W and 20° W (not shown).

In view of the above comparison, the SCDA and the error inflation (version V5s) have a positive impact on the predictive skill in the SPNA, especially in the Labrador Sea. However, this version has a negative effect on the predictive skill in the NS, specifically in the Norwegian and Lofoten basins, where a strong anti-correlation area is formed and develops early in the hindcast. Furthermore, the skill in the Barents Sea also degrades in this version.

5 Discussion

In this study, we have assessed the sensitivity of decadal predictive skill in the Subpolar North Atlantic (SPNA) and in the Norwegian Sea (NS) to the use of different types of

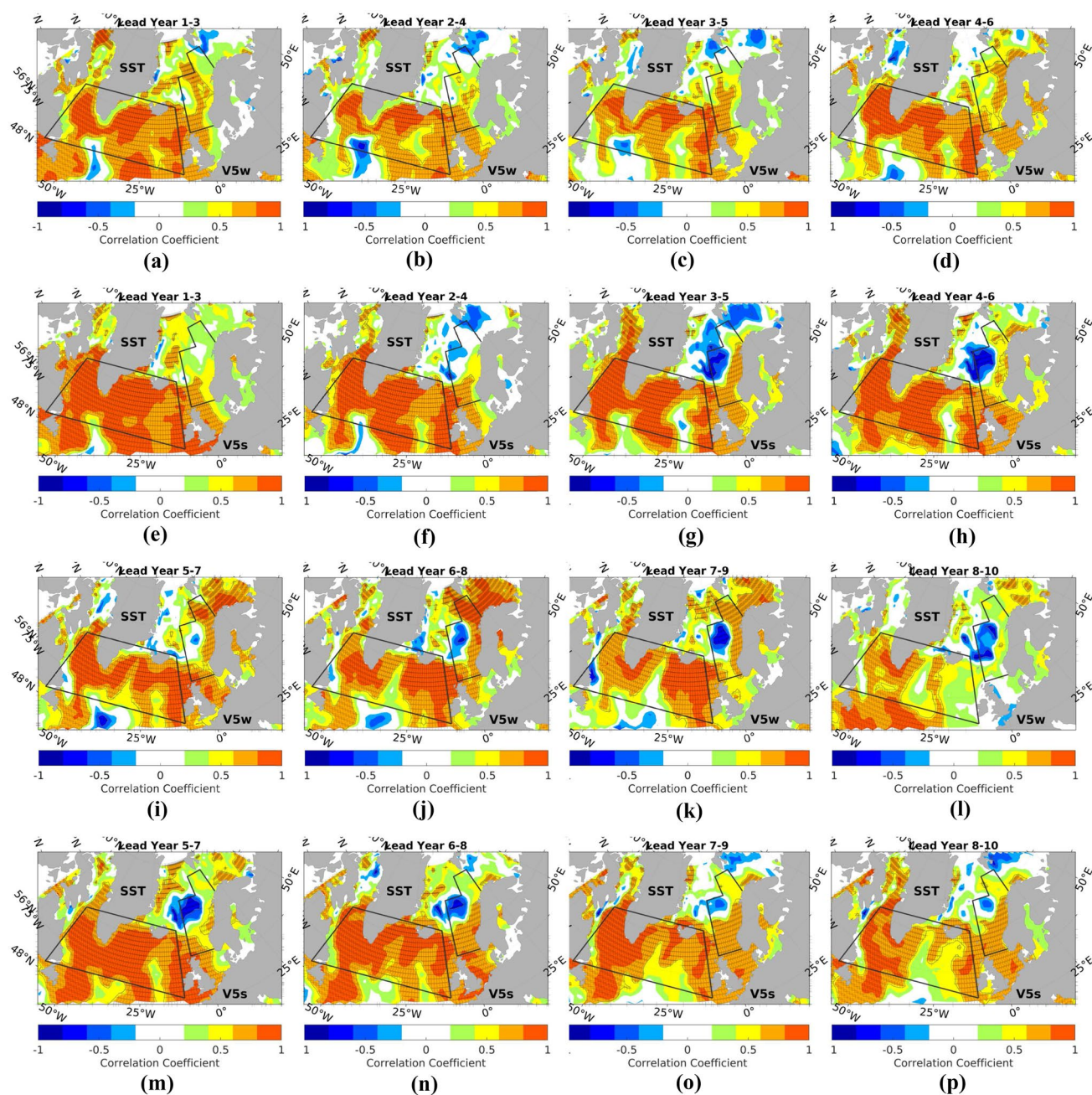


Fig. 10 Spatial anomaly correlation coefficient of SST in each grid point for V5w: 1–3 years (a), 2–4 years (b), 3–5 years (c), 4–6 years (d), 5–7 years (i), 6–8 years (j), 7–9 years (k), 8–10 years (l). For V5s: 1–3 years (e), 2–4 years (f), 3–5 years (g) and 4–6 years (h), 5–7

years (m), 6–8 years (n), 7–9 years (o), 8–10 years (p). The SPNA and NS regions are shown by the black boxes. Significant values are indicated by the hatched regions

data assimilation techniques in NorCPM. The effect of the external forcings does not influence the comparison between the different versions, however, gives overall less skill compared to the results including the linear trends (Fig. 14 upper panel, lower panel).

The SPNA is one of the areas with highest predictive skill at decadal time scales in climate prediction systems (Yeager and Robson 2017). It is also an important source region of

predictability to the NS (Årthun et al. 2017; Langehaug et al. 2017), and the correct representation of mechanisms underlying decadal variability is thus important for climate predictability in the Arctic–Atlantic region.

The predictability in the SPNA and in the NS has been associated with specific physical mechanisms that act on different time scales. The Atlantic Meridional Overturning Circulation (AMOC) is pointed out as a key source

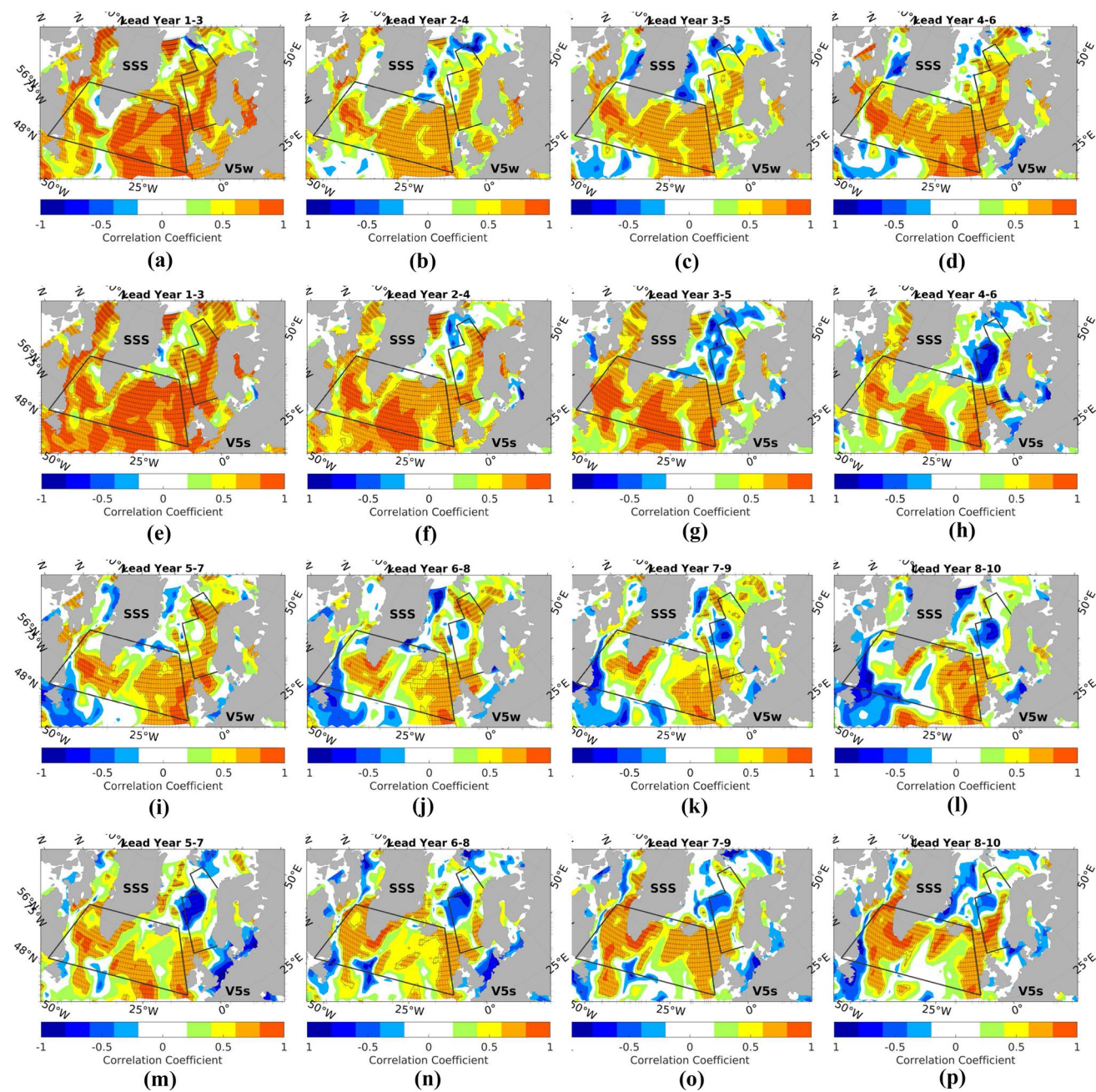


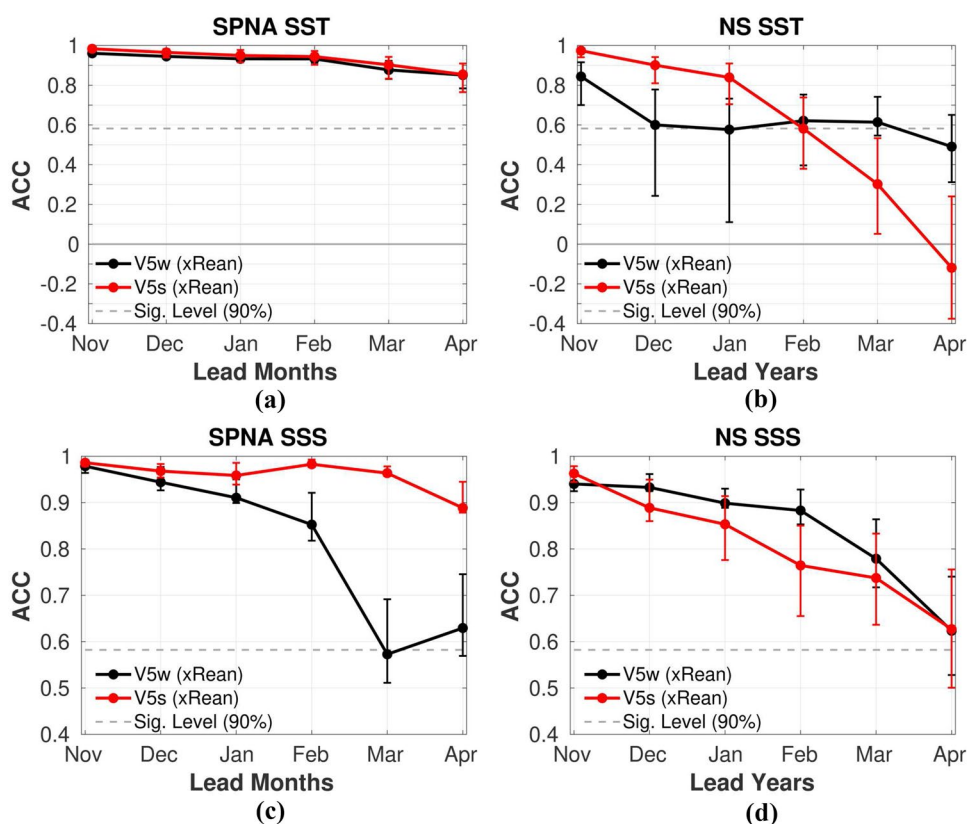
Fig. 11 Spatial anomaly correlation coefficient of SSS in each grid point for V5w: 1–3 years (a), 2–4 years (b), 3–5 years (c), 4–6 years (d), 5–7 years (i), 6–8 years (j), 7–9 years (k), 8–10 years (l). For V5s: 1–3 years (e), 2–4 years (f), 3–5 years (g) and 4–6 years (h), 5–7

years (m), 6–8 years (n), 7–9 years (o), 8–10 years (p). The SPNA and NS regions are shown by the black boxes. Significant values are indicated by the hatched regions

of decadal to multidecadal climate predictability in the SPNA. However, recently the Labrador Sea Water thickness anomalies have been indicated as a precursor of upper ocean heat content predictability in this area, as well as an important driver influencing the southward ocean circulation (Ortega et al. 2020; Yeager 2020). The Labrador Sea Water is formed by convective processes driven by high oceanic heat losses during winter. The intense heat

loss is associated with sea ice formation (Yeager 2020) and reduce the density stratification in the Labrador Sea. Dense Labrador Sea Water is carried southward as part of the deep branch of AMOC. Thus, the representation of sea ice plays an important role in the predictive skill of the SPNA. On a shorter time scale, SST variability with a period of 13–18 years has been found to dominate in the

Fig. 12 Anomaly correlation coefficient between NorCPM hindcasts V5w and V5s and the respective reanalysis for the first 6 months after the initialization (Nov–April) for SST from SPNA (a) and NS (b) and for SSS from SPNA (c) and NS (d). The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level



North Atlantic, and is suggested to contribute to recent cold anomaly in the SPNA (Årthun et al. 2021).

The pronounced decadal variability in the SPNA is linked to northward propagation of thermohaline anomalies and largely contributes to predictability in the NS (Årthun et al. 2017). The Greenland-Scotland Ridge (GSR) is where the warm and salty waters of the North Atlantic Current (NAC) enters the NS (Fig. 1). The representation of the flow across the GSR is thus important since it is the gateway of thermohaline anomalies coming from the SPNA. Heuzé and Årthun (2019) suggest, based on analysis of 23 CMIP5 models, that the model resolution and bathymetry in the GSR is a key factor for the oceanic heat transport from the North Atlantic to the Arctic. In addition, analysis from observations and ocean state estimate show that unrealistic eddy fluxes, and air-sea heat fluxes in the Norwegian Sea can limit the predictability carried by the poleward heat anomalies (Chafik et al. 2015; Asbjørnsen et al. 2019). The air-sea fluxes in this area are considerable; the Lofoten basin alone is responsible for 1/3 of the heat loss in the Nordic Seas, despite only covering 1/5 of its total area (Richards and Straneo 2015).

The NorCPM study herein shows that the increase of complexity in the initialization method in the CMIP5 versions (V5–V5s) results in a general skill increase for the SPNA, whereas the same is not achieved in the NS. Independent of the initialization approach used, there is a large

skill difference between the SPNA and NS, suggesting that there are difficulties in representing the physical processes in the NS or predictability might be more limited in this region compared to the SPNA. In NorCPM, one of the physical processes taking place in the SPNA, AMOC, is in relatively good agreement with observations (V5; Counillon et al. 2016). On the other hand, in the NS the propagation of SST anomalies is not properly represented by the model (V6w; Langehaug et al. 2021), and the surface currents are rather broader in the SPNA and NS (NorESM1-M; the underlying model of NorCPM; Langehaug et al. 2018). A similar result is found in the system MPI-ESM-LR; the use of oceanic EnKF improves the predictive skill in the SPNA, especially in the Labrador and Irminger Seas, while decreases the skill in the NS compared to the historical run (Brune and Baehr 2020). The spatial ACC maps in Fig. 10 show significant SST skill in a narrow region close to Norway, but in the frontal region where Atlantic Water meets Arctic Water the skill is largely reduced. The ACC maps showed that SST skill decreases in the Norwegian and Lofoten basins. These areas are dominated by intense eddy activity (Raj et al. 2019) and surface heat fluxes (Richards and Straneo 2015) and the non-significant SST skill can indicate the struggling of the prediction system to properly represent these processes. This will be further discussed below, when comparing the versions V5w and V5s.

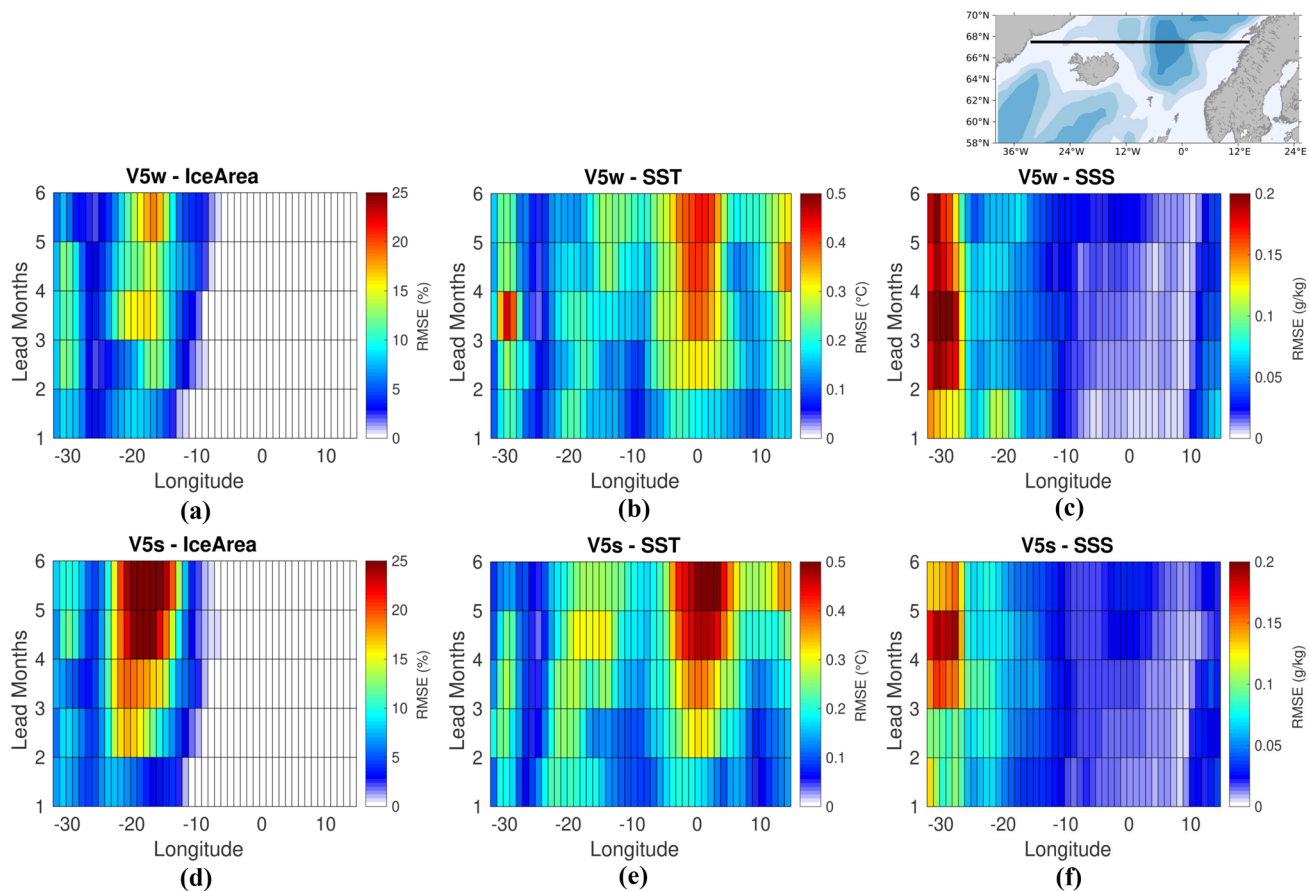


Fig. 13 RMSE calculated for the section North of Iceland (67.5° N). Longitude in the axis X and lead months after initialization in the axis Y. The calculation was made for versions V5w sea ice area (a),

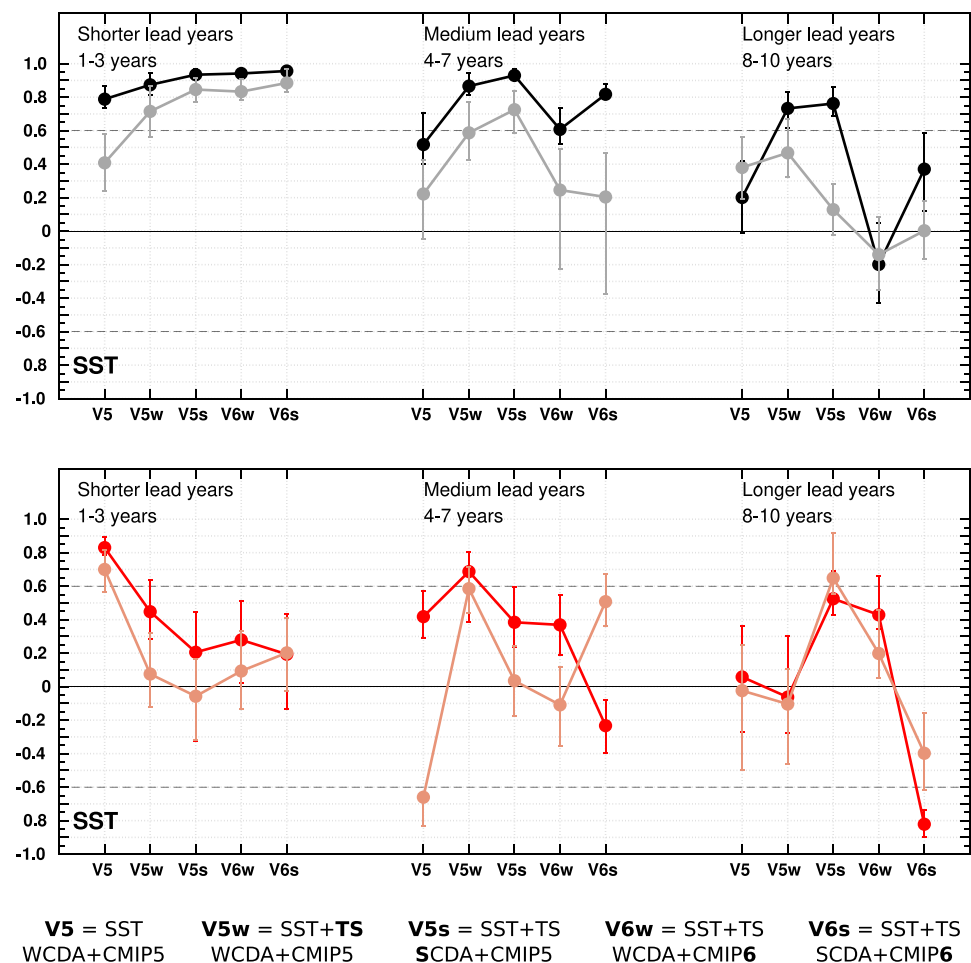
SST (b), SSS (c) and for V5s sea ice area (d), SST (e), SSS (f). In the right upper panel is the section position

The addition of subsurface data with the use of a different climatology reference period in the initialization method showed a general positive impact on the predictive skill in both regions. In the SPNA, this initialization method generated a considerable improvement of the skill in the surface ocean that lasted at medium and longer lead years, while for the NS the gain of skill was only related to SSS. Considering both ACC maps (Figs. 10, 11) and RMSE maps (Supplementary Figures 20, 21, 22), there is low or no skill in similar regions for SST and SSS: in the Newfoundland Basin, and in the Norwegian and Lofoten basins. The reason for gain in skill only for SSS in the NS could be due to the fact that SST is more influenced by local surface forcing than the SSS in this area (Asbjørnsen et al. 2019; Furevik et al. 2002). In addition to the inclusion of subsurface data, the different climatology reference periods used might also have an influence on the skill. The sparsity and quality of the observed data before the 1980s can increase the uncertainties in the mean climate calculation, which in turn can affect the uncertainties of anomalies used in the data assimilation. Unfortunately, with the available implementations it was not

possible to separate the effects of subsurface data assimilation and the different climatology reference periods.

Furthermore, we have looked at the impact of a jointly update of the ocean and sea ice (SCDA) and additional error inflation (comparison of versions V5w and V5s). When the sea-ice is corrected by the covariation with the ocean temperature, the surface ocean skill is slightly higher in the SPNA, while in the NS such increase is not verified. The reanalysis of V5s is the only one with strongly reduced bias in Arctic sea ice thickness, which grow back over lead years causing a strong drift in the hindcast (Bethke et al. 2021). The main differences between V5w (WCDA) and V5s (SCDA) in the NS are: significant SST skill in the V5s is only found in the southwest area at 1–3 lead years compared to a larger area in V5w that extends towards the Knipowich Ridge (10a); a significant anti-correlation area over the Norwegian and Lofoten basins appears earlier in V5s compared to V5w, at 2–4 lead years in the former (Fig. 10g). The anti-correlation area is also seen in the NorCPM-CMIP6 versions, being stronger in V6s than in V6w (Supplementary Figures 17, 18), and it is also seen in the historical run of

Fig. 14 Anomaly correlation coefficient for winter (Jan–Apr) SST in the SPNA (upper panel-black) and SST in the NS (lower panel-red) at shorter, medium and longer lead years in different versions of NorCPM. The vertical bars indicate the 25 and 75 percentiles and the dashed grey lines show the 90% significance level. The respective anomaly correlation coefficient calculated using the linearly detrended time series are in the upper panel (grey) for the SPNA and in lower panel (light red) for the NS



the respective model version (figure not shown). The anti-correlation area is also associated with high SST RMSE (Supplementary Figures 20, 21) extending from the Jan Mayen to the Mohn–Knipovich Ridges, the occurrence area of the Arctic Front. The Arctic Front is localized near the sea ice edge and is characterized by the interaction of the colder and fresher Arctic Water with the warmer and more saline Atlantic Water (Swift and Aagard 1981). Along the front, observations show active air-sea interaction (Raj et al. 2019); while high-resolution numerical simulations verified cross-ridge exchange on Mohn–Knipovich Ridge leading to a cooling and freshening of the Norwegian Atlantic Current (Ypma et al. 2020). It is uncertain, however, how the Arctic Front is represented in NorCPM, and is beyond the scope of this study. However, the results herein shows that this region and associated processes are difficult to represent in a coarse climate model.

In the SPNA the SCDA version, V5s, increased the skill and decreased the RMSE for both SST and SSS. The SST RMSE reduction is seen in the Labrador Sea and over the MAR (Supplementary Figures 20e–h, m–o), while the SSS RMSE decrease happens in the Labrador Sea and on the

Reykjanes Ridge, at 2–4 and 3–5 lead years (Supplementary Figures 22f, g, respectively). Despite the improvement in V5s version, in both V5s and V5w the model struggles to represent SSS over the MAR. Results from another dynamical prediction system, CESM-DPLE, show that the MAR plays an important role in the SPNA predictability, inducing a strong coupling between the AMOC and the SPNA circulation through the propagation of deep water mass anomalies (Yeager 2020). However, this mechanism has not been assessed in the present study.

Initialized decadal hindcasts from CMIP6 improved the representation of SST in the SPNA compared to CMIP5 according to Borchert et al. (2021). This study analyzed 6(7) prediction systems from CMIP5 (CMIP6) and showed that 88% of the SPNA SST variance at lead years 5–7 is explained by CMIP6 hindcasts, while CMIP5 explains only 42%. Borchert et al. (2021) attribute this difference to a good representation of the SPNA SST in CMIP6 historical simulations due to an increase of ensemble size, as well as to a high predictive skill in CMIP6 after the end of the CMIP5 period. In our study, ensemble size, number of initializations, and period analyzed were the same for all versions.

In NorCPM-CMIP6 the drop of skill is related to a wrong surface land temperature trend, in particular over the Canadian Arctic, that affects the atmospheric state and circulation over the North Atlantic and also the Nordic Seas (Bethke et al. 2021). Thus, the comparison here between versions NorCPM-CMIP5 and NorCPM-CMIP6 includes changes in the code and in the forcings, which makes it difficult to evaluate their isolated effects on the predictive skill.

Intercomparisons between prediction systems or versions can be an important approach to better understand how well some physical processes are represented in the model and might help to identify weaknesses and indicate potential directions for further development. To investigate the predictive skill in the Nordic Seas, Langehaug et al. (2017) analyzed the results of three prediction systems. The authors suggested a relationship between resolution and SST skill in the analyzed systems, in addition to highlighting that unrealistic sea ice cover could be a relevant bias in the Nordic Seas. Three initialization approaches were tested by Polkova et al. (2014), and the best results for sea surface height in the North Atlantic were verified in the full state initialization with heat and freshwater flux corrections. On the other hand, Polkova et al. (2019a) found improvements in some parts of the North Atlantic for surface temperatures and upper ocean heat content with a combination of ensemble Kalman filter and filtered anomaly initialization. Based on this, it is not possible to define an optimal technique for all areas and variables even using the same prediction system (Höschel et al. 2019), which is analogue to the results of this work.

6 Summary and conclusions

The goal of this study was to assess and quantify the impact of different initialization strategies on the development of the Norwegian Climate Prediction Model (NorCPM). The investigation was focused on the Subpolar North Atlantic (SPNA) and the Norwegian Sea (NS). The predictive skill was tested against NorCPM's own reanalysis, and for SST also against HadSST2.

The comparison among versions showed that the choice of Data Assimilation (DA) method appears to have largest impact on medium to long lead years. In the SPNA, increasing initialization complexity resulted in a general skill increase within NorCPM-CMIP5 versions, however, the same was not found for the NS. The additional assimilation of subsurface data in addition to a different climatology reference period gave a general skill improvement in both areas (Sect. 3.1). An improvement was found for both SST and SSS in the SPNA, while in the NS the skill for SST was maintained at medium and longer lead years and improved for SSS between 2–4 and 5–7 lead years.

The joint update of the ocean and sea ice (Strongly Coupled DA) with additional inflation error increased surface ocean skill in the SPNA and decreased it in the NS. This version presented an area with highly negative ACC over the Norwegian and Lofoten basins associated with high RMSE between the Jan Mayen and Mohn–Knipovich Ridges, showing that this version struggles in particular to represent ocean processes in these areas in the NS. The comparison between NorCPM-CMIP5 and NorCPM-CMIP6 showed a reduced ocean skill in the latter due to erroneous land surface conditions that affect the atmospheric state over the North Atlantic and Nordic Seas.

In general, the comparison of skill in this study shows that the NorCPM-CMIP5 version with the Strongly Coupled DA has the highest skill in the SPNA and without it (Weakly Coupled DA) for the NS. Furthermore, the results from this study show that despite the NS being directly influenced by thermohaline anomalies coming from the SPNA, the circulation of the Atlantic Water in the Norwegian and Lofoten basins and its interaction with colder and fresher waters over the Jan Mayen and Mohn–Knipovich Ridges are aspects that need to be improved in NorCPM. A better representation of the complex bathymetry close to the Greenland–Scotland Ridge and in the NS along with an eddy permitting grid might be a way to improve NorCPM results in the NS.

The NS is the path by which thermohaline anomalies coming from the North Atlantic influence the Arctic sea ice cover (Onarheim et al. 2015), cod stock (Årthun et al. 2018a) and the surface air temperature over Scandinavia (Årthun et al. 2017). The identification of the most suitable DA approach for the Arctic–Atlantic region contributes to the development of NorCPM. Identifying the main challenges to predictive skill in the NS furthermore contributes to a better representation of key processes not just in NorCPM, but also in other prediction systems. The results show that areas with processes as intense surface heat fluxes and eddy activity within the Norwegian–Lofoten Basins and in the area of the Arctic Front, might be key areas to improve NS skill. The investigation of the NS skill in other prediction systems might help to guide development for improvement in NorCPM.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00382-022-06437-4>.

Acknowledgements We thank François Counillon and Yiguo Wang for the help and clarifications about NorCPM versions. We thank Joan Mateu Horrach Pou for important discussions about NorCPM-CMIP6 versions. We acknowledge UNINETT Sigma2 (nn9039k, ns9039k) by providing the Computing and storage resources.

Author contributions LP coordinate the writing of this article. LP, HL and MÅ worked in the conceptualization of the article. LP, HL and IB contributed processing the results. TE, IB and MK contributed with

the results evaluation. All co-authors contributed to the writing of the article.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital). LP, HL, MÅ, TE, IB and MK received funding from the Trond Mohn Foundation through the Bjerknes Climate Prediction Unit (BFS2018TMT01). HRL, MÅ, and TE has also received funding from the Blue-Action Project (European Unions Horizon 2020 research and innovation program, Grant 727852).

Data availability The observation-based SST used HadISST2 is available at <https://www.metoffice.gov.uk/hadobs/hadsst2/>. NorCPM-CMIP5 reanalysis and hindcast and NorCPM-CMIP6 reanalysis are available upon request to the modeling group. The CMIP6-DCPP (V6w and V6s) output can be accessed at <https://doi.org/10.22033/ESGF/CMIP6.10844>

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexander MA, Deser C (1995) A mechanism for the recurrence of wintertime midlatitude SST anomalies. *J Phys Oceanogr* 25:122–137
- Årthun M, Eldevik T (2016) On anomalous ocean heat transport toward the Arctic and associated climate predictability. *J Clim* 29(2):689–704. <https://doi.org/10.1175/JCLI-D-15-0448.1>
- Årthun M, Eldevik T, Smedsrud LH, Skagseth Ingvaldsen RB (2012) Quantifying the influence of Atlantic heat on Barents sea ice variability and retreat. *J Clim* 25(13):4736–4743. <https://doi.org/10.1175/JCLI-D-11-00466.1>
- Årthun M, Eldevik T, Viste E, Drange H, Furevik T, Johnson HL, Keenlyside NS (2017) Skillful prediction of northern climate provided by the ocean. *Nat Commun*. <https://doi.org/10.1038/ncomms15875>
- Årthun M, Bogstad B, Daewel U, Keenlyside NS, Sandø AB, Schrum C, Ottersen G (2018) Climate based multi-year predictions of the Barents Sea cod stock. *PLoS One* 13(10):1–13. <https://doi.org/10.1371/journal.pone.0206319>
- Årthun M, Kolstad EW, Eldevik T, Keenlyside NS (2018) Time scales and sources of European temperature variability. *Geophys Res Lett* 45(8):3597–3604. <https://doi.org/10.1002/2018GL077401>
- Årthun M, Wills RC, Johnson HL, Chafik L, Langehaug HR (2021) Mechanisms of decadal North Atlantic climate variability and implications for the recent cold anomaly. *J Clim* 34(9):3421–3439. <https://doi.org/10.1175/JCLI-D-20-0464.1>
- Asbjørnsen H, Årthun M, Skagseth Eldevik T (2019) Mechanisms of ocean heat anomalies in the Norwegian Sea. *J Geophys Res Oceans*. <https://doi.org/10.1029/2018JC014649>
- Bentsen M, Bethke I, Debernard JB, Iversen T, Kirkevåg A, Selander H, Roelandt C, Seierstad IA, Hoose C, Kristjánsson JE (2013) The Norwegian earth system model, NorESM1-M Part 1: description and basic evaluation of the physical climate. *Geosci Model Dev* 6(3):687–720. <https://doi.org/10.5194/gmd-6-687-2013>
- Bethke I, Wang Y, Counillon F, Keenlyside N, Kimmritz M, Fransner F, Samuelsen A, Langehaug H, Svendsen L, Chiu PG, Leilane Goncalves dos Passos MB, Guo C, Tjiputra J, Kirkevåg A, Olivier D, Selander, Vågene JS, Fan Y, Lawrence P (2021) NorCPM1 and its contribution to CMIP6 DCPP
- Borchert LF, Menary MB, Swingedouw D, Sgubin G, Hermanson L, Mignot J (2021) Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6. *Geophys Res Lett* 48(3):1–10. <https://doi.org/10.1029/2020GL091307>
- Brune S, Baehr J (2020) Preserving the coupled atmosphere-ocean feedback in initializations of decadal climate predictions. *Wiley Interdiscip Rev Clim Change* 11(3):1–19. <https://doi.org/10.1002/wcc.637>
- Buckley MW, DelSole T, Susan Lozier M, Li L (2019) Predictability of North Atlantic Sea surface temperature and upper-ocean heat content. *J Clim* 32(10):3005–3023. <https://doi.org/10.1175/JCLI-D-18-0509.1>
- Chafik L, Nilsson J, Skagseth Lundberg P (2015) On the flow of Atlantic water and temperature anomalies in the Nordic Seas toward the Arctic Ocean. *J Geophys Res Oceans*. <https://doi.org/10.1002/2014JC010472>. Received
- Choudhury D, Sen Gupta A, Sharma A, Mehrotra R, Sivakumar B (2017) An assessment of drift correction alternatives for CMIP5 decadal predictions. *J Geophys Res Atmos* 122(19):10282–10296. <https://doi.org/10.1002/2017JD026900>
- Counillon F, Bethke I, Keenlyside N, Bentsen M, Bertino I, Zheng F (2014) Seasonal-to-decadal predictions with the ensemble kalman filter and the Norwegian earth system model: a twin experiment. *Tellus Ser A Dyn Meteorol Oceanogr* 66:1. <https://doi.org/10.3402/tellusa.v66.21074>
- Counillon F, Keenlyside N, Bethke I, Wang Y, Billeau S, Shen ML, Bentsen M (2016) Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model. *Tellus Ser A Dyn Meteorol Oceanogr*. <https://doi.org/10.3402/tellusa.v68.32437>
- Daniault N, Mercier H, Lherminier P, Sarafanov A, Falina A, Zunino P, Pérez FF, Ríos AF, Ferron B, Huck T, Thierry V, Gladyshev S (2016) The northern North Atlantic Ocean mean circulation in the early 21st century. *Prog Oceanogr* 146(June):142–158. <https://doi.org/10.1016/j.pocean.2016.06.007>
- Eldevik T, Nilsen JE, Iovino D, Anders Olsson K, Sandø AB, Drange H (2009) Observed sources and variability of Nordic seas overflow. *Nat Geosci* 2(6):406–410. <https://doi.org/10.1038/ngeo518>
- Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn* 53(4):343–367. <https://doi.org/10.1007/s10236-003-0036-9>
- Furevik T, Bentsen M, Drange H, Johannessen JA, Korabely A (2002) Temporal and spatial variability of the sea surface salinity in the Nordic Seas. *J Geophys Res Oceans* 107(12):1–16. <https://doi.org/10.1029/2001jc001118>
- Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, Worley PH, Yang ZL, Zhang M (2011) The community climate system model version 4. *J Clim* 24(19):4973–4991. <https://doi.org/10.1175/2011JCLI4083.1>
- Goddard L, Kumar A, Solomon A, Smith D, Boer G, Gonzalez P, Kharin V, Merryfield W, Deser C, Mason SJ, Kirtman BP,

- Msadek R, Sutton R, Hawkins E, Fricker T, Hegerl G, Ferro CA, Stephenson DB, Meehl GA, Stockdale T, Burgman R, Greene AM, Kushnir Y, Newman M, Carton J, Fukumori I, Delworth T (2013) A verification framework for interannual-to-decadal predictions experiments. *Clim Dyn* 40(1–2):245–272. <https://doi.org/10.1007/s00382-012-1481-2>
- Good SA, Martin MJ, Rayner NA (2013) EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J Geophys Res Oceans* 118(12):6704–6716. <https://doi.org/10.1002/2013JC009067>
- Heuzé C, Arthun M (2019) The Atlantic inflow across the Greenland–Scotland ridge in global climate models (CMIP5). *Elementa* 7:1. <https://doi.org/10.1525/elementa.354>
- Holland MM, Bitz CM, Eby M, Weaver AJ (2001) The role of ice-ocean interactions in the variability of the North Atlantic thermohaline circulation. *J Clim* 14(5):656–675. [https://doi.org/10.1175/1520-0442\(2001\)014<0656:TROI0I>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0656:TROI0I>2.0.CO;2)
- Höschel I, Illing S, Grieger J, Ulbrich U, Cubasch U (2019) On skillful decadal predictions of the subpolar North Atlantic. *Meteorol Z* 28(5):417–428. <https://doi.org/10.1127/metz/2019/0957>
- Keenlyside NS, Latif M, Jungclauss J, Kornblüeh L, Roeckner E (2008) Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature* 453(7191):84–88. <https://doi.org/10.1038/nature06921>
- Kimmritz M, Counillon F, Bitz CM, Massonnet F, Bethke I, Gao Y (2018) Optimising assimilation of sea ice concentration in an Earth system model with a multicategory sea ice model. *Tellus Ser A Dyn Meteorol Oceanogr* 70(1):1–23. <https://doi.org/10.1080/16000870.2018.1435945>
- Kröger J, Pohlmann H, Sienz F, Marotzke J, Baehr J, Köhl A, Modali K, Polkova I, Stammer D, Vamborg FS, Müller WA (2018) Full-field initialized decadal predictions with the MPI earth system model: an initial shock in the North Atlantic. *Clim Dyn* 51(7–8):2593–2608. <https://doi.org/10.1007/s00382-017-4030-1>
- Kushnir Y, Scaife AA, Arriitt R, Balsamo G, Boer G, Doblas-Reyes F, Hawkins E, Kimoto M, Kolli RK, Kumar A, Matei D, Matthes K, Müller WA, O’Kane T, Perlwitz J, Power S, Raphael M, Shimpo A, Smith D, Tuma M, Wu B (2019) Towards operational predictions of the near-term climate. *Nat Clim Change* 9(2):94–101. <https://doi.org/10.1038/s41558-018-0359-7>
- Langehaug HR, Matei D, Eldevik T, Lohmann K, Gao Y (2017) On model differences and skill in predicting sea surface temperature in the Nordic and Barents Seas. *Clim Dyn* 48(3–4):913–933. <https://doi.org/10.1007/s00382-016-3118-3>
- Langehaug HR, Sandø AB, Arthun M, Ilcak M (2018) Variability along the Atlantic water pathway in the forced Norwegian Earth System Model. *Clim Dyn* 52(1–2):1211–1230. <https://doi.org/10.1007/s00382-018-4184-5>
- Langehaug HR, Ortega P, Counillon F, Matei D, Maroon E, Keenlyside N, Mignot J, Wang Y, Swingedouw D, Bethke I, Yang S, Danabasoglu G, Bellucci A, Ruggieri P (2021) Propagation of thermohaline anomalies and their predictive potential along the Atlantic water pathway. *J Clim* 20:1–60. <https://doi.org/10.1175/jcli-d-20-1007.1>
- Liu W, Fedorov A, Sévellec F (2019) The mechanisms of the Atlantic meridional overturning circulation slowdown induced by Arctic sea ice decline. *J Clim* 32(4):977–996. <https://doi.org/10.1175/JCLI-D-18-0231.1>
- Lu F, Liu Z, Zhang S, Liu Y (2015) Strongly coupled data assimilation using leading averaged covariance (LACC). Part I: Simple model study. *Mon Weather Rev* 143(9):3823–3837. <https://doi.org/10.1175/MWR-D-14-00322.1>
- Lu F, Liu Z, Zhang S, Liu Y, Jacob R (2015) Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part II: CGCM experiments. *Mon Weather Rev* 143(11):4645–4659. <https://doi.org/10.1175/MWR-D-15-0088.1>
- Lurton T, Balkanski Y, Bastrikov V, Bekki S, Bopp L, Braconnot P, Brockmann P, Cadule P, Contoux C, Cozic A, Cugnet D, Dufresne JL, Éthé C, Foujols MA, Ghattas J, Hauglustaine D, Hu RM, Kageyama M, Khodri M, Lebas N, Levavasseur G, Marchand M, Ottlé C, Peylin P, Sima A, Szopa S, Thiéblemont R, Vuichard N, Boucher O (2020) Implementation of the CMIP6 forcing data in the IPSL-CM6A-LR model. *J Adv Model Earth Syst* 12(4):1–22. <https://doi.org/10.1029/2019MS001940>
- Matei D, Pohlmann H, Jungclauss J, Müller W, Haak H, Marotzke J (2012) Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. *J Clim* 25(24):8502–8523. <https://doi.org/10.1175/JCLI-D-11-00633.1>
- Matthes K, Funke B, Anderson ME, Barnard L, Beer J, Charbonneau P, Cilivert MA, Dudok de Wit T, Haberleiter M, Hendry A, Jackman CH, Kretschmar M, Kruschke T, Kunze M, Langematz U, Marsh DR, Maycock A, Misios S, Rodger CJ, Scaife AA, Seppälä A, Shangguan M, Sinnhuber M, Tourpali K, Usoskin I, van de Kamp M, Verronen PT, Versick S (2016) Solar forcing for CMIP6 (v3.1). geoscientific model development discussions 6(June):1–82. <https://doi.org/10.5194/gmd-2016-91>
- Meehl GA, Goddard L, Murphy J, Stouffer RJ, Boer GJ, Danabasoglu G, Dixon K, Giorgetta MA, Greene AM, Hawkins E, Hegerl G, Karoly D, Keenlyside N, Kimoto M, Kirtman B, Navarra A, Pulwarty R, Smith D, Stammer D, Stockdale T (2009) Decadal prediction can it be skillful? *Bull Am Meteorol Soc* 90(10):1467–1486. <https://doi.org/10.1175/2009BAMS2778.1>
- Meinshausen M, Vogel E, Nauels A, Lorbacher K, Meinshausen N, Etheridge DM, Fraser PJ, Montzka SA, Rayner PJ, Trudinger CM, Krummel PB, Beyerle U, Canadell JG, Daniel JS, Enting IG, Law RM, Lunder CR, O’Doherty S, Prinn RG, Reimann S, Rubino M, Velders GJ, Vollmer MK, Wang RH, Weiss R (2017) Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geosci Model Dev* 10(5):2057–2116. <https://doi.org/10.5194/gmd-10-2057-2017>
- Morioka Y, Doi T, Storto A, Masina S, Behera SK (2018) Role of subsurface ocean in decadal climate predictability over the South Atlantic. *Sci Rep* 8(1):8523. <https://doi.org/10.1038/s41598-018-26899-z>
- O’Mahony M (1986) Sensory evaluation of food: statistical methods and procedures. Food science and technology, Marcel Dekker Inc, New York. <https://doi.org/10.1177/001088048402400413>
- Onarheim IH, Eldevik T, Arthun M, Ingvaldsen RB, Smedsrud LH (2015) Skillful prediction of Barents Sea ice cover. *Geophys Res Lett* 42(13):5364–5371. <https://doi.org/10.1002/2015GL064359>
- Ortega P, Robson J, Menary M, Sutton R, Blaker A, Germe A, Hirschi J, Sinha B, Hermanson L, Yeager S (2020) Labrador Sea subsurface density as a precursor of multi-decadal variability in the North Atlantic: a multi-model study. *Earth Syst Dyn Discuss*. <https://doi.org/10.5194/esd-2020-83>
- Penny SG, Akella S, Alves O, Bishop C, Buehner M, Chevallier M, Counillon F (2017) Coupled data assimilation for integrated earth system analysis and prediction. *Bull Am Meteorol Soc* 98(7):ES169–ES172. <https://doi.org/10.1175/BAMS-D-17-0036.1>
- Penny SG, Bach E, Bhargava K, Chang CC, Da C, Sun L, Yoshida T (2019) Strongly coupled data assimilation in multiscale media: experiments using a quasi-geostrophic coupled model. *J Adv Model Earth Syst* 11(6):1803–1829. <https://doi.org/10.1029/2019MS001652>
- Polkova I, Köhl A, Stammer D (2014) Impact of initialization procedures on the predictive skill of a coupled ocean-atmosphere model. *Clim Dyn* 42(11–12):3151–3169. <https://doi.org/10.1007/s00382-013-1969-4>
- Polkova I, Brune S, Kadow C, Romanova V, Gollan G, Baehr J, Glowienka-Hense R, Greatbatch RJ, Hense A, Illing S, Köhl A, Kröger

- J, Müller WA, Pankatz K, Stammer D (2019) Initialization and ensemble generation for decadal climate predictions: a comparison of different methods. *J Adv Model Earth Syst* 11(1):149–172. <https://doi.org/10.1029/2018MS001439>
- Polkova I, Köhl A, Stammer D (2019) Climate-mode initialization for decadal climate predictions. *Clim Dyn* 53(11):7097–7111. <https://doi.org/10.1007/s00382-019-04975-y>
- Raj RP, Chatterjee S, Bertino L, Turiel A, Portebella M (2019) The Arctic Front and its variability in the Norwegian Sea. The Arctic Front and its variability in the Norwegian Sea, pp 1–22. <https://doi.org/10.5194/os-2018-159>
- Rayner NA, Brohan P, Parker DE, Folland CK, Kennedy JJ, Vanicek M, Ansell TJ, Tett SF (2006) Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. *J Clim* 19(3):446–469. <https://doi.org/10.1175/JCLI3637.1>
- Richards CG, Straneo F (2015) Observations of water mass transformation and eddies in the Lofoten basin of the Nordic seas. *J Phys Oceanogr* 45(6):1735–1756. <https://doi.org/10.1175/JPO-D-14-0238.1>
- Robson JI, Sutton RT, Smith DM (2012) Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s. *Geophys Res Lett* 39(19):1–6. <https://doi.org/10.1029/2012GL053370>
- Sellar AA, Walton J, Jones CG, Wood R, Abraham NL, Andrejczuk M, Andrews MB, Andrews T, Archibald AT, de Mora L, Dyson H, Elkington M, Ellis R, Florek P, Good P, Gohar L, Haddad S, Hardiman SC, Hogan E, Iwi A, Jones CD, Johnson B, Kelley DI, Kettleborough J, Knight JR, Köhler MO, Kuhlbrodt T, Liddicoat S, Linova-Pavlova I, Mizielinski MS, Morgenstern O, Mulcahy J, Neininger E, O'Connor FM, Petrie R, Ridley J, Rioual JC, Roberts M, Robertson E, Rumbold S, Seddon J, Shepherd H, Shim S, Stephens A, Teixiera JC, Tang Y, Williams J, Wiltshire A, Griffiths PT (2020) Implementation of U.K. earth system models for CMIP6. *J Adv Model Earth Syst* 12(4):1–27. <https://doi.org/10.1029/2019MS001946>
- Sluka TC, Penny SG, Kalnay E, Miyoshi T (2016) Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophys Res Lett* 43(2):752–759. <https://doi.org/10.1002/2015GL067238>
- Smith DM, Cusack S, Colman AW, Folland CK, Harris GR, Murphy JM (2007) Improved surface temperature prediction for the coming decade from a global climate model. *Science* 317(5839):796–799. <https://doi.org/10.1126/science.1139540>
- Smith DM, Eade R, Scaife AA, Caron LP, Danabasoglu G, DelSole TM, Delworth T, Doblas-Reyes FJ, Dunstone NJ, Hermanson L, Kharin V, Kimoto M, Merryfield WJ, Mochizuki T, Müller WA, Pohlmann H, Yeager S, Yang X (2019) Robust skill of decadal climate predictions. *NPJ Clim Atmos Sci* 2(1):1–10. <https://doi.org/10.1038/s41612-019-0071-y>
- Smith DM, Scaife AA, Eade R, Athanasiadis P, Bellucci A, Bethke I, Bilbao R, Borchert LF, Caron L, Counillon F, Danabasoglu G, Delworth T, Dunstone NJ, Flavoni S, Hermanson L, Keenlyside N, Kharin V, Kimoto M, Merryfield WJ, Mignot J, Mochizuki T, Modali K, Monerie P, Müller WA, Nicolí D, Ortega P, Pankatz K, Pohlmann H, Robson J, Ruggieri P, Swingedouw D, Wang Y, Wild S, Yeager S, Yang X, Zhang L (2020) North Atlantic climate far more predictable than models imply. *Nature* 583:20. <https://doi.org/10.1038/s41586-020-2525-0>
- Swift JH, Aagard K (1981) Seasonal transitions and water mass formation in the Iceland and Greenland seas. *Deep-Sea Res* 28A(10):1107–1129
- Tatebe H, Ishii M, Mochizuki T, Chikamoto Y, Sakamoto TT, Komuro Y, Mori M, Yasunaka S, Watanabe M, Ogochi K, Suzuki T, Nishimura T, Kimoto M (2012) The initialization of the MIROC climate models with hydrographic data assimilation for decadal prediction. *J Meteorol Soc Jpn* 90(A):275–294. <https://doi.org/10.2151/jmsj.2012-A14>
- Thomason LW, Ernest N, Millán L, Rieger L, Bourassa A, Vernier JP, Manney G, Luo B, Arfeuille F, Peter T (2018) A global space-based stratospheric aerosol climatology: 1979–2016. *Earth Syst Sci Data* 10(1):469–492. <https://doi.org/10.5194/essd-10-469-2018>
- Tjiputra JF, Roelandt C, Bentsen M, Lawrence DM, Lorentzen T, Schwinger J, Seland Heinze C (2013) Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). *Geosci Model Dev* 6(2):301–325. <https://doi.org/10.5194/gmd-6-301-2013>
- van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2012) Decadal prediction skill in a multi-model ensemble. *Clim Dyn* 38(7–8):1263–1280. <https://doi.org/10.1007/s00382-012-1313-4>
- Vera C, Barange M, Dube OP, Goddard L, Griggs D, Kobysheva N, Odada E, Parey S, Polovina J, Poveda G, Seguin B, Trenberth K (2010) Needs assessment for climate information on decadal timescales and longer. *Proced Environ Sci* 1(1):275–286. <https://doi.org/10.1016/j.proenv.2010.09.017>
- Vertenstein M, Craig T, Middleton A, Feddema D, Fischer C (2012) CESM1.0.4 Users Guide. Tech. rep., National Center for Atmospheric Research (NCAR). http://www.cesm.ucar.edu/models/cesm1.0/cesm/cesm_doc_1_0_4/book1.html
- Wang Y, Counillon F, Bethke I, Keenlyside N, Bocquet M, Mi Shen (2017) Optimising assimilation of hydrographic profiles into isopycnal ocean models with ensemble data assimilation. *Ocean Model* 114:33–44. <https://doi.org/10.1016/j.ocemod.2017.04.007>
- Watanabe M, Kimoto M (2000) On the persistence of decadal SST anomalies in the North Atlantic. *J Clim* 13(16):3017–3028. [https://doi.org/10.1175/1520-0442\(2000\)013<3017:OTPOD>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<3017:OTPOD>2.0.CO;2)
- Yang C, Masina S, Storto A (2017) Historical ocean reanalyses (1900–2010) using different data assimilation strategies. *Q J R Meteorol Soc* 143(702):479–493. <https://doi.org/10.1002/qj.2936>
- Yeager S (2020) The abyssal origins of North Atlantic decadal predictability. *Clim Dyn*. <https://doi.org/10.1007/s00382-020-05382-4>
- Yeager SG, Robson JI (2017) Recent progress in understanding and predicting Atlantic decadal climate variability. *Curr Clim Change Rep*. <https://doi.org/10.1007/s40641-017-0064-z>
- Yeager S, Karspeck A, Danabasoglu G, Tribbia J, Teng H (2012) A decadal prediction case study: late twentieth-century North Atlantic Ocean heat content. *J Clim* 25:5173–5189. <https://doi.org/10.1175/JCLI-D-11-00595.1>
- Ypma SL, Georgiou S, Dugstad JS, Pietrzak JD, Katsman CA (2020) Pathways and water mass transformation along and across the Mohn-Knipovich ridge in the Nordic Seas. *J Geophys Res Oceans* 125:9. <https://doi.org/10.1029/2020JC016075>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.