Impact of ocean data assimilation on climate predictions with ICON-ESM

Holger Pohlmann[1] · Sebastian Brune[2] · Kristina Fröhlich[3] · Johann H. Jungclaus[4] · Christine Sgoff[3] · Johanna Baehr[2]

Abstract

We develop a data assimilation scheme with the Icosahedral Non-hydrostatic Earth System Model (ICON-ESM) for operational decadal and seasonal climate predictions at the German weather service. For this purpose, we implement an Ensemble Kalman Filter to the ocean component as a first step towards a weakly coupled data assimilation. We performed an assimilation experiment over the period 1960–2014. This ocean-only assimilation experiment serves to initialize 10-year long retrospective predictions (hindcasts) started each year on 1 November. On multi-annual time scales, we find predictability of sea surface temperature and salinity as well as oceanic heat and salt contents especially in the North Atlantic. The mean Atlantic Meridional Overturning Circulation is realistic and the variability is stable during the assimilation. On seasonal time scales, we find high predictive skill in the tropics with highest values in variables related to the El Niño/Southern Oscillation phenomenon. In the Arctic, the hindcasts correctly represent the decreasing sea ice trend in winter and, to a lesser degree, also in summer, although sea ice concentration is generally much too low in both hemispheres in summer. However, compared to other prediction systems, prediction skill is relatively low in regions apart from the tropical Pacific due to the missing atmospheric assimilation. Further improvements of the simulated mean state of ICON-ESM, e.g. through fine-tuning of the sea ice and the oceanic circulation in the Southern Ocean, are expected to improve the predictive skill. In general, we demonstrate that our data assimilation method is successfully initializing the oceanic component of the climate system.

Keywords Ocean data assimilation · Seasonal climate prediction · Decadal climate prediction

1 Introduction

The German Meteorological Service "Deutscher Wetterdienst" (DWD) plans to update its currently used climate model for operational seasonal and decadal climate predictions from the Max-Planck-Institute Earth System Model (MPI-ESM) to the Icosahedral Non-hydrostatic Model (ICON; Zängl et al. 2015). ICON is a joint development between DWD, Max Planck Institute for Meteorology (MPI-M), the Karlsruhe Institute for Technology (KIT), the German Climate Computing Center (DKRZ) and other institutions in Germany and Switzerland. The ICON-Earth System Model (ICON-ESM; Jungclaus et al. 2022) has become available recently. We develop—as a first step towards a weakly coupled assimilation—a data assimilation scheme for the oceanic component.

Decadal climate prediction (Smith et al. 2019) is a relatively new field and research activities are supported by the Decadal Climate Prediction Project (DCPP; Boer et al. 2016; Merryfield et al. 2020) of the World Climate Research Program (WCRP), which is contributing to the Coupled Model Intercomparison Project (CMIP) phase 5 (CMIP5; Taylor et al. 2012) and phase 6 (CMIP6; Boer et al. 2016). Following an initiative of WCRP's Grand Challenge on Near Term Climate Prediction (Kushnir et al. 2019), decadal climate predictions are coordinated by the Lead Centre for Annual-to-Decadal Climate Prediction (LC ADCP) of the World Meteorological Organization (WMO). About a dozen global producing centers and other contributing centers publish

✉ Holger Pohlmann
   holger.pohlmann@dwd.de

1   Deutscher Wetterdienst, Hamburg, Germany

2   Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany

3   Deutscher Wetterdienst, Offenbach, Germany

4   Max Planck Institute for Meteorology, Hamburg, Germany

decadal climate predictions in this framework (Hermanson et al. 2022).

Seasonal climate predictions, on the other hand, are well established. The WMO Lead Centre for Long-Range Forecast Multi Model Ensemble (LC LRFMME) and the WCRP's Climate System Historical Forecast Project (CHFP; Tompkins et al. 2017) organize seasonal climate prediction activities. Another platform for coordinated seasonal predictions is established by the European Union's COPERNICUS program (https://climate.copernicus.eu). The Working Group on Seasonal to Interannual Prediction (WGSIP) has changed its name recently to the Working Group on Subseasonal to Interdecadal Prediction (WGSIP) to combine the efforts on the two time-scales.

Most of the seasonal and decadal climate prediction systems nowadays use a weakly coupled data assimilation, i.e. data assimilation is applied to each component of the climate system separately. Examples for ocean-only initializations are from the beginning of decadal climate predictions (e.g., Keenlyside et al. 2008; Pohlmann et al. 2009; Dunstone 2010; Swingedouw et al. 2013) and also more recent seasonal predictions (Zhu et al. 2017; Wang et al. 2019; Mu et al. 2020). Strongly coupled data assimilation, i.e. a common data assimilation for all climate components, might reduce the imbalances between the components caused by the weakly coupled data assimilation further (Penny et al. 2019; Tang et al. 2021). Eventually, we are planning a strongly coupled seasonal to decadal data assimilation system. The assimilation of the oceanic component with ICON-ESM is our first step to approach this aim.

A multi-initialization comparison (Polkova et al. 2019) with our previous climate model MPI-ESM has shown best results while using an oceanic implementation of the Ensemble Kalman Filter (EnKF) method with the Parallel Data Assimilation Framework (PDAF; Nerger and Hiller 2013). We transferred this setup from MPI-ESM to ICON-ESM, where we assimilate temperature and salinity profiles from observations over the past ca. 55 years to obtain initial conditions for the decadal hindcast simulations. Evidently, this model-consistent initialization strategy avoids some of the problems emerging from the previously used relaxation to reanalysis products, where in fact two different (ocean) models are combined (Brune et al. 2018), e.g. as has been the case in DWD's previous operational seasonal (Fröhlich et al. 2021) and decadal (Pohlmann et al. 2019) prediction systems, which were initialized from an oceanic reanalysis product.

Climate predictions may benefit from preserving large parts of the atmosphere–ocean feedback, a task that is not easy to sustain (Brune and Baehr 2020). While initializing only its oceanic part, we want to analyze how far the predictions with ICON-ESM benefit from this feedback and where they may need additional initialization input, e. g. from the atmospheric part. We organize this paper as follows: In chapter 2, we introduce the experiments and methods and we present our analysis in chapter 3. In chapter 4, we conclude with a summary and discussion of our results.
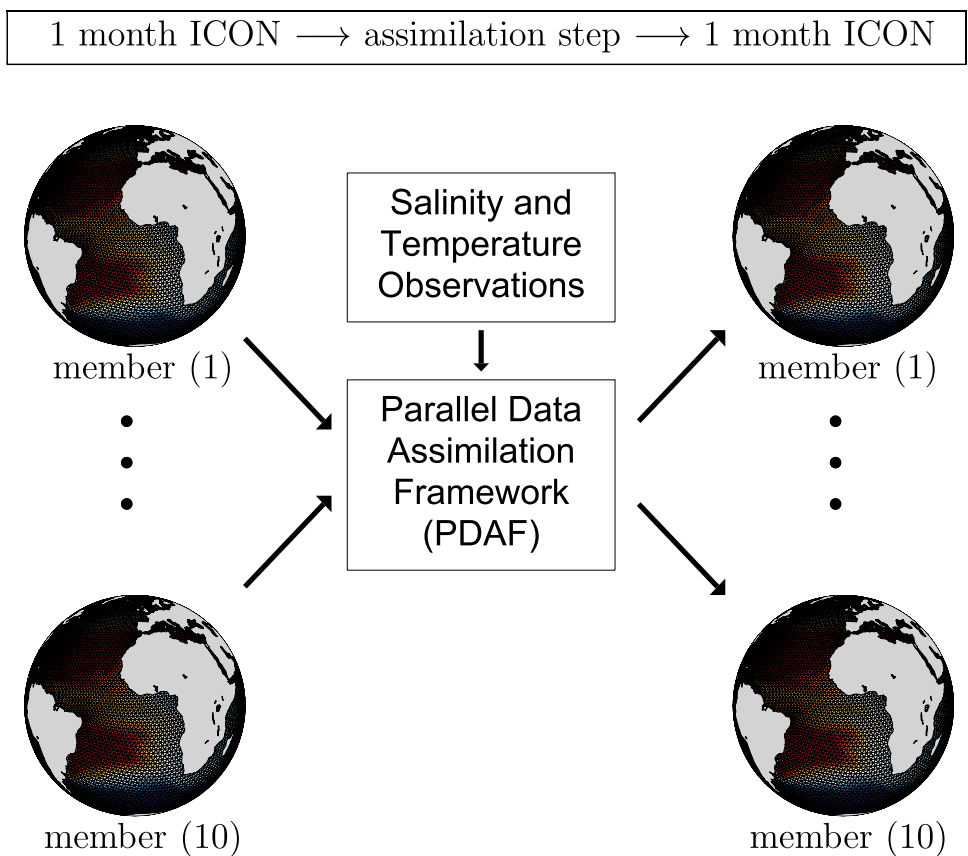
2 Method

We employ ICON-ESM in the configuration used for the CMIP6 historical simulations (Jungclaus et al. 2022) using transient external forcing from the CMIP6 (Eyring et al. 2016). ICON-ESM consists of the components ICON-Atmosphere (Giorgetta et al. 2018; Crueger et al. 2018), ICON-Ocean (Korn 2017; Korn et al. 2022), ICON-Land based on the Jena Scheme for Biosphere Atmosphere Coupling in Hamburg (JSBACH; Reick et al. 2021) and ICON-Biogeochemistry based on the Hamburg Ocean Carbon Cycle (HAMOCC; Maerz et al. 2020). Ocean and atmosphere are coupled with the "Yet-Another-Coupler" (YAC; Hanke et al. 2016). We use the ICON-ESM at a resolution of 160 km (R2B4) in the ICON-Atmosphere and 40 km (R2B6) in the ICON-Ocean. Jungclaus et al. (2022) have evaluated the Diagnosis, Evaluation, and Characterization of Klima (DECK) simulations with ICON-ESM against observations and find that the mean state and variability is in general similar to other climate models from the CMIP5 and CMIP6.

We produce retrospective decadal climate predictions (hindcasts) with ICON-ESM following the Decadal Climate Prediction Project protocol-A (DCPP-A; Boer et al. 2016). In a first step, we produce an assimilation run of 10 members with an oceanic EnKF implemented using PDAF (Nerger and Hiller 2013) over the period 1960–2014. Monthly ocean temperature and salinity profiles from the EN.4.2.1 data set (Good et al. 2013) are assimilated into ICON-ESM, which is then integrated to the next assimilation step a month later, when the cycle is repeated (Fig. 1). In our assimilation, we use sea surface observations of temperature and salinity only from oceanic profilers. That way, we retain sea surface temperature observations from satellites as one of our main validation references. In a second step, we initialize an ensemble of 10 decadal hindcast simulations from the (10) assimilation runs started in each year on the first of November.

In our oceanic assimilation implementation with PDAF we use a localized singular evolutive interpolated Kalman filter (SEIK; Pham 2001; Nerger and Hiller 2013). The horizontal localization is performed by considering only observations that fall into the localization range and the observation error determines the relative weight and impact of the observations. In the vertical, we suppress the use of cross-covariances across depth levels. With the relatively long assimilation step of one month, incremental updates are transported fast enough also to other depth levels. Also note that we do not use artificial inflation. Additionally, we

Fig. 1 Schematic of the data assimilation. The Ensemble Kalman Filter using PDAF assimilates once a month oceanic salinity and temperature profiles into ICON-ESM. The assimilation step is followed by a one-month ICON-ESM run with 10 ensemble members. The procedure is repeated in the next assimilation cycle. This way, the assimilation run is performed over the period 1960–2014 and provides the initial conditions for the decadal hindcast simulations

1 month ICON $\longrightarrow$ assimilation step $\longrightarrow$ 1 month ICON

Salinity and Temperature Observations

member (1)

Parallel Data Assimilation Framework (PDAF)

member (1)

member (10)

member (10)

rely on the pre-screening and quality checks performed by the EN4 team (Good et al. 2013). Based on our experience from previous experiments (Polkova et al. 2019), in the present study we use a local range of 5° and rather conservative observational errors of 1 K and 1 PSU for the temperature and salinity, respectively. We could envisage using smaller errors for salinity in the future, however, experience from previous studies with MPI-ESM showed a high risk for model inconsistencies and crashes (Brune et al. 2015).

In the following chapter, we analyze the hindcasts with respect to their lead-time dependent prediction skill. While many different skill metrics exist, we choose the Pearson's correlation coefficient (Wilks 2011) because of its independence on the bias of the hindcasts. Eventually, post-processing can correct the biases in the climate predictions (Pasternack et al. 2018). We define the lead-time l dependent correlation coefficient cor as

$$cor_l = \frac{\sum_{i=1}^{n} \left(x_{il} - \bar{x}_l\right)\left(y_{il} - \bar{y}_l\right)}{\sqrt{\sum_{i=1}^{n} \left(x_{il} - \bar{x}_l\right)^2}\sqrt{\sum_{i=1}^{n} \left(y_{il} - \bar{y}_l\right)^2}} \tag{1}$$

with x represents the ensemble mean hindcast variable of interest and y represents the observed value; i.e. their covariances divided by the product of their standard deviations. Correlation values close to one express a perfect prediction, values around zero mean no prediction skill and negative values indicate an anti-relation between the prediction and observation. We compare the correlation against the ensemble mean of the historical free runs and estimate the significance of the correlation values with a student's t-test (Siegert et al. 2017; Fröhlich et al. 2021). We use the one-sided t-test in the case of the hindcast correlation values to display the prediction skill and the two-sided t-test in the case of correlation differences that displays the effect of the initialization. To estimate the decadal prediction skill, we analyze annual means for different lead-years (ly). We start the decadal hindcasts on the first of November in each year. Ly1 represents the average over the following calendar year, starting January first hence the last complete year is ly9. For atmospheric variables, we analyze seasonal prediction skill for lead-months (lm) 2–4, which represents the average over December, January and February (DJF) of our hindcasts (assigned to the year in which the January falls) and additionally lm5 (March) and lm11 (September) for sea-ice variables.

3 Results

3.1 Decadal predictability

Sea surface temperature (SST) variability of ICON-ESM hindcasts exhibits high correlation values with observations from the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST; Rayner et al. 2003) over the globe for ly1 (Fig. 2a) and averages over ly1-5 (Fig. 2c). We find high prediction skill in the Atlantic, Indian Ocean and western Pacific, while in parts of the eastern Pacific predictability is low. The comparison of our results with the verification of other models from the LC ADCP (https://hadleyserver.metoffice.gov.uk/wmolc/; Hermanson et al. 2022) shows that the skill in our system is lower in the subtropical gyres of the Pacific and parts of the Southern Oceans and elsewhere competitive. The skill is increased by the initialization in most regions for ly1 and ly1-5, particularly in the tropical Pacific and northern North Atlantic (Fig. 2b, d).

Some of the regions with high SST prediction skill retain high correlation values also for the upper 700 m oceanic heat content (HC-700 m), referenced against observations from the Frontier Research System for Global Change (Ishii; Ishii et al. 2006), but the regions with significant predictive skill are much smaller for averages over ly1 as well as ly1-5 (Fig. 2e, g). The skill is increased by the initialization in the tropical Pacific and northern North Atlantic for ly1 and ly1-5, but there are also larger regions with a negative effect on the prediction skill like in the subtropical North and South Atlantic (Fig. 2f, h). At least some of the areas with low prediction skill of SST and HC-700 m can be attributed to the missing atmospheric data assimilation in our system. The wind has a strong influence on the predictability by its impact on ocean dynamics and mixed layer depths (Thoma et al. 2015). Another source of prediction skill from the atmosphere is the air temperature that is directly influencing the temperature of the ocean by heat fluxes. Other atmospheric sources of predictability stem from effects that are more indirect such as precipitation, cloud effects on radiation, evaporation, et cetera.

We further analyze the North Atlantic SST and HC-700 m as the average over the region 60°W-10°W, 50°N-60°N, where we found high predictive skill in our previous prediction system (Kröger et al. 2018). Note, this is also the region where the North Atlantic clearly stands out in terms of potential predictability (e.g. Pohlmann et al. 2004). The time series of the observed North Atlantic SST and HC-700 m show a low-frequency modulation with low values in the period 1970–1995 and high values thereafter (Fig. 3a, b). The hindcasts with ly1 and averages

over ly1-5 follow the observed signal and the correlation coefficients are statistically significant. For North Atlantic SST the correlation values are 0.78 and 0.86 for ly1 and ly1-5, respectively. The correlation values are lower for the ensemble mean of the historical free runs (0.33 and 0.41) for these lead-years, respectively. Additionally, for North Atlantic HC-700 m the correlation values are 0.77 and 0.74 for ly1 and ly1-5, respectively. Again, the correlation values are lower for the historical free runs (0.27 and 0.40) for these averaging periods, respectively.

Next, we give an overview of the correlation values for all possible periods with different start and end lead years (Fig. 3c, d). The idea to display the correlation values in this format was introduced by Athanasiadis et al. (2020) where it was applied to decadal predictability of blocking and North Atlantic Oscillation. In the diagram, the lowest row displays the correlation values for averages over ly1, ly1-2, ly1-3, …, ly1-9. The row above displays the correlation values for averages without the first lead-year, i.e. ly2, ly2-3, ly2-4, …, ly2-9, and the rows above show the values for averages with even later start-years, respectively. For North Atlantic SST, highest correlation values (cor = 0.87) are present for ly1-6 and ly1-7 and the correlation remains statistically significant up to ly6-8 and ly5-9. For North Atlantic HC-700 m the highest correlation value (cor = 0.77) is present for ly1 and the correlation is significant up to ly1-9.

The sea surface salinity (SSS) variability of the ICON-ESM hindcasts exhibits high correlation values with observations from the Frontier Research System for Global Change (Ishii et al. 2006) in the North Atlantic and central tropical and subtropical Pacific for averages over the ly1 (Fig. 4a) and ly1-5 (Fig. 4c) while elsewhere the prediction skill is low. These are the regions where the skill is increased by the initialization for both lead-year averages (Fig. 4b, d), respectively. The upper 700 m oceanic salt content (SC-700 m) shows high correlation values with observations from the Frontier Research System for Global Change (Ishii et al. 2006) only in the North Atlantic and tropical Atlantic (Fig. 4e, g). Similar to the SSS, these are the regions where the initialization has a positive effect on the prediction skill for both lead-year averages (Fig. 4f, h). Salinity observations are sparser than temperature observations and SSS observations from satellites are available only since 2009 (Olmedo et al. 2021) and too short for the evaluation of the prediction skill. We find SSS and SC-700 m predictability only in regions where salinity measurements are available over the whole assimilation period.

We analyze time series of area averaged SSS and SC-700 m for the same North Atlantic region as for SST/ HC-700 m (60°W-10°W, 50°N-60°N). The time series of the observed North Atlantic SSS and SC-700 m show a similar low-frequency signal as before with low values in the period 1970–1995 and high values thereafter (Fig. 5a,
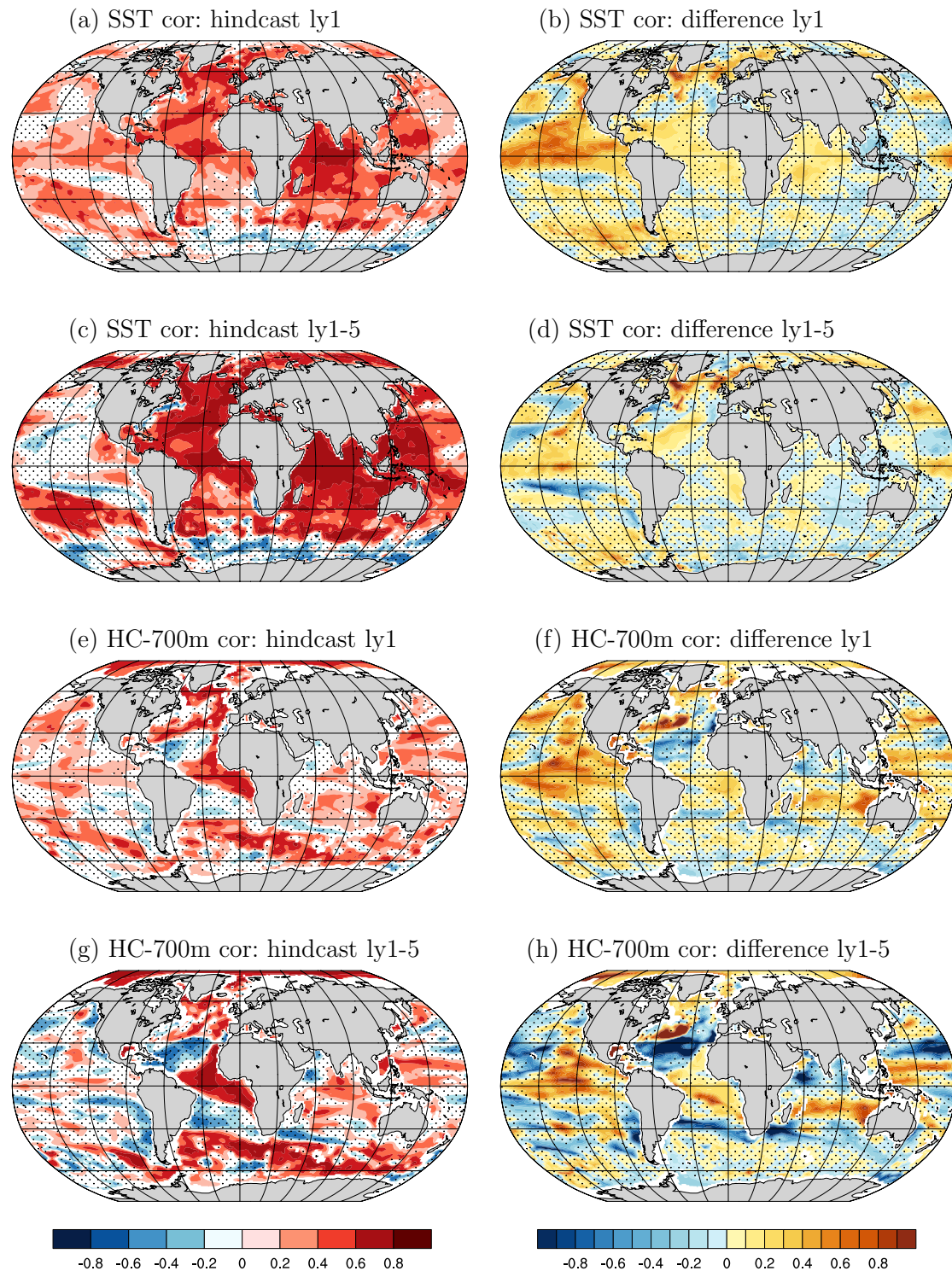
(a) SST cor: hindcast ly1

(b) SST cor: difference ly1

(c) SST cor: hindcast ly1-5

(d) SST cor: difference ly1-5

(e) HC-700m cor: hindcast ly1

(f) HC-700m cor: difference ly1

(g) HC-700m cor: hindcast ly1-5

(h) HC-700m cor: difference ly1-5

-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8

-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8

Fig. 2 Correlation of **a–d** sea surface temperature (SST) and **e–f** upper 700 m heat content (HC-700 m) from the ICON-ESM hindcasts with observations (HadISST and Ishii, respectively) for lead-years 1 (**a**, **e**) and 1–5 (**c**, **g**) and their differences to the correlation from the historical free runs (**b**, **d**, **f**, and **h**). The correlations are based on averages of 10 (5) hindcast (historical) ensemble members over the period 1961–2014 for SST and 1961–2012 for HC-700 m, respectively. Stippling indicates regions with non-significant values at the 95% level according to a t-test

(a) North Atlantic SST anomaly

(b) North Atlantic HC-700m anomaly
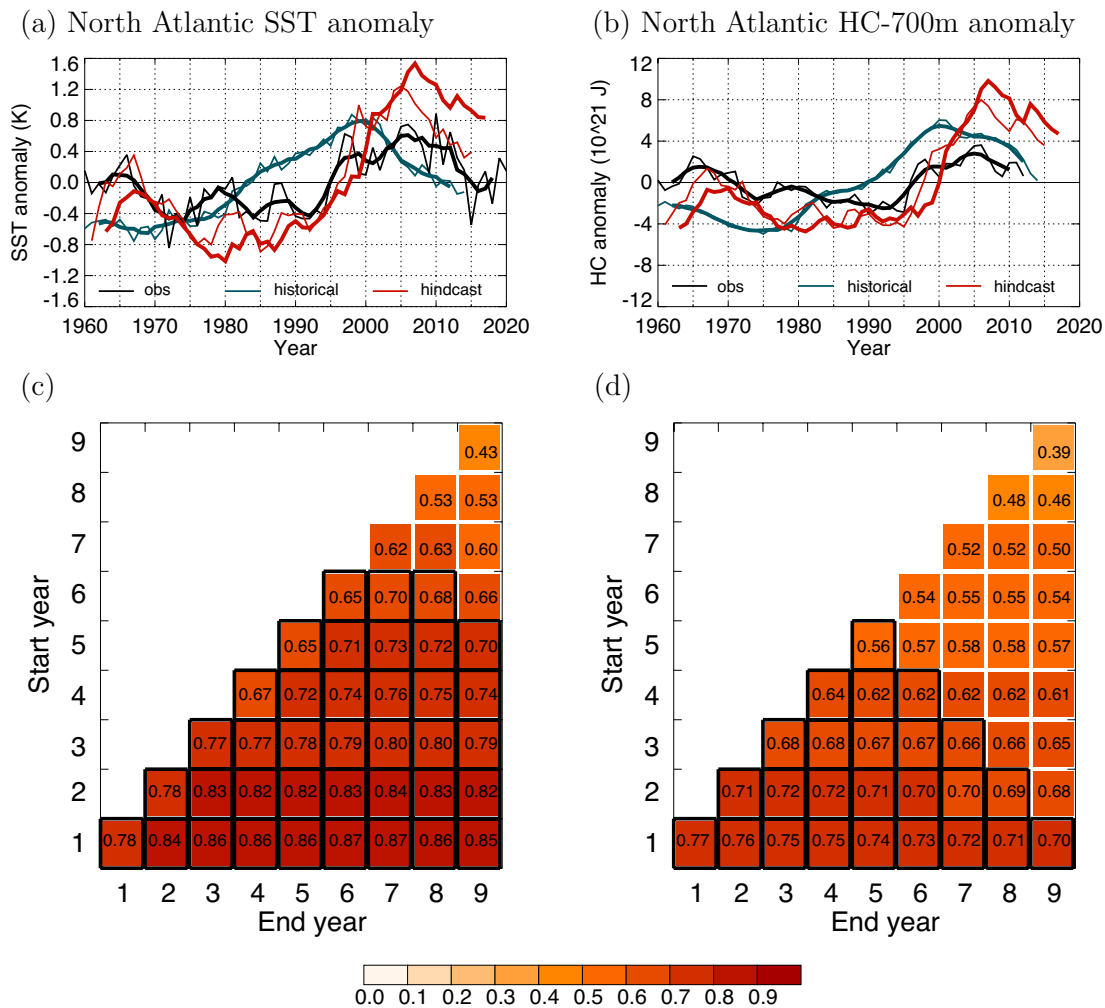
Fig. 3 Time series of ensemble and North Atlantic mean (60°W-10°W, 50°N-60°N) **a** sea surface temperature (SST) and **b** upper 700 m heat content (HC) anomalies from ICON-ESM hindcasts (red), the historical free runs (blue) and observations (HadISST and Ishii, respectively, black). The time-series of the observations are shown for annual (thin) and 5 year-running means (thick), and the hindcasts for lead-year 1 (thin) and lead-years 1–5 (thick). The diagrams below display the correlation coefficients for different lead-year ranges defined by the start and end years of the time-series for the North Atlantic SST (**c**) and HC-700 m (**d**). Framed tiles indicate significant values at the 95% level according to a *t*-test

b). The amplitudes of the simulated SSS and SC-700 m are larger than in the observations. For North Atlantic SSS the correlation values are 0.59 and 0.66 for ly1 and ly1-5, respectively. The historical free runs are not able to describe the North Atlantic SSS properly as shown by the negative correlation values for these lead-years (− 0.28 and − 0.38, respectively). Additionally, for North Atlantic SC-700 m the correlation values are 0.73 and 0.74 for ly1 and ly1-5, respectively. Again, the correlation values are lower for the historical free runs (− 0.12 and − 0.06, respectively).

The overview of the correlation values in the diagrams below the time series (Fig. 5c, d) shows that for North

Atlantic SSS, highest correlation values (cor = 0.79) are present for ly1-9 and ly2-9 and the correlation remains statistically significant up to ly6-8 and ly5-9. For North Atlantic SC-700 m the highest correlation value (cor = 0.79) is present for ly1-3 and ly2-3 and the correlation is significant up to ly1-7 and ly2-7. The predictability in the North Atlantic region is important via its teleconnections for example for the European temperature (e.g. Borchert et al. 2019). The predictability of salinity in the North Atlantic could be important for the prediction of fish larvae in the eastern North Atlantic (Miesner et al. 2022) and Barents Sea cod stock (Koul et al. 2021).
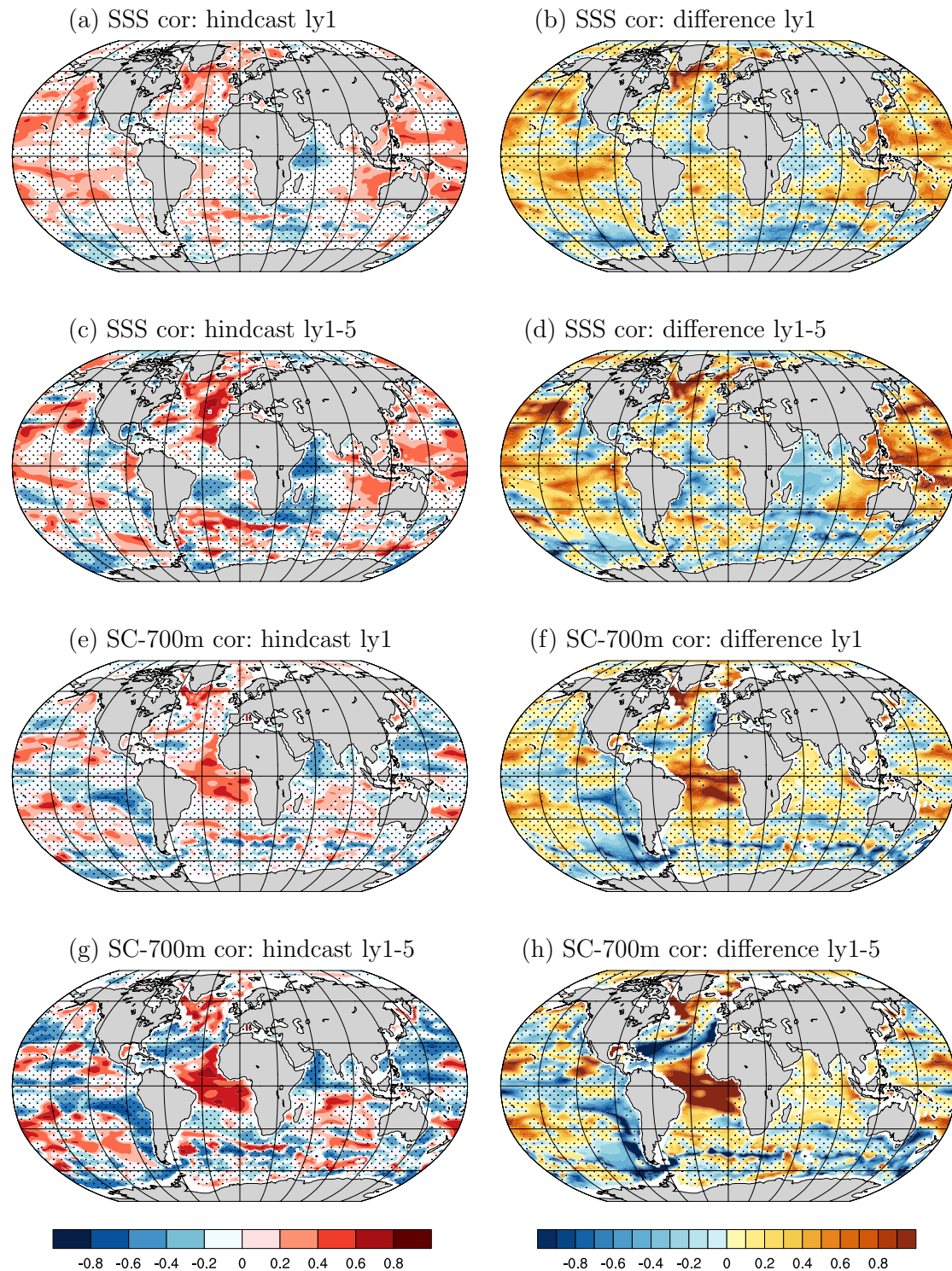
(a) SSS cor: hindcast ly1

(b) SSS cor: difference ly1

(c) SSS cor: hindcast ly1-5

(d) SSS cor: difference ly1-5

(e) SC-700m cor: hindcast ly1

(f) SC-700m cor: difference ly1

(g) SC-700m cor: hindcast ly1-5

(h) SC-700m cor: difference ly1-5

-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8

-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8

Fig. 4 Correlation of **a–d** sea surface salinity (SSS) and **e–f** upper 700 m salt content (SC-700 m) from the ICON-ESM hindcasts with observations (Ishii) for lead-years 1 (**a**, **e**) and 1–5 (**c**, **g**) and their differences to the correlation from the historical free runs (**b**, **d**, **f**, and **h**). The correlations are based on averages of 10 (5) hindcast (historical) ensemble members over the period 1961–2012. Stippling indicates regions with non-significant values at the 95% level according to a t-test

(a) North Atlantic SSS anomaly

(b) North Atlantic SC-700m anomaly

(c)

(d)
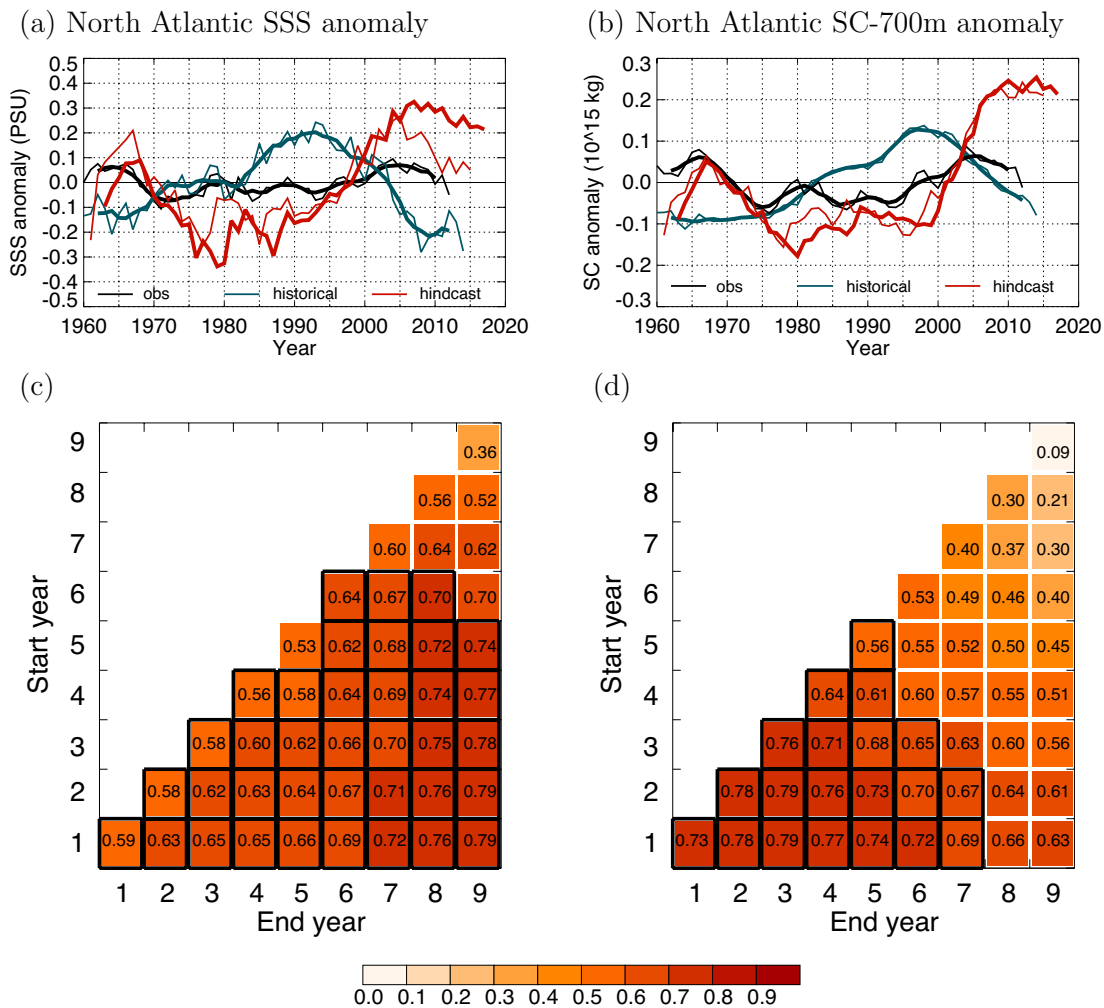
Fig. 5 Time series of ensemble and North Atlantic mean (60°W-10°W, 50°N-60°N) **a** sea surface salinity (SSS) and **b** upper 700 m salt content (SC) anomalies from ICON-ESM hindcasts (red), the historical free runs (blue) and observations (Ishii, black). The time-series of the observations are shown for annual (thin) and 5 year-running means (thick), and the hindcasts for lead-year 1 (thin) and lead-years 1–5 (thick). The diagrams below display the correlation coefficients for different lead-year ranges defined by the start and end years of the time-series for the North Atlantic SSS (**c**) and SC-700 m (**d**). Framed tiles indicate significant values at the 95% level according to a *t*-test

3.2 Mean state and variability of AMOC

The long-term mean of the Atlantic Meridional Overturning Circulation (AMOC) of the assimilation run over the period 1960–2014 shows the expected structure (e.g. Jackson et al. 2019) with an upper cell with a maximum of 18 Sv (1 Sverdrup = 10^6 m^3 s^{-1}) around 26°N in 1000 m depth and a weak counter-cell below (Fig. 6a). We show additionally the time series of AMOC at 26°N (Fig. 6b). The observed AMOC from the Rapid Climate Change Programme (RAPID) array (Moat et al. 2020) is of comparable strength. The AMOC from the assimilation has a positive trend in the 1960s and is thereafter relatively stable. However, the assimilation does not show the observed

decline in strength of more than 4 Sv around 2009. In the hindcasts, the AMOC strength at 26°N is generally lower for ly1 and even more for ly1-5 before the year 2000. This points to disturbed start conditions that causes the drift in the hindcasts. The Intergovernmental Report on Climate Change (IPCC) 6th assessment report (AR6; Arias et al. 2021) indicates that the AMOC at 30°N was relatively stable in the twentieth century and is expected to decrease over the twenty-first century. That way, the relatively stable AMOC during the period 1970–2000 looks reasonable but we do not see signs of a decreasing AMOC in our hindcasts in the first two decades of the twenty-first century. The ensemble spread evolves differently for different variables. Moreover, for variables in the deep ocean the
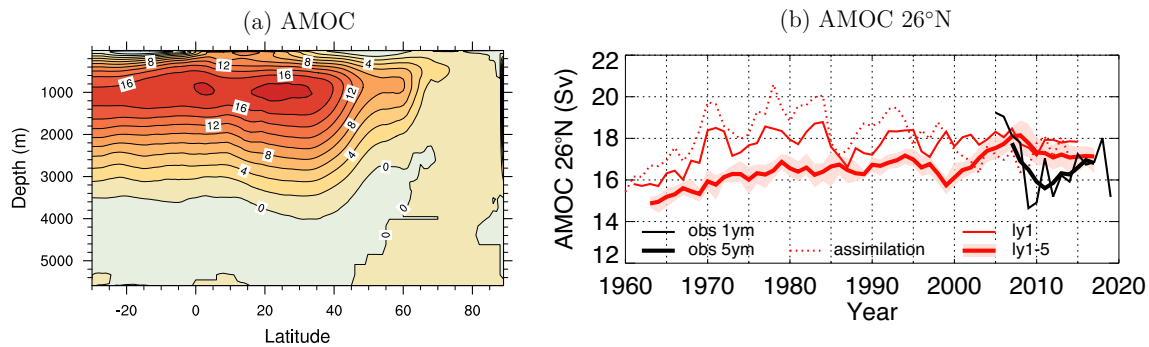
Fig. 6 a Ensemble mean of the Atlantic meridional overturning circulation (AMOC) averaged over the period 1960–2014 from the assimilation. b Time series of ensemble mean AMOC anomalies at 26°N in 1000 m depth from ICON-ESM assimilation (dotted red) and hindcasts for lead-year 1 (thin red) and lead-year 1–5 (thick red) and ensemble spread (shaded) and observations (RAPID) for 1 and 5 year means (thin and thick black, respectively)

ensemble spread rises slowly and for variables closer to the surface ensemble spread increases relatively fast. As an example, we show the ensemble spread of AMOC at 26°N for the hindcasts (Fig. 6b).

3.3 Seasonal predictability of surface variables

We analyze the seasonal predictability as the average of lm2-4. Since we started our hindcasts on every 1 November, this represents the average over December, January and February (DJF). Table 1 gives an overview of the correlation values of the seasonal hindcasts against observations for different variables. The correlation of sea surface height (SSH) of the hindcasts with satellite observations from the Archiving, Validation and Interpretation of Satellite Oceanographic Data project (AVISO; Fablet et al. 2018) is high in the tropical Pacific and Indian Ocean (Fig. 7a). The difference of correlation values between the historical free runs and the hindcasts in Fig. 7c shows that most of the skill is arising due to the oceanic initialization in these regions. The correlation of surface temperature (TS, i.e. SST over the ocean and land surface temperature over land) with observations from Goddard Institute for Space Studies Surface Temperature Analysis (GISTEMP; Lenssen et al. 2019) is generally high over the ocean and particularly over the tropical Pacific and northern North Atlantic (Fig. 7b). The difference of

correlation values between the historical free runs and the hindcasts in Fig. 7d demonstrates the strong impact of the initialization in these important regions.

Our prediction system is capable of predicting El Niño/Southern Oscillation (ENSO) events on seasonal time-scales. Important for ENSO is the variability in the NINO3.4 region, representing averaged values in the region 170°W–120°W, 5°S–5°N. The NINO3.4 SSH hindcasts of boreal winter (DJF) largely agree with satellite observations (cor = 0.83, Fig. 7e). The agreement of the NINO3.4 SST anomalies with observations from GISTEMP also lies in this range (cor = 0.79, Fig. 7f). The correlation values are lower for the historical free runs for NINO3.4 SSH and TS (− 0.21 and − 0.35, respectively). The prediction skill from other models is usually in the range of cor = 0.90 for the shorter period from 1980–2001 (Jin et al. 2008). Our NINO3.4 SST prediction lies also in this range for the shorter period (cor = 0.89).

Next, we analyze some atmospheric variables that are associated with ENSO. The correlation of precipitation of the hindcasts with observations from the Global Precipitation Climatology Project (GPCP; Adler et al. 2003) is significantly positive only in relatively small regions in the East and West Pacific (Fig. 8a). These are also the regions where the difference in skill is positive (Fig. 8c). ENSO teleconnections are biased in ICON-ESM in particular at the equator

Table 1 Overview about the seasonal boreal winter (DJF) prediction skill values for the different variables, regions and periods (see text)

Variable	SSH	TS	Precipitation		SLP	
Region	NINO3.4	NINO3.4	East Pacific	West Pacific	Tahiti	Darwin
Period	1993–2014	1961–2014	1979–2014	1979–2014	1961–2014	1961–2014
Hindcast correlation	**0.83**	**0.79**	**0.82**	**0.72**	**0.54**	**0.59**
Historical correlation	− 0.21	− 0.35	− 0.34	− 0.06	0.08	− 0.26
Reference	AVISO	GISTEMP	GPCP		HadSLP2	

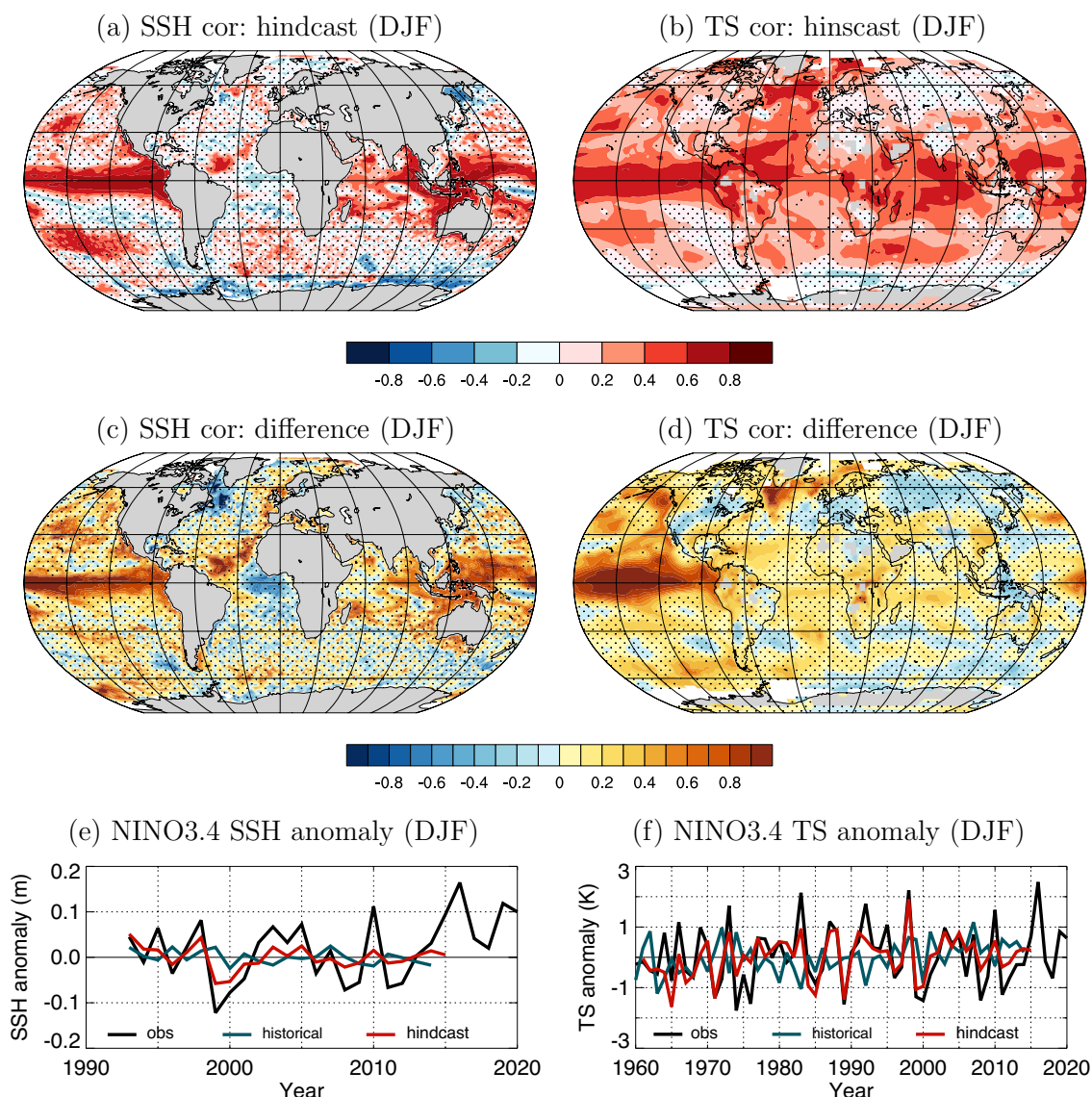Numbers in bold indicate significant values at the 95% level according to a *t*-test

Fig. 7 Correlation of **a** sea surface height (SSH) and (b) surface temperature (TS) from the ICON-ESM hindcasts with observations (AVISO and GISTEMP, respectively) for lead-month 2–4 (DJF) and **c**, **d** their differences to the correlation from the historical free runs. The correlations are based on averages of 10 (5) hindcast (historical) ensemble members over the periods **a**, **c** 1993–2014 and **b**, **d** 1961–2014. Stippling indicates regions with non-significant values at the 95% level according to a *t*-test. **e**, **f** Time series of **e** SSH and **f** TS anomalies in the NINO3.4 region (170°–120°W, 5°S-5°N) from observations (AVISO and GISTEMP, respectively, black), the historical free runs (blue) and hindcasts (red)

(Jungclaus et al. 2022). We define an East Pacific precipitation index as the average over the region 160°E-90°W, 10°S-10°N. The precipitation anomalies of the hindcasts in the East Pacific correspond with observations in this region and season (cor = 0.82, Fig. 8e). The correlation value is lower for the ensemble mean of the historical free runs (− 0.34). We define a West Pacific precipitation index as the average over the region 110°E-140°E, 5°N-25°N. The West Pacific precipitation anomalies of the hindcasts agree also well with observations in this region and season (cor = 0.72,

Fig. 8f). The correlation value is lower for the historical free run in this region (− 0.06).

For sea level pressure (SLP), we find high correlation values of the hindcasts against observations from the Hadley Centre Sea Level Pressure dataset (HadSLP2; Allan and Ansell 2006) in the East Pacific and West Pacific/Indian Ocean (Fig. 8b). These are also the regions with a positive skill difference to the historical free runs (Fig. 8d). The two regions are associated with the Southern Oscillation Index (SOI). We show two times series for SLP. We define the SLP index near Tahiti as the average over the region
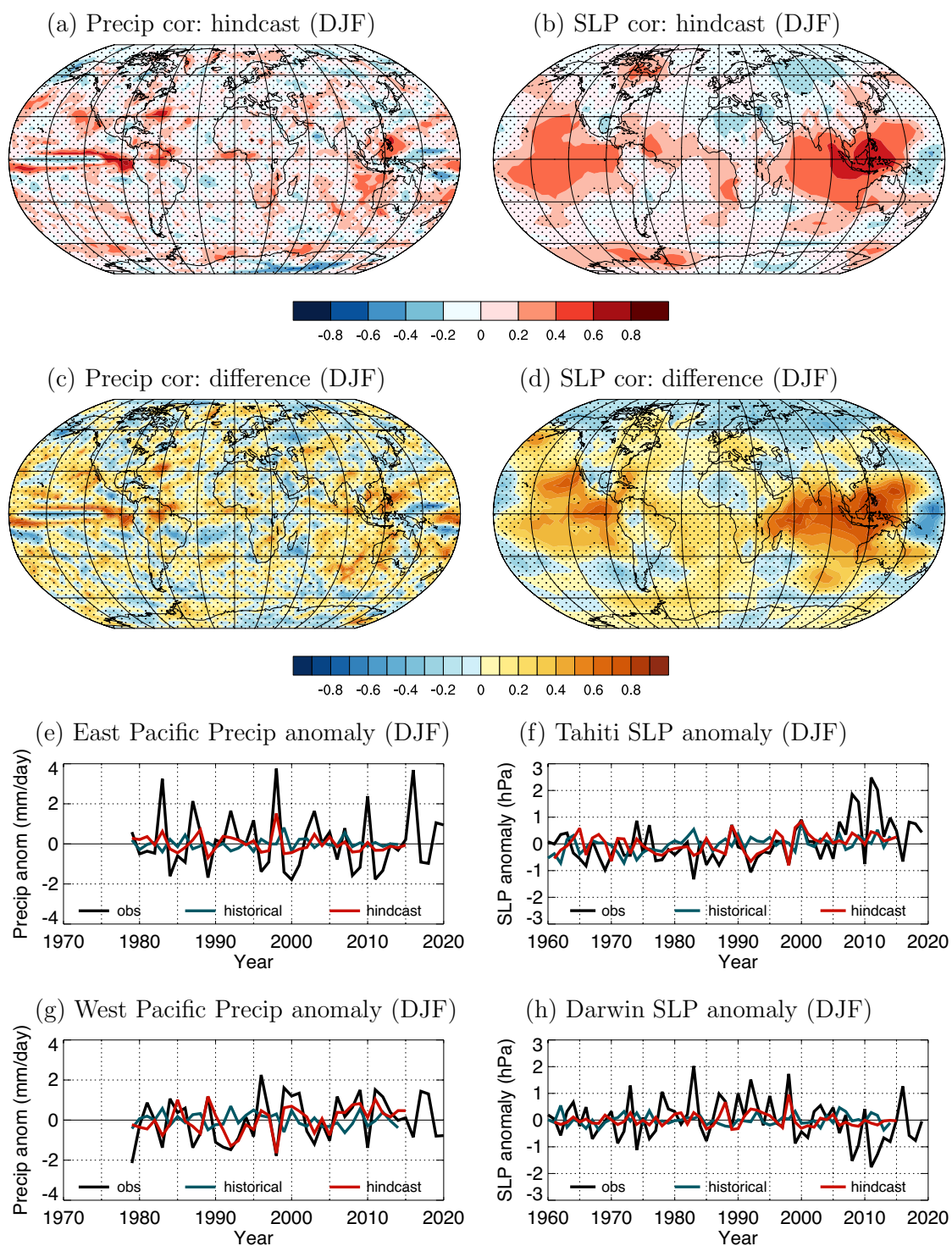
(a) Precip cor: hindcast (DJF)

(b) SLP cor: hindcast (DJF)

(c) Precip cor: difference (DJF)

(d) SLP cor: difference (DJF)

(e) East Pacific Precip anomaly (DJF)

(f) Tahiti SLP anomaly (DJF)

(g) West Pacific Precip anomaly (DJF)

(h) Darwin SLP anomaly (DJF)

Fig. 8 Correlation of **a** precipitation (Precip) and **b** sea level pressure (SLP) from the ICON-ESM hindcasts with observations (GPCP and HadSLP2, respectively) for lead-month 2–4 (DJF) and **c**, **d** their differences to the correlation from the historical free runs. The correlations are based on averages of 10 (5) hindcast (historical) ensemble members over the periods 1979–2014 for Precip and 1961–2014 for SLP. Stippling indicates regions with non-significant values at the 95% level according to a *t*-test. **e**, **g** Time series of Precip anomalies in the **e** East Pacific (160°E-90°W, 10°S-10°N) and **g** West Pacific (110°E-140°E, 5°N-25°N) from observations (black), the historical free runs (blue) and hindcasts (red). **f**, **h** Time series of SLP anomalies near **f** Tahiti (145°W-155°W, 0°-20°S) and **h** Darwin (125°E-135°E, 0°-15°S) from observations (black), the historical free runs (blue) and hindcasts (red)

145°W–155°W, 0°–20°S. The correlation of the Tahiti SLP data with observations is statistically significant (cor = 0.54, Fig. 8g). We define the SLP index near Darwin as the average over the region 125°E-135°E, 0°-15°S. The positive correlation of the Darwin SLP data with observations is also statistically significant (cor = 0.59, Fig. 8h). The correlation values are lower for the historical free run in these regions (0.08 and – 0.26, respectively).

3.4 Mean state and predictability of sea ice

We show averages of the Arctic and Antarctic sea ice concentration (SIC) as the mean over the period 1979–2014
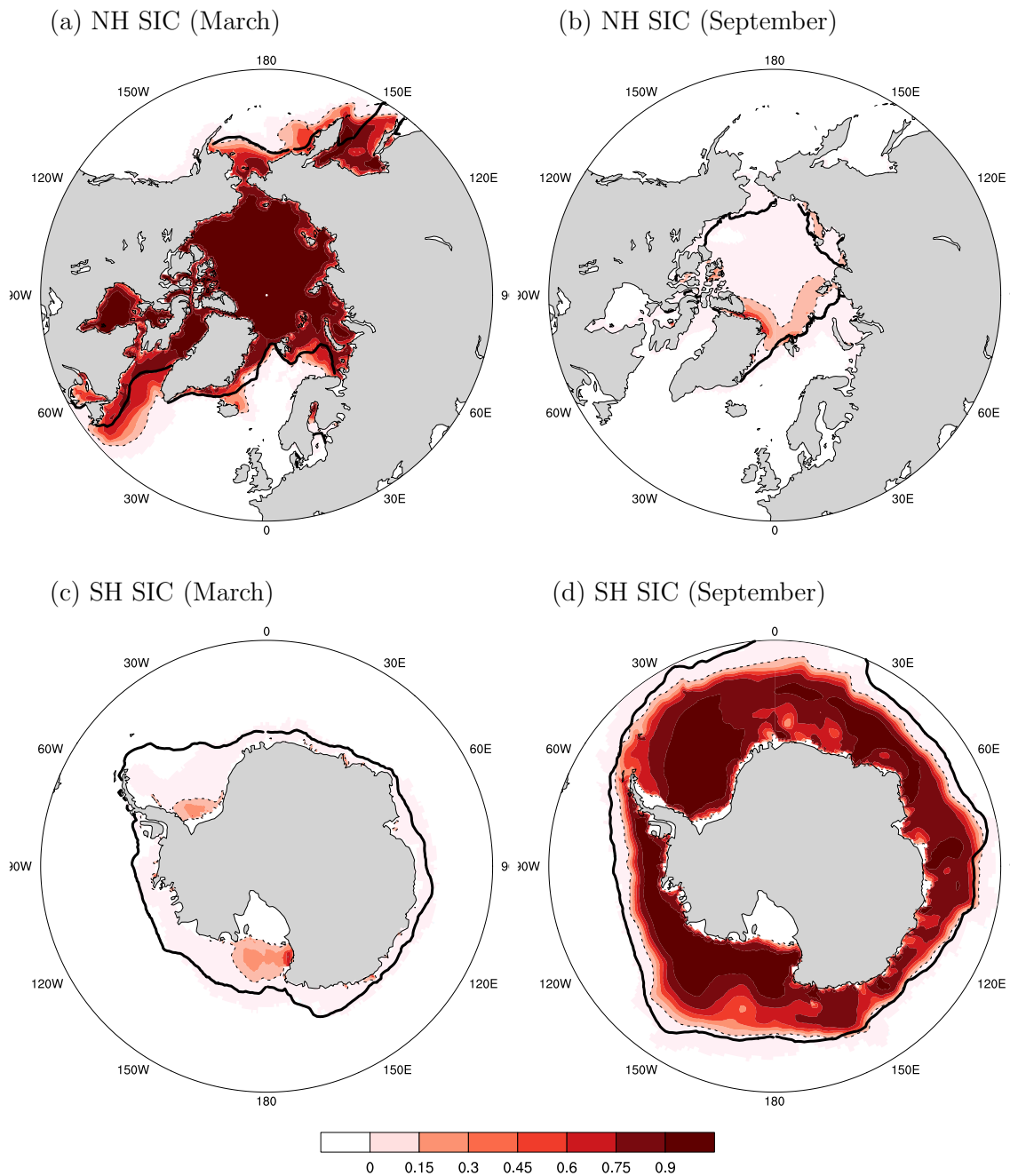
Fig. 9 Mean sea ice concentration (SIC) of the assimilation averaged over the period 1979–2014 in the **a**, **b** Northern Hemisphere (NH) and **c**, **d** Southern Hemisphere(SH) in **a**, **c** March and **b**, **d** September (colored). A dashed and full line indicates the sea-ice extent (area with at least 15% ice-concentration) from the assimilation and observations, respectively

together with the sea ice extent (SIE, i.e. the area with at least 15% SIC) from the assimilation and observations from HadISST (Rayner et al. 2003) in the respective summer and winter (Fig. 9). In the respective winter, the long-term mean of the SIE from the assimilation run shows relatively good agreement with observations in both hemispheres (Fig. 9a, d). In the Arctic, positive deviations from the observed long-term mean SIE are present in the Labrador and Bering Seas in winter (Fig. 9a). In the Antarctic, negative deviations from the observed SIE climatology are present almost circumpolar in winter (Fig. 9d). In summer, the SIE is much too low compared with observations in both hemispheres (Fig. 9b, c). In the Arctic, regions with SIC exceeding the 15% threshold can only be found in a relatively small area at the north coast of Greenland and extending further east, whereas in observations the Arctic remains almost completely ice covered in summer. The Arctic sea ice is usually multi-annual

and relatively thick at the Canadian archipelago averaged over the period 1979–2014 (Tschudi et al. 2020). In the Antarctic, regions with SIC exceeding the 15% threshold are limited to small areas in the Ross and Weddell Seas, while the observed sea ice cover remains much larger in summer. The sea ice biases in our assimilation experiment are similar to the biases in the ICON-ESM historical simulations (Jungclaus et al. 2022).

We analyze the SIE correlation of the hindcasts with observations from the National Snow and Ice Data Center (NSIDC; Fetterer et al. 2017) for the months with maximum and minimum SIE (i.e. March and September). Since our hindcasts are started on every 1 November this is for lead month 5 (March) and 11 (September), respectively. Table 2 gives an overview of the correlation values of the hindcasts against observations. The correlation of SIE of the hindcasts with observations is significantly positive only in the

Table 2 Overview about the prediction skill values of sea ice extent (SIE) for the Northern Hemisphere (NH) and Southern Hemisphere (SH) in winter and summer against observations from NSIDC (Fetterer et al. 2017) over the period 1979–2014

Variable	SIE			
Region	NH		SH	
Season (month of minimum and maximum)	March (winter)	September (summer)	March (summer)	September (winter)
Hindcast correlation	**0.80**	**0.76**	− 0.32	− 0.36
Historical correlation	0.24	**0.68**	− 0.40	− 0.46

Numbers in bold indicate significant values at the 95% level according to a *t*-test

(a) NH SIE anomaly (March)

(b) NH SIE anomaly (September)

(c) SH SIE anomaly (March)
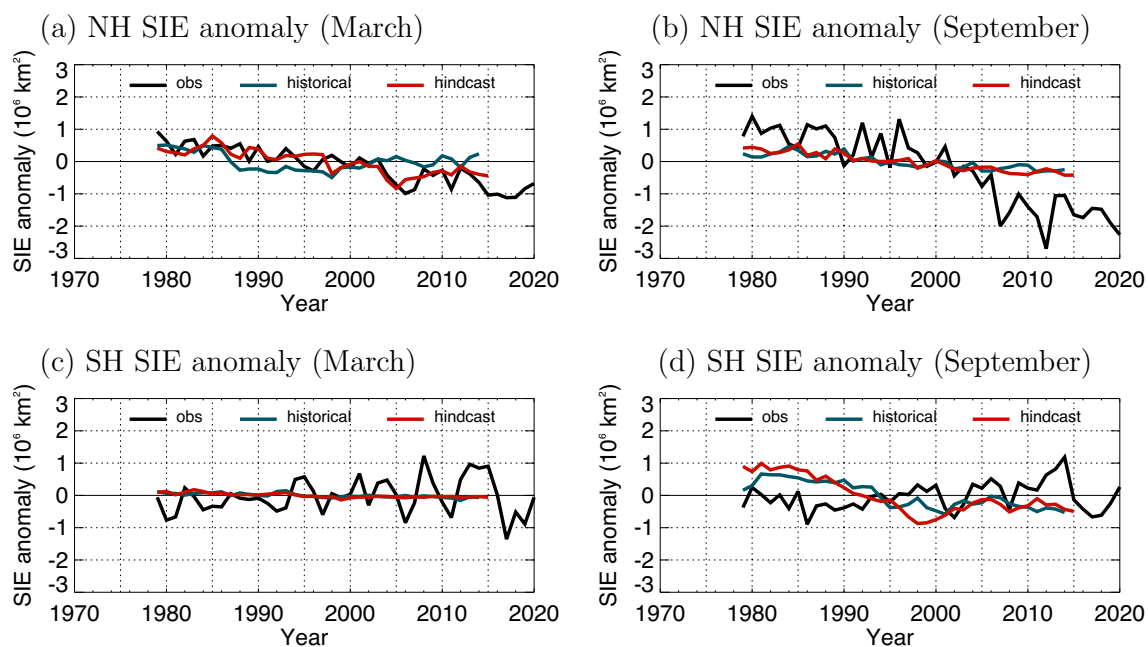
(d) SH SIE anomaly (September)

Fig. 10 Time series of sea ice extent (SIE) anomalies in the **a**, **b** Northern Hemisphere (NH) and **c**, **d** Southern Hemisphere (SH) in **a**, **c** March and **b**, **d** September from observations (NSIDC, black), the historical free runs (blue) and the hindcasts from ICON-ESM for lead months **a**, **c** 5 and **b**, **d** 11 (red)

Northern Hemisphere (NH) in both seasons, summer and winter, due to the agreement of the decreasing trend. For winter (NH, March), the strength of the trend agrees with the observed trend (Fig. 10a), but in summer (NH, September), the hindcasts underestimate the trend possibly due to the general underestimation of the SIE during this season (Fig. 10b). The skill in SIE is very similar to the skill from the historical free runs except for the NH SIE in March. This may also be the reason for the missing variability of SIE in the Southern Hemisphere (SH) in summer (Fig. 10c). In winter in the SH, the SIE trend of the hindcasts agrees with the observed trend only in the later period from the late 1990s (Fig. 10d). This may be also due to general problems with simulated variability in the Southern Ocean, e.g. the Antarctic circumpolar current is too weak in ICON-ESM compared to observations (Jungclaus et al. 2022).

4 Discussion and summary

We developed an oceanic initialization technique based on an oceanic EnKF assimilation as a first step towards a weakly coupled data assimilation in ICON-ESM. We performed an assimilation run over the period 1960–2014. The assimilation serves to initialize decadal hindcasts started on 1 November in each year. In general, oceanic temperature and salinity profile observations are successfully assimilated into ICON-ESM. With our approach of initializing only the oceanic part, we find—largely in agreement with expectations—high predictive skill in the following variables and regions:

We find multi-annual predictability of SST, SSS, HC-700 m and SC-700 m especially in the North Atlantic. Additionally, seasonal predictability is present in the tropics with highest values in variables related to ENSO. We find high predictive skill of SST and SSH especially in the tropical Pacific implicating a high predictive skill for precipitation and SLP in this region. ENSO predictability lies in the range of other models for DJF. However, compared to other prediction systems, prediction skill is relatively low in regions apart from the tropical Pacific due to the missing atmospheric assimilation. Note, that in our previous systems (Brune and Baehr 2020) we also did not assimilate SST directly but only indirectly via the atmosphere. Additionally, the hindcasts correctly represent the decreasing SIE trend in the Arctic in winter and to a lesser degree also in summer although the mean SIE in ICON-ESM is much too low in summer in both hemispheres. This, and additional general problems with simulating the variability in the Southern Ocean, causes the mismatch between simulated and observed SIE in the Antarctic in winter and summer.

We have used the ICON-ESM in our prediction system in relatively low resolution compared to other systems.

However, the advantage of ICON is its good performance at high resolution due to the scalability of the code and the use of non-hydrostatic equations for the atmosphere that allow high resolution convection permitting simulations (Stevens et al. 2019). Another advantage is the availability of adaptive grids as well for the atmosphere (Maurer et al. 2022) as for the ocean (Logemann et al. 2021; Korn et al. 2022). The role of such improvements for seasonal and near-term predictions will be investigated in forthcoming studies.

In general, we can confirm that our data assimilation method is successfully initializing the oceanic component of the climate system, even though the observational density for oceanic profile data is sparse in the first part of our assimilation period. The amount of observational profiles increased considerably from the early 2000s onwards with the advent of Array for Real-time Geostrophic Oceanography (ARGO; Roemmich and Owens 2000) data, which we expect to have a positive impact on future forecasts initialized by our system. We also expect that the second step towards a weakly coupled data assimilation—an additional atmospheric assimilation—will enhance the prediction skill further and will lead to high quality seasonal and decadal climate predictions. An additional fine-tuning of the climate model could also improve the prediction skill, e.g. with a more realistic Arctic and Antarctic sea ice climatology in summer and an improved oceanic circulation in the Southern Ocean. We are currently restructuring the ICON code by unifying the physical parameterization packages for numerical weather predictions and climate applications. The ICON-seamless project is working on advancing all aspects of the coupled climate system (atmosphere, land, ocean, cryosphere and data assimilation) for improved weather and climate predictions on time scales from days to centuries.

EN.4.2.1 data (Good et al. 2004) were obtained from https://www.metoffice.gov.uk/hadobs/en4/ (accessed 2018–01-19) and are © British Crown Copyright, Met Office, 2021, provided under a Non-Commercial Government License http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/ The HadISST data set (Rayner et al. 2003) was downloaded from https://www.metoffice.gov.uk/hadobs/hadisst (accessed 2021–12-10). The subsurface temperature and salinity analyses (Ishii et al. 2005) were downloaded from the research data archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory from https://rda.ucar.edu/dataset/ds285.3 (accessed 2022–02-02). Data from the RAPID AMOC monitoring project (Moat et al. 2020) is funded by the Natural Environment Research Council and are freely available from www.rapid.ac.uk/rapidmoc (accessed 2021–12-05). AVISO data (Fablet et al. 2018) were processed by SSALTO/DUACS and distributed by AVISO + (https://www.aviso.altimetry.fr) with support from CNES. The AVISO data set SEALEVEL_GLO_PHY_L4_REP_OBSERVA-TIONS_008_047 was downloaded from https://resources.marine.copernicus.eu (accessed 2021–12-10). The following data are provided by NOAA/OAR/ESRL PSL, Boulder, Colorado, USA: GPCP precipitation data (Adler et al. 2003) are obtained from https://psl.noaa.gov/data/gridded/data.gpcp.html (accessed 2021–12-10). GISTEMPv4 data (Lenssen et al. 2019) were downloaded from https://downloads.psl.noaa.gov/Datasets/gistemp/combined/250km/air.2x2.250.mon.anom.comb.nc (accessed 2021–12-10). The HadSLP2 data (Allen and Ansell 2006) are obtained from https://psl.noaa.gov/gcos_wgsp/Gridded/data.hadslp2.html (accessed 2022-08–31). NSIDC Sea ice extent data (Fetterer et al. 2017) are downloaded from https://psl.noaa.gov/data/timeseries/monthly/SHICE/ and https://psl.noaa.gov/data/timeseries/monthly/NHICE/ (accessed 2021–12-10).

Declarations

References

Adler RF, Huffman GJ, Chang A, Ferraro R, Xie P-P, Janowiak J et al (2003) The version 2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). J Hydrometeorol 4:1147–1167. https://doi.org/10.1175/1525-7541(2003)004%3c1147:TVGPCP%3e2.0.CO;2

Allan R, Ansell T (2006) A new globally complete monthly historical mean sea level pressure data set (HadSLP2): 1850–2004. J Clim 19:5816–5842. https://doi.org/10.1175/JCLI3937.1

Arias PA, Bellouin N, Coppola E, Jones RG, Krinner G, Marotzke J et al (2021) Technical summary. In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou B (eds) Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change. Cambridge University Press. In Press. https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_TS.pdf

Athanasiadis PJ, Yeager S, Kwon YO, Bellucci A, Smith DW, Tibaldi S (2020) Decadal predictability of North Atlantic blocking and the NAO. Npj Clim Atmos Sci. https://doi.org/10.1038/s41612-020-0120-6

Boer GJ, Smith DM, Cassou C, Doblas-Reyes F, Danabasoglu G, Kirtman B et al (2016) The decadal climate prediction project (DCPP) contribution to CMIP6. Geosci Model Dev 9:3751–3777. https://doi.org/10.5194/gmd-9-3751-2016

Borchert LF, Pohlmann H, Baehr J, Neddermann NC, Suarez-Gutierrez L, Müller WA (2019) Decadal predictions of the probability of occurrence for warm summer temperature extremes. Geophys Res Lett 46:14042–14051. https://doi.org/10.1029/2019GL085385

Brune S, Baehr J (2020) Preserving the coupled atmosphere-ocean feedback initializations of decadal climate predictions. Wires Clim Change 11:e637. https://doi.org/10.1002/wcc.637

Brune S, Nerger L, Baehr J (2015) Assimilation of oceanic observations in a global coupled Earth system model with the SEIK filter. Ocean Model 96:254–264. https://doi.org/10.1016/j.ocemod.2015.09.011

Brune S, Düsterhus A, Pohlmann H, Müller WA, Baehr J (2018) Time dependency of the prediction skill for the North Atlantic subpolar gyre in initialized decadal hindcasts. Clim Dyn 51:1947–1970. https://doi.org/10.1007/s00382-017-3991-4

Crueger T, Giorgetta MA, Brokopf R, Esch M, Fiedler S, Hohenegger C et al (2018) ICON-A, the atmosphere component of the ICON Earth system model: II. Model evaluation. J Adv Model Earth Syst 10:1638–1662. https://doi.org/10.1029/2017MS001233

Dunstone NJ, Smith DM (2010) Impact of atmosphere and sub-surface ocean data on decadal climate prediction. Geophys Res Lett 37:L02709. https://doi.org/10.1029/2009GL041609

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fablet R, Verron J, Mourre B, Chapron B, Pascual A (2018) Improving mesoscale altimetric data from a multitracer convolutional processing of standard satellite-derived products. IEEE Trans Geosci Remote Sens 56(5):2518–2525. https://doi.org/10.1109/TGRS.2017.2750491

Fetterer F, Knowles K, Meier W, Savoie M, Windnagel AK (2017) Updated daily. Sea Ice Index, Version 3. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5K072F8

Fröhlich K, Dobrynin M, Isensee K, Gessner C, Paxian A, Pohlmann H et al (2021) The German climate forecast system: GCFS. J Adv Model Earth Syst 13:e2020MS002101. https://doi.org/10.1029/2020MS002101

Giorgetta MA, Brokopf R, Crueger T, Esch M, Fiedler S, Helmert J et al (2018) ICON-A, the atmosphere component of the ICON Earth system model: I. Model description. Journal of Advances in Modeling Earth Systems 10:1613–1637. https://doi.org/10.1029/2017MS001242

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 118:6704–6716. https://doi.org/10.1002/2013JC009067

Hanke M, Redler R, Holfeld T, Yastremsky M (2016) YAC 1.2.0: new aspects for coupling software in Earth system modeling. Geosci Model Dev 9:2755–2769. https://doi.org/10.5194/gmd-9-2755-2016

Hermanson L, Smith D, Seabrook M, Bilbao R, Doblas-Reyes F, Tourigny E et al (2022) WMO global annual to decadal climate update: a prediction for 2021–2025. Bull Am Meteor Soc 103:1117–1129. https://doi.org/10.1175/BAMS-D-20-0311.1

Ishii M, Kimoto M, Sakamoto K, Iwasaki S (2005) Subsurface temperature and salinity analyses. Res Data Arch Natl Center Atmos Res Comput Inform Syst Lab. https://doi.org/10.5065/Y6CR-KW66

Ishii M, Kimoto M, Sakamoto K, Iwasaki SI (2006) Steric sea level changes estimated from historical ocean subsurface temperature and salinity analyses. J Oceanogr 62(2):155–170. https://doi.org/10.1007/s10872-006-0041-y

Jackson LC, Dubois C, Forget G, Haines K, Harrison M, Iovino D et al (2019) The mean state and variability of the North Atlantic circulation: a perspective from ocean reanalyses. J Geophys Res: Oceans 124:9141–9170. https://doi.org/10.1029/2019JC015210

Jin EK, Kinter JL, Wang B, Park C-K, Kang I-S, Kirtman BP et al (2008) Current status of ENSO prediction skill in coupled ocean–atmosphere models. Clim Dyn 31:647–664. https://doi.org/10.1007/s00382-008-0397-3

Jungclaus JH, Lorenz SJ, Schmidt H, Brovkin V, Brüggemann N, Chegini F et al (2022) The ICON Earth system model version 1.0. J Adv Model Earth Syst 14:e2021MS0028. https://doi.org/10.1029/2021MS002813

Keenlyside NS, Latif M, Jungclaus J, Kornblueh L, Roeckner E (2008) Advancing decadal-scale climate prediction in the North Atlantic sector. Nature 453:84–88. https://doi.org/10.1038/nature06921

Korn P (2017) Formulation of an unstructured grid model for global ocean dynamics. J Comput Phys 339:525–552. https://doi.org/10.1016/j.jcp.2017.03.009

Korn P, Brüggemann N, Jungclaus JH, Lorenz SJ, Gutjahr O, Haak H et al (2022) ICON-O: the ocean component of the ICON Earth system model - global simulation characteristics and local telescoping capability. J Adv Model Earth Syst 14:e2021MS002952. https://doi.org/10.1029/2021MS002952

Koul V, Sguotti C, Arthun M, Brune S, Düsterhus A, Bogstad B et al (2021) Skilful prediction of cod stocks in the North and Barents Sea a decade in advance. Commun Earth Environ 2:140. https://doi.org/10.1038/s43247-021-00207-6

Kröger J, Pohlmann H, Sienz F, Marotzke J, Baehr J, Köhl A et al (2018) Full-field initialized decadal predictions with the MPI earth system model: an initial shock in the North Atlantic. Clim Dyn 51:2593–2608. https://doi.org/10.1007/s00382-017-4030-1

Kushnir Y, Scaife AA, Arritt R, Balsamo G, Boer G, Doblas-Reyes F et al (2019) Towards operational predictions of the near-term climate. Nat Clim Chang 9:94–101. https://doi.org/10.1038/s41558-018-0359-7

Lenssen NJL, Schmidt GA, Hansen JE, Menne MJ, Persin A, Ruedy R, Zyss D (2019) Improvements in the GISTEMP uncertainty model. J Geophys Res: Atmos 124:6307–6326. https://doi.org/10.1029/2018JD029522

Logemann K, Linardakis L, Korn P, Schrumm C (2021) Global tide simulations with ICON-O: testing the model performance on highly irregular meshes. Ocean Dyn 71:43–57. https://doi.org/10.1007/s10236-020-01428-7

Maerz J, Six KD, Stemmler I, Ahmerkamp S, Ilyina T (2020) Microstructure and composition of marine aggregates as co-determinants for vertical particulate organic carbon transfer in the global ocean. Biogeosciences 17:1765–1803. https://doi.org/10.5194/bg-17-1765-2020

Maurer V, Früh B, Giorgetta MA, Steger C, Brauch J, Schnur R et al (2022) Domain nesting in ICON and its application to AMIP experiments with regional refinement. J Adv Model Earth Syst. https://doi.org/10.1002/essoar.10507778.1

Merryfield WJ, Baehr J, Batte L, Becker EJ, Butler AH, Coelho CAS (2020) Current and emerging developments in subseasonal to decadal prediction. Bull Am Meteor Soc 101:869–896. https://doi.org/10.1175/BAMS-D-19-0037.1

Miesner AK, Brune S, Pieper P, Koul V, Baehr J, Schrum C (2022) Exploring the potential of forecasting fish distributions in the North East Atlantic with a dynamic earth system model, exemplified by the suitable spawning habitat of blue whiting. Front Mar Sci 8:777427. https://doi.org/10.3389/fmars.2021.777427

Moat BI, Smeed DA, Frajka-Williams E, Desbruyeres DG, Beaulieu C, Johns WE et al (2020) Pending recovery in the strength of the meridional overturning circulation at 26°N. Ocean Sci 16:863–874. https://doi.org/10.5194/os-16-863-2020

Mu L, Nerger L, Tang Q, Loza SN, Sidorenko D, Wang Q et al (2020) Towards a data assimilation system for seamless sea ice prediction based on the AWI climate model. J Adv Model Earth Syst 12:e2019MS001937. https://doi.org/10.1029/2019MS001937

Nerger L, Hiller W (2013) Software for ensemble-based data assimilation systems - implementation strategies and scalability. Comput Geosci 55:110–118. https://doi.org/10.1016/j.cageo.2012.03.026

Olmedo E, Gonzales-Haro C, Hoareau N, Umbert M, Gonzales-Gambau V, Martinez J et al (2021) Nine years of SMOS sea surface salinity global maps at the Barcelona expert center. Earth Syst Sci Data 13:857–888. https://doi.org/10.5194/essd-13-857-2021

Pasternack A, Bhend J, Liniger MA, Rust HW, Müller WA, Ulbrich U (2018) Parametric decadal climate forecast recalibration (DeFoReSt 1.0). Geosci Model Dev 11:351–368. https://doi.org/10.5194/gmd-11-351-2018

Penny SG, Bach E, Bhargava K, Chang C-C, 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 11:1803–1829. https://doi.org/10.1029/2019MS001652

Pham DT (2001) Stochastic methods for sequential data assimilation in strongly nonlinear systems. Mon Weather Rev 129:1194–1207. https://doi.org/10.1175/1520-0493(2001)129%3c1194:SMF-SDA%3e2.0.CO;2

Pohlmann H, Botzet M, Latif M, Roesch A, Wild M, Tschuk P (2004) Estimating the decadal predictability of a coupled AOGCM. J Clim 17:4463–4472. https://doi.org/10.1175/3209.1

Pohlmann H, Jungclaus JH, Köhl A, Stammer D, Marotzke J (2009) Initializing decadal climate predictions with the GECCO oceanic synthesis: effects on the North Atlantic. J Clim 22:3926–3938. https://doi.org/10.1175/2009JCLI2535.1

Pohlmann H, Müller WA, Bittner M, Hettrich S, Modali K, Pankatz K, Marotzke J (2019) Realistic quasi-biennial oscillation variability in historical and decadal hindcast simulations using CMIP6 forcing. Geophys Res Lett 46:14118–14125. https://doi.org/10.1029/2019GL084878

Polkova I, Brune S, Kadow C, Romanova V, Gollan G, Baehr J et al (2019) Initialization and ensemble generation for decadal climate predictions: a comparison of different methods. J Adv Model Earth Syst 11:149–172. https://doi.org/10.1029/2018MS001439

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP et al (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res 108(D14):4407. https://doi.org/10.1029/2002JD002670

Reick CH, Gayler V, Goll D, Hagemann S, Heidkamp M, Nabel JEMS et al (2021) JSBACH 3 - The land component of the MPI Earth system model: documentation of version 3.2. Berichte zur Erdsystemforschung 240. https://doi.org/10.17617/2.3279802

Roemmich D, Owens W (2000) The Argo project: global ocean observations for understanding and prediction of climate variability. Oceanography 13:45–50. https://doi.org/10.5670/oceanog.2000.33

Siegert S, Bellprat O, Mengoz M, Stephenson DB, Doblas-Reyes FJ (2017) Detecting improvements in forecast correlation skill: statistical testing and power analysis. Mon Weather Rev 145:437–450. https://doi.org/10.1175/MWR-D-16-0037.1

Smith DM, Eade R, Scaife AA, Caron L-P, Danabasoglu G, DelSole TM et al (2019) Robust skill of decadal climate predictions. npj Clim Atmos Sci 2:13. https://doi.org/10.1038/s41612-019-0071-y

Stevens B, Satoh M, Auger L, Biercamp J, Bretherton CS, Chen X et al (2019) DYAMOND: the dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. Prog Earth Planet Sci 6:61. https://doi.org/10.1186/s40645-019-0304-z

Swingedouw D, Mignot J, Labetoulle S, Guilyardi E, Madec G (2013) Initialisation and predictability of the AMOC over the last 50 years in a climate model. Clim Dyn 40:2381–2399. https://doi.org/10.1007/s00382-012-1516-8

Tang Q, Mu L, Goessling HF, Semmler T, Nerger L (2021) Strongly coupled data assimilation of ocean observations into an ocean-atmosphere model. Geophys Res Lett 48:e2021GL094941. https://doi.org/10.1029/2021GL094941

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteor Soc 93(4):485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Thoma M, Greatbatch RJ, Kadow C, Gerdes R (2015) Decadal hindcasts initialized using observed surface wind stress: evaluation and prediction out to 2024. Geophys Res Lett 42:6454–6461. https://doi.org/10.1002/2015GL064833

Tompkins AM, de Zarate MI, Saurral RI, Vera C, Saulo C, Merryfield WJ et al (2017) The climate-system historical forecast project: providing open access to seasonal forecast ensembles from centers around the globe. Bull Am Meteor Soc 98:2293–2301. https://doi.org/10.1175/BAMS-D-16-0209.1

Tschudi MA, Meier WN, Stewart JS (2020) An enhancement to sea ice motion and age products at the national snow and ice data center (NSIDC). Cryosphere 14:1519–1536. https://doi.org/10.5194/tc-14-1519-2020

Wang Y, Counillon F, Keenlyside N, Svendsen L, Gleixner S, Kimmritz M et al (2019) Seasonal predictions initialised by assimilating sea surface temperature observations with the EnKF. Clim Dyn 53:5777–5797. https://doi.org/10.1007/s00382-019-04897-9

Wilks DS (2011) Statistical methods in the atmospheric sciences, 3rd edn. International geophysics series, Vol 100. Academic Press, Elsevier, Amsterdam, The Netherlands. ISBN: 9780123850225

Zängl G, Reinert D, Ripodas P, Baldauf M (2015) The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: description of the non-hydrostatic dynamical core. Q J R Meteorol Soc 141:563–579. https://doi.org/10.1002/qj.2378

Zhu J, Kumar A, Lee H-J, Wang H (2017) Seasonal predictions using a simple ocean initialization scheme. Clim Dyn 49:3989–4007. https://doi.org/10.1007/s00382-017-3556-6