Geoscientific
Model Development
Discussions

EGU
Open Access

Benefits of sea ice thickness initialization for the Arctic decadal climate prediction skill in EC-Earth3

Tian Tian[1], Shuting Yang[1], Mehdi Pasha Karami[2], François Massonnet[3], Tim Kruschke[2], and Torben Koenigk[2,4]

[1]Danish Meteorological Institute, Denmark
[2]Swedish Meteorological and Hydrological Institute, Sweden
[3]Georges Lemaitre Centre for Earth and Climate Research, Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[4]Bolin Centre for Climate Research, Stockholm University, Sweden

Correspondence: Tian Tian (tian@dmi.dk)

Abstract. A substantial part of Arctic climate predictability at interannual time scales stems from the knowledge of the initial sea ice conditions. Among all the variables characterizing sea ice, sea ice volume, being a product of sea ice area/concentration (SIC) and thickness (SIT), is the most sensitive parameter for climate change. However, the majority of climate prediction systems are only assimilating the observed SIC due to lack of long-term reliable global observation of SIT. In this study the

5 EC-Earth3 Climate Prediction System with anomaly initialization to ocean, SIC and SIT states is developed. In order to evaluate the benefits of specific initialized variables at regional scales, three sets of retrospective ensemble prediction experiments are performed with different initialization strategies: ocean-only; ocean plus SIC; and ocean plus SIC and SIT initialization. The increased skill from ocean plus SIC initialization is small in most regions, compared to ocean-only initialization. In the marginal ice zone covered by seasonal ice, skills regarding winter SIC are mainly gained from the initial ocean temperature anomalies.

10 Consistent with previous studies, the Arctic sea ice volume anomalies are found to play a dominant role for the prediction skill of September Arctic sea ice extent. Winter preconditioning of SIT for the perennial ice in the central Arctic Ocean results in increased skill of SIC in the adjacent Arctic coastal waters (e.g. the Laptev/East Siberian/Chukchi Seas) for lead time up to a decade. This highlights the importance of initializing SIT for predictions of decadal time scale in regional Arctic sea ice. Our results suggest that as the climate warming continues and the central Arctic Ocean might become seasonal ice free in the

15 future, the controlling mechanism for decadal predictability may thus shift from being the sea ice volume playing the major role to a more ocean-related processes.

1 Introduction

Summer sea ice in the Arctic Ocean has lost nearly three-quarters of its sea ice volume (SIV) since the 1970's (Kwok, 2018) caused by a reduction of both sea ice extent (SIE) and thickness (SIT). This sea ice melt, inducing ice-albedo feedback,

20 contributes to the larger warming of the atmosphere in the Arctic than the global mean, an effect known as polar amplification (Wadhams, 2012). Observations suggest that the Arctic has warmed at more than twice the rate of the globe (Holland and Bitz,

2003; Serreze and Barry, 2011; Dai et al., 2019). Moreover, the enhanced sea ice melt and associated transports of freshwater to the south weakens the Atlantic meridional overturning circulation and related poleward heat transport on decadal timescales (Sévellec et al., 2017). Therefore, a realistic representation of Arctic sea ice is an essential element of a coupled climate
25 prediction system.

Persistence has been recognized as a primary source of Arctic sea ice predictability in the last ten years (Blanchard-Wrigglesworth et al., 2011; Chevallier and Salas-Mélia, 2012; Chevallier et al., 2019): total SIV is the most persistent variable (∼4 yr) compared to local SIC (∼1 month) and total SIE (∼1 season); local SIT is the second persistent, ranging from seasons in the marginal ice zone (MIZ) to approximately a year in the Central Arctic Ocean (CAO). Moreover, the persistence timescale
30 is found to be variable in a climate system due to advection of sea ice anomalies, heat exchange between ocean and atmosphere, and changes in climate forcing, as reviewed by Guemas et al. (2016). For example, the memory of SIE can reemerge beyond its own persistence (e.g. 2 to 5 months) by storing memory in the upper-ocean heat content or SIT (Blanchard-Wrigglesworth et al., 2011), hence initialization of winter SIT anomalies provides some predictive capability for summer SIE with winter preconditioning in several studies (Holland et al., 2011; Chevallier and Salas-Mélia, 2012; Blanchard-Wrigglesworth and Bitz,
35 2014; Day et al., 2014). As another example, reemergence of sea ice anomalies follows modulation of the upper-ocean heat anomalies: if sea ice retreats anomalously early in spring, more heat is stored in the oceanic upper mixed layer that causes later freeze up (Blanchard-Wrigglesworth et al., 2011).

Sea ice in the Atlantic sector or the Barents Sea can be predicted from a few years to a decade, because of the strong role exerted by the ocean heat advected from upstream and converged on the position of the sea ice edge (Yeager et al., 2015;
40 Årthun et al., 2017; Dai et al., 2020). Some studies suggested that sea ice persistence in the central Arctic can be modulated by oscillation of the Arctic atmospheric circulation between predominantly cyclonic and anticyclonic circulation regimes over timescales of 5-7 years (Proshutinsky and Johnson, 1997; Armitage et al., 2020). From observations, a remarkable oscillation in SIE and SIV is featuring a pause or enhanced ice loss at a period of 7 years, corresponding to some prominent modes of internal variability, such as the winter North Atlantic Oscillation (Bitz et al., 1996; Swart et al., 2015; Gascard et al., 2019).
45 To date most seasonal prediction systems start from a SIT reanalysis data set (Collow et al., 2015; Dirkson et al., 2017), such as from the Pan-Arctic Ice Ocean Model and Assimilation System (PIOMAS, Schweiger et al., 2011). However, most ocean-sea ice reanalyses including PIOMAS do not explicitly assimilate SIT due to inadequate long-term observations of broad coverage. Chevallier et al. (2017) found that the representation of SIT can differ largely due to different local advective processes between models in the intercomparison of 14 state-of-the-art global ocean reanalyses. On the other hand, even fed
50 with identical reanalyzed SIT, there was much less agreement in predicting spatial pattern of SIC across dynamical models in the Arctic coasts than the central Arctic (Blanchard-Wrigglesworth et al., 2017). Such a degradation of forecast skill from the central Arctic to the coasts was found in other idealized multi-model experiments to be associated with advective sea ice processes rather than ice thermodynamics (Tietsche et al., 2014).

At decadal time-scales, some studies indicate that directly assimilating SIT can stimulate long-term forecast drift particularly
55 for the total SIV in the Arctic. Alternatively, anomaly initialization (AI), that is adding observed anomalies to the model climate state, can efficiently suppress the drift but not necessarily improve prediction skill due to the inconsistency between initialized

variables. In comparison of the two initialization methods with the EC-Earth2 climate model, Volpi et al. (2017) found that the Arctic SIV drifts towards a stabilized state (i.e. stable biases) with the observed full field of SIT for lead time longer than 5 years, whereas the drift is substantially reduced by initialization with SIT anomalies. Thus the prediction skill in near surface

60 temperature over the Arctic is improved with AI in comparison with full field initialization (FFI). However, the skill difference is found small in predicting the Arctic sea ice area between the two initialized experiments. Despite the established notion that initial information of SIT is a key source of predictability, assimilating it in operational systems actually remains a challenge.

The Arctic Ocean is changing (Jeffries et al., 2013) with a large loss of multiple year ice (MYI) and a rapid transition from thick perennial towards thinner seasonal sea ice (Wadhams, 2012; Onarheim et al., 2018; Kwok, 2018). This presents a

65 great challenge to predict decadal changes in Arctic sea ice with predictability varying in different regions at different time-scales (Guemas et al., 2016). To our knowledge, most studies on predicting regional Arctic sea ice conditions focused on timescales from a month to a few years (Bushuk et al., 2019; Cruz-García et al., 2019; Kimmritz et al., 2019). Few studies have comprehensively examined the impacts of ocean and sea ice initialization with observed anomalies in a climate prediction system with respect to decadal time scales. A recent study addressed this topic by only assimilating sea surface temperature

70 (SST) and hence had a special focus on increased skills by SST in the Arctic marginal ice zone (Dai et al., 2020).

The objective of this study is to investigate the decadal prediction skill of Arctic sea ice in the EC-Earth3 Climate Prediction System with anomaly initialization (EC-Earth3-CPSAI) to ocean, SIC and SIT states. We developed a novel method to assimilate local SIV anomalies by initializing both SIC and SIT. As the method is developed as a prototype for our initialization strategy implemented for the Coupled Model Intercomparison Project phase 6 (CMIP6) decadal climate prediction project

75 (DCPP) with EC-Earth3, the present study provides a documentation of the new climate prediction system with anomaly initialization including SIT in a multi-category sea ice model framework. It characterizes the performance with focus on the predictions in the Arctic. Three sets of ensemble hindcast experiments are performed and analyzed to evaluate the benefits of respective initialized variables and to quantify the added skill from SIT initialization.

This paper is structured as following: Section 2 introduces the climate prediction system EC-Earth3-CPSAI as well as

80 ensemble-experiment design; Section 3 characterizes the initialized climate prediction system with model bias, forecast drift as well as the imprint of initial conditions in the first year. Section 4 evaluates the benefits of specific initialized variables at inter-annual to decadal scales in terms of temporal smoothing and regional mean, respectively. Section 5 is the summary and discussion.

2 Model system and experiment design

2.1 The EC-Earth3 Climate Prediction System with Anomaly Initialization (EC-Earth3-CPSAI)

85

EC-Earth is a state-of-the-art Earth System Model developed by the EC-Earth consortium (Doescher and the EC-Earth Consortium, in preparation). The core of EC-Earth configures consists of component models for atmosphere, ocean and sea ice, so-called AOGCM. In this study we use the officially released AOGCM configuration of EC-Earth model for contributions to the CMIP6, EC-Earth3 (release v3.3.1.1). The atmospheric component is the Integrated Forecast System (IFS cycle 36r4)

90 developed by the European Centre for Medium Range Weather Forecasts (ECMWF). It uses the TL255 horizontal grid (i.e. triangular truncation at wavenumber 255 in spectral space with a linear reduced Gaussian grid, corresponding to a spacing of about 80 km) and 91 vertical model levels with the top level at 0.01 hPa. The ocean component is the Nucleus for European Modelling of the Ocean, version 3.6 (NEMO3.6) embedded coupled to the Louvain-la-Neuve sea Ice Model, version 3 (LIM3, Rousset et al., 2015). The NEMO-LIM3 is configured with the ORCA1 tri-polar nominal 1° grid and has 75 vertical levels. It

95 is worth noting that the sea ice model LIM3 applies an ice thickness distribution framework to deal with meter-scale variations in ice thickness (Rousset et al., 2015). Unlike its earlier version (e.g., LIM2), LIM3 allows for five ice thickness categories to account for the non-linear dependence of sea ice processes, in particular growth and melt, on ice thickness.

The EC-Earth3 has been used to generate a 25-member ensemble of the CMIP6 historical (1850-2014) and future (2015-2100) scenario simulations following the CMIP6 protocol (Eyring et al., 2016) by the EC-Earth consortium. The ensemble

100 is generated by starting the CMIP6 historical experiment from 25 different dates of a control run with pre-industrial forcing condition. We arbitrarily select one member (r5i1p1f1, hereafter referred to as FREE1) from the ensemble to obtain the model climatology for the ocean and sea ice used in the anomaly initialization. We combine FREE1 and other 4 members to compose a 5-member ensemble (hereafter referred to as FREE) for forecast skill assessment (see Table 1 and sections 2.2).

Table 1. List of experiments, reference data sets and variables in forecast skill assessment

Name	Data sources	Quantities evaluated	Ensemble size
REF	ERAI for atmosphere	TAS	1
	ORAS5 for ocean and sea ice	SIC(SIE), SIT(SIV)	5
FREE	CMIP6 historical, no anomaly init.(AI)	SIC(SIE), SIT(SIV), TAS	5[†]
AI0	AI to ocean	SIC(SIE), SIT(SIV), TAS	5[‡]
AI1	AI to ocean+SIC	SIC(SIE), SIT(SIV), TAS	5
AI2	AI to ocean+SIC+(SIT+SNT)	SIC(SIE), SIT(SIV), TAS	5

[†] FREE are five members from the 25-member ensemble of the CMIP6 historical simulations, initialized from different states (r1, r4, r5, r8 and r18) selected from the 500-year pre-industrial control with EC-Earth3 (piControl, r1i1p1f1 see Fig. S1a). r5 is one of the FREE ensemble, referred to as FREE1; [‡] AI0 (also AI1 and AI2) consists of 5 ensemble members, initialized by 5 sets of ocean and sea ice states, which are based on idential model climatology of FREE1 with observation anomalies from respective 5 ORAS5 ensemble members.

The AI method in decadal climate prediction was formulated by Pierce et al. (2004). This approach has already been applied

105 for the ocean component of the earlier EC-Earth decadal prediction experiments in CMIP5 based on EC-Earth2.3 and using the ORAS4 ocean reanalysis as observational basis (Hazeleger et al., 2013). Since ORAS4 does not provide sea ice data, sea ice anomalies for this previous exercise were obtained from a stand-alone simulation with the ocean-sea ice component of the EC-Earth2.3. For the current study, we apply ORAS5 reanalysis to derive the initial conditions (ICs). One of the advantages of using ORAS5 is that it provides a physically consistent ocean and sea ice states by using NEMO3.4 coupled with LIM2. Zuo

110 et al. (2019) show that, compared to ORAS4, ORAS5 better represents the ocean climate state and variability in terms of SST and sea level, mostly due to increased model resolution from 1° to 0.25° horizontally and from 42 to 75 levels vertically and

due to updates in the assimilated observation data sets, such as sea ice satellite data since 1979. A brief comparison of the two
generations of EC-Earth decadal prediction system with AI is provided in Table S1.

In anomaly initialization, the initial state is generated using the observed state but replacing the observed initial climatology
115 with the modelled initial climatology. Here the observed (modelled) initial climatology at a starting date (i.e., climatology of
November 1) is calculated as an average of the climatological monthly means between the two nearest months (i.e, October and
November), over the period 1979-2014 for ORAS5 (FREE1), as can be seen in Table 2. Horizontally ORAS5 data is bilinearly
interpolated from 0.25 ° to 1 ° ORCA grid. The initialized variables in the ocean model are three dimensional temperature and
salinity. To avoid initial inconsistency in the large-scale dynamics of the system, we do not initialize horizontal velocities. This
120 is a common approach for initialized hindcasts/predictions (see Table 1 in Polkova et al., 2019).

Table 2. Definition of initial/forecast- climatology and anomaly

Categories	Climatology	Anomaly
Initialization	REF[a]: Nov. 1, averaged over the period 1979-2014, obtained by linear interpolation using the monthly mean of October and November	REF[a]: Nov. 1, daily$^{Y=1979,...,2018}$ - climatology^{REFa}
	FREE1: the same as REF[a], but for FREE1	REF[a]
Forecast skill assessment	REF[b]: monthly mean, 20-year mean over the period 1997-2016	REF[b]: monthly$^{Y=1997,...,2016}$ - climatology^{REFb}
	FREE: the same as REF[b], but for FREE	FREE: monthly$^{Y=1997,...,2016}$ - climatology^{FREE}
	AIs: the same as REF[b], but for forecast at lead years [1-10][†], respectively	AIs: as FREE, but for respective forecast lead years

REF[a] and REF[b] are taken from the same REF dataset as in Table 1, but covering different periods.

[†] The first year forecast climatology for the period 1997-2016, denoted as Y1, are calculated using the first-year forecasts from hindcast experiments initiated every year on
November 1, for 1996-2015 (marked by the two red triangles in Fig. 3, while the climatology of 10-year lead time (Y10) are calculated using experiments initiated on
November 1, for 1987-2006, which means the forecasts were initialized 10 years prior to 1997-2016.

The sea ice variables initialized in EC-Earth3 are ice concentration, ice and snow volume (denoted as A^{ice}, V^{ice} and V^{sn})
in five thickness categories at a grid-cell level. However, SIC and SIT from the ORAS5 reanalysis are single values for the
grid-box mean as they are assimilation products using the sea ice model LIM2. Therefore, the volume from ORAS5 is first
calculated by multiplying thickness for ice and snow (SIT and SNT) with SIC. A comparison to climatology yields its anomaly.
125 If added to the model climatology, A^{ice} can in some cases violate the valid range (0-100 %). Therefore, a few adjustments are
made: 1) if $A^{ice} \leq 0$ but $A^{FREE1} > 0$, then A^{ice}=5 % and the ice (snow) thickness H^{ice}=0.1 m (H^{sn}=0.01 m); 2) if $A^{ice} \geq 100$
%, then A^{ice}=99.7 %; 3) adjust $V^{ice,sn}$ again using A^{ice} and $H^{ice,sn}$ obtained in step (1-2). In this study, sea ice initialization
with AI is limited to the region north of 30° N and leaves no modification to the southern area in FREE1.

A major challenge is to distribute the sea ice variable given as one category produced by LIM2 in ORAS5 into five cate-
130 gories in LIM3 in EC-Earth3. Previous studies have explored different solutions in this regard. In sea ice seasonal forecasts,

some assimilate satellite SIC in a multivariate data assimilation scheme so as to update SIT instead of directly assimilating it (Massonnet et al., 2015; Kimmritz et al., 2018), while others use forecast tendencies, namely maintaining the distribution of volume between the categories, by multiplying each category volume with the ratio of observed over modelled mean SIT or similarly by nudging towards observations across each category (Allard et al., 2018; Blockley and Peterson, 2018). However,

135 there is no unique solution, since the multiple sub-grid scale configurations can be compatible with one total SIV. For decadal prediction with EC-Earth3-CPSAI, we develop a novel method with 1) a weighting function mapping single-category SIC onto multiple categories; 2) a multi-category thickness distribution depending on concentration levels; 3) both are used when converting the initial volume (i.e. $V^{ice,sn}$) at the grid-cell level to its sub-grid, while thickness in the last category is determined with a constraint of $V^{ice,sn}$ being conservative.

140   To split A^{ice} into different categories g_l^{ice} (where $l = 1 - L$, denotes the ice category with a total number of categories L=5 in LIM3), the weight-likelihood function is derived from the modelled SIC in a 300-year pre-industrial control run with EC-Earth3 (denoted with superscript "ctrl" hereafter). Data for November 1 was calculated by averaging October and November monthly means. For SIC at a specific grid point (i.e. A^{ctrl} has only time dimension), tn is the time index over 300 years, at which $A^{ctrl}(tn)$ has the minimal difference from A^{ice}. We assume that based on the same EC-Earth3 model version, g_l^{ice} is

145 likely regulated by the weighting function $weight_l^{ctrl}$. Thus for each grid point,

$$A^{ice} = \sum_{l=1}^{L} g_l^{ice}, \tag{1}$$

$$weight_l^{ctrl} = g_l^{ctrl}(tn)/A^{ctrl}(tn), \tag{2}$$

150 $$g_l^{ice} = weight_l^{ctrl} A^{ice}. \tag{3}$$

We aim at imposing local SIV anomalies to LIM3 while keeping the sum of volume over all thickness categories unchanged as in Eq.(4). We use the 300-year control run to establish the relationship between ice thickness in the five ice categories (h_l^{ctrl}) and the total ice concentration A^{ctrl}. Different from $weight_l^{ctrl}$ in Eq.(2), Fig. 1 plots the h_l^{ctrl}-A^{ctrl} histogram for all grid points in the control run, with A^{ctrl} ranging from 10 to 100 % at intervals of 10 %. It can be seen that the thickness distributions for

155 h_l^{ctrl} ($l = 1 - 4$) are quite robust within each bin, suggesting that the more ice-covered (e.g. A^{ice}>70 %), the lower variance in thickness, in other words, lower probability to melt and shift to neighboring bins. We neglect SIT initialization when A^{ctrl} is below 10 %, both because statistically these grid points only account for 8 % of total ice-covered ones and because an observation error of 10 % is often assumed while assimilating SIC (Mathiot et al., 2012).

$$V^{ice} = \sum_{l=1}^{L} g_l^{ice} h_l^{ice}. \tag{4}$$

160 $$v_L^{ice} = V^{ice} - \sum_{l=1}^{L-1} g_l^{ice} h_l^{ice}. \tag{5}$$

Following the h_l^{ctrl}-A^{ctrl} histogram, except for the last category, h_l^{ice} ($l \leq 4$) are determined by the corresponding bin of A^{ice}, while the volume in each category is determined by $v_l^{ice} = g_l^{ice} h_l^{ice}$. Then the residual of V^{ice} will be accommodated in the last category ($l = L = 5$) as v_L^{ice} in Eq.(5) and h_L^{ice} is contingently resolved. This method imposes the combined anomalous signals of SIC and SIT to sea ice initialization. The same method is applied to discretize snow volume with a multi-category

165   snow thickness-A^{ctrl} histogram (not shown).

A^{ice}-h_l^{ice} **relationship**
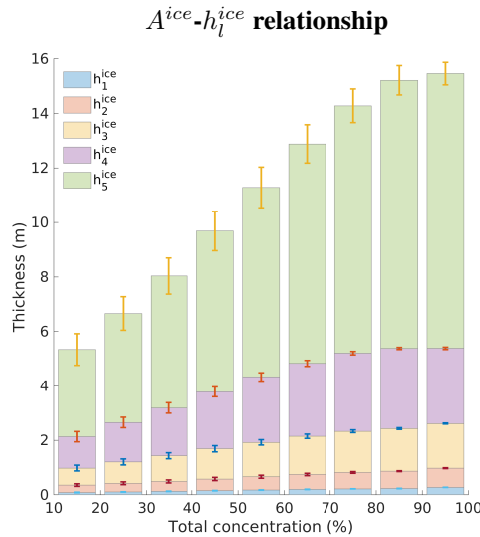
Figure 1. Multi-category ice thickness h_l^{ice} ($l =$1-5) distribution on total ice concentration A^{ice} at intervals of 10% with mean (column) and standard deviation (error bar), based on the knowledge from a 300-year climate simulation with pre-industrial forcing. It is used to discretize h_l^{ice}, if the total ice concentration and volume (A^{ice} and V^{ice}) at a grid cell is know from anomaly initialization. F. ex. for grid cells with A^{ice} between 10-20 %, h_l^{ice} ($l =$1-5) is 0.08, 0.27, 0.62, 1.15 and 3.20, respectively. The ice volume at respective category ($l =$1-4) is determined by $v_l^{ice} = g_l^{ice} h_l^{ice}$, where g_l^{ice} is splitted from A^{ice} to each category by using a weight-likelihood function (see Eq.1-3). For the 5th category, h_5^{ice} is determined by v_5^{ice}/g_5^{ice}, where $v_5^{ice} = V^{ice} - \sum_{l=1}^{4} g_l^{ice} h_l^{ice}$ in Eq.(5). Note that the grid-box mean thickness $H_m = V^{ice}/A^{ice} = \sum_{l=1}^{5} v_l^{ice} / \sum_{l=1}^{5} g_l^{ice} \neq \sum_{l=1}^{5} h_l^{ice}$. Therefore, H_m is regulated by both V^{ice} and g_l^{ice} at the local grid.

A consistency check, also called "sanity check", is carried out for the ICs of the ocean and sea ice model, in order to adjust water and heat flux-relevant variables in the surface boundary fields to the initialized g_l^{ice}, v_l^{ice} and v_l^{sn} in a physically consistent way. The method has been used in different sea ice prediction systems (Massonnet et al., 2015; Kimmritz et al., 2018). To complete the initialization, an one-day spin up integration is performed with a reduced time-step of 100s and introducing a mask

170   of coastal water (< 100 m deep) to neglect SIT initialization there (indicated in Fig. 2b). This is to ensure dynamical consistency in the initial states and avoid numerical instabilities. The practice of masking some locations out in SIT initialization has been used in many studies, for example by introducing a threshold of SIC>40 % to implement full-field initialization (Blockley and Peterson, 2018) or confining modification of SIT to the Arctic basin with a geographic weighting mask to discard initialization changes in the marginal ice zone (Blanchard-Wrigglesworth et al., 2017).

2.2 Sensitivity experiments with sea ice initialization

The above described AI approach for both ocean and sea ice ICs is hereafter referred to as AI2. A five-member ensemble of the AI2 experiment is performed in this study, with 5 sets of ocean and sea ice ICs generated using the anomalies of five individual members of ORAS5, respectively. For the atmosphere a full-field initialization based on ERA-Interim (hereafter as ERAI, Dee et al., 2011) is implemented. Initialized ensembles of predictions (re-forecasts) start yearly on November 1 for the period from 1979 to 2018 (a total of 40 start dates) and run 2 months plus 10 years long with the external forcing following the CMIP6 DCPP protocol for dcppA-hindcast and dcppB-forecast experiments (Boer et al., 2016). We generated 10 additional members by means of perturbed atmospheric ICs. This whole ensemble with a total of 15 members states a contribution to CMIP6 DCPP with EC-Earth3-CPSAI (see Table S1). However, for this study only the first 5 ensemble members (with unperturbed atmospheric ICs) are used and compared to complementary 5-member ensemble predictions described in the following.

Our primary interest is the impact of Arctic sea ice on decadal prediction skill in the last two decades, which is expected to be more representative for the coming decades than an Arctic with large amounts of thick MYI in the last century. In order to investigate the sensitivity of decadal prediction skill to SIT initialization in EC-Earth3-CPSAI, two more initialized ensemble experiments are performed with ocean-only initialization (AI0) and ocean plus SIC initialization (AI1). We only performed 5 members with AI0 and AI1 for sensitivity analysis, and thus for a fair comparison this study focuses only on 5 members for AI2 and FREE, too. For AI2, our assessment here uses the five members initialized with the five members of OARS5. FREE consists of FREE1 and other four members from the 25 member ensemble of the EC-Earth CMIP6 historical (1979-2014) and the corresponding SSP2-4.5 (2015-2017). These four members are deliberately selected to represent the wide range of natural variability in the EC-Earth3 CMIP6 control experiments from which the ensemble of EC-Earth historical simulations starts (Fig. S1a). An assessment of the overall feature of FREE shows no significant difference between FREE and the full ensemble of 25 members (e.g. Fig.S1b), even though the regional differences could be large. A summary of all experiments is given in Table 1. The ensemble-mean of the AI experiments versus that of FREE are used to evaluate the impact of the respective initialization approach. The difference between AI2 and AI0 is used to assess the added prediction skill with sea ice initialization, while the comparison between AI2 and AI1 is used to evaluate the skill gained by SIT initialization.

2.3 Skill assessment

As reference fields (REF), the climatological annual and seasonal means are calculated as 20-year averages for the period from 1997 to 2016 based on monthly means of ORAS5 for sea ice and ERAI for the near surface air temperature (TAS), respectively. We note that the reference data is not from another independent observation product, but assimilation of observations using models which are close to those in EC-Earth3. There are several reasons: 1) ORAS5 represents Arctic sea ice reasonably well compared to other ocean reanalysis, thanks to direct assimilation of SST and SIC observations with 5-d assimilation window (Tietsche et al., 2017; Uotila et al., 2019; Zuo et al., 2019); 2) the aggregated quantities (e.g. SIE and SIV) are not sensitive to models, and the errors of ORAS5 fall within the range of observation uncertainties (up to 10% in SIC and SIV, Schweiger et al., 2011; Chevallier et al., 2017); 3) ORAS5 and FREE are both in ORCA grids, which avoid spatial errors potentially being

either masked or enhanced by remapping from observation- to model-grid; 4) Similar to Volpi et al. (2017), our main focus is to evaluate the relative skills between different initialization methods with EC-Earth3-CPSAI, compared to the skill of FREE.

210   Our assessment is based on forecast anomalies (Table 2), rather than absolute errors. For skill assessment, it always uses the full 20-year period 1997-2016 for the hindcasts of each lead year, e.g. the assessment for lead year 1 (Y1) includes the first year re-forecasts initialized from 1996,..., 2015, while the assessment for lead year 10 (Y10) includes the 10th year re-forecasts started from 1987,..., 2006. The 20-year forecast climatology is calculated for individual lead year, therefore the forecast anomalies vary depending on the respective forecast lead year (following the recommendation for CMIP6-DCPP, see

215   Boer et al., 2016). This data selection process guarantees the use of all reforecasts data available for the period of interest and at the same time a consistent estimation of the model and reference climatologies (García-Serrano and Doblas-Reyes, 2012). The metrics of anomaly correlation coefficient (ACC) and the Root Mean Square Error (RMSE) with respect to REF are computed for specific lead time, following the method by Volpi et al. (2017). The confidence interval is calculated with a t-distribution for the ACC, and with a χ^2 distribution for the RMSE. The assessment of temporal development is provided at different lead

220   times: year 1, years 2-5, years 6-9, and years 2-9. The latter three averaging periods are selected as basic choices to verify skill dependence on averaging and lead time (Goddard et al., 2013).

Our analysis focuses on three key variables, SIC, SIT and TAS in the Arctic. Our assessment for total SIE and SIV is applied over the Northern Hemisphere region (NH) with a typical threshold 15 % SIC and 0.15 m SIT, respectively, to exclude the extensive areas of open water (Schweiger et al., 2011). TAS index is computed as field average over the polar cap domain,

225   namely the region north of the 70° N circle (see Fig. 2). The data analyzed for this study are briefly summarized in Table 1. Regional skills are assessed for ten sub-regions (shown in Fig. 2b), representing the sector for the CAO and its adjacent waters, namely the central Arctic (80°N north) and the Laptev/E. Siberian/Chukchi/Beaufort Seas, and the sector for the Arctic MIZ and the transition waters, namely the Kara/Barents/GIN (Greenland/Iceland/Norwegian) Seas as well as the Baffin/Hudson Bays. The CAO sector is coincidently confined by the climatological September ice edge (see Fig. 2a and 2b), implying

230   the dominant influence of thick MYI due to geometric constraints of Arctic coastlines. By contrast, the MIZ sector adjoins the North Atlantic Ocean, climatologically covered by thinner seasonal ice in winter (except for the east coast of Greenland receiving thick ice transported from CAO through Fram Strait). Canadian Archipelago and the Pacific side of MIZ are not presented in regional assessment, however, they are counted in the assessment for the total Arctic SIE and SIV.

3   Characterization of the initialized climate predictions with EC-Earth3-CPSAI

3.1   Components of sea ice initialization

The model bias for the initial dates (i.e. November 1) is calculated over the period 1979-2014 between the modelled and reference mean state that are used for obtaining the sea ice anomaly initialization (formulated in Table 2). SIC is generally overestimated (i.e. blue areas dominate over the red ones in Fig. 2a), and particularly there is up to 20 % more SIC in FREE1 than in REF in the Baffin Bay and from the eastern Greenland coasts towards Svalbard. However, in the Pacific Arctic, e.g. the

240   outlet of Chukchi Sea, and along the coast of the E. Siberian/Laptev/Kara Seas, SIC is 15-20 % less than in REF. The largest

biases are mostly located outside of the observed summer sea ice cover, illustrated by the climatological September sea ice edge of ORAS5 (i.e. REF). As a result, the modelled climatology of September sea ice is also overestimated over the GIN/Barents Seas, consistent with 1-2 m thicker ice as seen in Fig. 2b. This indicates that the sea ice biases over rather large areas in the GIN Seas originate from the modelled MYI. Compared to FREE1, FREE (i.e. ensemble mean) shows similar patterns of model

245 bias but with increased magnitudes (due to a large spread in the FREE ensemble, not shown).

Model bias (FREE1-REF)
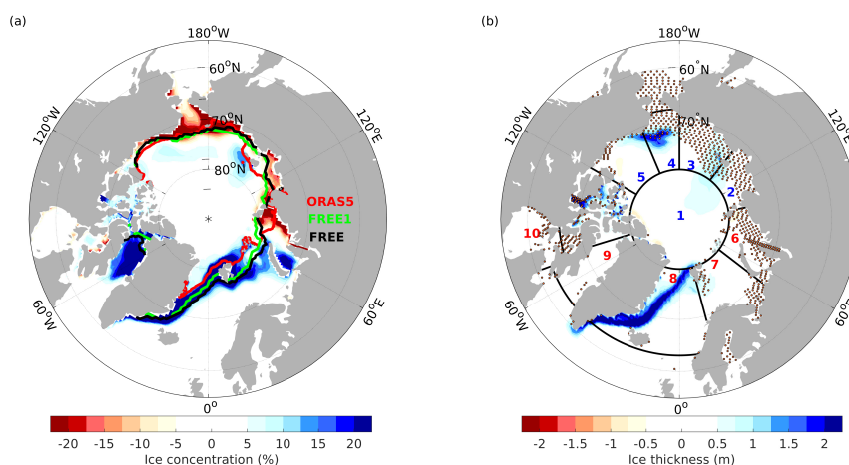
Figure 2. Model bias with respect to ORAS5 (REF) on 1 November for the period 1979-2014 for the single-member FREE1. (a) SIC. Color lines indicate the September sea ice extent (15 % SIC) climatology for ORAS5 (red), FREE1 (green) and FREE (black). (b) SIT with dots indicating a mask for area with water depth <100 m where no SIT initialization is applied. The maps have the bounding latitude at 56°N. Blue areas represent regions with more SIC and SIT in the model than REF, reflecting cold bias in surface temperature. Numbers 1-5 in blue denote the sector for the Central Arctic Ocean (CAO) and its adjacent waters, including (1) central Arctic (80°N north), (2) Laptev Sea, (3) East Siberian Sea, (4) Chukchi Sea and (5) Beaufort Sea. Numbers 6-10 in red denote the sector for the Arctic Marginal Ice Zone (MIZ) and the transition waters, including (6) Kara Sea, (7) Barents Sea, (8) GIN (Greenland/Iceland/Norwegian) Seas, (9) Baffin Bay and (10) Hudson Bay.

Figure 3 shows the initial anomalies from REF on November 1 in each year during 1979-2018 aggregated for the NH SIE and SIV. The positive anomaly of SIV has reduced by 2/3 from ∼15 to ∼5 thousand km^3 in the early 1980s. Since then, the anomalies of SIV and SIE both declined linearly with years and shifted from positive to negative values in the early 2000s, with respect to the mean of the whole period 1979-2014 used for initialization. The 20-year averages of initial anomalies from

250 1996 to 2015 are -0.11 million km^2 for SIE and -3 thousand km^3 for SIV, respectively.

In the next section, we focus on the impact of Arctic sea ice initialization with negative anomalies on decadal prediction skill during recent two decades. As the climate warming continues and the CAO might become seasonally ice free in the future. Our goal is to shed light on how the key mechanism governs the decadal predictability of the Arctic sea ice in the coming decades.
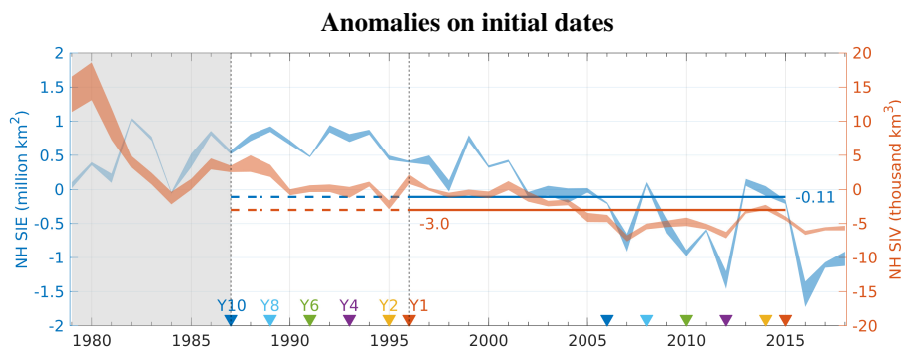
Figure 3. Initial anomalies of SIC (SIT) on 1 November from REF aggregated for the NH sea ice extent (volume) in blue (red) during 1979-2018. The ensemble spread (between min and max) is presented by filled area. The 20-year averages of initial anomalies for lead year 1 (Y1, between two red triangles) are -0.11 million km^2 for SIE and -3.0 thousand km^3 for SIV (red and blue straight lines), respectively. An backwards extension to 1987 are shown in dashed lines. Note that the skill assessment period is from 1997 to 2016 (see Table 2), where Y1 denotes forecast initialized from 1996, ..., 2015 and Y10 denotes the 10th year forecast initialized from 1987, ..., 2006. Thus the initial anomalies are more positive for predictions with longer lead years assessed, as can be seen when compared to those in Y1.

3.2 Forecast drift

255 Figure 4 shows the monthly forecast error of SIE, SIV and TAS for different initialization methods as a function of forecast lead time in the 10-year prediction. The monthly forecast error is determined by subtracting the climatological monthly mean of REF for the period of January 1997 to December 2016, rather than forecast anomaly as defined in Table 2. To compare, the monthly biases of FREE (FREE1) for the respective variables were shown as gray (black dashed) lines and filled areas in Fig. 4a-c, which are simply repeated annual cycles averaging for the same period (1997-2016) with annual mean bias of 0.4

260 (-0.2) million km^2, 5.6 (2.3) thousand km^3 and -0.7 (0.1) K for SIE, SIV and the Arctic cap TAS, respectively. It is by chance that FREE1 (black dashed line) has smaller bias in SIE, SIV and TAS over the Arctic than the ensemble mean of FREE (thick gray line in Fig. 4, also see FREE bias in annual min/max sea ice and mean TAS in Fig. S2 and S3).

A general feature of all three variables is that biases in all initialized experiments (AIs) vary in the range between FREE1 and the ensemble mean of FREE, so that initialization results in a positive annual mean bias to the SIE and SIV (i.e. larger SIE

265 and SIV) and negative annual mean bias to TAS with respect to FREE1, taking the annual mean bias of AI2 (blue dots) as an example. There is a slight tendency that the forecast error increases for longer lead time. The larger biases in AIs experiments than FREE1 result from anomaly initialization where the model bias of FREE1 is not removed from the initial state, but surplus with more positive anomalies of sea ice for prediction with longer lead years assessed, compared to those in the first year forecast over the period 1997-2016 (i.e. Y1 in Fig. 3). The differences between AIs are generally small after Y1, indicating

270 the forecasts are drifted toward the model climate (as represented by the ensemble mean). And the forecast error for SIE is relatively less sensitive to different initialization than that of SIV and TAS. On average the long-term forecast drift is small as indicated by the annual mean errors in AI2 (blue dots in Fig. 4). There is a slight decline in both SIE and SIV from Y1 to Y3 followed by a return between Y5-Y7, and a tendency of larger SIE and SIV for longer lead time. Correspondingly, TAS in the

Geoscientific
Model Development
Discussions

initialized hindcasts is gradually pulled to the colder climate over the Arctic cap domain, showing a tendency of negative bias
275 in all AIs in both winter and summer for longer lead time, i.e. Y8-Y10.

The model biases show rather strong seasonal cycles, with generally smaller biases in winter but much larger in summer for all three variables in AIs than FREE. These results emphasize the importance of drift-correction via correcting the lead-time dependent bias for multi-year prediction skill assessment. Among all AIs, AI1 (pink line) shows the least magnitude of positive seasonal bias of SIE and SIV from Y6 afterwards, and correspondingly the warmest TAS. It seems that, for longer lead times
280 (>Y5), AI1 performs closest to FREE1. It suggests that AI1 with ocean plus SIC initialization imposes less strong constraints on the development of the sea ice forecast than ocean-only (AI0) and all (AI2) initialization at decadal timescales.
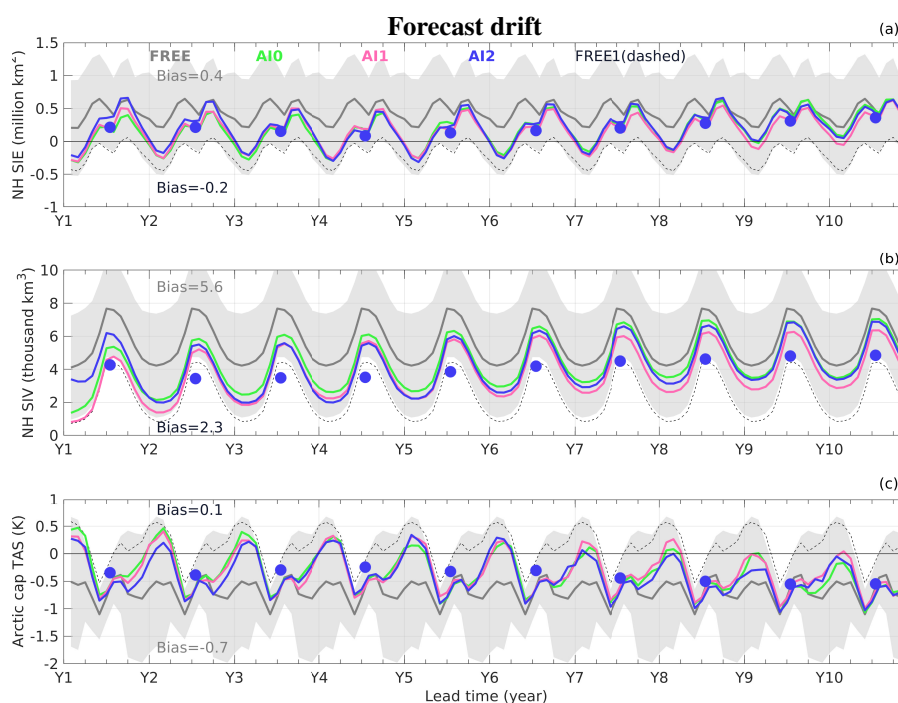
Figure 4. The development of forecast drift for hindcast experiments using different initialization methods for the NH sea ice extent (a), volume (b) and TAS over the polar cap domain (c), respectively. The drift is computed as difference of monthly climatology over 1997-2016 between forecast for a specific lead time and REF. The time series is shown as 3-month moving average and therefore the first three months after initialization (i.e. Nov, Dec and Y1-Jan) are masked out. The colored lines plots the ensemble means of the experiments initialized with ocean-only (i,e, AI0, Green), ocean plus SIC (AI1, pink) and ocean plus SIC and SIT (AI2, blue), respectively. The un-initialized experiments FREE are shown as repeating annual cycles for ensemble mean (gray line) and spread ($\pm 1\sigma$, gray shading), as well as the member used for initialization (FREE1 in black dashed line). The annual mean of AI2 are marked with solid dots for each forecast years.

Geoscientific
Model Development
Discussions

3.3 Imprint of initial conditions in the first year

In this section we examine the immediate benefit (degradation) from initial anomalies (ICs inconsistency) due to different initialization strategies.

285 #### 3.3.1 First winter forecast

The ACC and RMSE of SIC, SIT and TAS with respect to REF are calculated at each grid-point for forecast of the first winter (December-January-February) of 1996-97 to 2015-16. Figure 5 depicts the ACC for the first winter predictions with areas stippled if the Pearson correlations are not significant (p=0.05). Effects from the ocean-only initialization (AI0) are seen as an increased correlation of SIC and TAS in the Barents Seas and the Hudson Bay (Fig. 5, first row compared to the second 290 one). The improvement from AI0 to AI1 by adding SIC initialization is found minor. AI0 increases ACC significantly in the Barents Sea and the Hudson Bay, while AI1 seems counteracting the benefits from AI0 in the two regions with lower ACC and smaller area with significant positive correlation. Comparing AI2 to AI1, the SIT initialization significantly enhances the high correlation areas for SIC in some parts of the Baffin Bay and east of the Kara Sea. For SIT, AI2 results in the largest area of significant positive correlation in the CAO as well as in the East Greenland Current (EGC). For TAS predicted by FREE, 295 it shows the largest area with significant positive correlation over the Arctic Ocean than these with AIs. The major benefits of AI2 is seen outside of the polar cap domain, manifested as significantly enhanced correlations over the North Atlantic Ocean including small areas in the Baffin Bay and the northernmost part of Greenland.

Figure 6 illustrates the RMSE of FREE with respect to REF, and the RMSE skill scores for the forecast SIC, SIT and TAS of the first winter. In the first row the largest RMSE (up to 25 % of SIC and 1 m of SIT) as simulated in FREE coincides 300 with regions where there are largest initial biases shown in Fig. 2. In order to identify the benefit of initialization for specific model components, we evaluate the RMSE skill score (RMSESS), which compares the RMSE of AI2 ensemble to the other experiments, i.e.

$$RMSESS = 1 - RMSE_{AI2}/RMSE_{INIT} \tag{6}$$

where *INIT* denote experiments with different initialization methods, i.e. FREE, AI0 or AI1, respectively. A positive RM-305 SESS indicates better accuracy (smaller RMSE) of AI2 compared to *INIT*. Specifically, a positive RMSESS relative to FREE (AI2/FREE hereafter) indicates benefit from both ocean and sea ice initialization, the RMSESS relative to AI0 (AI2/AI0 hereafter) indicates benefit from sea ice (SIC and SIT) initialization, while the RMSESS relative to AI1 (AI2/AI1 hereafter) singles out the benefit of SIT initialization.

For SIC (Fig. 6, left panels), the skill score (~0.05) appears to be unaffected by any initialization in the central Arctic, 310 because it is constantly covered by sea ice in wintertime. AI2/FREE is generally positive (benefit from all init.) in the outskirt of the central Arctic, whereas negative (less skillful) in the Bering Strait and EGC, coinciding with the maximum RMSE of SIC in FREE. The relatively higher score in AI2/FREE compared to AI2/AI0 in the Hudson Bay and along the ice edge in the North Atlantic sector indicates a dominant contribution from AI0 (i.e. ocean-init.).
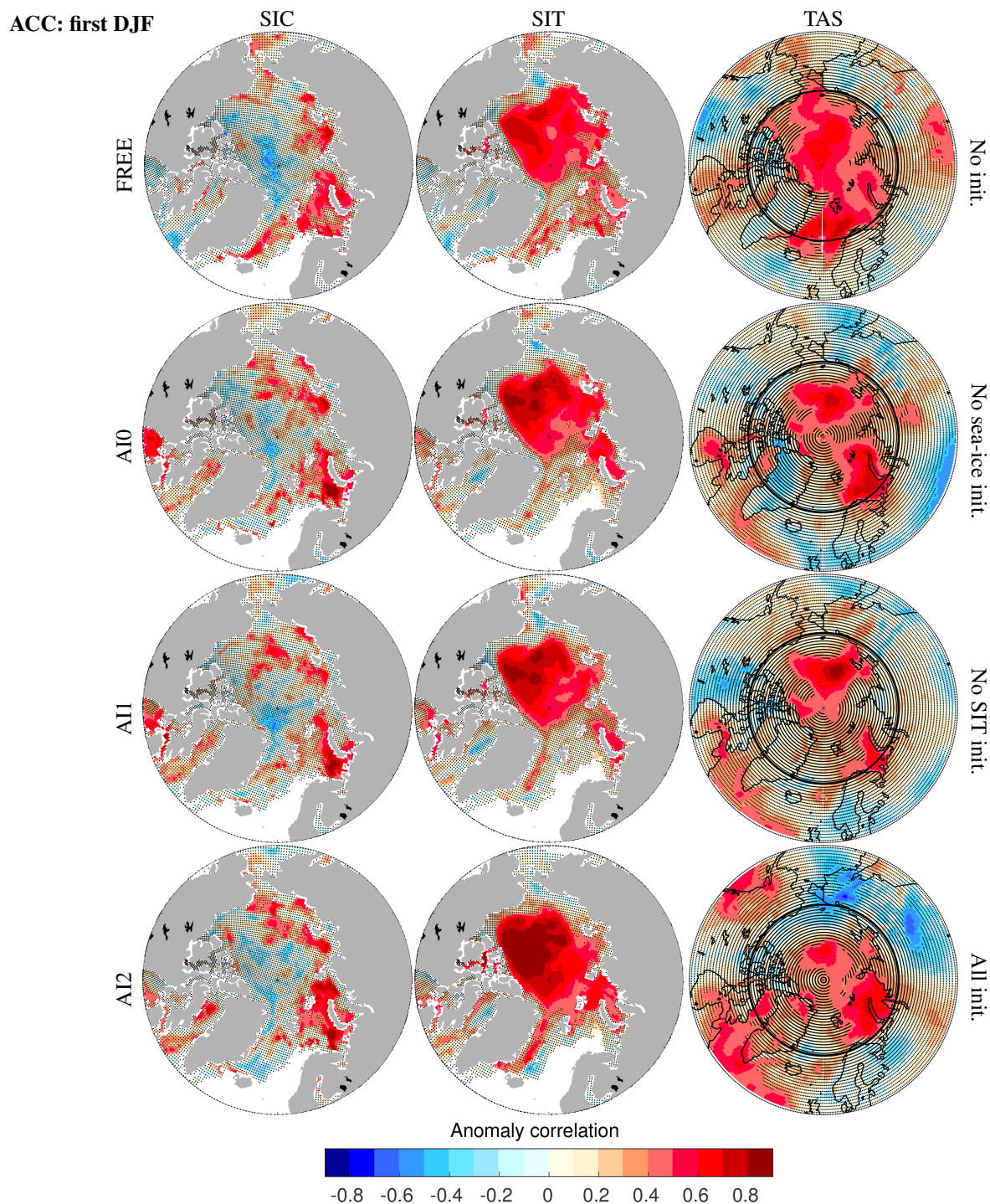
Figure 5. Anomaly correlation of the first winter SIC (left), SIT (middle) and TAS (right), respectively for FREE, AI0, AI1 and AI2 experiments from top to bottom. The first winter forecast evaluated here is the DJF-mean after initialization on 1 November, on each year of 1996-2015. The reference data are taken from ORAS5 or ERAI. Regions are stippled if not significant (p = 0.05). The black line illustrates the polar cap domain (north of 70° N).
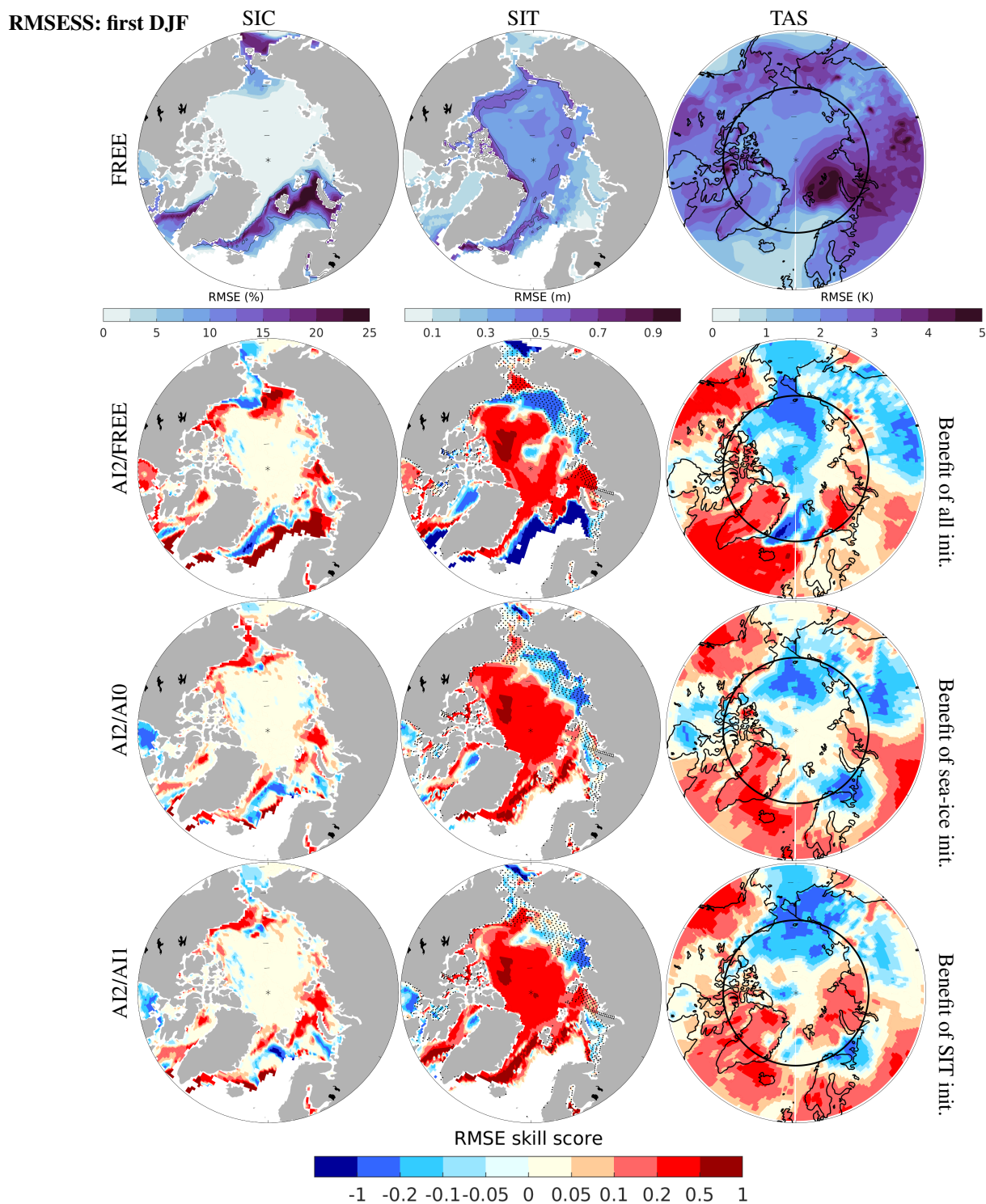
Figure 6. RMSE of FREE with respect to REF (top row) and RMSE skill score (lower rows) for the first wintertime (DJF) forecast: SIC (left), SIT (middle) and TAS (right). The contour lines mark RMSE≥10% or 0.5 m for SIC and SIT, respectively. The RMSE skill score is calculated as 1-(RMSE$_{AI2}$/RMSE$_{INIT}$), where *INIT* denotes FREE (no init.), AI0 (no sea-ice init.) and AI1 (no SIT init.). Scores above 0 denote more accurate in AI2 than *INIT*, and vice versa. The regions discard SIT initialization in AI2 are stippled.

Comparing skill scores for SIT (Fig. 6, middle panels), RMSESS>0.2 is commonly found poleward of the September sea ice

315 edge determined by FREE (Fig. 2a), suggesting that the higher skill of AI2 compared to FREE, AI0 and AI1 in the region is
related to higher persistence of thick MYI (Kwok, 2018). Low skill in the E. Siberian/Laptev Seas is because SIT initialization
is not applied in AI2 to these regions due to the shallowness (stippled in Fig. 6), while worse skill in the Laptev Sea than AI1
suggests inconsistency between the local SIC initialization and SIT advected from the neighboring CAO regions. There are two
regions showing opposite signs of RMSESS between AI2/FREE and the others. One along the ice edge in the North Atlantic

320 sector shows consistently high score (>0.5) in both AI2/AI0 (sea-ice init.) and AI2/AI1 (SIT init.), indicating the enhanced skill
by initializing SIT, whereas negative RMSESS (<-1) in AI2/FREE (all init.), suggesting counteracting between ocean and SIT
anomalies of the initial states in AI2. The other region is in the Hudson Bay, where, on the contrary, improvements in SIC and
SIT are only seen in AI2/FREE (RMSESS>0) but not in AI2/AI0 and AI2/AI1, indicating a dominant role of ocean dynamics
in shaping the growth of sea ice in this region.

325 The impact of initialization on reduction of TAS RMSE is mostly outside of the polar cap domain, with RMSESS >0.2 in
the North Atlantic Sector of the Arctic including Greenland and the Alaska Peninsula, whereas RMSESS<-0.2 in the Pacific
Sector of the Arctic and eastward to the Siberian region. Similarities between all RMSESS figures suggest that the TAS skills
are mostly attributable to the SIT initialization. Differences between AI2/FREE and AI2/AI1 indicate that SIT initialization is
beneficial to the Kara Sea whereas the Barents Sea is more regulated by ocean processes. FREE is best for TAS in large parts

330 of the CAO with respect to both ACC and RMSE, presumable because the atmospheric large scale circulation is undertaking
adjustment to the initialized states in the first few months.

In summary, the direct improvement in RMSESS for the first winter prediction from AI0 is SIC in the MIZ sector, from the
Hudson Bay eastwards to the Barents Sea. The added value from the ocean plus SIC initialization (i.e. AI1) is small and the
only way to improve the SIT skill is to impose local SIV anomalies to the initial sea ice state (i.e. AI2 methods). Taking the sea

335 ice model bias from FREE for March (in Fig. S2a and b) into account, the SIT initialization seems more effective on improving
the first winter prediction skills for regions where the model tends to have too much ice (which is mainly the North Atlantic)
while it works less well in regions where the model has too little ice (the Pacific Arctic sector).

3.3.2 The first 12 months forecast

As the persistence timescales of upper-ocean heat content and SIT can be longer than one season, we continue to identify

340 the relative contribution from ocean and sea ice initial constraints as the dominant source of predictability in the first forecast
year. The prediction skill in the first 12 months, counted from the start month (i.e. November), is evaluated by analyzing the
temporal development of ACC and RMSE relative to REF (as Eq.1-2 in Volpi et al., 2017). The definition of forecast and
reference anomalies is given in Table 2. The monthly mean of total SIE and SIV are calculated and extracted from the first
12 months forecast initialized on November 1 from 1996-2015 as indicated in Fig. 3. In the same way, the first 12 months

345 forecasts of TAS is averaged over the polar cap domain (see Fig. 2a) in order to investigate the direct impact from MYI in the
CAO on the atmosphere, as most of the region north of 70 °N is sea ice covered year-around in the climatology (Fig. S2c and

d). When the correlations of AIs meet, or even fall below FREE in Fig. 7, it means that there is no skill from the initialization any more. On the contrary, when the RMSE of AIs exceeds the FREE one, it indicates no benefit from initialization afterwards.

For SIE, all AIs show significantly smaller RMSE than FREE until June, but with similar ACC throughout of the year 350 (Fig. 7a, d). For SIV, AI2 shows significant reduction of RMSE from the first months to March-April, and remains low until August, even though not much difference in ACC except the first month (Fig. 7b, e). For TAS, there are hardly any differences among the initialized or uninitialized runs for both RMSE and ACC (Fig. 7c, f). In summary, with respect to SIV, only AI2 (all init.) outperforms FREE for the first 8-10 forecast months. This suggests the importance of SIT initialization with constraint on SIV anomaly in EC-Earth3-CPSAI. The increased skill for SIE by initializing SIT is only prominent in the first winter and 355 very little for longer lead time into the year. The similarity between AI0, AI1 and AI2 suggests the major contribution of ocean initialization to reduce RMSE until lead month 11. Similar results are also seen in the study by Dai et al. (2020) in which the coupled climate prediction system NorCPM only assimilating SST can skillfully predict pan-Arctic SIE up to 12 months.
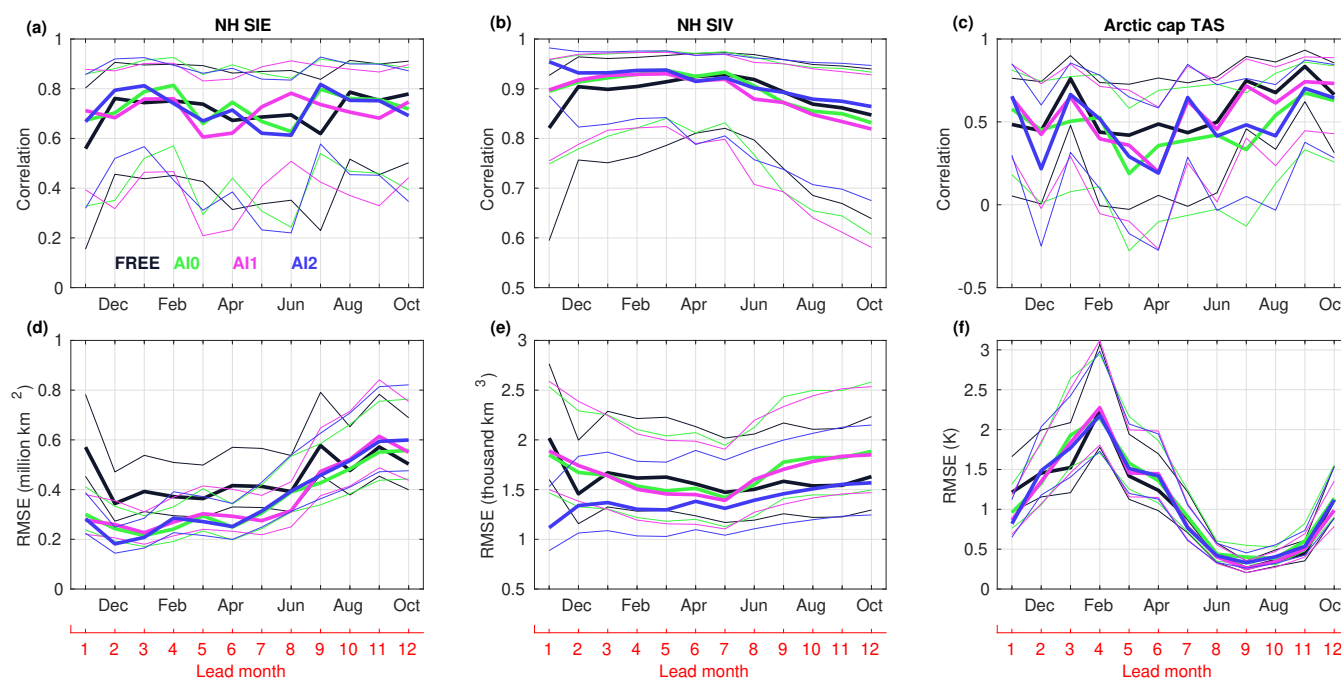
Figure 7. Time evolution of anomaly correlation (a-c) and root mean square error (RMSE, d-f) for the NH SIE (million km^2), SIV (thousand km^3) and TAS over the polar cap domain (K), against REF for the first 12 months forecasts started on November 1 from 1996-2015. It is calculated for each lead time with respect to its monthly climatology. The thin lines represent the 95 % confidence interval obtained with a t-distribution for the correlation and a χ^2 distribution for the RMSE.

17

Geoscientific
Model Development
Discussions

EGU

4 Decadal-scale skill assessment

In this section we focus on understanding the origin of decadal predictability and the relative contribution from the upper-
360 ocean heat content and local SIT. For simplicity, the results of AI1 are not discussed in the following as the improvement
by SIC initialization is seen limited in comparison with ocean-only initialization (Fig. S4-S7). The initial year is excluded in
this section because the imprint of initial conditions is strong and discussed above. Following the verification framework for
interannual-to-decadal prediction experiment (Goddard et al., 2013), the temporal information is smoothed at different scales:
years 2-5, years 6-9, and years 2-9.

365 ### 4.1 Year 2-9 average predictions

Prediction skill is assessed based on the average over forecast years 2-9 for AI0 and AI2 and compared to the respective FREE
projection. Figure 8 shows that the areas with significantly positive correlation (>0.5) are larger in the year 2-9 than that for
the first winter (Fig. 5) for all three variables and all experiments. It is because temporal smoothing will typically reduce
higher frequency noise and increase signals resulting from external forcings such as increase of greenhouse gas concentrations.
370 Therefore, FREE in the top panels of Fig. 8 shows high anomaly correlation in both SIT and TAS in the CAO region, likely
associated with the trends in local SIV under Arctic warming over the recent two decades. For SIC, AI0 presents a dominant
role of ocean initialization. Both AI0 and AI2 show similar spatial patterns with higher positive correlations than in FREE in
eastern Arctic between the E. Siberian Sea to the Barents Sea. For SIT, the correlation in AI2 is significantly enhanced up to 0.7
between the central Arctic and the E. Siberian Sea, compared to AI0. Although SIT initialization was not locally implemented
375 to the E. Siberian Sea, sea ice is moving driven by advective processes or winds into the E. Siberian Sea (Guemas et al., 2016),
which results in enhanced correlations of SIC and SIT in both AI0 and AI2 at longer lead time. The correlation of TAS in AI0
is significantly increased in the Baffin Bay and the Eastern Arctic seas as well as its landward vicinity, relative to FREE. The
positive effect is further enhanced in AI2 to the outside of the polar cap domain, i.e. from the coastal area of the Barents Sea to
the E. Siberian. However, the significant skill in the EGC via Fram Strait in FREE is degraded by sea ice initialization.
380 With respect to RMSE in FREE (Fig. 9, first row), the spatial patterns derived from the year 2-9 average are very similar to
those of first winter means shown in Fig. 6 but the magnitude is smaller. There is a strong contrast of forecast errors between
the CAO and the GIN/Barents Seas in Fig. 9, which is coherent with the positive (negative) bias in the annual sea ice maximum
(minimum) in the GIN (Barents) Seas over 1997-2016 (Fig. S2). Both observations and climate models have suggested that in
these regions the variability of sea ice is largely influenced by the oceanic heat flux convergence, therefore, the prediction errors
385 can be greatly reduced by advection of ocean temperature anomalies whereas little benefit from SIT initialization is expected
(Koenigk and Mikolajewicz, 2009; Årthun et al., 2017; Onarheim et al., 2018; Bushuk et al., 2019; Dai et al., 2020).

The improvements in AI2 (all init.) compared to FREE for SIC decadal prediction are presented in the skill map of AI2/FREE
(Fig. 9, left). The high RMSESS (>0.2) seen in the MIZ from the first winter mean (Fig. 6) is not only maintained here in years
2-9 but also becomes prominent in the central Arctic and the Pacific Arctic sector. However, AI2 shows limited improvement
390 for regions where RMSE is highest in FREE (as mentioned in the paragraph above). With respect to SIT, the regions with

improved skill (RMSESS>0) in AI2 compared to FREE are shifted from the western Arctic in the first winter (Fig. 6) to the eastern Arctic in years 2-9. Comparing the lower panels of Fig. 9, the RMSESS of AI2/FREE for both SIC and SIT in the North Atlantic sector are much higher than those in AI2/AI0. This suggests that ocean initialization is the most important source of predictive skill at decadal time-scale, however, sea ice initialization holds some additional value in this respect. The

395 positive effect of initialization (seen in both AI2/FREE and AI2/AI0) on sea ice skill between Chukchi Sea and Laptev Sea is only prominent in the year 2-9 average, but absent in the first-winter mean. This manifests as the regional benefit on longer time-scales, rather than a direct impact from sea ice initialization in the first winter prediction.
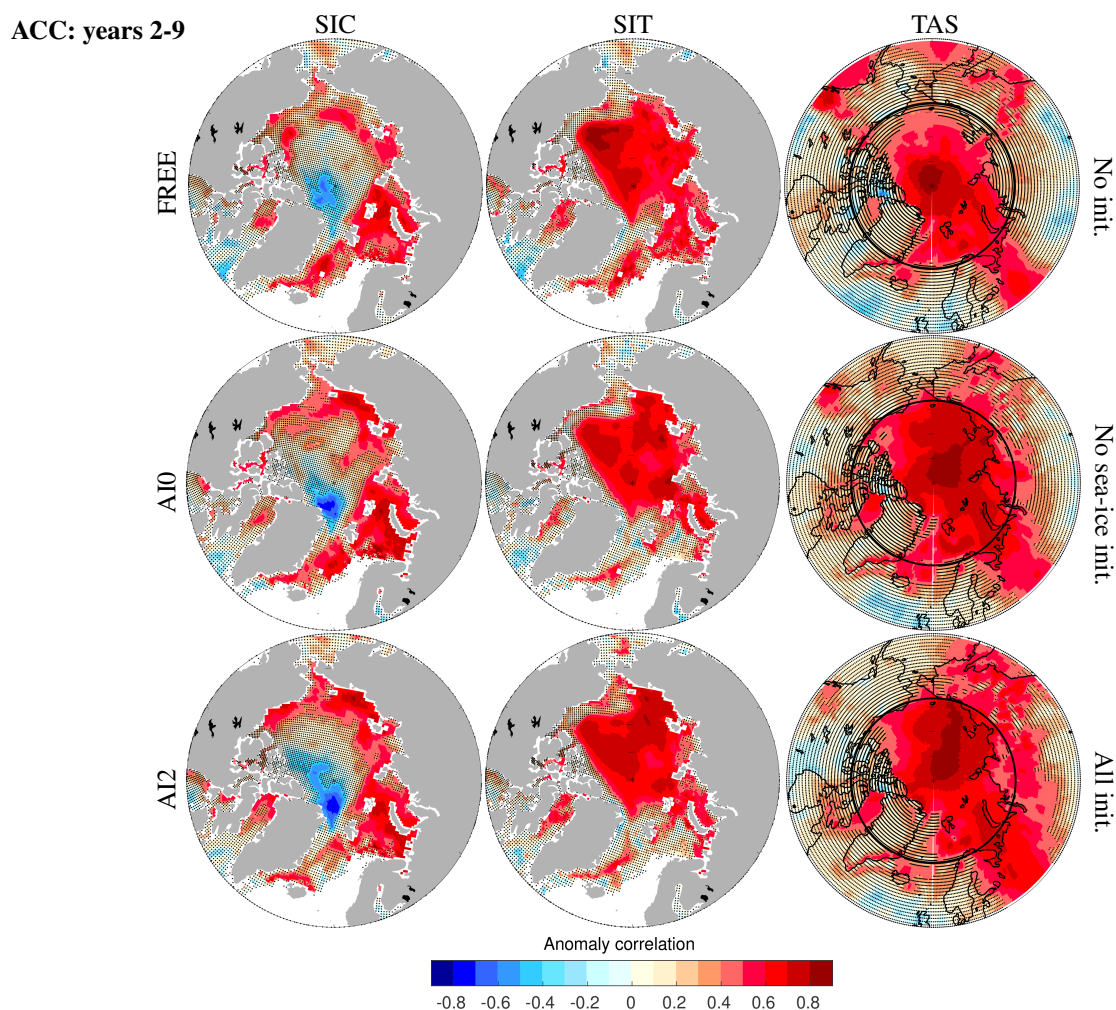
Figure 8. Same as Fig. 5, but for years 2-9.

The positive skill score (AI2/FREE in red) for TAS indicates that hindcasts with AI2 (all init.) are more skillful than unini-tialized ones both inside and outside the polar cap domain, except the Pacific Arctic and the sea ice outflows from the central

400 Arctic through EGC. With respect to the north Atlantic subpolar gyre (SPG) region, TAS over the western SPG (the Labrador

Basin) is seen to benefit more from the initialization than the east side of SPG. Such a zonally asymmetric pattern is consistent with the leading mode of observed sea surface height (SSH) variability in the North Atlantic (Koul et al., 2020), suggesting large scale atmospheric response to ocean initialization. The added skill by sea ice initialization represented by RMSESS of AI2/AI0 appears also prominent in the landward vicinity of the Barents Sea and the E. Siberian Sea. This is consistent with the

405   enhanced correlation of AI2 relative to AI0 (Fig. 8).
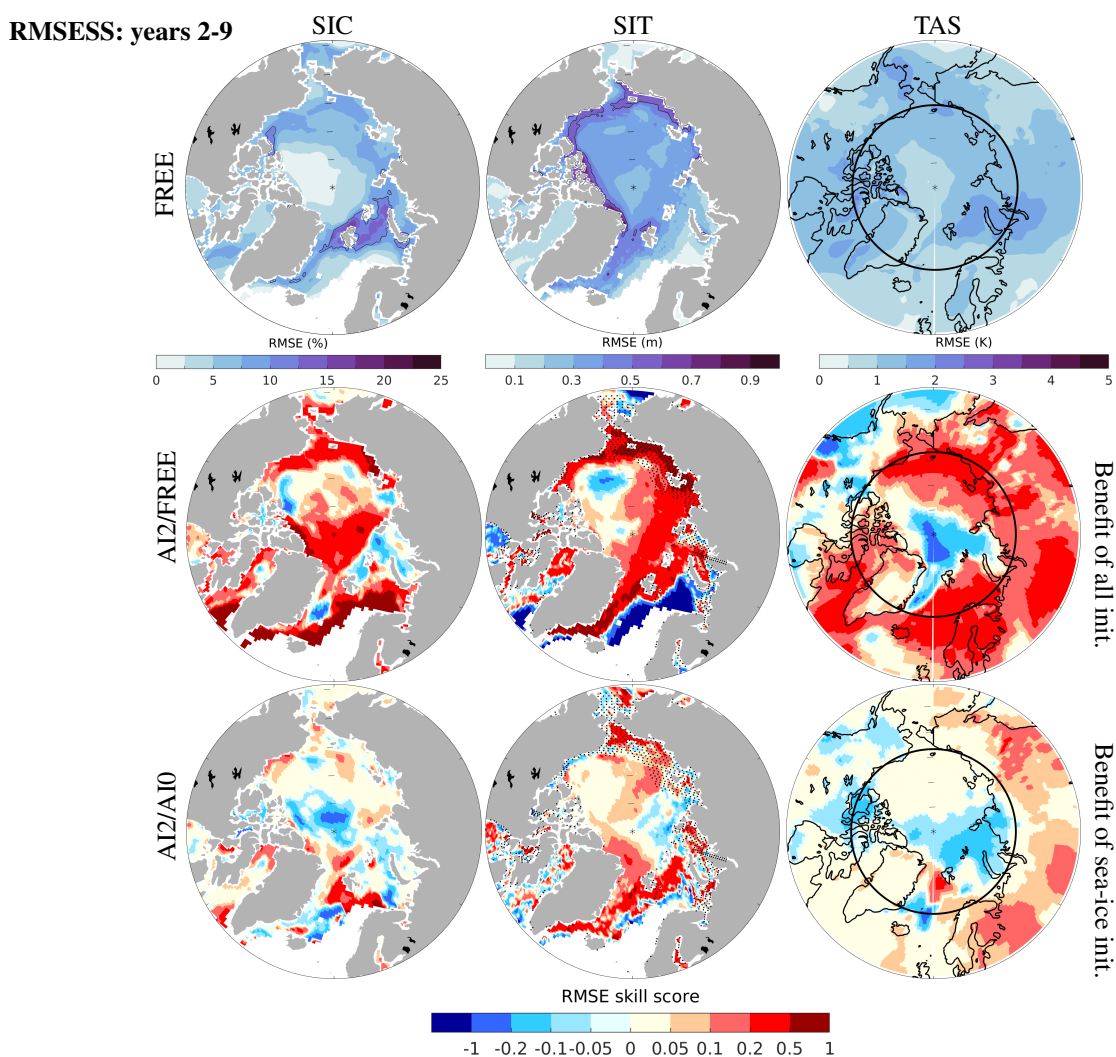
Figure 9. Same as Fig. 6, but for years 2-9.

The skill of years 2-5 (see Fig. S4-S5) are very similar to those of years 2-9 (Fig. 8-9). After forecast year 1, the skills gained by AI2 initialization relative to FREE seem not to be sensitive to the averaging time-scales from multi-year to a decade. In examining the RMSESS (AI2/AI0) dependence on lead time via comparison of skills for years 2-5 versus 6-9 (Fig. S6-S7), it is found that the added skill for SIC and TAS is more prominent in years 6-9 than in years 2-5 in the North Atlantic sector (e.g.

410 the GIN/Barents Seas), whereas it is the opposite in the Beaufort Sea. However, these small differences of regional skill are averaged out at the decadal time-scale.

4.2 Regional-mean skill matrix

The results from sections 3.3.1 and 4.1 provide important insights into the regional variability of sea ice prediction skill at different time-scales. It is clear that the benefit from SIT initialization in shaping the local sea ice development do not last

415 beyond the first summer, but its impact can on remote regions which may last for several forecast years resulting from accurate representation of ocean dynamics. In the final analysis, we aim at providing details on the temporal evolution of the skills of initialized hindcasts for the ten sub-regions of the Arctic. The RMSESS are calculated on monthly basis for each grid points in years 1-9 and then the area means are used to represent its regional skill for the ten regions in the Arctic as defined in Fig. 2b.

In Fig. 10, regional skill AI2/FREE is assessed for SIC for the CAO (left column) and MIZ (left second column) sectors,

420 respectively. The right two columns show the skill score difference between AI2/FREE and AI0/FREE, to single out the contribution of sea ice initialization to the prediction skill of AI2 in the corresponding regions. Red (blue) colors denote positive (negative) value of results, indicating better (worse) accuracy of AI2 than the reference hindcasts (i.e. FREE or AI0). Among the ten regions, the Barents Sea is the most beneficial region due to AI2, with the highest score (> 0.5) from July to next February in several forecast years ahead. The positive difference in the right column indicates enhanced skills from AI0

425 to AI2 between August and October in particular for longer lead time of 5-6 years. The second most beneficial region is the GIN Seas showing similar temporal patterns as the Barents Sea but the maximum score is evident in year 1. Conversely both Baffin Bay and Hudson Bay show the worst skill of AI2 (RMSESS<-0.5) in autumns. AI2/FREE is less skillful than AI0/FREE (Fig. 10, right column), indicating a negative effect of sea ice initialization in the region. Both the most positive and negative effects of sea ice initialization on top of ocean initialization are found in the MIZ regions.

430 In the CAO sector, there is little benefit from initialization in SIC prediction from November to next March, as expected for the ice growth months when the regions are fully ice-covered. But a slight increase in skill (RMSESS∼0.1) is found in the central Arctic and the Laptev Sea in the melt season of August and September. Consistently in the neighboring Kara Sea (the transition between the Laptev and Barents Seas), good skill (RMSESS>0.2) is gained between September and October at a lag of one month. In the Pacific Arctic sector, namely the E. Siberian/Chukchi/Beaufort Seas, AI2 generally shows moderate skills

435 (0.2>RMSESS>0) in November and December. With respect to the skill difference (AI2/FREE - AI0/FREE), a minor role of the sea ice initialization for the skill is found in the CAO and its adjacent waters.

In contrast to SIC, the skill of SIT is high (RMSESS∼0.5) in both CAO and MIZ regions for several years (Fig. 11). This suggests the direct improvement from SIT initialization. Overall, sea ice initialization (AI2) adds more skill than ocean-only initialization (AI0) in predicting SIT in all years in the GIN Seas. In the CAO region, the added skill (AI2/FREE-AI0/FREE,

440 right panels in Fig. 11) emerges in years 1-3 in the central Arctic and years 1-4 in the Beaufort and Chukchi Seas, and then reemerges in years 5 and 7, respectively. Interestingly AI2 exhibits 12-month lag of added skill in the Laptev/E. Siberian Seas. It corroborates with the finding in the RMSESS maps with no skill in first winter but high skill in years 2-9 (see AI2/AI0 in

Fig. 6 versus Fig. 9). Their regional SIT predictability may be attributed to advective processes or winds (Guemas et al., 2016), because SIT initialization is neglected due to the shallowness (see Fig. 2b).
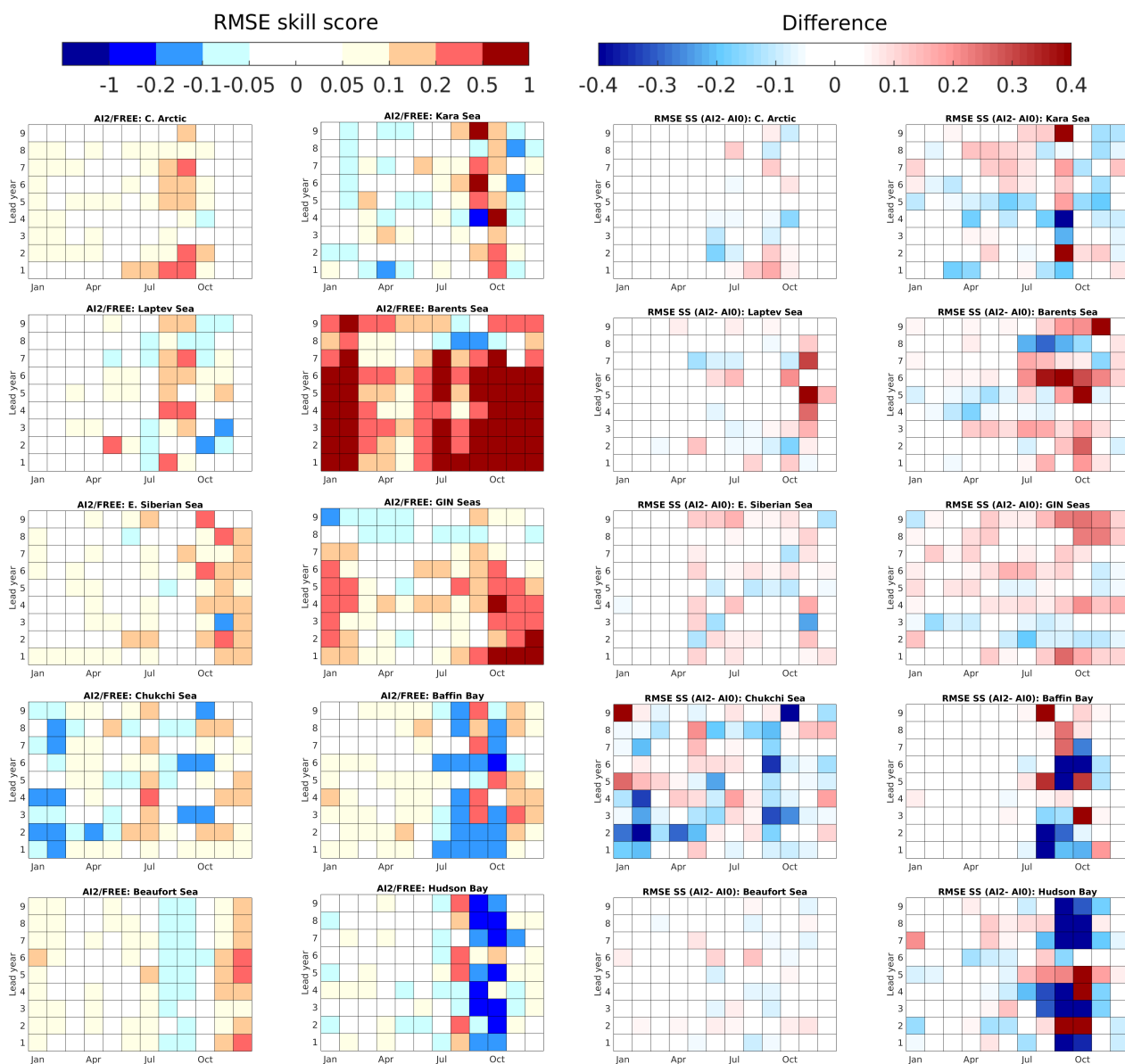

Figure 10. Regional Arctic SIC prediction on decadal time-scales: RMSE skill score of AI2 relative to FREE (AI2/FREE, two columns on the left) and the difference of skill score between AI2 and AI0 (two columns on the right). The ratio of RMSE in AI2 to FREE is averaged over regions, and then AI2/FREE is calculated as $1 - (\mathrm{RMSE}_{AI2}/\mathrm{RMSE}_{FREE})$. White colors denote 0 score, which means RMSEs in AI2 and FREE are equal. The difference is calculated as AI2/FREE - AI0/FREE. Red (blue) colors denote higher (lower) skill score than the reference.
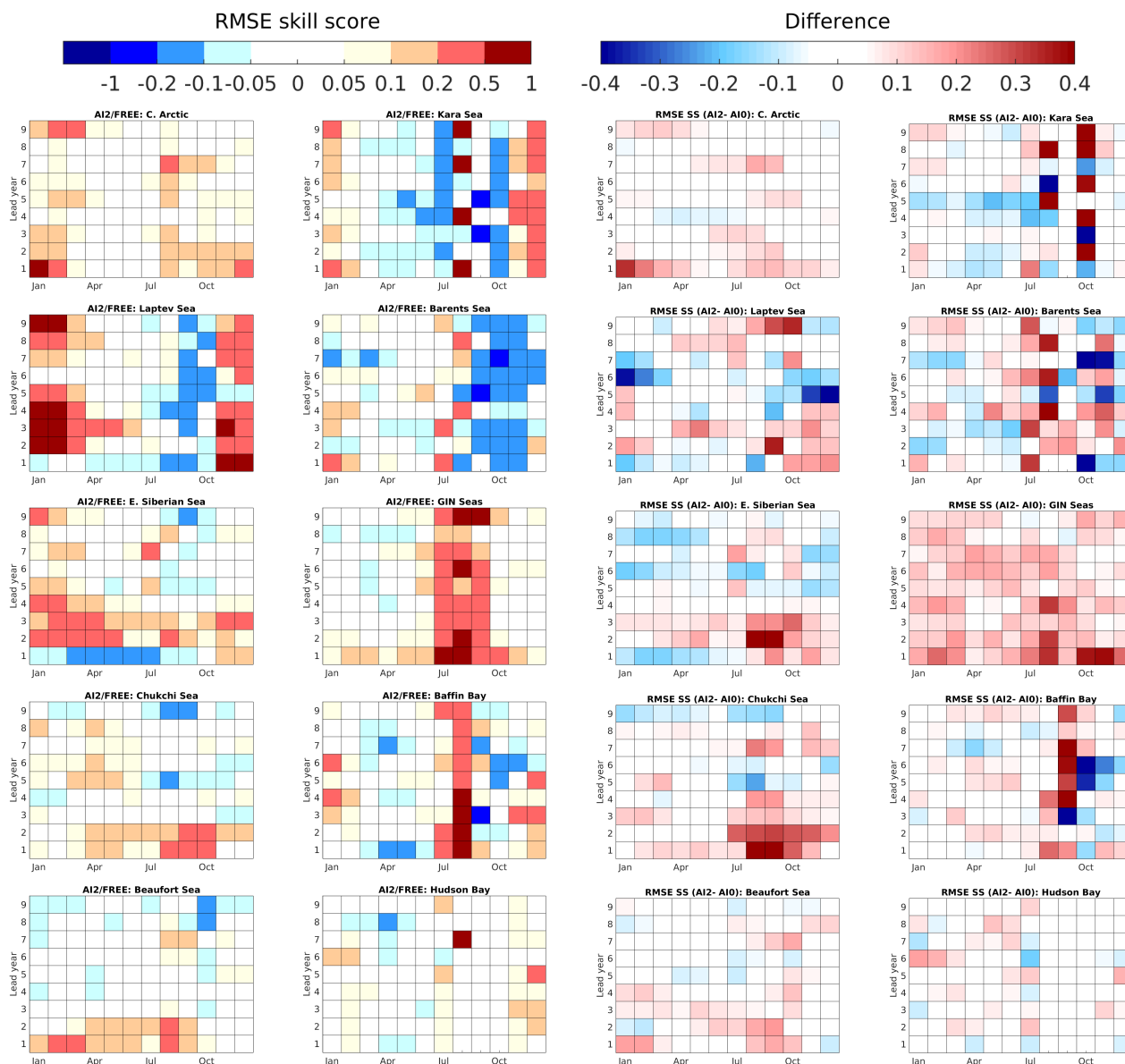
Figure 11. As Fig. 10, but for SIT.

In summary the difference of RMSESS between AI2 and AI0 for both SIC and SIT indicate that the impact of sea ice initialization on reducing sea ice forecast errors is not just limited to the first few years locally, but can also reemerge after 5-7 forecast years in the CAO regions extending to the transition zone of the Kara Sea, suggesting a combined effect of ocean advection and sea ice initialization. The skill gained by initializing ocean can be enhanced by SIT initialization on timescales ranging from few years to a decade. Additionally, the RMSESS matrix for the regional SIE and SIV are provided for each lead month in Fig. S8-S9 and the results from RMSESS are generally consistent with the skill assessment in Fig. 10-11. There is

generally significant skill in regional SIV in the CAO regions in the ice-growth months during the forecast years 1-9. Only the Barents Seas shows significant skill in both SIE and SIV from January to June in years 1-9. Otherwise there is almost no significant skill in the MIZ regions and in regional SIE.

5 Discussion and conclusions

455   This study addresses the following questions using the global climate model EC-Earth3: can SIT initialization improve the Arctic decadal prediction skill? Where and when may the prediction of regional seas benefit from SIT initialization? Three predictability regimes are classified according to added skill by ocean (AI0-FREE) and SIV (AI2-AI0) anomalies in the initialization:

– The local SIV anomalies (anomalous signal combined from SIC and SIT) in the central Arctic imposed by AI2 (with
460   SIT) initialization provide winter preconditioning for the pan-Arctic September SIE, resulting in the long-lasting persistence of the SIV that can play a role for a decade. This corroborates the findings by Cruz-García et al. (2019) that the central Arctic SIV and the pan-Arctic SIE are correlated in September over three continuous years in all six global climate models used in their "idealized" experiments. Our results show that the regions with significant skill of SIT are consistently confined to the interior of the Arctic Basin in all hindcasts (including FREE in Fig. 6 versus Fig. 9), namely
465   the CAO sector inside the September climatological ice edge (Fig. 2a). This highlights the persistence of local thick ice. Theoretically, the variability of thick ice has little connection with the upper ocean, due to the insulating role played by the sea ice cover during most of the year (Flato, 1995). From observations, the variability and trend of perennial SIE are found to be the largest in the September sea ice minimum, in contrast to seasonal SIE which characterizes as ice-free in summer and by the largest variability and trend in winter (Onarheim et al., 2018). Here, we found a clear evidence of
470   a winter-preconditioning mechanism in the central Arctic with increased skill for SIC in the following summer, which drops in years 3-4 and reemerges in years 5-7. Furthermore, the emergence of skill in SIC and SIT in the Laptev/E. Siberian Seas from the second winter onward may demonstrate the role of the SIT anomalies in the central Arctic over the CAO and a combined effect of advection processes on reshaping the spatial pattern of sea ice in these two regions.

– In the Atlantic side of MIZ with predominant seasonal ice, variability and trend of SIE is largely influenced by oceanic
475   heat flux in winter. Therefore, the errors in predictions can be greatly reduced by advection of ocean temperature anomalies. Many global climate models have shown that the upper ocean heat content significantly contributes to the prediction skill of sea ice in the MIZ (Bushuk et al., 2019; Cruz-García et al., 2019; Dai et al., 2020). Therefore, little improvement from local SIT can be expected in these regions. Consistent with their results, we found that there is significant degradation of SIC skill in the Hudson Bay in the first winter prediction in AI2 compared to AI0. Similarly SIT initialization
480   seems not improve the local SIC prediction in the Baffin Bay. Among these regions, the prediction skill of SIC in the Barents Sea is the highest and up to 7 years during the melt-to-growth seasons (July to next February in Fig. 10). This is likely attributable to the persistence of SST anomalies and advection of ocean temperature anomalies from the North

Atlantic Ocean, as there is no direct increase in skill from SIT (AI2-AI0) in the first winter months (Fig. 6). In terms of ACC for years 2-9, significant skill in predicting SIC is found mostly in the eastern Arctic Basin from E. Siberian Sea to the Barents Sea at decadal-scales (Fig. 8). All hindcasts show generally higher skill of SIC in the Arctic MIZ in the Atlantic Arctic sector than in the Pacific Arctic sector, consistent with the results in Dai et al. (2020).

– In the GIN Seas, best skill in September SIT ($\sim$ 9 years) and winter SIC ($\sim$ 7 years) are associated with advection of SIT anomalies from the central Arctic and SST anomalies from the Atlantic Ocean, respectively. In summer, there is only ice in the EGC coming from the north through Fram Strait. In winter, the ice thickness in the EGC is still determined by the ice transport through Fram Strait but the extension of the ice cover might be dominated by upper surface temperature and ocean circulation. Regarding the temporal evolution of regional means, AI2 improves the accuracy of SIT predictions in all years, particularly for the September minimum, suggesting a remote origin of MYI from the central Arctic in Fig. 11. In contrast, the best skills of SIC in the GIN Seas are gained in the winter months from October to February (Fig. 10), appearing in the marginal ice zone adjacent to CAO. The GIN Seas may actually be divided into a region in the west at the Greenland coast, which is dominated by MYI coming from CAO and a region in the east, which behaves rather similar to the Barents Sea: ice melts in summer and winter ice depends on winds and ocean currents.

The added value of SIC initialization is limited in comparison with ocean-only initialization (i.e. AI1-AI0). As SIV is calculated as a product of SIC and SIT, only assimilating SIC will result in a synchronous correlation between the total sea ice area and SIV (of the same month) in the initial states, which is found to be weak both from a multi-centennial preindustrial simulation (Chevallier and Salas-Mélia, 2012) and from the observed anomaly (i.e. ORAS5 reanalysis in Fig. 3). In turn, the incorrect initial SIV fields can affect the forecast errors of sea ice through the melting season and thereby degrade prediction skill of SIE (Blockley and Peterson, 2018; Kimmritz et al., 2018). Therefore, assimilating SIC alone in EC-Earth3-CPSAI may not be sufficient to constrain winter SIT variability to predict decadal variations in SIE and SIV under Arctic warming.

To conclude, our sensitivity experiments with EC-Earth3-CPSAI by imposing different initialized model components demonstrate that AI2 (all init.) yields an improved performance for decadal prediction for the Arctic regions, as it provides an improvement in predicting SIE and SIV anomalies and reducing errors in regional sea ice states. As climate warming continues, the central Arctic that is covered mostly by MYI will likely become seasonal ice free in the future. The controlling mechanism for decadal predictability in the region may thus shift from the current SIV persistence dominated regime to a more ocean-related processes. These findings state the foundation for the AI2-approach being the choice for a full contribution to CMIP6-DCPP covering 60 initializations (1960-2019) with 15 ensemble members each. A more general assessment of this system's predictive skill beyond the Arctic is currently in preparation.

Code and data availability. The EC-Earth model (version 3.3.1.1) with its standard coupled model configuration (T255L91-ORCA1L75) is used for the experiments here. The entire code of EC-Earth is not available due to restrictions in the distribution of the atmosphere component IFS. Confidential access to the entire code can be granted for editors and reviewers; please use the contact form at http://www.ec-earth.org/about/contact. For the methods of anomaly initialization to ocean and sea ice, we followed the approach described by Hazeleger

et al. (2013), namely by adding observation anomaly to model climatology; this can be implemented at one command line with the utility of Climate Data Operator (CDO). The programmes used to convert one-category sea ice initial states (i.e. SIC, SIT and SNT) to 5 categories and the scripts used to produce the figures are available at https://doi.org/10.5281/zenodo.4297603 (Tian et al., 2020). Data used in this paper are available at https://doi.org/10.5281/zenodo.4297926 (Tian, 2020). Links to model output of sensitivity experiments can be found from

520   the aforementioned URL. The CMIP6 data (e.g. FREE and AI2) can also be downloaded from any Earth System Grid Federation (ESGF) data portal.

References

Allard, R. A., Farrell, S. L., Hebert, D. A., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A., et al.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system, Adv. Space Res., 62, 1265–1280, 2018.

Armitage, T. W., Manucharyan, G. E., Petty, A. A., Kwok, R., and Thompson, A. F.: Enhanced eddy activity in the Beaufort Gyre in response to sea ice loss, Nat. Commun., 11, 1–8, 2020.

Årthun, M., Eldevik, T., Viste, E., Drange, H., Furevik, T., Johnson, H. L., and Keenlyside, N. S.: Skillful prediction of northern climate provided by the ocean, Nat. Commun., 8, 15 875, 2017.

Bitz, C. M., Battisti, D., Moritz, R., and Beesley, J.: Low-frequency variability in the Arctic atmosphere, sea ice, and upper-ocean climate system, J. Climate, 9, 394–408, 1996.

Blanchard-Wrigglesworth, E. and Bitz, C. M.: Characteristics of Arctic sea-ice thickness variability in GCMs, J. Climate, 27, 8244–8258, 2014.

Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M., and DeWeaver, E.: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations, J. Climate, 24, 231–250, 2011.

Blanchard-Wrigglesworth, E., Barthélemy, A., Chevallier, M., Cullather, R., Fučkar, N., Massonnet, F., Posey, P., Wang, W., Zhang, J., Ardilouze, C., et al.: Multi-model seasonal forecast of Arctic sea-ice: forecast uncertainty at pan-Arctic and regional scales, Clim. Dyn., 49, 1399–1410, 2017.

Blockley, E. W. and Peterson, K. A.: Improving Met Office seasonal predictions of Arctic sea ice using assimilation of CryoSat-2 thickness, Cryosphere, 12, 3419–3438, https://doi.org/10.5194/tc-12-3419-2018, 2018.

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., et al.: The decadal climate prediction project (DCPP) contribution to CMIP6, Geosci. Model Dev., 9, 3751–3777, 2016.

Bushuk, M., Msadek, R., Winton, M., Vecchi, G., Yang, X., Rosati, A., and Gudgel, R.: Regional Arctic sea–ice prediction: potential versus operational seasonal forecast skill, Clim. Dyn., 52, 2721–2743, 2019.

Chevallier, M. and Salas-Mélia, D.: The role of sea ice thickness distribution in the Arctic sea ice potential predictability: A diagnostic approach with a coupled GCM, J. Climate, 25, 3025–3038, 2012.

Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., et al.: Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project, Clim. Dyn., 49, 1107–1136, 2017.

Chevallier, M., Massonnet, F., Goessling, H., Guémas, V., and Jung, T.: The role of sea ice in sub-seasonal predictability, in: Sub-Seasonal to Seasonal Prediction, pp. 201–221, Elsevier, https://doi.org/10.1016/B978-0-12-811714-9.00010-3, 2019.

Collow, T. W., Wang, W., Kumar, A., and Zhang, J.: Improving Arctic sea ice prediction using PIOMAS initial sea ice thickness in a coupled ocean–atmosphere model, Mon. Weather Rev., 143, 4618–4630, 2015.

Cruz-García, R., Guemas, V., Chevallier, M., and Massonnet, F.: An assessment of regional sea ice predictability in the Arctic ocean, Clim. Dyn., 53, 427–440, 2019.

Dai, A., Luo, D., Song, M., and Liu, J.: Arctic amplification is caused by sea-ice loss under increasing CO_2, Nat. Commun., 10, 1–13, 2019.

Dai, P., Gao, Y., Counillon, F., Wang, Y., Kimmritz, M., and Langehaug, H. R.: Seasonal to decadal predictions of regional Arctic sea ice by assimilating sea surface temperature in the Norwegian Climate Prediction Model, Clim. Dyn., pp. 1–16, https://doi.org/10.1007/s00382-020-05196-4, 2020.

Day, J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, Geophys. Res. Lett., 41, 7566–7575, 2014.

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, d. P., et al.: The
575   ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly J. R. Meteorol. Soc., 137, 553–597, 2011.

Dirkson, A., Merryfield, W. J., and Monahan, A.: Impacts of sea ice thickness initialization on seasonal Arctic sea ice predictions, J. Climate, 30, 1001–1017, 2017.

Doescher, R. and the EC-Earth Consortium: The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6, Geosci.
580   Model Dev. Discuss., in preparation.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, 2016.

Flato, G. M.: Spatial and temporal variability of Arctic ice thickness, Ann. Glaciol., 21, 323–329, 1995.

García-Serrano, J. and Doblas-Reyes, F.: On the assessment of near-surface global temperature and North Atlantic multi-decadal variability
585   in the ENSEMBLES decadal hindcast, Clim. Dyn., 39, 2025–2040, 2012.

Gascard, J.-C., Zhang, J., and Rafizadeh, M.: Rapid decline of Arctic sea ice volume: Causes and consequences, The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-2, 2019.

Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., Kharin, V., Merryfield, W., Deser, C., Mason, S. J., et al.: A verification framework for interannual-to-decadal predictions experiments, Clim. Dyn., 40, 245–272, 2013.

590 Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., Fučkar, N. S., Germe, A., Hawkins, E., Keeley, S., et al.: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales, Quarterly J. R. Meteorol. Soc., 142, 546–561, 2016.

Hazeleger, W., Guemas, V., Wouters, B., Corti, S., Andreu-Burillo, I., Doblas-Reyes, F., Wyser, K., and Caian, M.: Multiyear climate predictions using two initialization strategies, Geophys. Res. Lett., 40, 1794–1798, 2013.

595 Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in coupled models, Clim. Dyn., 21, 221–232, 2003.

Holland, M. M., Bailey, D. A., and Vavrus, S.: Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3, Clim. Dyn., 36, 1239–1253, 2011.

Jeffries, M. O., Overland, J. E., and Perovich, D. K.: The Arctic shifts to a new normal, Phys. Today, 66, 35, 2013.

Kimmritz, M., Counillon, F., Bitz, C., Massonnet, F., Bethke, I., and Gao, Y.: Optimising assimilation of sea ice concentration in an Earth
600   system model with a multicategory sea ice model, Tellus A, 70, 1–23, 2018.

Kimmritz, M., Counillon, F., Smedsrud, L., Bethke, I., Keenlyside, N., Ogawa, F., and Wang, Y.: Impact of ocean and sea ice initialisation on seasonal prediction skill in the Arctic, J Adv Model Earth Syst, https://doi.org/10.1029/2019MS001825, 2019.

Koenigk, T. and Mikolajewicz, U.: Seasonal to interannual climate predictability in mid and high northern latitudes in a global coupled model, Clim. Dyn., 32, 783, 2009.

605 Koul, V., Tesdal, J.-E., Bersch, M., Hátún, H., Brune, S., Borchert, L., Haak, H., Schrum, C., and Baehr, J.: Unraveling the choice of the north Atlantic subpolar gyre index, Scientific reports, 10, 1–12, 2020.

Kwok, R.: Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018), Environ. Res. Lett., 13, 105 005, https://doi.org/10.1088/1748-9326/aae3ec, 2018.

Massonnet, F., Fichefet, T., and Goosse, H.: Prospects for improved seasonal Arctic sea ice predictions from multivariate data assimilation,
610   Ocean Model., 88, 16–25, 2015.

Mathiot, P., König Beatty, C., Fichefet, T., Goosse, H., Massonnet, F., and Vancoppenolle, M.: Better constraints on the sea-ice state using
        global sea-ice data assimilation, Geosci. Model Dev., 5, 1501–1515, 2012.

Onarheim, I. H., Eldevik, T., Smedsrud, L. H., and Stroeve, J. C.: Seasonal and regional manifestation of Arctic sea ice loss, J. Climate, 31,
        4917–4932, 2018.

615   Pierce, D. W., Barnett, T. P., Tokmakian, R., Semtner, A., Maltrud, M., Lysne, J., and Craig, A.: The ACPI project, element 1: Initializing a
        coupled climate model from observed conditions, Clim. Change, 62, 13–28, 2004.

Polkova, I., Brune, S., Kadow, C., Romanova, V., Gollan, G., Baehr, J., Glowienka-Hense, R., Greatbatch, R. J., Hense, A., Illing, S., et al.:
        Initialization and ensemble generation for decadal climate predictions: A comparison of different methods, J Adv Model Earth Syst, 11,
        149–172, 2019.

620   Proshutinsky, A. Y. and Johnson, M. A.: Two circulation regimes of the wind-driven Arctic Ocean, J. Geophys. Res. O, 102, 12 493–12 514,
        1997.

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Levy, C., Masson, S., and
        Vivier, F.: The Louvain-La-Neuve sea ice model LIM3. 6: global and regional capabilities, Geosci. Model Dev., 8, 2991–3005, 2015.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic sea ice volume, J. Geophys. Res. O,
625   116, 2011.

Serreze, M. C. and Barry, R. G.: Processes and impacts of Arctic amplification: A research synthesis, Global Planet. Change, 77, 85–96,
        2011.

Sévellec, F., Fedorov, A. V., and Liu, W.: Arctic sea-ice decline weakens the Atlantic Meridional Overturning Circulation, Nat. Clim. Change,
        7, 604–610, 2017.

630   Swart, N. C., Fyfe, J. C., Hawkins, E., Kay, J. E., and Jahn, A.: Influence of internal variability on Arctic sea-ice trends, Nat. Clim. Change,
        5, 86, 2015.

Tietsche, S., Day, J., Guemas, V., Hurlin, W., Keeley, S., Matei, D., Msadek, R., Collins, M., and Hawkins, E.: Seasonal to interannual Arctic
        sea ice predictability in current global climate models, Geophys. Res. Lett., 41, 1035–1043, 2014.

Tietsche, S., Balmaseda, M. A., Zuo, H., and Mogensen, K.: Arctic sea ice in the global eddy-permitting ocean reanalysis ORAP5, Clim.
635   Dyn., 49, 775–789, 2017.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., et al.: An
        assessment of ten ocean reanalyses in the polar regions, Clim. Dyn., 52, 1613–1650, 2019.

Volpi, D., Guemas, V., and Doblas-Reyes, F. J.: Comparison of full field and anomaly initialisation for decadal climate prediction: towards
        an optimal consistency between the ocean and sea-ice anomaly initialisation state, Clim. Dyn., 49, 1181–1195, 2017.

640   Wadhams, P.: Arctic ice cover, ice thickness and tipping points, Ambio, 41, 23–33, 2012.

Yeager, S. G., Karspeck, A. R., and Danabasoglu, G.: Predicted slowdown in the rate of Atlantic sea ice loss, Geophys. Res. Lett., 42, 10–704,
        2015.

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