Enhancing sea ice knowledge through assimilation of sea ice thickness from ENVISAT and CS2SMOS

Nicholas Williams[1], Yiguo Wang[1], and François Counillon[12]

[1]Nansen Environmental and Remote Sensing Center, Bergen, Norway
[2]Bjerknes Centre for Climate Research, Bergen, Norway

Correspondence: Nicholas Williams (nicholas.williams@nersc.no)

Abstract. Arctic sea ice extent has declined significantly over the past four decades, opening up the Arctic to shipping and resource extraction while also impacting wildlife and local communities. This has led to an increasing need for skillful sea ice predictions. We focus on furthering the understanding of the role that sea ice thickness plays in the skilfulness of seasonal Arctic sea ice predictions. We look at how observations of sea ice thickness can improve both sea ice reanalyses and predictions. We use the Norwegian Climate Prediction Model (NorCPM) with 1° horizontal resolution for the ocean and sea ice components and approximately 2° for the atmosphere and land components, which has previously assimilated ocean and sea ice concentration observations. We additionally assimilate two sea ice thickness products: CS2SMOS, and, for the first time in any study, ENVISAT. This allows us to produce a two-decade (2003-2023) reanalysis with sea ice thickness assimilation focusing on the Arctic Ocean. This reanalysis is then used to initialise and generate a series of year-long seasonal hindcasts for each season of the reanalysis. The reanalysis and hindcasts are compared to observations and other reanalyses to assess the impact of sea ice thickness observations. Assimilation of sea ice thickness data strongly improves the representation of sea ice thickness and volume, primarily in the central Arctic as well as the ice edge location. Although ENVISAT observations have greater uncertainties, the dataset still provides a useful impact on the model. For prediction, sea ice thickness initialisation reduces the model biases of thickness throughout the year as well as errors in the detrended anomalies. Ice thickness bias correction results in improvements in the representation of the ice edge location, i.e., the timing and extent of the summer melting. Thickness initialisation has little improvements for detrended sea ice extent anomalies, but yields some skill in the Beaufort Sea and Central Arctic during summer. Overall, we show the impact of sea ice thickness assimilation has a positive effect on prediction skill in NorCPM.

1 Introduction

In the Arctic, both the atmosphere and ocean have undergone dramatic warming trends over the satellite era (Przybylak and Wyszyński, 2020; Steele et al., 2008), and continue to experience warming much faster than the rest of the globe (Cohen et al., 2014), leading directly to sea ice loss (Dai et al., 2019). The sea ice loss in this region is having a strong impact on Arctic wildlife and habitats (Descamps et al., 2017). There are also wide-reaching effects on tourism, resource extraction and communities living in the Arctic (Arruda and Krutkowski, 2017). This has led to increasing interest in the study of Arctic sea

ice and, in particular, of seasonal predictions, which aim to predict the sea ice on seasonal time scales. As the Arctic climate has warmed, the sea ice has not only retreated (Comiso et al., 2008; Stroeve et al., 2014b), but also thinned (Nghiem et al., 2007; Giles et al., 2008; Kwok, 2018). Much of the Arctic sea ice has gone from being perennial year-round to seasonal, with an increase in the length of the open-water season (Nghiem et al., 2007; Sumata et al., 2023). Such radical change also makes seasonal Arctic sea ice prediction more challenging.

The field of seasonal Arctic sea ice prediction emerged and rapidly developed within the past two decades, focusing on identifying and improving our understanding of the physical properties of sea ice, in order to improve the prediction of Arctic sea ice under a changing climate. This began with the first seasonal sea ice predictions using global climate models (GCM) and continued with prediction studies from a variety of research centres and institutes around the world (Wang et al., 2013; Chevallier et al., 2013; Peterson et al., 2015; Guemas et al., 2016; Bushuk et al., 2017; Kimmritz et al., 2019). The sea ice prediction network (SIPN) is at the centre of this field. The SIPN collects September sea ice extent (SIE) forecasts initialised from June, July and August and September into the Sea Ice Outlook (SIO) report (Stroeve et al., 2014a; Bushuk et al., 2024). At present, over 30 different research groups of polar scientists contribute to the SIPN, using dynamical or statistical approaches to produce these predictions. Idealistic model experiments in GCMs estimate that sea ice area and extent in the Arctic could be predictable between 12-48 months in advance, though predictability beyond 36 months is dominated by atmospheric forcing and not the initial conditions (Blanchard-Wrigglesworth et al., 2011a, b). Sea ice thickness (SIT) may be predictable up to 2 years in advance (Holland et al., 2011; Koenigk et al., 2012), and can be important for SIE predictions up to 2 years ahead (Tietsche et al., 2014) in an idealised framework.

The evolution of the Arctic sea ice is governed by the evolution of SIT distribution, and these changes come from two sources: dynamical and thermodynamical. Sea ice dynamics govern the movement of the sea ice either in space, or within the thickness distribution (commonly known as ridging). Thermodynamic changes in the sea ice are composed of melting (lateral, bottom and top melt) and freezing (congelation and frazil ice formation). SIT therefore is a key variable for modelling the sea ice state. SIT observations can substantially improve reanalysis estimates of thickness, with smaller impacts on other model variables such as sea ice concentration (SIC) and sea ice drift (SID), which has been shown in a number of sea ice reanalysis studies using different observational sources of thickness (Xie et al., 2018; Fritzner et al., 2019; Williams et al., 2022), and freeboard (Sievers et al., 2023). For prediction of SIE, correct SIT initialisation and its advection through the Arctic is believed to be one of the key mechanisms for skill within the Arctic (Blanchard-Wrigglesworth et al., 2011b; Ordoñez et al., 2018; Giesse et al., 2021; Zhang et al., 2021; Min et al., 2023; Zhang et al., 2023), and has been shown recently for CryoSat-2 (CS2) observations when using the CICE model standalone (Sun and Solomon, 2024). This is particularly true during the growth season which has been shown to translate to improved skill in SIE during the summer melt season (Blockley and Peterson, 2018; Holland et al., 2019). Additional mechanisms identified as important in the Arctic for SIE prediction are initialisation, persistence and re-emergence of SIC (Blanchard-Wrigglesworth et al., 2011b; Ordoñez et al., 2018; Giesse et al., 2021), ocean heat transport (OHT), ocean heat content (OHC) and melt onset timing (Schröder et al., 2014; Serreze et al., 2016; Kwok, 2018; Zheng et al., 2021). Atmospheric processes, particularly wind, which determine the dominant Arctic ocean currents of the Transpolar drift stream and the Beaufort Gyre, are also important but have a short memory, so its predictability can deteriorate quickly

(Msadek et al., 2014; Serreze and Meier, 2019). Though several mechanisms have been identified in general, many are still poorly understood. In this paper, we will focus on improving thickness initialisation, and in doing so, using a longer period of SIT observations to study Arctic sea ice predictability.

In the past 15 years, significant work has been done on developing and producing new SIT observation datasets. At the forefront of this field are SIT datasets from CS2 (Laxon et al., 2013), available since October 2010, the combined CS2 and Satellite Moisture and Salinity (SMOS) dataset (Ricker et al., 2017) known as CS2SMOS. These datasets are the longest running available Arctic SIT observations, but until recently summer observations of SIT were not available due to the presence of melt ponds, though this has now changed with the work of Landy et al. (2022) to produce a year-round CS2 thickness dataset. The launch of ICESat-2 in 2018 also led to a winter SIT dataset using laser observations (Petty et al., 2020). However, before these satellites were launched, there were the forerunner satellites of ICESat (Kwok et al., 2004) and ENVISAT (Louet and Bruzzi, 1999). As they were not primarily designed for SIT retrieval, the datasets have issues, particularly with spatial and temporal coverage, and also the instruments used. Monthly observations of SIT have been derived from ENVISAT for the winter months not available during summer due to melt ponds, as with CS2 (Connor et al., 2009), which spans from October 2002 to 2012. In this work, we use the ENVISAT dataset to lengthen our reanalysis and increase the number of predictions possible for our statistical analysis.

In this study, we will investigate how assimilation of SIT observations from ENVISAT and CS2SMOS can first benefit our sea ice reanalysis and then summer Arctic sea ice predictions by using seasonal hindcasts (i.e. retrospective predictions) started from 2003 to 2023. To our knowledge, ENVISAT SIT observations have not been assimilated before in a GCM, so this study will investigate their use and feasibility of inclusion for assimilation in GCMs for the first time. As we use ENVISAT observations, this also means we have a longer time series of SIT observations to assimilate, and thus a longer reanalysis and more hindcasts with which we can analyse, study and verify its usefulness within these fields. We investigate not only SIE/SIT predictions but also the prediction of the sea ice edge location using the integrated ice edge error (IIEE).

The paper is organised as follows. In Section 2, we outline the climate model, observations and experimental design. In Section 3, we outline the metrics and independent observations used for evaluating and validating the performance of the model. In Section 4.1, we show the results of the sea ice reanalysis from 2003 to 2023, including verification of the reanalysis with independent data from the Beaufort Gyre Exploration Project (BGEP) moorings. In Section 4.2, we show the results of the hindcast experiments. In Section 5, we discuss the key results and conclude this study.

2 NorCPM

We use the Norwegian Climate Prediction Model (NorCPM, Counillon et al., 2016; Wang et al., 2019), which combines the Norwegian Earth System Model (NorESM Bentsen et al., 2012) and the Ensemble Kalman Filter (EnKF, Evensen, 1994) data assimilation method. The version of NorCPM is as in Kimmritz et al. (2019), which contributed to the SIPN (Bushuk et al., 2024), but we are testing, in addition, a separate version that assimilates ice thickness data.

2.1 NorESM

The NorESM version used here is NorESM1-ME (Bentsen et al., 2013). It is based on the Community Earth System Model (CESM, Hurrell et al., 2013). However, the ocean component is replaced with an isopycnal coordinate ocean general circulation model (BLOM, Bentsen et al., 2012), and the Community Atmosphere Model version 4 (CAM4, Neale et al., 2010) with the original prescribed aerosol formulation is replaced by the atmospheric model CAM4-OSLO with a prognostic aerosol life cycle formulation using emissions and new aerosol-cloud interaction schemes (Kirkevåg et al., 2013). As in CESM1.0.4, NorESM1-ME uses the Los Alamos Sea Ice Model version 4 (CICE, Hunke et al., 2015) and the Community Land Model (CLM) version 4 (Lawrence et al., 2011). These are coupled using version 7 of the coupler designed for the CESM (Craig et al., 2012).

The atmosphere and land model has an approximately 2° finite volume grid, with horizontal resolutions of 1.9° in latitude and 2.5° in longitude, while the ocean and sea ice have approximately a $1° \times 1°$ horizontal resolution. BLOM uses 51 isopycnal layers plus 2 additional layers for the bulk mixed layer, with time-evolving thicknesses and densities. This version of NorESM is run with CMIP5 historical forcings (Taylor et al., 2012) and the RCP8.5 (Moss et al., 2010) beyond 2005. A similar version of NorCPM has contributed to CMIP6 Decadal Climate Prediction Project (Bethke et al., 2021). However, upgrading to CMIP6 forcings degraded NorCPM's baseline climate and hindcast performance (Bethke et al., 2021; Passos et al., 2023), and we use CMIP5 forcings in this study.

The sea ice model CICE uses five categories in its thickness distribution, optimal for representing the sea ice cover at reasonable computing power (Bitz et al., 2001; Massonnet et al., 2019). The horizontal transport of sea ice is solved using an incremental remapping scheme (Lipscomb and Hunke, 2004), and solving for the sea ice stresses using the elastic-viscous-plastic rheology (Hunke and Dukowicz, 1997). The one-dimensional vertical Bitz and Lipscomb model (Bitz and Lipscomb, 1999), is used to solve the thermodynamic equations, with melt pond, aerosol (Holland et al., 2012) and radiation transfer parameterisations (Briegleb and Light, 2007). Sea ice transported in thickness space is solved using a remapping scheme (Lipscomb, 2001).

2.2 Assimilation implementation

The EnKF is a sequential ensemble data assimilation (DA) method using Monte Carlo integration followed by a linear analysis update (Evensen, 1994). The method is multivariate and updates the model state variables based on their ensemble covariances with the observations. Specifically, we use a deterministic formulation of the EnKF (Sakov and Oke, 2008), which solves the analysis without needing to perturb the observations. The Deterministic EnKF outperforms the standard EnKF, particularly for small ensemble sizes (Sakov and Oke, 2008).

We assimilate the monthly average observations in the middle of the model month (i.e., the 15th day) and update the instantaneous model state based on all observations (Counillon et al., 2016; Kimmritz et al., 2019). The update of the model state during the assimilation is split into two steps: the ocean model variables and the model variable SIC are updated jointly by the assimilation of oceanic and SIC observations, and then the model variable SIC is again updated by the assimilation of SIT

observations. The atmosphere and land components are not updated by the assimilation but adjusted dynamically via coupling in between the monthly assimilation cycles.

Assimilation of ocean temperature and salinity profiles, sea surface temperature (SST), and SIC observations is performed as described in Kimmritz et al. (2019). We employ anomaly-field assimilation, using a monthly reference climatology calculated from 1982 to 2016. We update both the ocean and sea-ice components based on the observations from both components, so called strongly coupled ocean-sea ice DA (Laloyaux et al., 2016; Kimmritz et al., 2018). Strongly coupled ocean-sea ice DA in NorCPM was shown to be more effective than weakly coupled DA in which sea ice observations are used to only update the sea ice variables (Kimmritz et al., 2018). We update the full ocean physics state vector in isopycnal coordinates (i.e., 3D temperature, salinity, velocities and layer thickness) and update the multicategory SIC in the sea ice state vector (i.e., the multicategory aicen within the 5 categories, see DEPTH HI_PRESERVE in Kimmritz et al. (2018)). The sea ice volume in each thickness category is changed proportionally so that the thickness of each thickness category remains identical to that of the prior ensemble (i.e., the multicategory hicen before assimilation). This prevents the need to reshuffle ice to a different thickness category in the post-analysis, which proved to be optimal in an idealised twin experiment (Kimmritz et al., 2018). The post-processing step ensures that sea ice state variables remain within physical ranges and recompute the energy budget of each of the multicategory sea ice quantities (Appendix 1 in Kimmritz et al. (2018) for further details). However, in Kimmritz et al. (2019), assimilation was carried out in 2 steps with SST and SIC updating the mixed layer depth ocean and the sea ice state, followed by assimilation of ocean profiles that update the full 3D ocean. This degraded performance in the lower latitudes where there are few profiles data to constrain the ocean interior. Therefore, we carry out a single assimilation step based on all observation products, which updates the full ocean and sea ice state jointly. This handled the degradation reported in Kimmritz et al. (2019) and has no impact on performance at high latitudes (not shown).

NorESM has a large SIT bias (Bentsen et al., 2012), and while assimilation of ocean observation reduces it partially, some of the bias remains. Bethke et al. (2021), compared two versions of NorCPM assimilating ocean observations, one that updates only the ocean component and one that updates the ocean and sea ice components. The latter yields a strong reduction of the bias of SIT and provides enhanced predictions. Note also that it takes about ten years for the model to rebuild the SIT bias once assimilation is stopped (their Figure S15). We, therefore, use full-field assimilation to correct the SIT bias that can influence the variability. In the first attempt, we used anomaly-field assimilation (Carrassi et al., 2014). However, the assimilation impact of SIT anomalies was inconclusive, with no added skill for predictions (not shown).

When assimilating SIT observations, we only update the individual category sea ice fraction, which can change the sum of the ice fraction. In the post-processing of the assimilation, the sea ice volume in each thickness category is changed proportionally so that the thickness of each thickness category remains identical to that of the prior. We do not update the ocean component, as the covariances between SIT and the ocean are very small, and may cause more harm than benefit because of sampling error.

We use a series of ad-hoc techniques to handle issues related to sampling error, typically used in NorCPM (Counillon et al., 2016; Bethke et al., 2021). First, we used the R-factor (Sakov et al., 2012), which inflates the observation error by a factor of 2 for the update of the ensemble anomaly. We also use the K-factor formulation (Sakov et al., 2012), which inflates the observation error so that the analysis remains within m times the standard deviation of the ensemble ($m = 2$ in this study).

This avoids producing too strong updates. Finally, we use a local analysis framework (Evensen, 2003), which only uses the local observations and limits the risk of spurious covariance. Tapering with a smooth distance-weighted Gaspari and Cohn function (Gaspari and Cohn, 1999) ensures continuity in the update. The localisation radius for the ocean variables is a function of latitude (Wang et al., 2017), and for the SIC and SIT observations we use a localisation radius of 800 km as used previously (Kimmritz et al., 2018; Massonnet et al., 2015).

Our DA implementation assumes the observation errors to be independent. However, since observations of SST, SIC and SIT are gridded products, the observation error may be correlated due to heavy post-processing during data production. To mitigate this fact, we only retain the nearest observation of each observation type in the local analysis (Counillon et al., 2016; Kimmritz et al., 2018). For the hydrographic profiles, all data within the local window are used (Wang et al., 2017).

2.3 Assimilated datasets

SST and SIC monthly average data are from the NOAA OISSTV2 dataset (Reynolds et al., 2007; Huang et al., 2021) available on a $1^\circ \times 1^\circ$ global grid. The data is originally produced as weekly fields, NOAA then produces monthly fields using a linear interpolation of the weekly fields to daily fields then averaging those daily values over each month. The SST data are produced by combining both in-situ and satellite observations and SST's simulated by sea ice cover (Reynolds et al., 2007). Note that SST data in the regions covered by sea ice are not assimilated in this study (Kimmritz et al., 2019; Wang et al., 2019). However, the ocean underneath the sea ice is updated based on the ensemble covariance. Monthly SST observation errors are estimated using the weekly error estimation provided by the dataset. We experience that error in SST tends to be very low and we imposed a minimum threshold of 0.1 °C. SIC observation errors are not provided with the dataset and thus we use a 20% constant value (Kimmritz et al., 2019), which is the generally agreed upon value for the highest uncertainties in the summer melt season (Ivanova et al., 2015; Cavalieri and Parkinson, 2012; Comiso, 2017).

The temperature and salinity profile data from October 2003 to October 2021 are taken from EN4.2.1 dataset (Gouretski and Reseghetti, 2010; Good et al., 2013). The data from then onwards are from EN4.2.2 dataset (Gouretski and Cheng, 2020). The EN4 hydrographic dataset is split into different categories depending on its quality. In this study, we only use data in category 1 (i.e., the highest quality). Their associated observation errors are determined as in Wang et al. (2017).

For the SIT observations, we use two datasets: the ESA CCI dataset that includes SIT retrieved from ENVISAT (Connor et al., 2009), and the AWI CS2SMOS dataset (V2.6) (Ricker et al., 2018) that retrieves SIT from the SMOS satellite and the CS2 satellite. The ESA CCI dataset covers the period from October 2002 to April 2012, however we only use data up to March 2010 because the more accurate CS2SMOS product becomes available after this. For both datasets, data is only available outside the melt season between October and April. Melt ponds on top of sea ice make it challenging to identify leads, which is crucial for the estimation of freeboard and SIT (Laxon et al., 2013). Additionally, we do not use October or April data due to certain quality issues associated with the fringe months marking the change between the melting and growth seasons in the Arctic sea ice. The observation error statistics are provided by the datasets.

2.4 Experiment design

In this work, we want to test the added value of SIT assimilation. For that, we have run several experiments. We use an ensemble model simulation without data assimilation and two reanalyses as follows:

FREE: a 30-member ensemble run without data assimilation integrating from 1850 to December 2023 with CMIP5 historical forcings and with RCP8.5 beyond 2005. The ensemble is initialized on 01 January 1850 by randomly selecting 30 states on 01 January in different years from a stable pre-industrial run (with one single member).

CTRL: A 30-member reanalysis started in 1982 branched from FREE (30 members), using a similar setting as in Kimmritz et al. (2019); Bushuk et al. (2024). It assimilates hydrographic profiles, SST, and SIC data (section 2.3) with an anomaly-field assimilation framework.

+SIT: a 30-member reanalysis is branched off from CTRL on the 15th of October 2002. It assimilates SIT data (section 2.3) in addition to that assimilated in CTRL.

FREE allows us to estimate the skill related to external forcings (Kimmritz et al., 2019). We use CTRL to compare the model with and without the assimilation of SIT data, while still assimilating the ocean and SIC observations. We use an ensemble size of 30 members primarily due to limited computational resources. In addition, many of the parameters in NorCPM (localisation, inflation) have been tuned to work ideally with an ensemble size of 30 which we found large enough to provide robust results for ocean and sea ice update with NorCPM (Counillon et al., 2014; Kimmritz et al., 2018).

Three sets of seasonal hindcasts (i.e., retrospective predictions) are created from FREE, CTRL and +SIT. They start on the 15th of January, April, July and October each year between 2003 and 2023 and run for 1 year with 10 members. This comprises 84 hindcasts in total (21 years and 4 hindcasts per year) for each set of hindcast experiments.

3 Validation metrics and datasets

We validate the reanalyses and hindcasts using their ensemble means. For the hindcast, we assess the performance for different start and lead months. Note that our hindcasts start in the middle of the month, but it has assimilated the monthly average of that month – e.g., our hindcast started on the 15th of April has assimilated April monthly average data. Lead month 1 starts after 15 days of model integration (e.g., May is lead month 1 of April-initialised prediction).

We define the SIE as the total sum of the area of grid cells where the ensemble mean of SIC is at or above 15%.

3.1 Validation metrics

We use bias and bias-free root-mean-square error (bfRMSE), defined as

$$\mathrm{bias} = \sum_{i=1}^{N} W_i (x_i - y_i), \tag{1}$$

$$\mathrm{bfRMSE} = \sqrt{\sum_{i=1}^{N} W_i (x_i - y_i)^2}, \tag{2}$$

where N is the total number of data points, W_i is a normalising area-based weight constant –i.e. when the metric is computed considering grid cells that do not have the same area. For point-wise bias and bfRMSE calculations, W_i is $\frac{1}{N}$. x_i is the model ensemble mean values, and y_i is the observed values, which are both anomalies from their respective climatology (model or observation). These are usually averaged either over time or space (where each grid cell is then weighted by its area). Note that root-mean-square error (RMSE) is the quadratic sum of bias and bfRMSE.

We additionally use anomaly correlation coefficient (ACC) to test the variability of the reanalyses/hindcasts and the observations, defined as follows:

$$\mathrm{ACC} = \frac{\sum_{i=1}^{N} x_i' y_i'}{\sqrt{\sum_{i=1}^{N} x_i'^2} \sqrt{\sum_{i=1}^{N} y_i'^2}}, \tag{3}$$

where x_i' and y_i' are model and observation values. For the reanalysis, we use the standard ACC and bfRMSE, but for the predictions, we use the detrended values – i.e. both time series are detrended before computing the metric. The reason is that the trend in reanalysis is part of the signal that one aims to represent and can be challenging due to discontinuity in the observation data set or drift in the system. In prediction, the trend is often removed as it is considered to be a trivial predictor (Bushuk et al., 2020). The statistical significance of the Pearson correlation coefficient is tested by using the Student's t-test with a significant level of 5% with degrees of freedom calculated as in Von Storch and Zwiers (2002).

The Degrees of Freedom for Signal (DFS) is calculated for each set of assimilated observations in the analysis to test the impact of their assimilation (Wahba et al., 1995; Cardinali et al., 2004).

$$\mathrm{DFS} = \sum_{i=1}^{L} \frac{\partial x_a}{\partial y}, \tag{4}$$

where L is the total number of observations and x_a is the model posterior. The DFS metric quantifies how the assimilation of observations has reduced the dimension or rank of the ensemble (Sakov et al., 2012). A larger DFS value implies that the assimilation has more change into the system, i.e., reducing the number of degrees of freedom (the unit of the metric). The DFS can be between 0 (no impact) and the total number of degrees of freedom minus one (where all members collapse to a single member)[1]. A well-balanced data assimilation system aims to make minimal changes necessary to comply with observations, and as such, one should reach neither the lower nor the upper DFS value. As a consequence, DFS is often used to calibrate

[1]Note that the total number of degrees of freedom is the minimum between the ensemble size and the number of observations used in the local analysis.

the strength of the data assimilation system when observation error is poorly known (Sakov et al., 2012) -prevent too strong or too weak assimilation. DFS can also be used to isolate the relative influence of each observation on the total impact of the assimilation, which is our aim in this study. More specifically, DFS is used here to diagnose the relative influence of the SIT assimilation compared to other datasets, for instance, where it is most beneficial, as well as to quantify the impact of ENVISAT SIT versus C2SMOS.

Finally, we use the Integrated Ice Edge Error (IIEE) to assess the error in the location of the ice edge (Goessling et al., 2016). It is defined as

$$\text{IIEE} = \int_A \max(c_x - c_y, 0) dA + \int_A \max(c_y - c_x, 0) dA, \tag{5}$$

where A is the area, $c = 1$ where the SIC is above 15% and 0 elsewhere, with subscripts x and y denoting the model and observations respectively. This works as essentially the sum of all areas where SIE is overestimated (first term on right hand side (RHS) of the above equation) and underestimated (second term). The IIEE can also be decomposed in a different way, using a mean absolute error (MAE) and a displacement (DISP). This is formulated in Goessling et al. (2016) as

$$\text{IIEE} = \text{MAE} + \text{DISP} \tag{6}$$

$$\text{MAE} = |\text{O} - \text{U}| \tag{7}$$

$$\text{DISP} = 2 \cdot min(\text{O}, \text{U}), \tag{8}$$

where O is the area where SIE is overestimated, and U is the area where SIE is underestimated. The IIEE is a useful verification metric because it is conceptually simple. It is well verified with existing long-term SIC satellite data, is an important characteristic of the sea ice cover and is more useful to potential forecast users than total SIE (Goessling et al., 2016).

3.2 Independent validation datasets

In the paper, a part of the validation will be carried out with the assimilated dataset. In the reanalysis, such data cannot be considered independent, but still represents a baseline verification (that DA works as expected).

We use observations of sea ice draft from the BGEP moorings for independent validation. BGEP includes upward-looking sonar (ULS) measuring instruments (Krishfield et al., 2014). There are a total of 4 moorings, 2 of which (A and B) have been in operation since August 2003, C was in operation until 2008, and D has been in operation since 2005. All moorings are located in the Beaufort Sea (their location in the Arctic is shown in Figure 3c). The ULS instruments measured sea ice draft every 2 seconds before 2014, and every second after. The sea ice draft, which is the thickness of the sea ice under sea level, is found by subtracting the range measured by the ULS instrument from the known depth of the instrument. The draft measurements have a stated error of $+/-5-10$ cm (Krishfield and Proshutinsky, 2006). To compare this data to our model monthly mean thickness, we first convert the sea ice draft into SIT using the method of Rothrock et al. (2008), and then average all the measurements over each month or year to convert to monthly or yearly averages respectively, which we can compare to the model. If there was missing data in a month, we discard the ULS data for that month for that mooring, if there were more than thirty days of missing continuous data in a year, we do not compute a yearly average for that year.

The Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) reanalysis was first introduced in 2003 by Zhang and Rothrock (2003). PIOMAS is a coupled ice-ocean model that assimilates SST data in ice-free grid cells using optimal interpolation, and SIC using a relatively simple nudging technique which aims to move the model variables closer to their observed counterparts by use of a weighting factor. PIOMAS has been substantially validated with a large number of SIT observations (Schweiger et al., 2011), so is often used as a comparison dataset for other SIT and sea ice volume datasets. The PIOMAS reanalysis is used for validation in this study.

4 Results

4.1 Reanalysis

In this section, we analyse our 21-year reanalyses over the period where satellite SIT observations are available.

In Figure 1, we present the time evolution of the assimilation diagnostic. We can first notice that the bias has a seasonal signal that relates to the lack of observation during summer, when the bias increases. For ENVISAT, the system is too thick at the start but gets too thin at the end of the season, while with C2S the too thick bias remains positive until the end of the seasonal observation period. In an ensemble data assimilation system, one can use the ensemble spread as a measure of the system's error. A first check to assess the reliability of the system is to ensure that the quadratic sum of the ensemble standard deviation and observation error (here denoted as total error) matches the bias-free error of the ensemble mean (RMSE, Rodwell et al., 2016). We can notice that our system exhibits too high dispersion during the ENVISAT period, but that the reliability is very good in the C2S period. The overdispersion during the ENVISAT may relate to the observation error, which is very high. We can also notice that the ensemble spread and RMSE covary in time very well (both seasonally and inter-annually). There is a discrepancy at the start of the assimilation season that relates to the bias being large (not to be accounted for in the reliability budget analysis).

We start by analysing the relative influence of each observation type in constraining errors in the system with the DFS (Section 3.1). The observations assimilated complement each other quite well and dominate in different regions (Figure 2). The ocean observations dominate in the sea ice-free regions of the Arctic. SIC assimilation is most impactful at the sea ice edge, where SIC shows the largest variability. SIT assimilation is mainly effective in the central Arctic, where the ice is thicker. As expected, the SIT assimilation becomes more effective (higher DFS) when we switch over from assimilating ENVISAT to CS2SMOS. However, ENVISAT assimilation is still having a substantial impact in the central Arctic according to the DFS, where other observations are few, or have limited impact.

We first verify performance against the SIT observations (Figure 3), which is assimilated in +SIT. We only present the validation against the CS2SMOS period due to the differing regional coverage between ENVISAT and CS2SMOS, but the conclusions are similar for ENVISAT (not shown). FREE has a clear thick bias (up to 4 m), with the spatial pattern of the bias increasing as the ice gets thicker — largest close to the Canadian Arctic Archipelago (CAA). In CTRL, the biases in the central Arctic Ocean have been reduced, but there is still a thin strip of too-thick sea ice pushed against the CAA. In +SIT the

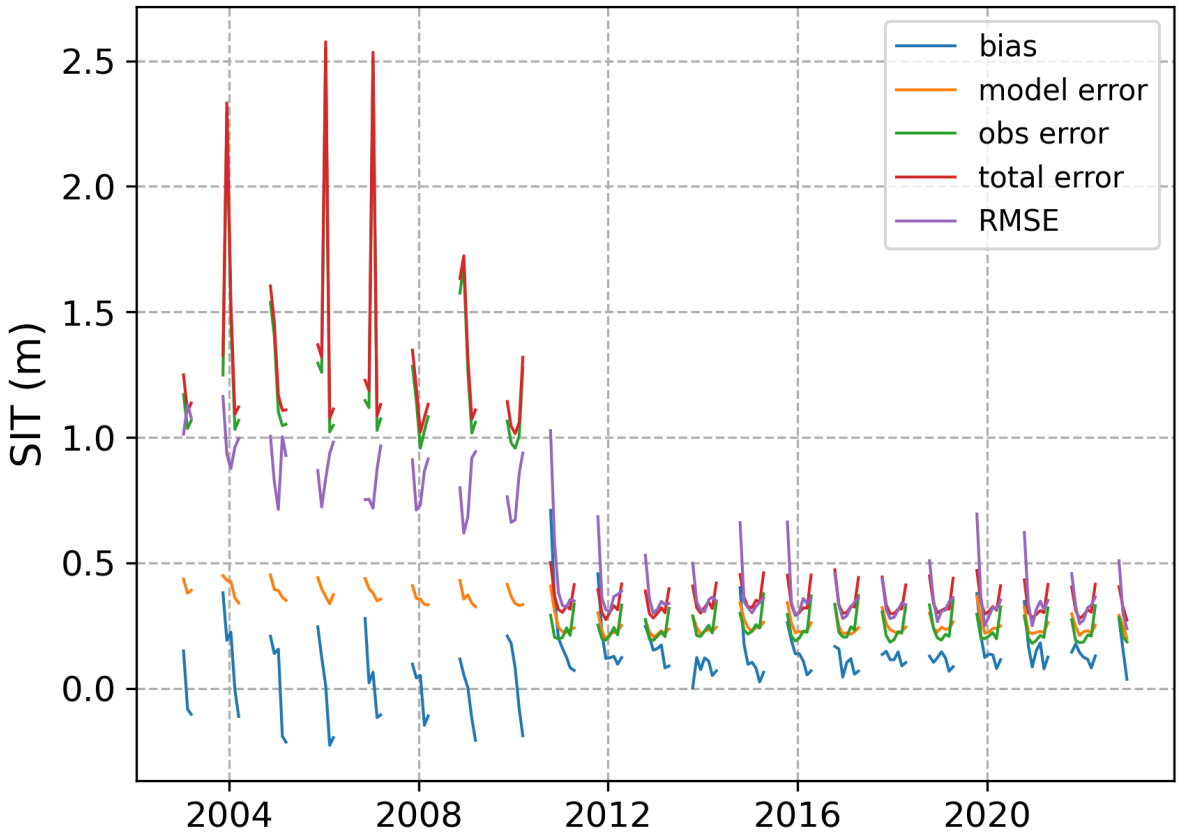

Figure 1. The time evolution of five key metrics of the EnKF for SIT assimilation (computed on the innovation vector – the difference between the observation and the model at the observation location): which is the bias, model error (ensemble standard deviation), observation error (standard deviation), total error (quadratic sum of model error and observation error) and RMSE. Note that there are gaps in the figure due to the lack of SIT observations outside winter.

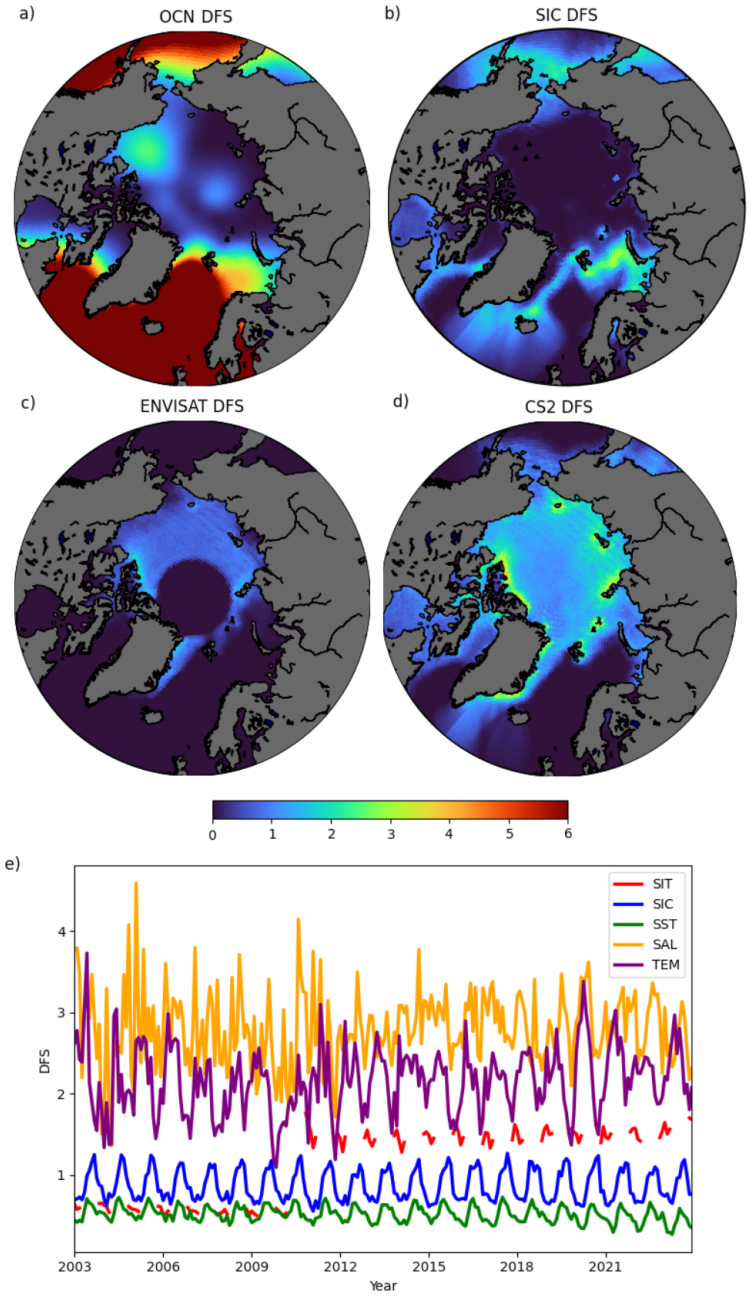

Figure 2. (a) Combined mean DFS of salinity, ocean temperature and sea surface temperature. (b) Mean DFS of SIC. (c) Mean DFS of ENVISAT SIT. (d) Mean DFS of CS2 SIT. (e) Pan-Arctic monthly mean DFS for each observation assimilated in our +SIT reanalysis. TEM refers to (ocean) temperature and SAL refers to salinity.

Table 1. Bias, ACC and bfRMSE for the SIT of BGEP ULS A, B, C and D in comparison with the FREE, CTRL and +SIT reanalyses. The statistics are based on monthly mean detrended data. The statistics with yearly data are given in brackets. The unit of bias and bfRMSE is the metre. The system with the best performance are highlighted in bold.

Metric	Experiment	ULS A	ULS B	ULS C	ULS D
Bias	FREE	1.51 (1.49)	1.51 (1.66)	1.03 (1.56)	0.43 (1.57)
	CTRL	0.96 (0.56)	0.62 (0.87)	0.58 (0.84)	0.34 (0.49)
	+SIT	**0.08 (-0.1)**	**-0.14 (-0.11)**	**-0.37 (-0.55)**	**-0.22 (-0.15)**
bfRMSE	FREE	0.11 (0.35)	0.13 (0.37)	0.15 (0.21)	0.06 (0.34)
	CTRL	0.08 (0.15)	0.06 (0.20)	0.11 (0.12)	0.05 (0.12)
	+SIT	**0.03 (0.05)**	**0.03 (0.06)**	**0.08 (0.08)**	**0.04 (0.05)**
ACC	FREE	0.87 (0.87)	0.77 (0.88)	0.73 (0.74)	0.80 (0.80)
	CTRL	0.85 (0.85)	0.79 (0.79)	0.55 (0.56)	0.76 (0.78)
	+SIT	**0.92 (0.92)**	**0.91 (0.91)**	**0.87 (0.88)**	**0.89 (0.89)**

thickness biases in the marginal seas surrounding the central Arctic have almost been completely removed, which showcases the influence of assimilating ice thickness data.

For bfRMSE, FREE features errors of around 20 cm in a majority of places, reaching errors up to 1 m close to the CAA. Surprisingly, the SIT bfRMSEs are increased in CTRL in much of the central Arctic compared to FREE. This suggests that the assimilation of SIC and ocean observations has led to increased bfRMSEs for SIT. However, as the bias is much larger than bfRMSEs, the increased bfRMSEs result more from the evolution of the SIT bias in between and during the assimilation (the total RMSE is reduced). In the +SIT reanalysis, almost all grid cells feature low bfRMSE, especially in the central Arctic and CAA. There remain higher errors in the Fram Strait – albeit weaker than for FREE and CTRL – that are typically driven by sea ice export (Sumata et al., 2015). In this region, SIT observation errors are very large, and SIT assimilation effectiveness is reduced (Figure 2). EXP-OCT performs better than EXP-OC against SMOS for the April-June period, while EXP-OC performs better against SMOS for the July-September period, with little differences in October.

The previous comparison was done against assimilated observations of CS2SMOS, and we now compare with SIT derived from the independent BGEP moorings (Section 3.2). The moorings are all located in the Beaufort Gyre (Figure 3c), so do not provide an assessment over the whole pan-Arctic region. The comparison (Figure 4 and Table 1) is well in line with the above CS2SMOS validation. FREE has the largest biases (as high as 1.5 m); CTRL has smaller biases than FREE due to ocean and SIC assimilation. +SIT has the smallest biases among the three reanalyses. FREE has bfRMSE from 0.06 m to 0.15 m, varying with the mooring and ACC around 0.7-0.8 as it captures the decreasing trends well with yearly data and the seasonal cycle with monthly outputs. The assimilation of the ocean and SIC data reduces bfRMSE (CTRL in Table 1) but yields a slight degradation of ACC at stations C and D. Assimilating SIT further reduces bfRMSE consistently with Figure 3 and improves correlation (about 0.9). Overall, +SIT shows the best performance for all four moorings and performance is better during the CS2SMOS observations period than during the ENVISAT period.

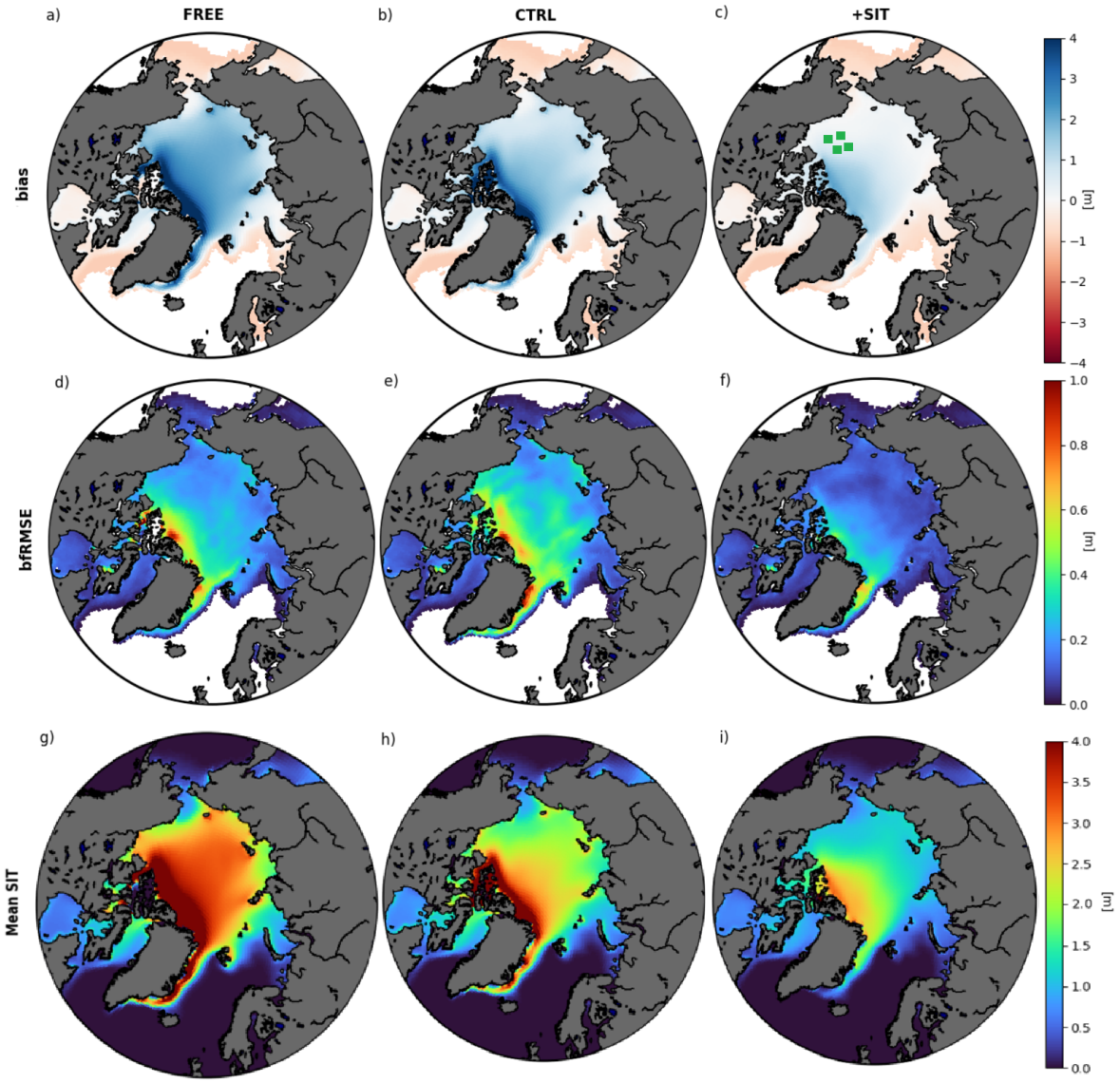

Figure 3. Bias (a - c), bfRMSE (d - f) and mean (g - i) of SIT in each of our NorCPM experiments in comparison to SIT observations from CS2SMOS between 2010 and 2023. In panel (c) we also show the locations of the BGEP ULS moorings as green squares.

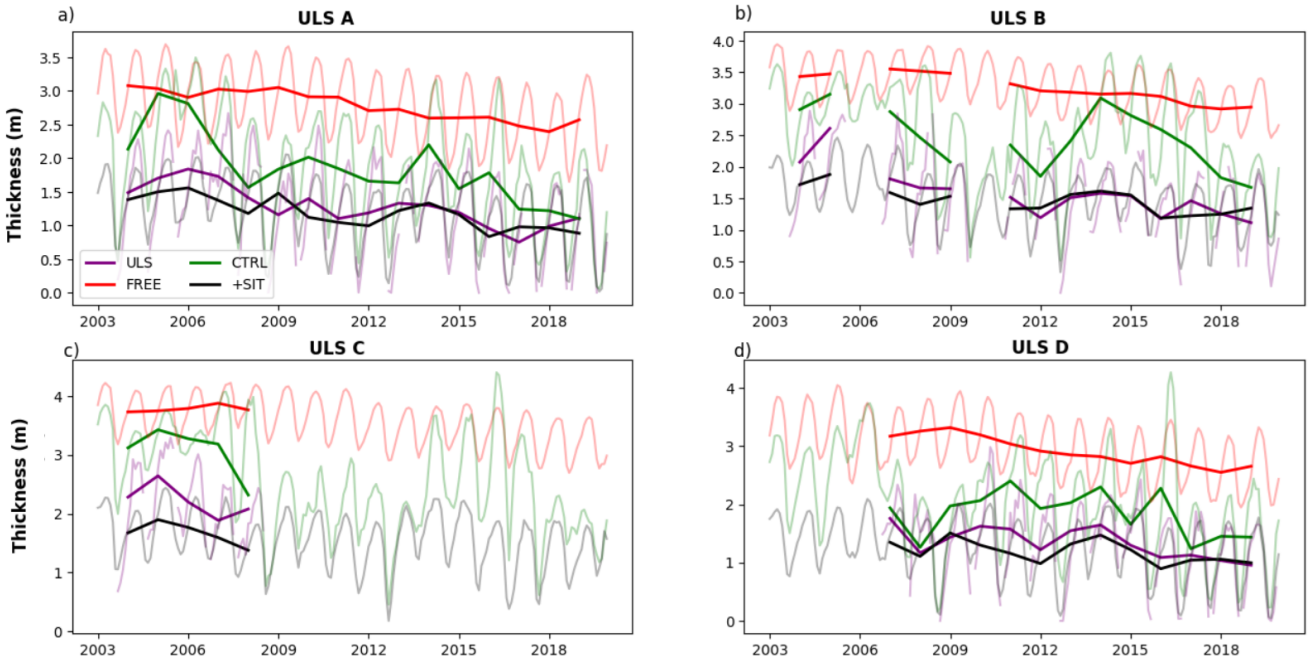

Figure 4. Yearly (thick line) and monthly (thin line) average sea ice thickness at ULS A (top left), B (top right), C (bottom left) and D (bottom right) and the respective SITs in our reanalyses during our experimental period. In some years, there are not enough observations (more than 30 days continuous of observations missing) from a ULS mooring to formulate a yearly average, so these years are masked.

We further analyse the seasonal variability of error of SIT and SIE (Figure 5). There is little difference between the reanalyses in the bias of SIE (Figure 5a) as we use an anomaly-field assimilation framework for ocean and SIC, which takes the climatology of FREE as an attractor. CTRL shows lower SIE bias than FREE in the ice-retreat season but slightly larger biases in the ice-advance season. SIE bias is substantially reduced in +SIT between April and September, and similar to FREE in the rest of the year.

For bfRMSE of SIE (Figure 5b), +SIT and CTRL have lower errors than FREE, showing the positive impact of the assimilation of ocean and SIC data. However, the bfRMSE in SIT+ is larger than CTRL from January to August.

CTRL reduces the SIT bias compared to FREE uniformly throughout the year, and +SIT nearly removes entirely the SIT biases (Figure 5c), even outside of the assimilated season when compared to PIOMAS. The bfRMSE is slightly increased in CTRL compared to FREE, likely for the same reason explained above, i.e. bfRMSE being lower than bias can be misleading. bfRMSE in SIT+ is interesting because it is very low when computed against CS2SMOS (Figure 5d), but larger than FREE when computing bfRMSE against PIOMAS. PIOMAS tends to underestimate the interannual variability of SIT because it overestimates the thickness of thin ice and underestimates the thickness of the thick ice (Schweiger et al., 2011; Wang et al., 2016). As such, the ensemble mean of FREE, which is nearly indistinguishable from the linear decline, compares favourably with PIOMAS for bfRMSE.

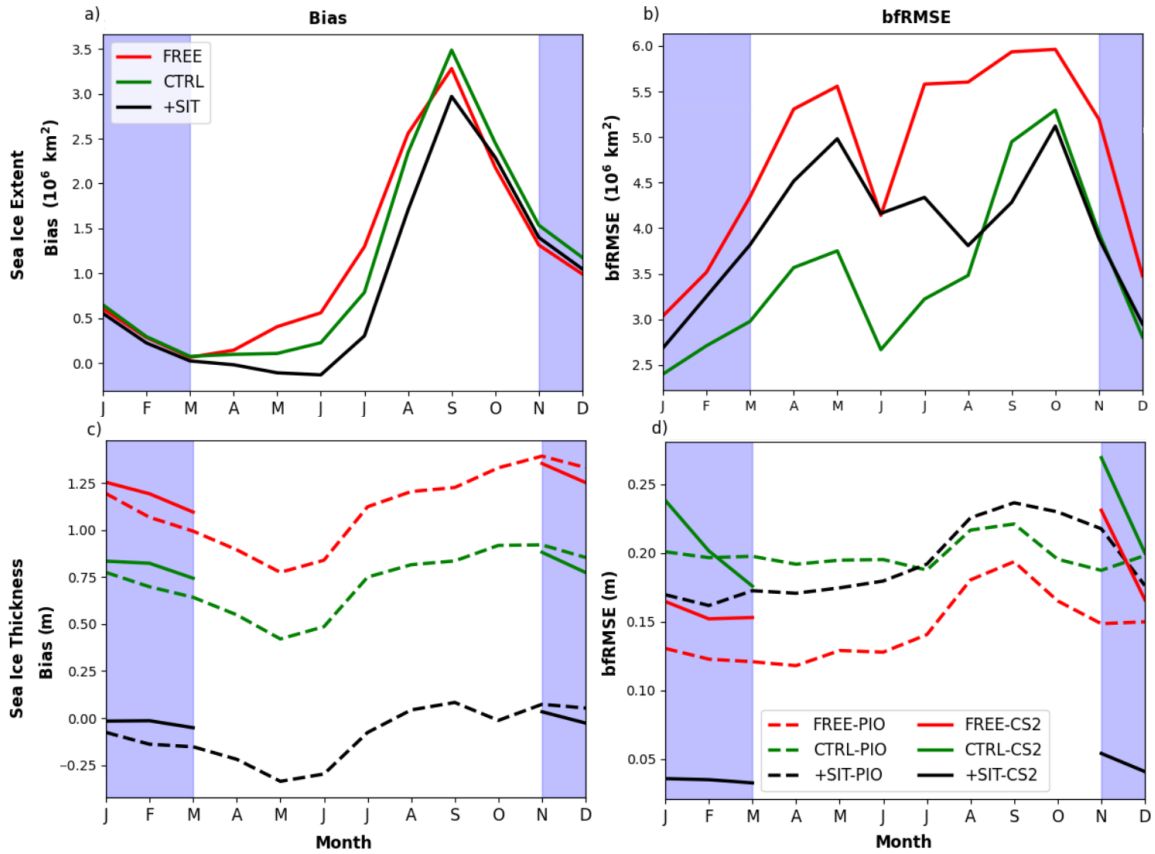

Figure 5. Bias (left-hand side) and bfRMSE (right-hand side) for SIE (top) and SIT (bottom) for FREE, CTRL and +SIT. SIC observations compared to SIC observations from NOAA, and the SIT from the PIOMAS reanalysis between 2003 and 2023 and CS2SMOS observations between 2010-2023. The blue-shaded area shows the time when SIT observations are assimilated.

We then investigate the time evolution of the SIE and SIV throughout the reanalysis for March and September (Figure 6 and Table 2) when SIE reaches a maximum and minimum.

In March, FREE shows a weak decreasing trend in agreement with the observation but no year-to-year variability (as internal variability is not synchronised). CTRL shows good agreement with the assimilated NOAA observation estimate. Interannual variability is also much larger, implying that the individual members of CTRL are better constrained. +SIT also shows improved agreement with observations (amplitude of internal variability and ACC). The SIE bias is reduced from 7.4×10^{10} km^2 to 2.3×10^{10} km^2 by the SIT assimilation. ACC and bfRMSE of SIE are slightly degraded compared to CTRL (albeit better than FREE). This primarily relates to a spurious increasing trend until 2010 during the ENVISAT SIT observation period (Figure 6a). The performance of ACC and bfRMSE for +SIT is comparable to CTRL in the C2SMOS era (not shown). All systems capture the decreasing trend in SIV well. Interannual variability is stronger in CTRL than in FREE and even more pronounced in +SIT (Figure 6d). Despite some visual agreement in CTRL, ACC and bfRMSE are not improved compared to

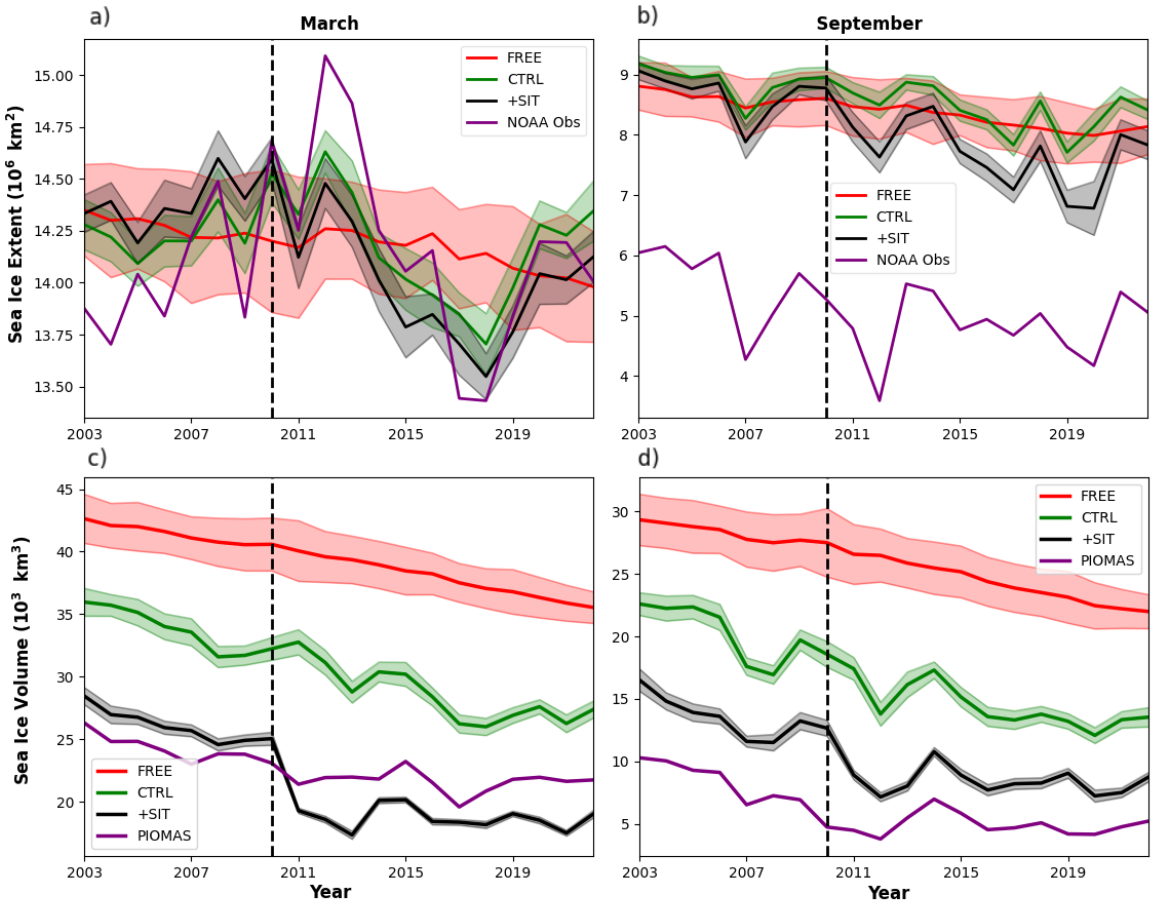

Figure 6. Arctic SIE (top) and SIV (bottom) in March (left-hand side) and September (right-hand side) for validation datasets and our experiments (FREE, CTRL and +SIT). Shaded region shows the ensemble mean plus/minus one standard deviation for each experiment. SIC observations from NOAA, and the SIV from the PIOMAS reanalysis between 2003 and 2023. The dashed vertical lines split the whole period to the ENVISAT and C2SMOS eras.

Table 2. Bias, bfRMSE and ACC for the SIE and SIV in March and September in the FREE, CTRL and +SIT experiments. The unit of bias and bfRMSE in SIE (SIV) is km^2 (km^3). Observed SIE is computed from NOAA OISSTV2 and we use SIV computed from PIOMAS. The system with the best performance is highlighted in bold.

SIE	March			September		
Experiment	Bias	bfRMSE	ACC	Bias	bfRMSE	ACC
FREE	6.5×10^{10}	9.8×10^{10}	0.1	3.3×10^{10}	2.4×10^{10}	0.6
CTRL	7.4×10^{10}	$\mathbf{2.2} \times 10^{10}$	**0.8**	3.4×10^{10}	3.9×10^{10}	0.7
+SIT	$\mathbf{2.3} \times 10^{10}$	2.9×10^{10}	0.6	$\mathbf{2.9} \times 10^{10}$	$\mathbf{3.8} \times 10^{10}$	**0.8**
SIV	March			September		
Experiment	Bias	bfRMSE	ACC	Bias	bfRMSE	ACC
FREE	16×10^{12}	$\mathbf{2.1} \times 10^{12}$	0.8	19×10^{12}	$\mathbf{2.4} \times 10^{12}$	0.7
CTRL	7.0×10^{12}	3.25×10^{12}	0.8	10.0×10^{12}	3.3×10^{12}	0.9
+SIT	$\mathbf{-1.0} \times 10^{12}$	3.72×10^{12}	**0.9**	$\mathbf{4.0} \times 10^{12}$	2.5×10^{12}	**0.9**

FREE. +SIT reduces the bias but has a larger bfRMSE than CTRL when compared to PIOMAS. Still, ACC is improved, and Figure 5d suggests that PIOMAS underestimates interannual variability (quarrelling results for bfRMSE depending on whether we use C2SMOS or PIOMAS). +SIT has a strong discontinuity in 2010 for SIV during the transition between ENVISAT and C2SMOS.

In September, all systems have high positive SIE biases as a direct consequence of anomaly assimilation, as seen in Figure 6. Table 2 shows that CTRL and +SIT show higher ACC values than FREE. The agreement is better in +SIT, with ACC increasing from 0.7 to 0.8, and slightly reduced bfRMSE. In Figure 5, the ensemble mean of FREE shows, again, nearly no interannual variability. In the same figure, we see that +SIT better captures the amplitude of the peaks and, in particular, the minimum in 2007 and 2012. This shows a positive effect from the SIT assimilation, whereby a bias reduction in SIT at the end of winter leads to more grid cells becoming ice-free at the end of summer. For SIV, FREE captures the trend well. CTRL and +SIT show a good agreement of interannual variability with PIOMAS, albeit with +SIT showing overall the best agreement. +SIT also reduces the bias effectively.

The ensemble spread in extent and volume is shown in all experiments (Figure 6). The spread in FREE is the largest of the three experiments by far, particularly for the SIE. This is expected as FREE does not assimilate any observations, so the spread is less constrained. The ensemble spreads in CTRL and +SIT are comparable for the SIE in March and September, whereas +SIT has a lower spread for the SIV, particularly in March. Again, this is no surprise, because SIT observations are not assimilated in summer, so we expect +SIT to have a larger spread at the end of summer in September.

Finally, we investigate the mean climatology of IIEE (Section 3.1). The results (Figure 7) resemble the bias of SIE (Figure 5), which has the dominant contribution. However, +SIT and FREE had comparable SIE bias in summer, while for IIEE, SIT+ is superior, indicating it is not the sole contributor for IIEE. As such, FREE has the largest IIEE throughout the year, except in September, October and November (where it is equal to CTRL, Figure 7). +SIT performs best for all months. It yields

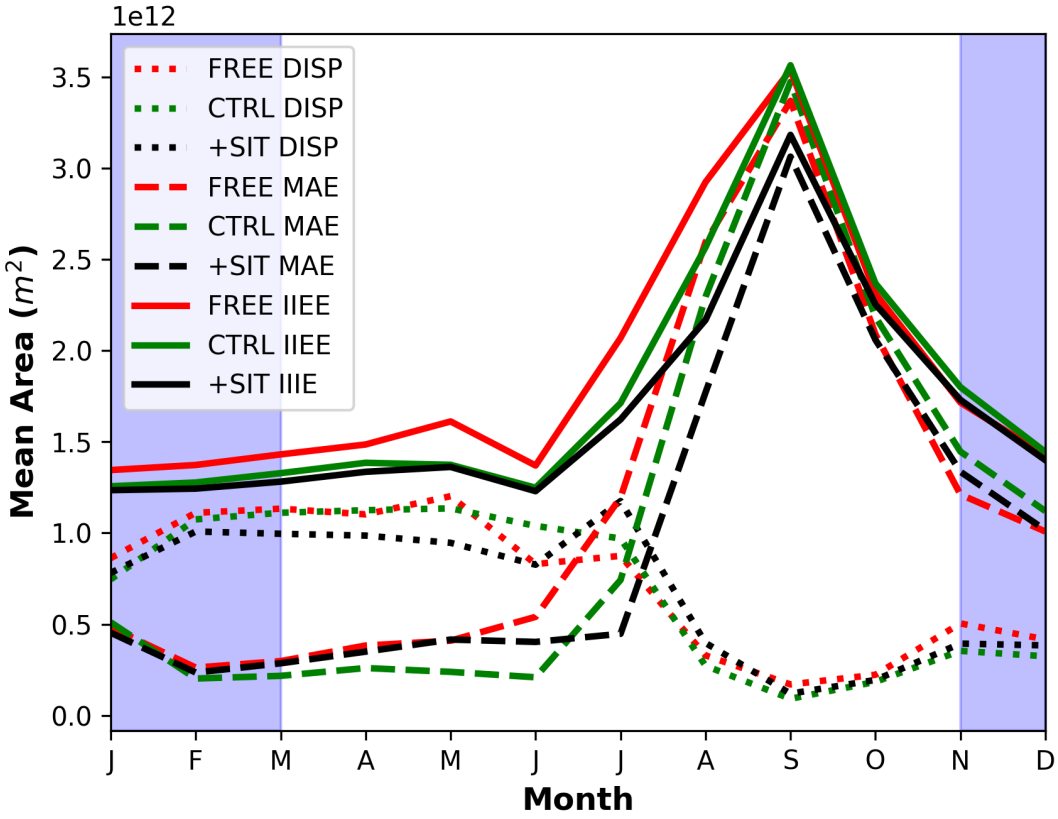

Figure 7. Climatology of IIEE, MAE and displacement (DISP) of three NorCPM experiments averaged over 2003-2023. The blue-shaded area shows the time when SIT observations are assimilated.

substantial improvement over FREE in the winter months (with some minor improvements over CTRL) and a substantial improvement over CTRL in the summer months. This implies that the location of the ice edge is most improved in summer by SIT assimilation. In winter, the location of the ice edge is controlled by the growth of new ice for which SIT plays a minor role, whereas in summer, the location of the ice edge is influenced by the melting of ice – i.e. when a grid cell becomes ice-free in summer (Section 5 for more detailed discussions) — for which the initial volume of ice is crucial. Thus, we can anticipate that SIT initialisation can bring added value to the prediction of the SIE minimum.

4.2 Prediction

We evaluate the prediction skill of our seasonal hindcasts for SIT, SIE, and IIEE using the metrics outlined in Section 3.1, and building on the reanalysis evaluation. For all quantities, we remove the linear trend that is highly predictable. The FREE experiment serves as a baseline for assessing skill driven by external forcings but shows minimal skill for detrended ACC

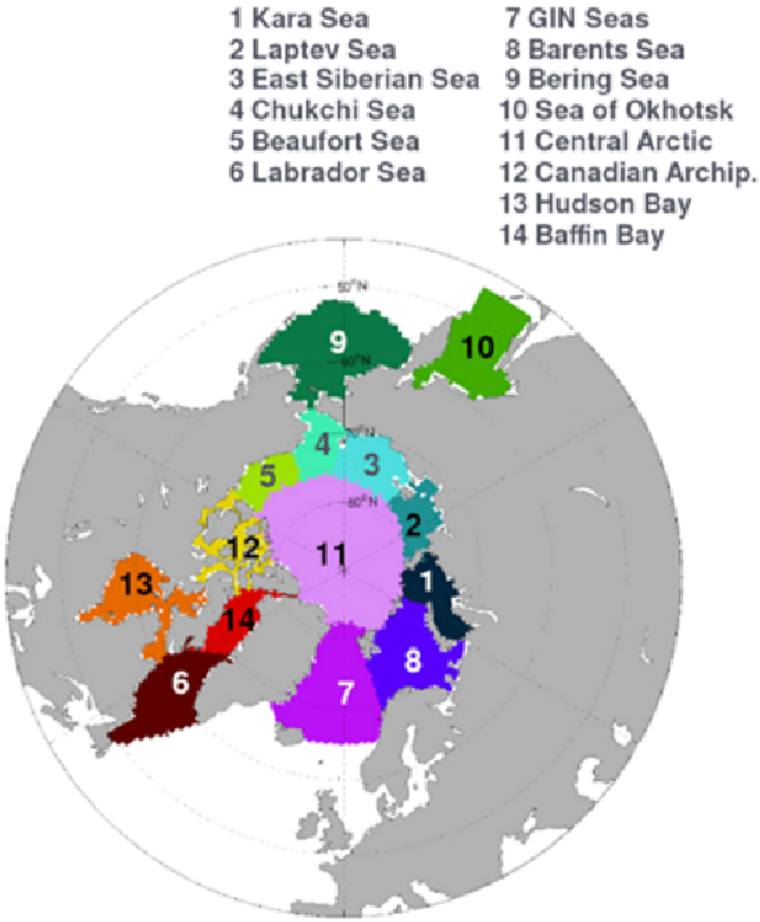

Figure 8. A map showing the regions used in this study. The regions are defined by Bushuk et al. (2017).

(Kimmritz et al., 2019), so its results are not discussed here. Prediction skill and underlying mechanisms vary significantly by region (Bushuk et al., 2024). Our assessment spans both the pan-Arctic region and specific basins, following the approach of Bushuk et al. (2017). We use the same regions as defined in Bushuk et al. (2017), shown in Figure 8. We compare the performance of CTRL and +SIT to highlight the role of SIT in improving predictability across different regions. We present the central Arctic, Beaufort Sea, Barents Sea, Bering Sea, and pan-Arctic regions, where the largest differences in SIE, SIT and IIEE skill between CTRL and +SIT are observed, but an assessment for other regions is also available in the supplementary material. Even if the time span of our experiment covers an unprecedentedly long period to assess the impact of SIT on seasonal predictions, the limited sample size of 21 hindcasts per season makes it challenging to perform the significance test of the differences between the systems.

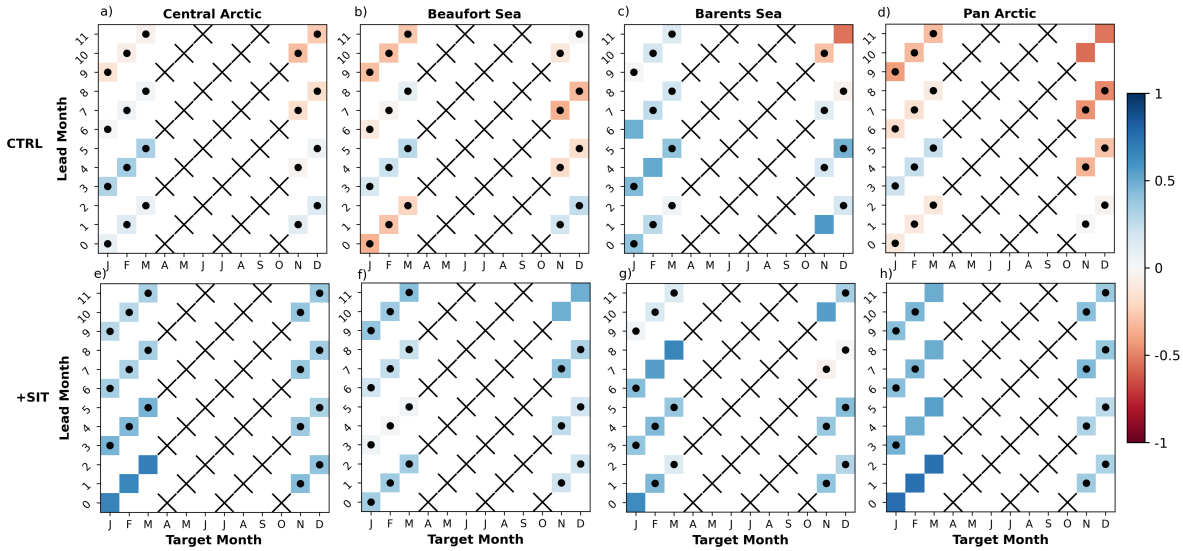

Figure 9. Detrended ACCs of our seasonal hindcasts for SIT from CTRL and +SIT with observations of SIT from CS2SMOS. Crosses are shown when comparison with observations is not possible due to lack of observations. The dots represent the ACC values that are not statistically significant.

4.2.1 Prediction of sea ice thickness

The assimilation of SIT observations improves detrended ACC and detrended bfRMSE of SIT prediction (Figures 9 and 10). In each region, ACCs of +SIT are positive and higher than the ones of CTRL, which are mostly not statistically significant (the top row of Figure 9). However, in most cases, ACC of +SIT does not pass the student's t-test due to the small sample size (the bottom row of Figure 9). Results are clearer with bfRMSE. +SIT shows lower bfRMSE than CTRL in almost all four regions and target months (Figure 10).

In the central Arctic, there is an overall improvement of detrended ACCs and bfRMSE for all lead months, but significant detrended ACCs +SIT are only found for +SIT in January-March. This coincides with times when detrended bfRMSE is also most reduced from CTRL. There is also a small reduction of detrended bfRMSE in winter. This suggests that the year-to-year anomaly of SIT can be predicted beyond 12 lead months.

In the Beaufort Sea, there is again an overall increase in detrended ACC (as for the central Arctic), but it is only significant in November-December, for hindcast initialised in January. Detrended bfRMSE is also strongly reduced for this hindcast. A clear improvement is also found in January-March for long lead times. It is somewhat surprising to see a degradation, in January for hindcasts started in October (at lead month 3), but an improvement for hindcasts started in March (at lead month 9). Both hindcasts assimilated the last SIT in March, but the October hindcast assimilated SIC and ocean observation from April to September in addition. This implies that the SIC assimilation degraded the accuracy of the SIT. Hence, we have seen in Section

4.1, that CTRL reduces error compared to FREE, but still performs substantially poorer than SIT+ for SIT. This indicates that the persistence of SIT initialisation across summer is more accurate than using SIC for updating SIT.

The benefit is much less clear in the Barents Sea. For the July-initialised prediction, there was significant ACC up to lead month 8 in February-March. There is also a significant correlation associated with a strong reduction of bfRMSE at lead month 12 in December. This is associated with a good climatology for SIT, SIE and (possibly) SID in November, December and January.

The pan-Arctic is the region for which we see the largest benefit from initialising SIT. The detrended ACC values are improved and the correlations are significant in January-March up to lead months 12. The associated detrended bfRMSE values are in agreement and also strongly reduced. Improvements in November-December are comparatively smaller than that of the later winter months, which can again be attributed to the lack of SIT observations in summer.

Overall, the SIT prediction results are quite promising and consistent with some previous studies (Blanchard-Wrigglesworth et al., 2011b, a). The impact is largest in the central Arctic and Beaufort Sea, where the observations are most accurate and where SIT anomalies persist longest. Our analysis did not assess prediction skills beyond 12 months lead time, but significant prediction skills of SIT were found to reach this limit — e.g. in the Beaufort and the Pan-Arctic.

4.2.2 Prediction of sea ice extent and sea ice edge

For the predictions of SIE, we focus on bfRMSE and IIEE, as results from the ACC are less clear. ACC differences for detecting benefits from initialization can be problematic if the skill from the externally forced climate trend is high and the ACC differences are small – the normalisation step can yield misleading results (Smith et al., 2019).

The differences in bfRMSE between CTRL and +SIT are more nuanced than for SIT, as shown in Figure 11 and somewhat disappointing. In the central Arctic, there is little change except in the September SIE prediction for the April and July initialisations, which shows some improvements in +SIT. In the Barents Sea, the results mixed, with some improvements in predictions from January, but degradation in the other start months. In the Beaufort Sea, there are quite strong improvements around August-October for every lead time, but it does also appear to lead to some degradations in the months preceding. Finally, in the Bering Sea, improvements are more homogeneous for the first few months, with January initialised hindcast showing more substantial improvements. There are some improvements in bfRMSE in some other Pacific-side marginal regions.

As we did for the reanalysis, we investigate the IIEE of the hindcasts in Figure 12. The IIEE shows substantial improvement centered around September for all lead months, with improvement generally occurring between August to November. The result for the full RMSE of SIE looks very similar (not shown). As such, the location of the ice edge is improved primarily due to the bias changes in SIT, which improved the SIE bias during that period (Figure 5). In order to better exemplify that we will look in more detail at the Beaufort Sea has consistently shown the most positive results for predictions of SIE in our study so far.

We look at the SIE climatological mean of the hindcast in the different systems and for different start months in the Beaufort Sea (Figure 13). The +SIT system climatological minimum is in better agreement with the NOAA observations for each start month than the CTRL and FREE predictions. The improvements are largest during the melt season – even for the October

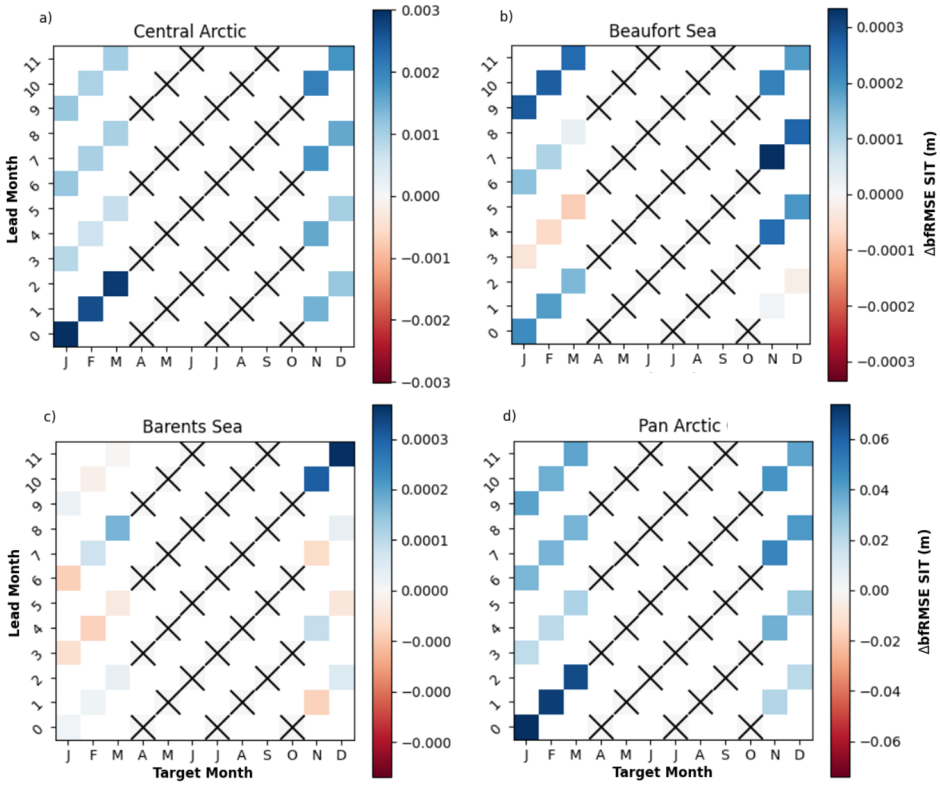

Figure 10. Differences in detrended bfRMSE of SIT from CTRL and +SIT experiments against observations from CS2SMOS. As Figure 9, we show bfRMSE only in four of the regions defined by Bushuk et al. (2017). Blue (red) colour indicates +SIT is better (worse) than CTRL. Crosses are shown when comparison with observations is not possible due to lack of observations.

initialised hindcast. This is in good agreement with the improvements we saw for the IIEE and confirms that the main reason for improvements in IIEE relates to improving the SIT bias that reduces the bias in SIE during the melting season.

5 Discussions and conclusions

In this study, we have used the NorCPM coupled global climate model to assimilate SIT data alongside SIC, SST and ocean profiles of temperature and salinity data. We compared the performance of our system that assimilated SIT in addition to the other datasets, to highlight the role of SIT in improving predictability across different regions. We first produced a reanalysis and then used this reanalysis to initialise seasonal hindcasts (+SIT). We validated our results not only using the standard metrics (bias, bfRMSE and ACC), but also IIEE, which provides unique insights on skill at the ice edge. We also measured our results against independent SIT measurements from the BGEP ULS moorings.

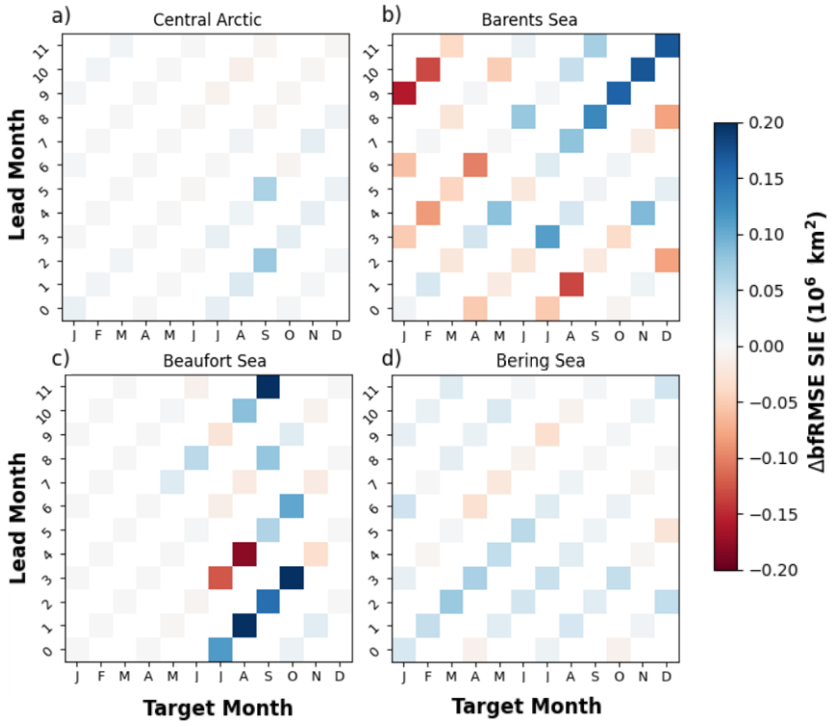

Figure 11. Difference in SIE detrended bfRMSE of CTRL and +SIT predictions computed with observations of SIE from NOAA in four Arctic regions. Blue (red) indicates that +SIT has lower (higher) bfRMSE than CTRL).

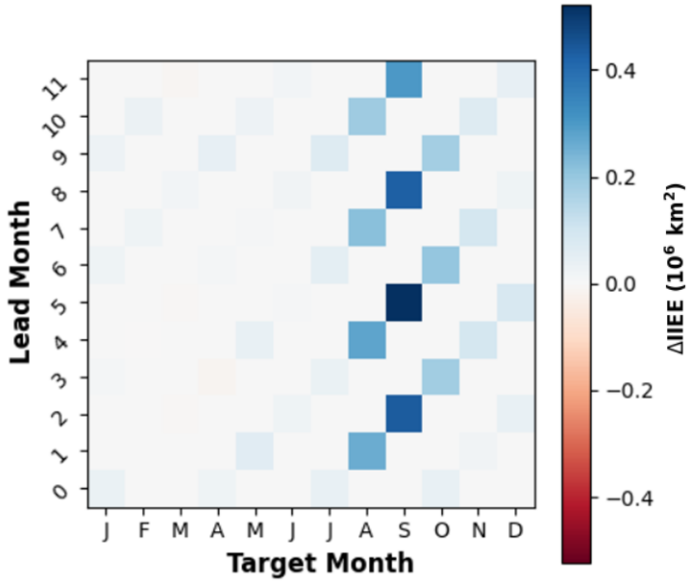

Figure 12. Difference of IIEE between CTRL and +SIT over the whole Arctic. Blue (red) means +SIT is better (worse) than CTRL.

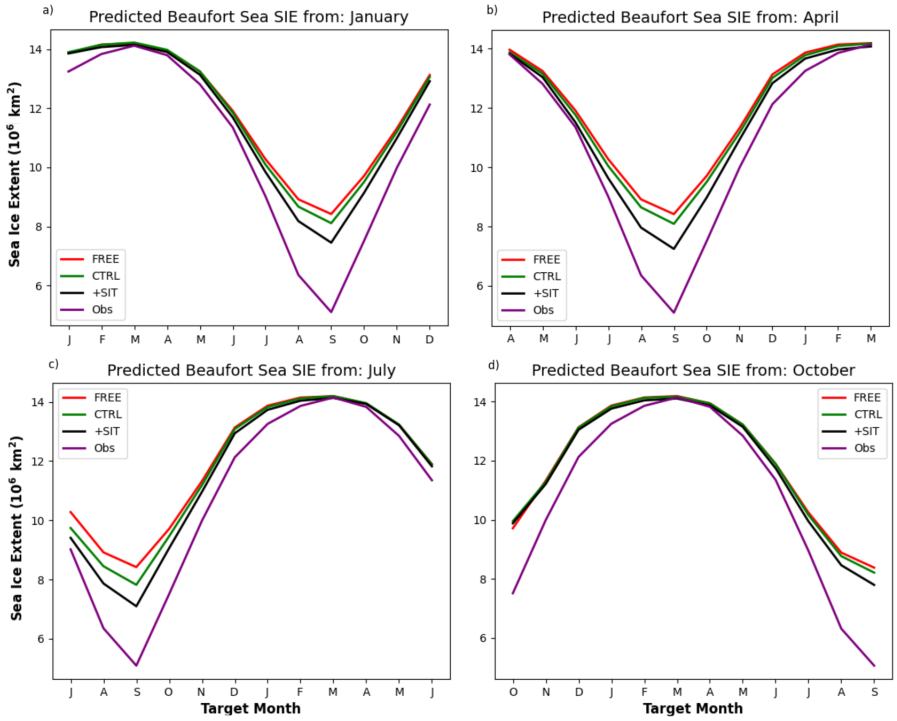

Figure 13. Climatological mean of SIE predictions from our NorCPM experiments and observations from NOAA in the Beaufort Sea. The region is as defined by Bushuk et al. (2017)

We evaluated the assimilation of ENVISAT and CS2SMOS SIT data into NorCPM. While ENVISAT has higher uncertainties than CS2SMOS, it extended the reanalysis period and improved SIT and ice edge hindcast estimates in the central Arctic, particularly during the melt season through the SIT bias reduction in winter.

CS2SMOS provided more accurate data and greater reductions in biases in comparison to ENVISAT. Despite its limitations, ENVISAT data proved useful for reanalyses – highlighted by the DFS metrics that quantify the influence of each observational product – and provides meaningful observations in a period and location where observations are crucially lacking. We experienced some challenges in the transition period between ENVISAT and CS2SMOS, noticeable by discontinuities in the time series of SIE and SIV. It is because ENVISAT has a high uncertainty due to the instrumentation, which disproportionately observes thicker ice (Louet and Bruzzi, 1999; Schwegmann et al., 2016; Tilling et al., 2018). There has been some initiatives to harmonize the two products by correcting for the bias in ENVISAT using CS2 and the period where the two products observational period overlapped between 2010–2012 (Tilling et al., 2019). Using a bias-corrected ENVISAT together with CS2 could lead to further improvements in reanalysis and prediction for SIT and SIE. There will also be a possibility in the future to assimilate a recently developed ice thickness dataset dating back to 1994 (Bocquet et al., 2024), substantially lengthening the reanalysis by a further ten years and ensuring the consistency of the different sources of observations.

Our newest reanalysis +SIT showed substantial improvements compared to previous NorCPM versions (i.e., CTRL and FREE) with regard to ice thickness and ice volume. Bias and bfRMSE in SIT were both significantly reduced in +SIT, as our FREE and CTRL reanalyses have too thick ice all over the central Arctic and then large positive thickness biases in some places of up to 5 metres in comparison to CS2SMOS. The positive effects of the thickness assimilation in the +SIT reanalysis thus led to improvements in the ice volume estimates. The improvements in the volume estimates are difficult to completely validate due to the lack of sea ice volume observations, but compared well with the PIOMAS reanalysis, which has been extensively validated with thickness measurements over a long period (Schweiger et al., 2011). The +SIT reanalysis also showed a reduced bfRMSE of SIV, highlighting that the assimilation does correct more than just the SIT bias. Validation with independent BGEP ULS measurements of SIT showed that the +SIT reanalysis had substantially smaller bias and bfRMSE than CTRL and FREE, including during the ENVISAT period from 2003 to October 2010. The improvement in SIT also yielded improvement in SIE during the melting season and, therefore, also for IIEE.

We found that thickness prediction was substantially improved by the ice thickness assimilation across almost all regions. The largest part of the error reduction was from correcting model bias, which persisted throughout the year in good agreement with previous studies (Bushuk et al., 2017; Schröder et al., 2019). However, our results also showed that the SIT assimilation improved the detrended bfRMSE up to 12 lead months, with the most notable improvement in the Pan-Arctic, Beaufort Sea and the central Arctic. Skill was not as clear in the Barents Sea. The results showed that we could predict SIT anomalies for at least up to a year in some regions. It would be interesting to extend the length of the hindcast to asses up to which time scale year-to-year anomalies of SIT are predictable. We could also clearly identify the detrimental impact of the lack of SIT observations during summer, highlighting the need to consider a product that extrapolates SIT estimate during that season – as for example Landy et al. (2022).

Correcting the SIT bias and anomaly had a lower-than-expected impact on the Arctic SIE predictions. Correcting the SIT bias yielded improvements in the climatology of SIE in the melting season and, as a consequence for the IIEE metric. This improvement in late summer agreed with previous sea ice prediction studies with thickness initialisation (Bushuk et al., 2017; Schröder et al., 2019), which have found SIT assimilation can help to reduce thickness biases and thus improve SIE prediction. The reason is that SIT reduced the positive bias in our model and, as a consequence, increased the area and the transition pace towards a sea ice minimum. However, if one removes the trend and the mean bias, the SIT brings some but little added value. Some skills were identified in the Beaufort Sea during summer, central Arctic and Bering Sea. This was when and where NorCPM (CTRL) tends to perform the poorest compared to other dynamical models submitted to SIPN (Bushuk et al., 2024). Still, SIT assimilation also yielded some degradation in the other basins (e.g., Barents Sea). This was likely related to the discontinuity in the observation period (summer observations missing and transition from ENVISAT to CS2SMOS). Furthermore, it should be remembered that even if our study covers an unprecedentedly long period (21 years), this is still too short to assess robustly year-to-year variability, especially considering the large trend in Arctic sea ice that can module internal variability.

Overall, this study advances the prediction capabilities of sea ice in NorCPM through additionally assimilating SIT data alongside SST, ocean profiles of salinity and temperature and SIC data. Here, we showed some improvements in SIT, SIE and

IIEE prediction using SIT initialisation. While the improvement of SIT was extended for the whole year and all lead time, SIE and IIEE were improved primarily during summer and in the central Arctic, where the improvement relates to a reduction of bias. We also found that the ENVISAT dataset can be useful for sea ice reanalyses and prediction (for monthly averages), which, to our knowledge, have been assimilated into a global climate model for the first time. Still, the SIT improvement on detrended SIE anomalies is lower than expected. As SIT plays an important role in the dynamics and thermodynamics of sea ice, a possible reason may be that the sea ice model parameters used in this study have been calibrated to compensate for model bias. This can be bias in SIT, but also bias in the other components. A new ensemble-based parameter estimate was developed in NorCPM to tune model parameters efficiently (Singh et al., 2022, 2024). For our next steps, we aim to test the use of sea ice drift assimilation for refining three key sea ice parameters: air-ice stress, ocean-ice stress, and ice strength in a version of the model state that is sustained to a low level.

Code and data availability. EN4.2.1 and EN4.2.2 salinity and temperature profile data can be downloaded from https://www.metoffice.gov.uk/hadobs/en4/download-en4-2-1.html, NOAA SST and SIC data is available from ftp://ftp.cdc.noaa.gov/Datasets/noaa.oisst.v2/, ENVISAT ESA CCI SIT data is available from http://catalogue.ceda.ac.uk/uuid/f4c34f4f0f1d4d0da06d771f6972f180, C2SMOS v2.6 data is available from ftp://ftp.awi.de/sea_ice/product/cryosat2_smos/v206/, PIOMAS data is available from https://psc.apl.uw.edu/research/projects/arctic-sea-ice-volume-anomaly/data/. BGEP ULS mooring data is available from the Woods Hole Oceanographic Institution website https://www2.whoi.edu/site/beaufortgyre/data/mooring-data/. Reanalysis and Hindcast Data used to produce the figures are available upon request.

Author contributions. FC, YW and NW all wrote code pertaining to sea ice thickness assimilation in NorCPM and designed the experiments. NW conducted the experiments and produced the figures. NW produced the paper with assistance, feedback and edits from YW and FC.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. This study was supported by the Research Council of Norway (Grant Nos. 328886, 350390, 309562, 352204). It was also partly funded by ObsSea4Clim "Ocean observations and indicators for climate and assessments" funded by the European Union. Grant Agreement number: 101136548. DOI: 10.3030/101136548. Contribution nr. 38. This work also received grants for computer time from the Norwegian Program for supercomputer (NN9039K) and storage grants (NS9039K). We would also like to acknowledge the reviewers for their helpful comments and suggestions to improve the manuscript.

References

Arruda, G. M. and Krutkowski, S.: Social impacts of climate change and resource development in the Arctic: Implications for Arctic governance, Journal of Enterprising Communities: People and Places in the Global Economy, 11, 277–288, 2017.

Bentsen, M., Bethke, I., Debernard, J., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I., Hoose, C., et al.: The Norwegian earth system model, NorESM1-M–Part 1: Description and basic evaluation, Geosci. Model Dev. Discuss, 5, 2843–2931, 2012.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. A., Hoose, C., et al.: The Norwegian Earth System Model, NorESM1-M–Part 1: description and basic evaluation of the physical climate, Geoscientific Model Development, 6, 687–720, 2013.

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H., Svendsen, L., Chiu, P.-G., et al.: NorCPM1 and its contribution to CMIP6 DCPP, Geoscientific Model Development, 14, 7073–7116, 2021.

Bitz, C., Holland, M., Weaver, A., and Eby, M.: Simulating the ice-thickness distribution in a coupled climate model, Journal of Geophysical Research: Oceans, 106, 2441–2463, 2001.

Bitz, C. M. and Lipscomb, W. H.: An energy-conserving thermodynamic model of sea ice, Journal of Geophysical Research: Oceans, 104, 15 669–15 677, 1999.

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, Journal of Climate, 24, 231–250, 2011a.

Blanchard-Wrigglesworth, E., Bitz, C., and Holland, M.: Influence of initial conditions and climate forcing on predicting Arctic sea ice, Geophysical Research Letters, 38, 2011b.

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

Bocquet, M., Fleury, S., Rémy, F., and Piras, F.: Arctic and Antarctic sea ice thickness and volume changes from observations between 1994 and 2023, Journal of Geophysical Research: Oceans, 129, e2023JC020 848, 2024.

Briegleb, B. and Light, B.: A Delta-Eddington multiple scattering parameterization for solar radiation in the sea ice component of the Community Climate System Model, NCAR technical note, pp. 1–108, 2007.

Bushuk, M., Msadek, R., Winton, M., Vecchi, G. A., Gudgel, R., Rosati, A., and Yang, X.: Skillful regional prediction of Arctic sea ice on seasonal timescales, Geophysical Research Letters, 44, 4953–4964, 2017.

Bushuk, M., Winton, M., Bonan, D. B., Blanchard-Wrigglesworth, E., and Delworth, T. L.: A mechanism for the Arctic sea ice spring predictability barrier, Geophysical Research Letters, 47, e2020GL088 335, 2020.

Bushuk, M., Ali, S., Bailey, D. A., Bao, Q., Batté, L., Bhatt, U. S., Blanchard-Wrigglesworth, E., Blockley, E., Cawley, G., Chi, J., et al.: Predicting September Arctic Sea Ice: A Multi-Model Seasonal Skill Comparison, Bulletin of the American Meteorological Society, 2024.

Cardinali, C., Pezzulli, S., and Andersson, E.: Influence-matrix diagnostic of a data assimilation system, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 130, 2767–2786, 2004.

Carrassi, A., Weber, R., Guemas, V., Doblas-Reyes, F., Asif, M., and Volpi, D.: Full-field and anomaly initialization using a low-order climate model: a comparison and proposals for advanced formulations, Nonlinear Processes in Geophysics, 21, 521–537, 2014.

Cavalieri, D. J. and Parkinson, C. L.: Arctic sea ice variability and trends, 1979–2010, The Cryosphere, 6, 881–889, 2012.

Chevallier, M., y Mélia, D. S., Voldoire, A., Déqué, M., and Garric, G.: Seasonal forecasts of the pan-Arctic sea ice extent using a GCM-based seasonal prediction system, Journal of Climate, 26, 6092–6104, 2013.

Cohen, J., Screen, J. A., Furtado, J. C., Barlow, M., Whittleston, D., Coumou, D., Francis, J., Dethloff, K., Entekhabi, D., Overland, J., et al.: Recent Arctic amplification and extreme mid-latitude weather, Nature geoscience, 7, 627–637, 2014.

Comiso, J.: Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, Version 3, boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2017.

Comiso, J. C., Parkinson, C. L., Gersten, R., and Stock, L.: Accelerated decline in the Arctic sea ice cover, Geophysical research letters, 35, 2008.

Connor, L. N., Laxon, S. W., Ridout, A. L., Krabill, W. B., and McAdoo, D. C.: Comparison of Envisat radar and airborne laser altimeter measurements over Arctic sea ice, Remote Sensing of Environment, 113, 563–570, 2009.

Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, L., and Zheng, F.: Seasonal-to-decadal predictions with the ensemble Kalman filter and the Norwegian Earth System Model: a twin experiment, Tellus A: Dynamic Meteorology and Oceanography, 66, 21 074, 2014.

Counillon, F., Keenlyside, N., Bethke, I., Wang, Y., Billeau, S., Shen, M. L., and Bentsen, M.: Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model, Tellus A: Dynamic Meteorology and Oceanography, 68, 32 437, 2016.

Craig, A. P., Vertenstein, M., and Jacob, R.: A new flexible coupler for earth system modeling developed for CCSM4 and CESM1, The International Journal of High Performance Computing Applications, 26, 31–42, 2012.

Dai, A., Luo, D., Song, M., and Liu, J.: Arctic amplification is caused by sea-ice loss under increasing CO2, Nature communications, 10, 1–13, 2019.

Descamps, S., Aars, J., Fuglei, E., Kovacs, K. M., Lydersen, C., Pavlova, O., Pedersen, Å. Ø., Ravolainen, V., and Strøm, H.: Climate change impacts on wildlife in a High Arctic archipelago–Svalbard, Norway, Global Change Biology, 23, 490–502, 2017.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10 143–10 162, 1994.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean dynamics, 53, 343–367, 2003.

Fritzner, S., Graversen, R., Christensen, K. H., Rostosky, P., and Wang, K.: Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system, The Cryosphere, 13, 491–509, 2019.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, 1999.

Giesse, C., Notz, D., and Baehr, J.: On the origin of discrepancies between observed and simulated memory of Arctic sea ice, Geophysical Research Letters, 48, e2020GL091 784, 2021.

Giles, K. A., Laxon, S. W., and Ridout, A. L.: Circumpolar thinning of Arctic sea ice following the 2007 record ice extent minimum, Geophysical Research Letters, 35, 2008.

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

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, Journal of Geophysical Research: Oceans, 118, 6704–6716, 2013.

Gouretski, V. and Cheng, L.: Correction for systematic errors in the global dataset of temperature profiles from mechanical bathythermographs, Journal of Atmospheric and Oceanic Technology, 37, 841–855, 2020.

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

Guemas, V., Chevallier, M., Déqué, M., Bellprat, O., and Doblas-Reyes, F.: Impact of sea ice initialization on sea ice and atmosphere prediction skill on seasonal timescales, Geophysical Research Letters, 43, 3889–3896, 2016.

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, Climate dynamics, 36, 1239–1253, 2011.

Holland, M. M., Bailey, D. A., Briegleb, B. P., Light, B., and Hunke, E.: Improved sea ice shortwave radiation physics in CCSM4: The impact of melt ponds and aerosols on Arctic sea ice, Journal of Climate, 25, 1413–1430, 2012.

Holland, M. M., Landrum, L., Bailey, D., and Vavrus, S.: Changing seasonal predictability of Arctic summer sea ice area in a warming climate, Journal of Climate, 32, 4963–4979, 2019.

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., and Zhang, H.-M.: Improvements of the daily optimum interpolation sea surface temperature (DOISST) version 2.1, Journal of Climate, 34, 2923–2939, 2021.

Hunke, E., Lipscomb, W., Turner, A., Jeffery, N., and Elliott, S.: CICE: The Los Alamos sea ice model documentation and software user's manual 1568 version 5.1, Tech. rep., Los Alamos National Laboratory, 2015.

Hunke, E. C. and Dukowicz, J. K.: An elastic–viscous–plastic model for sea ice dynamics, Journal of Physical Oceanography, 27, 1849–1867, 1997.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., et al.: The community earth system model: a framework for collaborative research, Bulletin of the American Meteorological Society, 94, 1339–1360, 2013.

Ivanova, N., Pedersen, L. T., Tonboe, R., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., et al.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, The Cryosphere, 9, 1797–1817, 2015.

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

Kimmritz, M., Counillon, F., Smedsrud, L. H., Bethke, I., Keenlyside, N., Ogawa, F., and Wang, Y.: Impact of ocean and sea ice initialisation on seasonal prediction skill in the Arctic, Journal of Advances in Modeling Earth Systems, 11, 4147–4166, 2019.

Kirkevåg, A., Iversen, T., Seland, Ø., Hoose, C., Kristjánsson, J., Struthers, H., Ekman, A. M., Ghan, S., Griesfeller, J., Nilsson, E. D., et al.: Aerosol–climate interactions in the Norwegian Earth System Model–NorESM1-M, Geoscientific Model Development, 6, 207–244, 2013.

Koenigk, T., König Beatty, C., Caian, M., Döscher, R., and Wyser, K.: Potential decadal predictability and its sensitivity to sea ice albedo parameterization in a global coupled model, Climate dynamics, 38, 2389–2408, 2012.

Krishfield, R. and Proshutinsky, A.: BGOS ULS data processing procedure, Woods Hole Oceanographic Institution, 2006.

Krishfield, R. A., Proshutinsky, A., Tateyama, K., Williams, W. J., Carmack, E. C., McLaughlin, F. A., and Timmermans, M.-L.: Deterioration of perennial sea ice in the Beaufort Gyre from 2003 to 2012 and its impact on the oceanic freshwater cycle, Journal of Geophysical Research: Oceans, 119, 1271–1305, 2014.

Kwok, R.: Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018), Environmental Research Letters, 13, 105 005, 2018.

Kwok, R., Zwally, H. J., and Yi, D.: ICESat observations of Arctic sea ice: A first look, Geophysical Research Letters, 31, 2004.

Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., and Janssen, P.: A coupled data assimilation system for climate reanalysis, Quarterly Journal of the Royal Meteorological Society, 142, 65–78, 2016.

Landy, J., Dawson, G., Tsamados, M., and et al: A year-round satellite sea-ice thickness record from CryoSat-2, Nature, 2022.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., et al.: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model, Journal of Advances in Modeling Earth Systems, 3, 2011.

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., et al.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophysical Research Letters, 40, 732–737, 2013.

Lipscomb, W. H.: Remapping the thickness distribution in sea ice models, Journal of Geophysical Research: Oceans, 106, 13 989–14 000, 2001.

Lipscomb, W. H. and Hunke, E. C.: Modeling sea ice transport using incremental remapping, Monthly weather review, 132, 1341–1354, 2004.

Louet, J. and Bruzzi, S.: ENVISAT mission and system, in: IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293), vol. 3, pp. 1680–1682, IEEE, 1999.

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

Massonnet, F., Barthélemy, A., Worou, K., Fichefet, T., Vancoppenolle, M., Rousset, C., and Moreno-Chamarro, E.: On the discretization of the ice thickness distribution in the NEMO3. 6-LIM3 global ocean–sea ice model, Geoscientific Model Development, 12, 3745–3758, 2019.

Min, C., Yang, Q., Luo, H., Chen, D., Krumpen, T., Mamnun, N., Liu, X., and Nerger, L.: Improving Arctic sea-ice thickness estimates with the assimilation of CryoSat-2 summer observations, Ocean-Land-Atmosphere Research, 2, 0025, 2023.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., et al.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, 2010.

Msadek, R., Vecchi, G. A., Winton, M., and Gudgel, R. G.: Importance of initial conditions in seasonal predictions of Arctic sea ice extent, Geophysical Research Letters, 41, 5208–5215, 2014.

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., Conley, A. J., Garcia, R., Kinnison, D., Lamarque, J.-F., Marsh, D., Mills, M., Smith, A. K., Tilmes, S., Vitt, F., Morrison, H., Cameron-Smith, P., Collins, W. D., Iacono, M. J., Easter, R. C., Ghan, S. J., Liu, X., Rasch, P. J., and Taylor, M. A.: Description of the NCAR Community Atmosphere Model (CAM 4.0), Technical Report NCAR/TN-485+STR, National Center for Atmospheric Research (NCAR), chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www2.cesm.ucar.edu/models/ccsm4.0/cam/docs/description/cam4_desc.pdf, accessed: 2024-07-25, 2010.

Nghiem, S., Rigor, I., Perovich, D., Clemente-Colón, P., Weatherly, J., and Neumann, G.: Rapid reduction of Arctic perennial sea ice, Geophysical Research Letters, 34, 2007.

Ordoñez, A. C., Bitz, C. M., and Blanchard-Wrigglesworth, E.: Processes controlling Arctic and Antarctic sea ice predictability in the Community Earth System Model, Journal of Climate, 31, 9771–9786, 2018.

Passos, L., Langehaug, H. R., Årthun, M., Eldevik, T., Bethke, I., and Kimmritz, M.: Impact of initialization methods on the predictive skill in NorCPM: an Arctic–Atlantic case study, Climate Dynamics, 60, 2061–2080, 2023.

Peterson, K. A., Arribas, A., Hewitt, H., Keen, A., Lea, D., and McLaren, A.: Assessing the forecast skill of Arctic sea ice extent in the GloSea4 seasonal prediction system, Climate Dynamics, 44, 147–162, 2015.

Petty, A. A., Kurtz, N. T., Kwok, R., Markus, T., and Neumann, T. A.: Winter Arctic sea ice thickness from ICESat-2 freeboards, Journal of Geophysical Research: Oceans, 125, e2019JC015 764, 2020.

Przybylak, R. and Wyszyński, P.: Air temperature changes in the Arctic in the period 1951–2015 in the light of observational and reanalysis data, Theoretical and Applied Climatology, 139, 75–94, 2020.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily high-resolution-blended analyses for sea surface temperature, Journal of climate, 20, 5473–5496, 2007.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607–1623, 2017.

Ricker, R., Girard-Ardhuin, F., Krumpen, T., and Lique, C.: Satellite-derived sea ice export and its impact on Arctic ice mass balance, The Cryosphere, 12, 3017–3032, 2018.

Rodwell, M., Lang, S. T., Ingleby, N., Bormann, N., Hólm, E., Rabier, F., Richardson, D., and Yamaguchi, M.: Reliability in ensemble data assimilation, Quarterly Journal of the Royal Meteorological Society, 142, 443–454, 2016.

Rothrock, D., Percival, D., and Wensnahan, M.: The decline in arctic sea-ice thickness: Separating the spatial, annual, and interannual variability in a quarter century of submarine data, Journal of Geophysical Research: Oceans, 113, 2008.

Sakov, P. and Oke, P. R.: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus A: Dynamic Meteorology and Oceanography, 60, 361–371, 2008.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K., Oke, P., and Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic, Ocean Science, 8, 633–656, 2012.

Schröder, D., Feltham, D. L., Flocco, D., and Tsamados, M.: September Arctic sea-ice minimum predicted by spring melt-pond fraction, Nature Climate Change, 4, 353–357, 2014.

Schröder, D., Feltham, D. L., Tsamados, M., Ridout, A., and Tilling, R.: New insight from CryoSat-2 sea ice thickness for sea ice modelling, The Cryosphere, 13, 125–139, 2019.

Schwegmann, S., Rinne, E., Ricker, R., Hendricks, S., and Helm, V.: About the consistency between Envisat and CryoSat-2 radar freeboard retrieval over Antarctic sea ice, The Cryosphere, 10, 1415–1425, 2016.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic sea ice volume, Journal of Geophysical Research: Oceans, 116, 2011.

Serreze, M. C. and Meier, W. N.: The Arctic's sea ice cover: trends, variability, predictability, and comparisons to the Antarctic, Annals of the New York Academy of Sciences, 1436, 36–53, 2019.

Serreze, M. C., Crawford, A. D., Stroeve, J. C., Barrett, A. P., and Woodgate, R. A.: Variability, trends, and predictability of seasonal sea ice retreat and advance in the C hukchi S ea, Journal of Geophysical Research: Oceans, 121, 7308–7325, 2016.

Sievers, I., Rasmussen, T. A., and Stenseng, L.: Assimilating CryoSat-2 freeboard to improve Arctic sea ice thickness estimates, The Cryosphere Discussions, 2023, 1–23, 2023.

Singh, T., Counillon, F., Tjiputra, J., Wang, Y., and Gharamti, M. E.: Estimation of ocean biogeochemical parameters in an earth system model using the dual one step ahead smoother: A twin experiment, Frontiers in Marine Science, 9, 775 394, 2022.

Singh, T., Counillon, F., Tjiputra, J. F., and Wang, Y.: Ensemble-based parameter estimation for improving ocean biogeochemistry in an Earth system model, Authorea Preprints, 2024.

Smith, D., Eade, R., Scaife, A., Caron, L.-P., Danabasoglu, G., DelSole, T., Delworth, T., Doblas-Reyes, F., Dunstone, N., Hermanson, L., et al.: Robust skill of decadal climate predictions, Npj Climate and Atmospheric Science, 2, 13, 2019.

Steele, M., Ermold, W., and Zhang, J.: Arctic Ocean surface warming trends over the past 100 years, Geophysical Research Letters, 35, 2008.

Stroeve, J., Hamilton, L. C., Bitz, C. M., and Blanchard-Wrigglesworth, E.: Predicting September sea ice: Ensemble skill of the SEARCH sea ice outlook 2008–2013, Geophysical Research Letters, 41, 2411–2418, 2014a.

Stroeve, J. C., Markus, T., Boisvert, L., Miller, J., and Barrett, A.: Changes in Arctic melt season and implications for sea ice loss, Geophysical Research Letters, 41, 1216–1225, 2014b.

Sumata, H., Gerdes, R., Kauker, F., and Karcher, M.: Empirical error functions for monthly mean Arctic sea-ice drift, Journal of Geophysical Research: Oceans, 120, 7450–7475, 2015.

Sumata, H., de Steur, L., Divine, D. V., Granskog, M. A., and Gerland, S.: Regime shift in Arctic Ocean sea ice thickness, Nature, 615, 443–449, 2023.

Sun, S. and Solomon, A.: Suitability of the CICE sea ice model for seasonal prediction and positive impact of CryoSat-2 ice thickness initialization, The Cryosphere, 18, 3033–3048, 2024.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American meteorological Society, 93, 485–498, 2012.

Tietsche, S., Day, J. 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, Geophysical Research Letters, 41, 1035–1043, 2014.

Tilling, R., Ridout, A., and Shepherd, A.: Assessing the impact of lead and floe sampling on Arctic sea ice thickness estimates from Envisat and CryoSat-2, Journal of Geophysical Research: Oceans, 124, 7473–7485, 2019.

Tilling, R. L., Ridout, A., and Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data, Advances in Space Research, 62, 1203–1225, 2018.

Von Storch, H. and Zwiers, F. W.: Statistical analysis in climate research, Cambridge university press, 2002.

Wahba, G., Johnson, D. R., Gao, F., and Gong, J.: Adaptive tuning of numerical weather prediction models: Randomized GCV in three-and four-dimensional data assimilation, Monthly Weather Review, 123, 3358–3370, 1995.

Wang, W., Chen, M., and Kumar, A.: Seasonal prediction of Arctic sea ice extent from a coupled dynamical forecast system, Monthly Weather Review, 141, 1375–1394, 2013.

Wang, X., Key, J., Kwok, R., and Zhang, J.: Comparison of Arctic sea ice thickness from satellites, aircraft, and PIOMAS data, Remote Sensing, 8, 713, 2016.

Wang, Y., Counillon, F., Bethke, I., Keenlyside, N., Bocquet, M., and Shen, M.-l.: Optimising assimilation of hydrographic profiles into isopycnal ocean models with ensemble data assimilation, Ocean Modelling, 114, 33–44, 2017.

Wang, Y., Counillon, F., Keenlyside, N., Svendsen, L., Gleixner, S., Kimmritz, M., Dai, P., and Gao, Y.: Seasonal predictions initialised by assimilating sea surface temperature observations with the EnKF, Climate Dynamics, 53, 5777–5797, 2019.

Williams, N., Byrne, N., Feltham, D., Van Leeuwen, P. J., Bannister, R., Schroeder, D., Ridout, A., and Nerger, L.: The effects of assimilating a sub-grid scale sea ice thickness distribution in a new Arctic sea ice data assimilation system, EGUsphere, pp. 1–35, 2022.

Xie, J., Counillon, F., and Bertino, L.: Impact of assimilating a merged sea-ice thickness from CryoSat-2 and SMOS in the Arctic reanalysis, The Cryosphere, 12, 3671–3691, 2018.

Zhang, J. and Rothrock, D. A.: Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, Monthly Weather Review, 131, 845–861, 2003.

Zhang, Y.-F., Bushuk, M., Winton, M., Hurlin, B., Yang, X., Delworth, T., and Jia, L.: Assimilation of satellite-retrieved sea ice concentration and prospects for september predictions of Arctic sea ice, Journal of Climate, 34, 2107–2126, 2021.

Zhang, Y.-F., Bushuk, M., Winton, M., Hurlin, B., Gregory, W., Landy, J., and Jia, L.: Improvements in September Arctic Sea Ice Predictions Via Assimilation of Summer CryoSat-2 Sea Ice Thickness Observations, Geophysical Research Letters, 50, e2023GL105 672, 2023.

Zheng, L., Cheng, X., Chen, Z., and Liang, Q.: Delay in Arctic Sea ice freeze-up linked to early summer sea ice loss: evidence from satellite observations, Remote Sensing, 13, 2162, 2021.