Assimilation of sea ice thickness derived from CryoSat-2 along-track freeboard measurements into the Met Office's Forecast Ocean Assimilation Model (FOAM)

Emma K. Fiedler[1], Matthew J. Martin[1], Ed Blockley[1], Davi Mignac[1], Nicolas Fournier[1], Andy Ridout[2], Andrew Shepherd[3], and Rachel Tilling[4,5]

[1]Met Office, Exeter, EX1 3PB, UK
[2]Centre for Polar Observation and Modelling, University College London, London, WC1E 6BT, UK
[3]Centre for Polar Observation and Modelling, University of Leeds, Leeds, LS2 9JT, UK
[4]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[5]Earth System Interdisciplinary Center, University of Maryland, College Park, MD, USA

Correspondence: Emma Fiedler (emma.fiedler@metoffice.gov.uk)

Abstract. The feasibility of assimilating sea ice thickness (SIT) observations derived from CryoSat-2 along-track measurements of sea ice freeboard is successfully demonstrated using a 3D-Var assimilation scheme, NEMOVAR, within the Met Office's global, coupled ocean-sea ice model, Forecast Ocean Assimilation Model (FOAM). The CryoSat-2 Arctic freeboard measurements are produced by the Centre for Polar Observation and Modelling (CPOM) and are converted to SIT within FOAM using modelled snow depth. This is the first time along-track observations of SIT have been used in this way, with other centres assimilating gridded and temporally-averaged observations. The assimilation leads to improvements in the SIT analysis and forecast fields generated by FOAM, particularly in the Canadian Arctic. Arctic-wide observation-minus-background assimilation statistics for 2015-2017 show improvements of 0.75 m mean difference and 0.41 m root-mean-square difference (RMSD) in the freeze-up period, and 0.46 m mean difference and 0.33 m RMSD in the ice break-up period. Validation of the SIT analysis against independent springtime in situ SIT observations from NASA Operation IceBridge (OIB) shows improvement in the SIT analysis of 0.61 m mean difference (0.42 m RMSD) compared to a control without SIT assimilation. Similar improvements are seen in the FOAM 5-day SIT forecast. Validation of the SIT assimilation with independent Beaufort Gyre Exploration Project (BGEP) sea ice draft observations does not show an improvement, since the assimilated CryoSat-2 observations compare similarly to the model without assimilation in this region. Comparison with airborne electromagnetic induction (Air-EM) combined measurements of SIT and snow depth shows poorer results for the assimilation compared to the control, despite covering similar locations to the OIB and BGEP datasets. This may be evidence of uncertainty in the Air-EM validation observations, sampling error, noise in the SIT analysis, or uncertainties in the modelled snow depth or the assimilated SIT observations. The SIT analysis could be improved by upgrading the observation uncertainties used in the assimilation. Despite the lack of CryoSat-2 SIT observations available for assimilation over the summer due to the detrimental effect of meltponds on retrievals, it is shown that the model is able to retain improvements to the SIT field throughout the summer months due to prior, wintertime SIT assimilation. This also results in regional improvements to the July modelled sea ice concentration (SIC) of 5% RMSD in the European sector, due to slower melt of the thicker sea ice.

1 Introduction

In recent decades, Arctic sea ice cover has undergone a considerable reduction in both thickness and extent (e.g. Comiso et al., 2008; Kwok et al., 2009; Lindsay and Schweiger, 2015; Stroeve and Notz, 2018; Meredith et al., 2019), which has the potential to impact weather and climate at lower latitudes (e.g. Koenigk et al., 2016; Screen, 2017), alter the ecosystem and living environment (e.g. Meier et al., 2014), and to change the nature of Arctic shipping by opening up new sea routes (e.g. Smith and Stephenson, 2013; Wei et al., 2020; Zeng et al., 2020). Therefore, predictions of sea ice cover on timescales of days, seasons and beyond are becoming increasingly important.

It has long been understood that the initialisation of numerical weather prediction (NWP) models using data assimilation is vital for producing accurate forecasts over a range of timescales. In the case of sea ice forecasting, assimilation of satellite sea ice concentration (SIC) observations into coupled ocean-sea ice models is well established and routine at almost all centres producing operational sea ice and ocean forecasts (e.g. Bertino and Lisaeter, 2008; Blockley et al., 2014; Smith et al., 2015; Posey et al., 2015; Lemieux et al., 2016). Timeseries of well-homogenised satellite observations of SIC are available dating back to 1979 (e.g. Lavergne et al., 2019) allowing for accurate monitoring of SIC and extent, and use in hindcasts and reanalyses.

Large-scale observations of sea ice thickness (SIT) have become available much more recently than SIC, with the launch of the first dedicated polar satellite mission, ESA's CryoSat-2, in April 2010. CryoSat-2 is able to observe sea ice freeboard (the height of sea ice above the ocean surface) and, based on the assumption that the sea ice is floating in hydrostatic equilibrium, the freeboard observations can be converted to SIT. Although observations of sea ice freeboard have been successfully obtained from previous radar and laser altimetry missions, namely ESA's European Remote Sensing (ERS) and Envisat satellites, and NASA's Ice, Cloud and land Elevation Satellite (ICESat; e.g. Laxon et al., 2003; Kwok and Rothrock, 2009), these have considerably larger unobserved areas over the poles than CryoSat-2. Observations from ICESat were additionally limited by cloud cover and, for the radar missions, instrument footprint sizes were notably larger than for CryoSat-2, with a greater contribution of noise in the observations from radar speckle (Laxon et al., 2013). Consequently, the processing of satellite freeboard observations is still very much an ongoing and active area of research and development, particularly regarding the accuracy of observations, the characterisation of uncertainties and the timeliness of data delivery (e.g. Ricker et al., 2014, 2016; Tilling et al., 2016, 2018). The field is additionally much less mature than the use of altimetry to measure sea level anomaly (SLA), which is readily assimilated into operational ocean forecasting models (e.g. Blockley et al., 2014).

As a complement to radar altimetry, which can be used to retrieve SIT greater than around 1 m thick, observations of thin SIT can be obtained from the Soil Moisture and Ocean Salinity (SMOS) satellite, launched in November 2009. Using a passive microwave radiometer, the thickness of sea ice less than 1 m thick is inferred from L-band brightness temperature measurements (Kaleschke et al., 2010; Tian-Kunze et al., 2014). The only additional estimates of SIT are sparse in situ observations from surface and submarine platforms.

A number of studies have emphasised the importance of accurate initialisation of SIT fields for seasonal predictions of sea ice concentration and extent: Day et al. (2014); Massonnet et al. (2015); Collow et al. (2015); Dirkson et al. (2017); and CryoSat-2

and/or SMOS SIT observations have been used to initialise seasonal sea ice forecasts by Blockley and Peterson (2018); Yang et al. (2019); Allard et al. (2020). Several studies have demonstrated the impact of using satellite SIT observations in addition to SIC to initialise short-term operational sea ice forecasts: e.g. Yang et al. (2014); Mu et al. (2018); Xie et al. (2018); Liang et al. (2020). All of these previous studies have made use of gridded, temporally-averaged satellite SIT datasets (e.g. weekly, monthly) and not the along-track CryoSat-2 data used herein. It was not originally envisioned that using CryoSat-2 observations without a certain degree of spatial and temporal averaging would be possible, due to noise in the freeboard retrievals (Wingham et al., 2006). However, the nature of operational ocean forecasting at the Met Office, and developments towards a coupled ocean-ice-atmosphere NWP framework, with short assimilation time windows of the order of 6 hours, means that along-track rather than temporally-averaged observations are required. Additionally, the use of observations without gridding or temporal averaging aims to improve analysis quality by reducing spatially-correlated uncertainties in the measurements, and allowing accurate uncertainty estimates, vital for data assimilation, to be more easily determined than would be possible for data with more processing applied.

Therefore, in this study, we investigate the feasibility of assimilating Arctic SIT observations derived from along-track CryoSat-2 radar altimeter sea ice freeboard measurements into a global, coupled ocean-sea ice model. We demonstrate that the assimilation system (including prior observation quality control) successfully reduces the effect of noise in the observations such that initial gridding and temporal averaging are not required. Work is currently underway at the Met Office to assimilate SMOS observations in conjunction with those from CryoSat-2, and this will be reported in a future publication. A validation of the SIT and SIC analyses, and 1- and 5-day forecasts generated using the FOAM system with the assimilation of CryoSat-2 SIT, in addition to all the standard observation types, is presented. Analysis and forecast performance compared to a control without SIT assimilation are assessed. The paper is structured as follows: Sect. 2 provides descriptions of the FOAM system, observations used in this study, and assimilation methods. Results from the SIT assimilation experiment and the control are shown in Sect. 3, and validation using assimilation statistics and independent in situ observations follows in Sects. 4 and 5 respectively. A final discussion and conclusions are presented in Sect. 6.

2 Methods

2.1 The FOAM test system

The Forecast Ocean Assimilation Model (FOAM; Blockley et al., 2014) is the Met Office's global, coupled ocean-sea ice model. It is forced at the surface using output from the Met Office NWP system. The ocean and sea ice components of FOAM are also used in an ocean-sea ice-atmosphere-land coupled short-range forecasting system (Guiavarc'h et al., 2019). Analyses and 5-day forecasts of ocean and sea ice variables are produced operationally from the coupled system, and are disseminated through the Copernicus Marine Environment Monitoring Service (CMEMS; marine.copernicus.eu). FOAM analyses are also used operationally to initialise the Met Office's seasonal forecasting system, GloSea (MacLachlan et al., 2014). Here we focus on the forced ocean and sea ice FOAM system, but the implementation of any developments will also be of benefit to the coupled short-range and seasonal prediction systems.

The ocean model component of FOAM, Nucleus for European Modelling of the Ocean (NEMO; Madec, 2017) is coupled to the Los Alamos Sea Ice Model, CICE (Hunke et al., 2015). The tripolar grid used by the FOAM system has recently been upgraded to 1/12 degree resolution (ORCA12) for both ocean and sea ice components (Barbosa Aguiar et al., 2021), but the version of FOAM used herein has a 1/4 degree grid (ORCA025). The ocean model has 75 vertical levels (Storkey et al., 2018) and the CICE configuration includes 5 thickness categories (plus open water), multi-layer thermodynamics and prognostic melt ponds (Ridley et al., 2018). The data assimilation scheme, NEMOVAR (Waters et al., 2015), is used in FOAM in a 3D-Var First Guess at Appropriate Time (FGAT) configuration to assimilate observations of ocean and sea ice variables.

The observation data types assimilated in the FOAM test system used in this study are the same as for the operational system, namely temperature, salinity, sea level anomaly (SLA) and SIC. No SIT observations are currently assimilated operationally in FOAM. For this study, sea surface temperature (SST) observations are obtained from ships, moored and drifting buoys, AVHRR sensors onboard NOAA and MetOp satellites, and the VIIRS sensor onboard the Suomi-NPP satellite. Argo floats, moored buoys, gliders, and research Conductivity, Temperature and Depth (CTD) instruments provide temperature and salinity profiles. Temperature profiles are also obtained from instrumented marine mammals (Carse et al., 2015) and Expendable Bathythermographs (XBTs). SLA data are provided by altimetry from the Jason-2, Jason-3, Sentinel-3A, CryoSat-2 and AltiKa satellites. SIC measurements from the Special Sensor Microwave Imager/Sounder (SSMIS) instruments onboard the DMSP series of satellites are also assimilated. A control FOAM system using these observations and an experiment system with the additional assimilation of SIT observations derived from CryoSat-2 measurements are employed for this study. Both the SIT assimilation experiment and control systems use a 24-hour assimilation time window. The systems are forced at the surface using hourly wind fields and three-hourly temperature, humidity, precipitation and radiative fluxes from the Met Office NWP system, at a horizontal resolution of 17 km.

The SIT assimilation experiment and control systems have been used to generate daily analysis and 1-day forecasts of ocean and sea ice variables for the three-year period from January 2015 to December 2017. This follows a three-month spin-up period from October 2014, initialised using a previous FOAM reanalysis. The analysis fields were used to initialise 5-day forecasts for each day in selected periods: March - April, June - July, and September - November, for each of the three years in the study. These months were chosen to cover the Arctic ice break-up and freeze-up periods, as well as including March and September to coincide with the annual Arctic maximum and minimum sea ice extents respectively. It should be noted that CryoSat-2 SIT observations are only available between October and April each year, owing to the detrimental impact of summertime melting on the satellite retrievals (Tilling et al., 2016). Therefore June - July was also selected to assess the behaviour of the sea ice forecasts over the summer months when SIT observations are not available for assimilation.

2.2 Observations used for SIT assimilation

2.2.1 Conversion from freeboard to thickness

The satellite data used in this study are along-track, CryoSat-2 radar altimeter observations of Arctic sea ice freeboard, processed by the Centre for Polar Observation and Modelling (CPOM). The data have an along-track resolution of 300 m and

have undergone extensive independent validation (Tilling et al., 2015, 2018). This dataset was selected in part as it is available in near-real-time (Tilling et al., 2016), and this timeliness is necessary for future implementation into the operational FOAM
125 system.

It would be possible to assimilate freeboard observations directly into the model, rather than converting to SIT first. However, the assimilation of further SIT datasets is planned, including those with direct observations of SIT (e.g. from SMOS). For simplicity, a single additional state vector for the assimilation, to cover all ice thickness data types, was selected. SIT was chosen since it is a model prognostic, whereas freeboard is a diagnostic. Freeboard observations (f_i) are converted to SIT (h_i)
130 in FOAM as part of the observation operator (the model run used to compute the model equivalent of the observations), using Eq. 1. Following Tilling et al. (2016) and assuming the ice is floating in hydrostatic equilibrium:

$$h_i = \frac{f_i \rho_w + h_s \rho_s}{(\rho_w - \rho_i)} \tag{1}$$

135 where h_s is the FOAM modelled snow depth (in m) at the observation time and interpolated to the freeboard observation location. ρ_w, ρ_s and ρ_i are the densities of water, snow and ice respectively. These are assumed constant, and are set to 1026.0, 330.0 and 917.0 kg m^{-3} respectively, as used in the CICE sea ice model component of FOAM.

Prior to the conversion to SIT, the modelled snow depth is also used to provide a correction to the freeboard observation (the "radar" freeboard; f_{i_radar}) to obtain the "true" or "corrected" freeboard (f_i). This accounts for the reduction in speed of
140 the altimeter radar pulse due to the presence of snow on the sea ice (Tilling et al., 2016). The correction is an addition of the quantity $0.25h_s$ to the radar freeboard observations, based on:

$$f_i = f_{i\_radar} + [(C_o/C_s) - 1.0]h_s = f_{i\_radar} + 0.25h_s \tag{2}$$

145 where h_s is the snow depth as in Eq. 1, $C_o = 3.0 \times 10^8 \ m \ s^{-1}$ is the speed of light in a vacuum, and $C_s = 2.4 \times 10^8 \ m \ s^{-1}$ is the speed of light in snow.

Currently, CPOM makes use of a modified snow depth climatology, based on Warren et al. (1999) and halved over first-year ice, for processing CryoSat-2 sea ice freeboard retrievals and conversion to SIT (Tilling et al., 2015). This approach is also used by other centres processing CryoSat-2 freeboard observations: Alfred Wegener Institute (AWI; Ricker et al., 2014) and
150 NASA (Kwok and Cunningham, 2015). Instead, here the FOAM modelled snow depth is used. Modelled snow depth has a greater spatial and temporal variability than can be obtained from a climatology, as demonstrated by Mallett et al. (2021) and illustrated on Fig. 1. Using this method also maintains consistency between SIT and snow depth within the FOAM model. A preliminary validation indicates that the FOAM snow depth is somewhat thinner than the modified climatology of Warren et al.
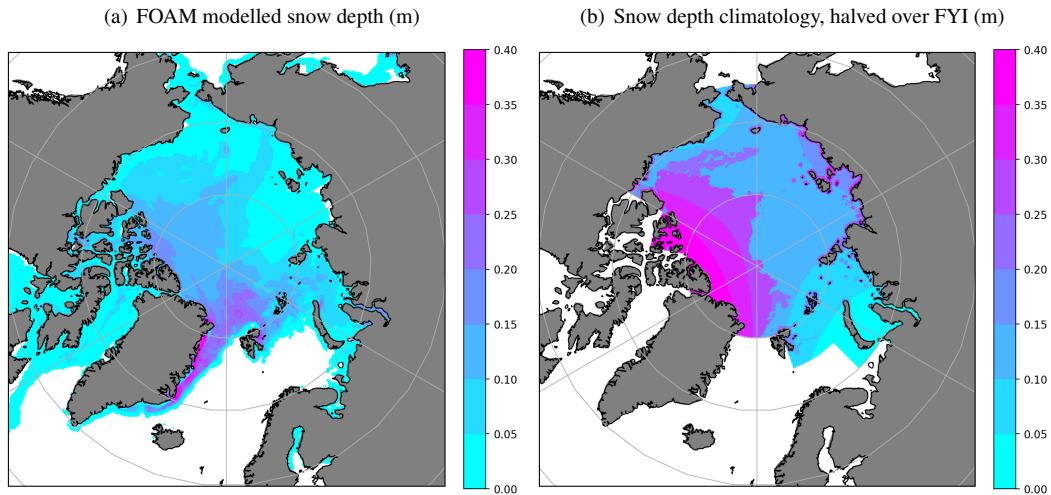
Figure 1. Comparison of snow depth from FOAM model and Warren et al. (1999) climatalogy. (a) FOAM daily mean modelled Arctic snow depth (m) for example date 15 January 2015, and (b) Warren et al. (1999) snow depth climatology (m) for January, halved over first year ice (FYI) using EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF) sea ice type observations from 15 January 2015 (Aaboe et al., 2020).

(1999), as shown on Fig. 1, particularly over multi-year ice. Tuning experiments demonstrate that simply increasing the snow
155   depth in the model does not result in better evaluation of the SIT analysis against independent observations, owing to feedbacks in the model and between the SIT assimilation and the snow depth itself.

Snow depth uncertainty is a large source of error in radar altimetry sea ice measurements, both in the retrievals of freeboard and the subsequent conversion to SIT (e.g. Giles et al., 2007; Ricker et al., 2015). Due to the linear relationship between SIT and snow depth (Eqns. 1 and 2), an underestimation of the snow depth would lead to an underestimate in the SIT. Large
160   uncertainties in the snow depth may apply whether it has been modelled or taken from climatology. Additional uncertainty is also introduced in Eqn. 1 through lack of knowledge of the snow and sea ice densities which, although constants in the CICE model used here, are spatially and temporally varying in reality (e.g. Alexandrov et al., 2010; Kern et al., 2015). Uncertainties due to variations in water density can be neglected (Ricker et al., 2014; Kurtz et al., 2014). In order to quantify and reduce the uncertainty in the FOAM modelled snow depth, future plans will include the assimilation of satellite snow depth observations.

2.2.2   Quality control and pre-processing

165   Since retrievals of sea ice freeboard from CryoSat-2 are noisy (e.g. Laxon et al., 2013), it is important to apply quality control and pre-processing prior to assimilating the data. As an initial check, following Tilling et al. (2018), freeboard observations below -0.3 m and above 3 m are rejected. Any negative values remaining after the conversion to SIT (see Sect. 2.2.1) are also removed. A Bayesian background check for removal of poor-quality observations (Ingleby and Huddleston, 2005), as currently

6

170 used for ocean data types in FOAM, was investigated for SIT. In this method, any observations that deviate too far from the model background field are rejected, taking into account the model background and observation error variances. However, since the SIT observations are so different from the modelled SIT prior to any SIT assimilation (as will be shown in Sect. 3), it is difficult to avoid rejecting large numbers of observations when beginning the assimilation. This leads to issues in the analysis quality and excessively long spin-up periods for the assimilation. Use of a background check for subsequent months

175 once the assimilation is established could be investigated further, noting that model drift in the absence of SIT observations over the summer months may once again lead to the rejection of good quality observations come autumn. Nevertheless, it will be demonstrated in this study that acceptable results can be obtained without including this check.

After the initial gross quality control checks, "super-observations" of the median freeboard observation within a specified radius are created. Use of median averaging rather than mean prevents outliers from influencing the result. Super-obbing is often

180 used in data assimilation to subsample satellite observations to the model grid size. It is also an established method of reducing the correlated uncertainty in the observations (e.g. Janjic et al., 2017), which is assumed to be negligible in the assimilation system used here (Sect. 2.4.1). The super-obbing radius was selected to be 10 km, as this is similar to the model grid size in the Arctic on the 1/4 degree tripolar (ORCA025) grid used here by NEMO and CICE. The radius could be increased to include more observations per super-ob, but 10 km allows grid-scale variability to be preserved in the data. For a representative

185 example date of 1 March 2015, the mean number of freeboard observations used to create each super-ob is 18.4 (minimum 2, maximum 101), with the greatest number of observations available at high latitudes where the orbit tracks converge. The mean difference of the observation-minus-background decreased from 0.12 m to 0.08 m pre- and post-superobbing, and the root-mean-square difference (RMSD) from 1.05 m to 0.68 m. This indicates a substantial reduction of noise in the data.

2.3 Observations used for SIT validation

190 Independent in situ observations suitable for validation of the SIT assimilation experiment, and which cover the experiment time period, are available from the field campaigns NASA Operation IceBridge (OIB; Kurtz et al., 2013), from airborne electromagnetic induction (Air-EM) observations as part of the Pan-Arctic Measurements and Arctic Regional Climate Model Simulations Project (PAM-ARCMIP; Haas et al., 2009), and from Beaufort Gyre Exploration Project (BGEP; Krishfield et al., 2014) moorings that measure sea ice draft using upward-looking sonar. The following sections provide details of these datasets.

195 ### 2.3.1 Operation IceBridge (OIB)

The OIB field campaign used an aircraft equipped with a scanning lidar altimeter, a snow radar and cameras to obtain springtime measurements of Arctic sea ice freeboard, thickness and snow depth, for several weeks each year between 2009 and 2018. The accuracy of the SIT measurements is estimated to be 0.4 m (Farrell et al., 2012). Data for 2015 to 2017 overlap with the dates of the SIT assimilation experiment carried out in this study, and are available for the dates shown in Table 1.

200 The dataset used here is from the QuickLook V1 product (Kurtz et al., 2019), as the more reliable V2 product was not available for the SIT assimilation experiment period at the time of assessment. The dataset was processed by the OIB project and comprises 50 km cluster averages of SIT point measurements, at a spacing of approximately 25 m. Data from multiple

	March	April
2015	20, 22, 24-31	3-4
2016	-	20, 28
2017	8-14, 17, 19-20, 24	4

Table 1. Dates of Operation IceBridge data used in this study.

flights over the same locations were combined into the same cluster, for flights fewer than 10 days apart. Only point observations

with low uncertainties were selected for the clusters: less than 1 m uncertainty for very thin ice, up to a maximum uncertainty of

205 2 m for ice thicker than 4 m. Clusters were required to have at least 500 point samples, and they have an average of 1670 points

with a maximum of 7000 points. As part of this study, further processing was carried out to remove any cluster observations

with associated standard deviations above 3 m. It is necessary to use spatially-averaged rather than point observations for

validation, due to the very high level of noise in the point measurements and for more appropriate comparison to a model

designed to simulate average sea ice behaviour without sub-gridscale heterogeneity.

210 Figures 2(a) and (b) show the 2015-2017 OIB cluster average SIT observations and the standard deviations of the clusters.

The figure illustrates the large amount of valuable in situ data available from this project.

2.3.2 Airborne electromagnetic combined snow and ice thickness (Air-EM) observations

Field campaigns for PAM-ARCMIP were conducted over the Arctic for several weeks each year between 2001 and 2015. An

aircraft equipped with an instrument known as an "EM-Bird" was used to measure combined SIT and snow depth by airborne

215 electromagnetic (EM) induction (Haas et al., 2009). Over level ice, the accuracy of Air-EM measurements of combined SIT

and snow depth is 0.1 m (Pfaffling et al., 2007; Haas et al., 2009). Only data for the 2015 field campaign overlap with the dates

of the SIT assimilation experiment in this study, for the following dates: 7-9, 11, 22-23 April.

Similar to the OIB data (Sect. 2.3.1), 50 km cluster averages have been produced by the Air-EM data providers. The clusters

include contributions from multiple flights if these spanned a few days over a small region. Further processing was carried

220 out as part of this study to remove any cluster observations with standard deviations over 3 m. Figures 2(c) and (d) show the

Air-EM cluster average observations and the standard deviations of the clusters for the 2015 field campaign.

2.3.3 Beaufort Gyre Exploration Project (BGEP)

The BGEP dataset consists of observations of sea ice draft (thickness of ice below the ocean surface) from upward-looking

sonar (ULS) instruments attached to bottom-anchored moorings in the Beaufort Sea. The locations of the moorings providing

225 data used in this study are shown on Fig. 2(e). Data are available continuously between September 2003 and August 2017,

with an estimated accuracy of 0.05 - 0.10 m (Krishfield and Proshutinsky, 2006). Data used in this study are monthly means

of sea ice draft, processed by the data providers from the raw 2-second observations, and which do not include any open water

estimates. In the same way as for the other validation datasets, any averaged sea ice draft observations with standard deviations
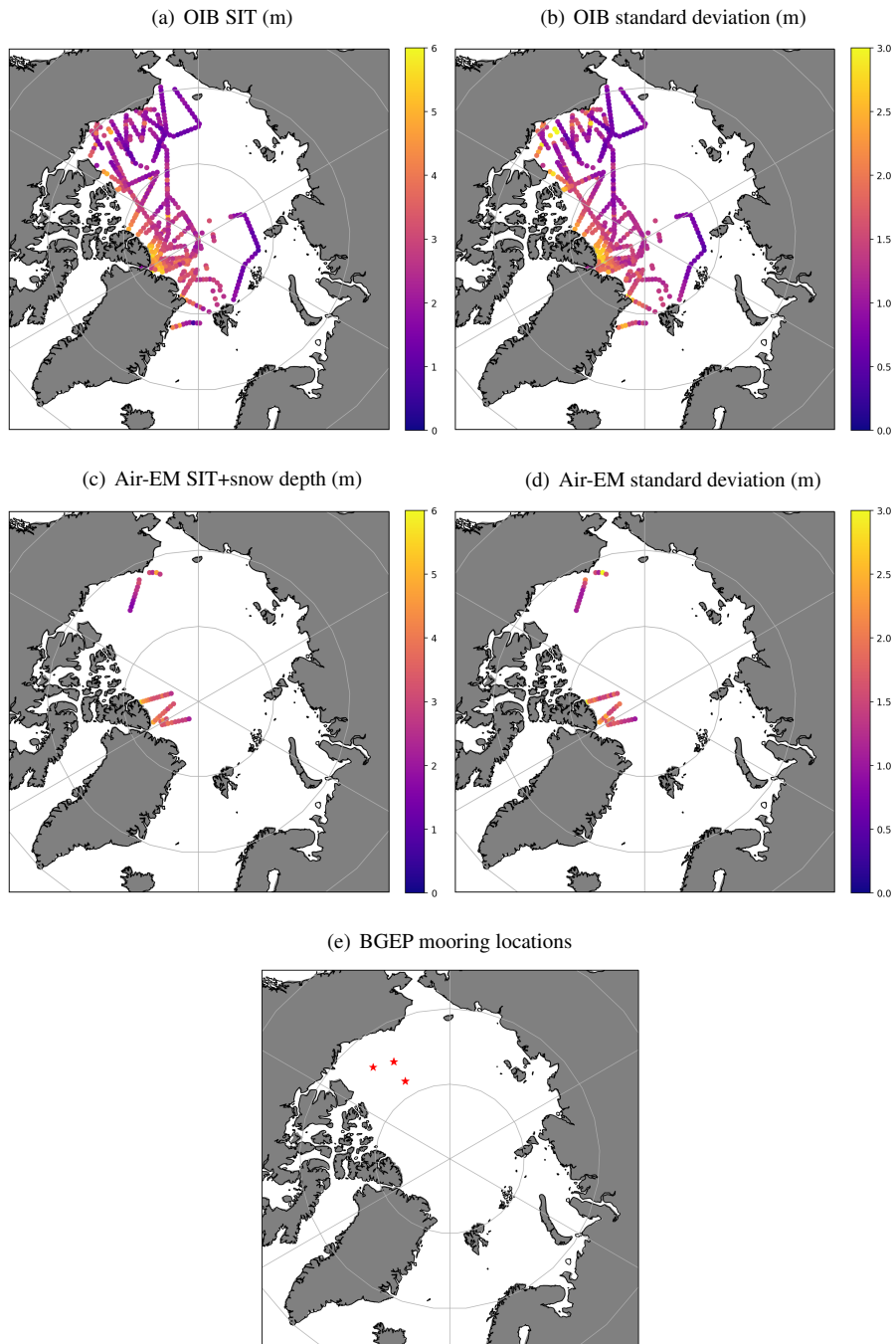
Figure 2. Observations (in m) and standard deviations (in m) of validation data used in this study. (a) OIB cluster average SIT observations and (b) standard deviations of OIB clusters, for 2015 - 2017 field campaigns; (c) Air-EM cluster average observations of combined SIT and snow depth, and (d) standard deviations of Air-EM clusters, for 2015 field campaign; (e) locations of BGEP moorings used in this study.

above 3 m were excluded from this study. Further processing was also undertaken here to convert the measurements of sea ice draft to SIT by dividing observations by 0.89, following Rothrock et al. (2003). However, it should be noted that this method does not take into account the presence of snow on the surface of the ice, which will likely affect the derived SIT observations outside of the summer months.

2.4 Assimilation of SIT observations

The NEMOVAR assimilation system is used to calculate increments of SIT to be applied to the model, in the same way as is done for the other assimilated data types in FOAM. The inputs to NEMOVAR are the SIT observations and, for each observation, the model SIT value aggregated over all thickness categories and interpolated to the location of the observation at the closest model time-step to the observation time, during a one-day model forecast. The uncertainties associated with the observations and model forecast are also provided to NEMOVAR, in the form of error covariances, as described in the following subsections. NEMOVAR outputs the changes required to bring the model SIT into line with observations from that day, taking into account their respective uncertainties, as a field of increments on the model grid. These SIT increments are applied to the CICE model using an Incremental Analysis Update (IAU) method (Bloom et al., 1996), as is used for the other variables in FOAM. In this method, a fraction of the increments is added at each time-step during a 24-hour period, such that the total increment is applied by the end of the model day. Following Blockley and Peterson (2018), SIT increments are added to each of the five sub-grid SIT categories, if the ice concentration within that category is above 1%. The initial fraction of the contribution of each category to the gridbox mean ice volume is calculated, and that fraction of the SIT increment is then added to that category. This maintains the initial volume distribution of ice (and ice area) across each sub-grid SIT category. Changes resulting from SIT increments are only made where the grid cell aggregate SIC is greater than a conservatively-chosen value of 40%. This means that only SIC increments are able to add new ice, since this data is deemed the more reliable, particularly for the generally thinner ice of the marginal ice zone. The SIT increments are applied after any SIC assimilation changes at each time-step and, similar to SIC, no balancing is performed with the other variables.

2.4.1 Observation error covariance

The assimilation system requires estimates of observation uncertainties for SIT. The higher the magnitude of the observation uncertainty, the smaller the impact the observation will have on the analysis. Observation uncertainties are provided as observation error variances (OBE), with the assumption that the observation error is uncorrelated. This is not necessarily true, but is a standard simplification (e.g. Stonebridge et al., 2018). The OBE is a combination of measurement uncertainty and representation uncertainty. Measurement uncertainty includes the raw measurement error and uncertainties due to the retrieval algorithm, including effects such as surface roughness (Landy et al., 2020), ice salinity (Nandan et al., 2017) and variability in the densities of snow and ice (Ricker et al., 2014). Representation uncertainty results from unresolved scales and processes in the model, observation operator uncertainty and quality control uncertainty (e.g. Janjic et al., 2017). In this study, the representation uncertainty is set to 0.05 m standard deviation as an initial estimate, being of a similar magnitude to the minimum measurement uncertainty (see below). Further work is required to refine this specification. Since estimates of measurement

uncertainty are not provided with the CPOM data used in this study, this has instead been determined using a simple function. This is based on Fig. 2 of Ricker et al. (2017), who derived relationships between the magnitude of CryoSat-2 ice thicknesses and their associated measurement uncertainties, in terms of % error. Further details of these methods are given in Ricker et al.
265 (2014). The SIT and uncertainty relationships determined by Ricker et al. (2017) have slightly different characteristics at different times of the year, but using a single, general function in this study as a first parameterisation of the measurement uncertainty is a reasonable choice. The function used here is shown on Fig. 3(a), and is specified in Eq. 3 as follows, in m for thickness h_i and standard deviation σ, with a cap of 8 m, and using the mathematical constant, e:

$$\sigma = \begin{cases} 8 & \text{for } h_i < 0.7 \text{ m} \\ [7e(\frac{-1}{0.3-h_i^e}) + 1] * \frac{h_i}{100} & \text{for } 0.7 \leq h_i < 3.0 \text{ m} \\ \{[(h_i - 3.0) * 5] + [7e(\frac{-1}{0.3-3.0^e}) + 1]\} * \frac{h_i}{100} & \text{for } h_i \geq 3.0 \text{ m} \end{cases} \quad (3)$$

270 The uncertainty standard deviation associated with each SIT observation is calculated using the function given in Eq. 3. 8 m is an arbitrarily-selected value of the maximum uncertainty for thin ice observations, being a very large value compared to the SIT observations themselves. This ensures that these data do not influence the analysis. 0.7 m is where the function reaches this value. For SIT between ∼1.5-3 m the uncertainty is at its minimum, meaning these observations are given stronger weighting in the analysis, and uncertainty begins to increase again for observations thicker than 3 m.

275 Sensitivity tests were conducted to produce the optimum SIT analysis by finding an appropriate balance between the OBE and the model background error variance (BGE; Sect. 2.4.2). This allowed the final form of the OBE function to be tuned and the minimum uncertainty assigned to the most reliable observations to be set. The minimum OBE is rather small compared to previous studies (e.g. Tilling et al. (2018) gives the accuracy of CryoSat-2 SIT as 13 cm) and this likely due to an underestimate in the model BGE (Sect. 2.4.2).

280 Figs. 3(b) and (c) respectively show the monthly mean observed CryoSat-2 SIT for March 2015, and the measurement uncertainty standard deviation for the data, derived from the parameterisation shown in Fig. 3(a). Data for both Figs. 3(b) and (c) have been binned onto a 1/4 degree regular grid. The figures show the largest uncertainties occuring in regions of very thick ice, as there are limited observations of thin ice in the dataset following the gross error check (Sect. 2.2.2).

There is a high level of random uncertainty in the CryoSat-2 along-track freeboard observations, owing to speckle noise in
285 the range measurements, and inaccuracies in the observations of sea surface height (Ricker et al., 2014). The OBE estimate does not take this into account. These random uncertainties can be greater than 1 m, and are usually removed by spatial averaging (e.g. Ricker et al., 2017). The super-obbing process described in Sect. 2.2.2 will mitigate this issue somewhat, but the uncertainty in the observations will be higher around the ice edge since there are fewer overlapping orbits at these latitudes. This results in as few as 2 observations included in a super-observation, compared with up to 100 or so at higher latitudes.
290 Work is currently underway to also assimilate SMOS SIT data to improve the representation of thin ice in FOAM, and this will improve the SIT analysis in the regions most affected by this issue. Other possible improvements include filtering out the affected observations from the analysis by increasing the OBE in the CryoSat-2 data below a specified latitude or below a

minimum number of observations used in the super-observation. Additionally, a check of the observation quality against the model SIT background could be used, although, as discussed in Sect. 2.2.2, this method relies on the model being unbiased.

2.4.2 Background error covariance

Uncertainties in the 1-day model forecast prior to the assimilation of observations (the "background") are parameterised for the SIT assimilation as spatially and seasonally-varying fields of background error variance (BGE) with associated spatial error correlation length scales. The seasonal BGEs are interpolated to the model date within the FOAM system. This method is also used for SIC assimilation in FOAM (Blockley et al., 2014). For SIT, the BGE for each season and the correlation length scale were estimated using the "Canadian Quick" covariance method (Polavarapu et al., 2005) with 3 years of FOAM hindcast SIT data (1 June 2015 to 31 May 2018). In this method, differences between daily model fields are used as a proxy for the model forecast error. The spatial correlations in the results were assessed and yielded an estimate of 50 km for the minimum SIT correlation length scale. This value was used as a constant length scale everywhere, except at very high northern latitudes where it was extended to compensate for the data gap north of 88.0 N resulting from the orbital inclination of the CryoSat-2 satellite (the "pole hole"). Here, the length scale was increased to 100 km for observations at latitudes north of 87.5 N, making the assumption that the model uncertainties near the pole are highly correlated with the uncertainties at 87.5 N. This has the effect of filling the pole hole using increments spread from the surrounding area, and allows the SIT analysis to vary in this region. A sensitivity test was used to select the magnitude of the extended length scale and the latitude threshold. The values chosen appear to give a good result (Fig. 4): the pole hole is filled, but with minimal impact on the increments nearby. A quantitative validation is not possible, owing to the absence of satellite or suitable in situ SIT observations at this location.

Figure 4 also demonstrates that the 50 km length scale allows information from the observations to propagate spatially over the domain, which is helpful owing to the sparseness of the daily CryoSat-2 observations in regions where the orbit tracks do not overlap (Fig. 4(d)). A dual length scale correlation formulation is available in NEMOVAR (Mirouze et al., 2016; Fiedler et al., 2019) and use of this for SIT remains an avenue for future investigation.

SIT modelled by the FOAM system is too thin without SIT assimilation (as will be shown in Sect. 3), so the model BGE calculated using the Canadian Quick method is likely to be an underestimate. However, it provides a starting point for iterations of BGE calculations using different methods, once the SIT assimilation is in place. Examples of such methods are given in Bannister (2008) and include those based on innovation (observation-minus-background) correlations (Hollingsworth and Lönnberg, 1986), differences between forecasts of varying lengths (the "NMC" method; Parrish and Derber, 1992), or ensemble data assimilation methods (Houtekamer et al., 1996).

3 SIT assimilation results

SIT derived from along-track CryoSat-2 freeboard observations can be successfully assimilated into the FOAM forecasting system. Figures 5(a) and (b) show daily mean SIT analysis fields for the SIT assimilation experiment and the control respectively, for 15 January 2015 as an example date. Gridded CPOM CryoSat-2 SIT observations for January 2015 are also included
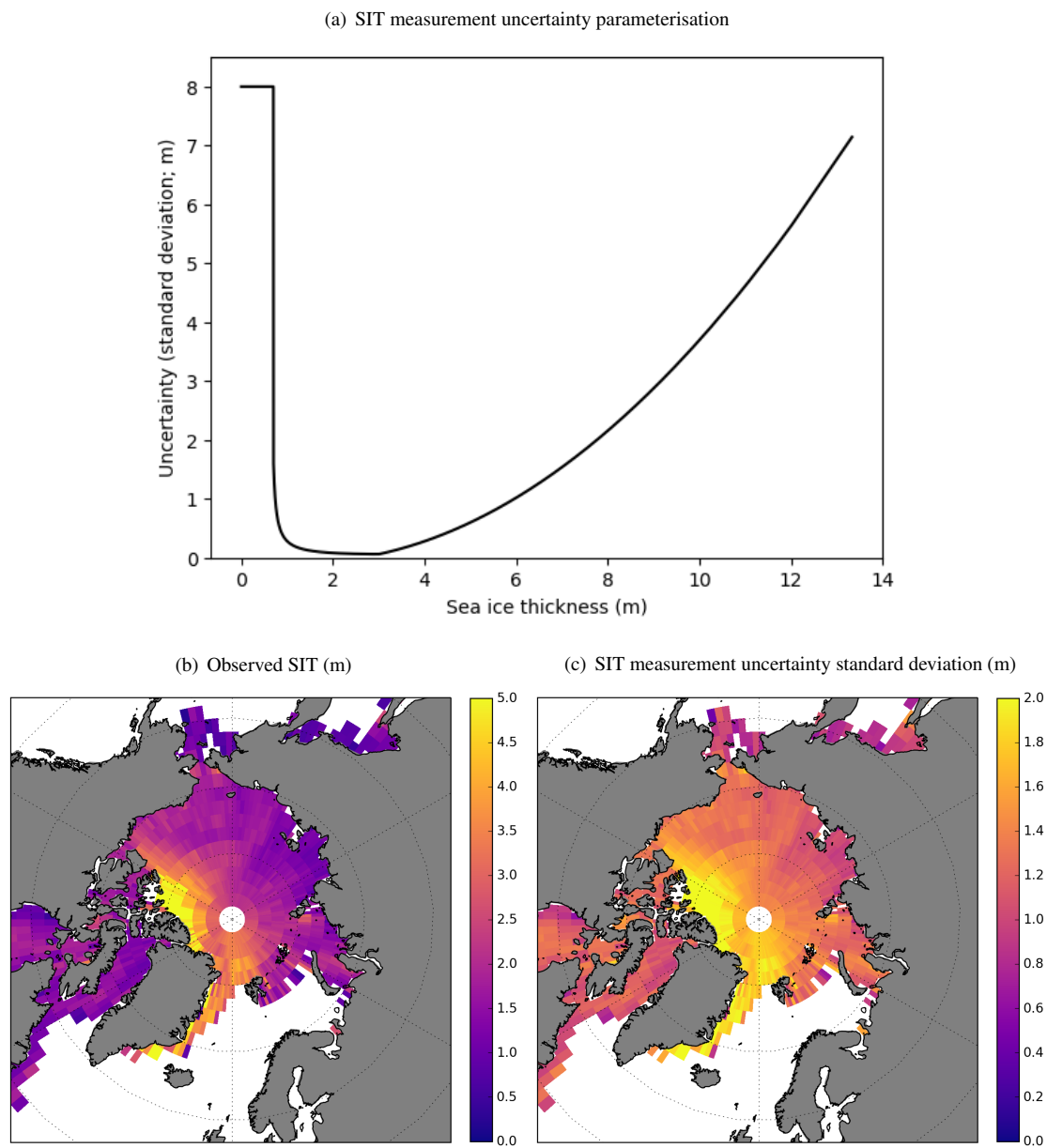
(a) SIT measurement uncertainty parameterisation

(b) Observed SIT (m)

(c) SIT measurement uncertainty standard deviation (m)

Figure 3. Observation measurement uncertainty estimates for SIT. (a) Parameterisation for SIT measurement uncertainty in terms of standard deviation (m), as a function of SIT (m). March 2015 mean binned (b) observed CryoSat-2 SIT (m) and (c) observation measurement uncertainty standard deviation (m), calculated using the parameterisation shown in (a), for the observations shown in (b). Observations have undergone initial quality control and pre-processing.
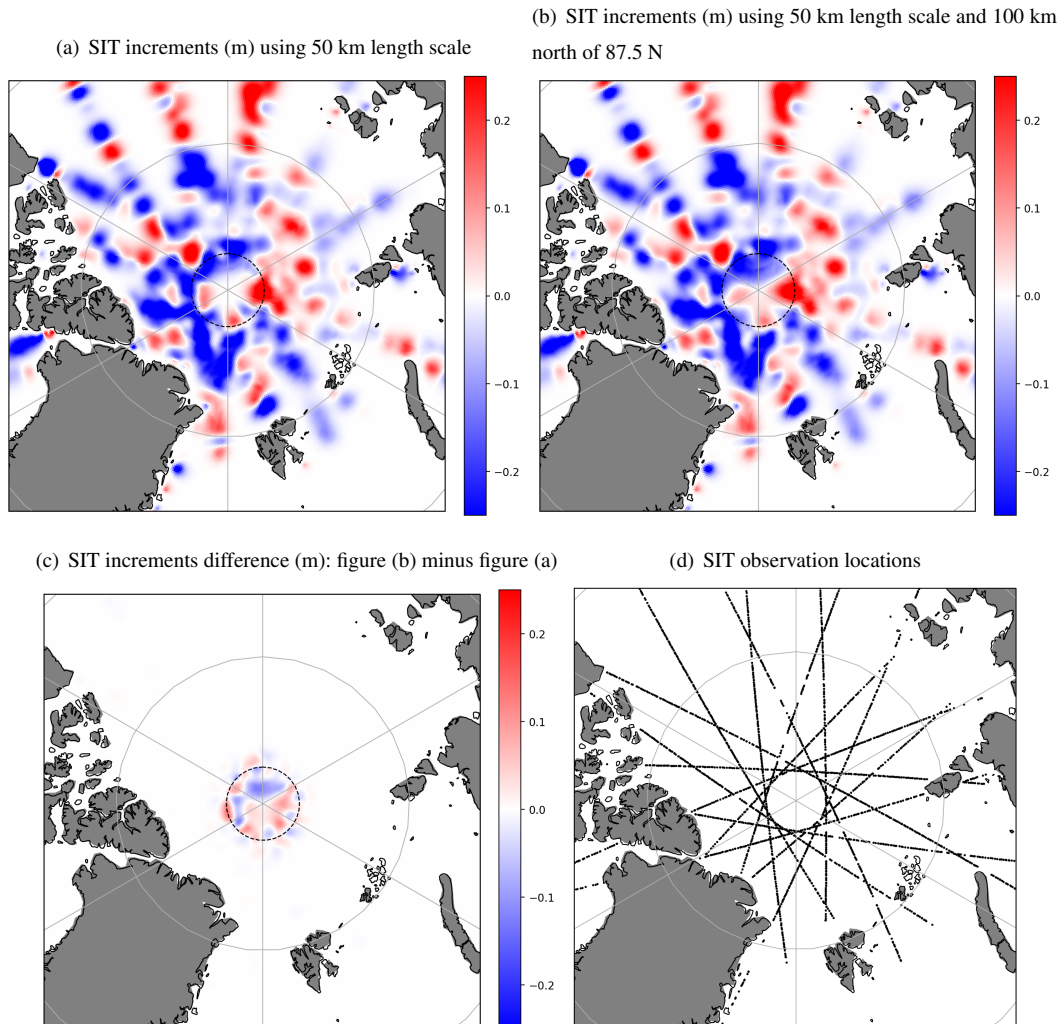
Figure 4. Demonstrating filling of the satellite observation data gap at the north pole using SIT information from the surrounding area. 87.5 N gridline shown as a dashed black line. SIT increments (m) produced using different background error correlation length scales, example date 30 January 2015 for (a) 50 km length scale throughout, (b) 50 km length scale, with 100 km north of 87.5 N, and (c) difference between (a) and (b). (d) Locations of assimilated SIT observations for the domain shown, 30 January 2015.

325   for comparison (Fig. 5(c)). The results demonstrate that the FOAM system is able to reproduce the monthly, gridded obser-
vation field by assimilating daily along-track data. The control SIT model field is smoother than the observations, and much
thinner in the Canadian Arctic region. The SIT assimilation experiment has produced a thicker sea ice field, particularly in the
Canadian Arctic and around Greenland, and introduced more spatial variability throughout.

Figure 5(a) demonstrates that the background error correlation length scale for the SIT assimilation is suitable, since the satellite tracks of the day's observations are not obvious in the analysis field (length scale too short), and there is no excessive smoothing (length scale too long). Despite the absence of observations, the pole hole has a realistic-looking SIT, which is thicker than in the control (Fig. 5(b)). This demonstrates that the method of using SIT information spread from the surrounding area, as described in Sect. 2.4.2, is successful.

Figure 6 shows the difference between monthly averaged SIT fields for the SIT assimilation experiment minus the control, for the months of the Arctic sea ice maximum and minimum extents, March and September respectively. 2016 is shown as an example year. The main impact of the SIT assimilation in March is to increase the SIT in the Canadian Arctic and European sectors (Fig. 6(a)). Some thinning of the sea ice is also seen outside of these regions, which for March 2016 occurs around the East Siberian, Chukchi and Laptev Seas. Although there are no SIT observations available for assimilation between May and September inclusive each year, differences between the SIT assimilation experiment and control are still apparent in the September SIT field (Fig. 6(b)). This demonstrates that the effect of the SIT assimilation on the model in the months prior to May persists throughout the summertime to September. This effect was also shown in the experiments of Blockley and Peterson (2018) and Allard et al. (2018), and illustrates that the SIT assimilation is having the desired impact on the FOAM model. The SIT assimilation has a minimal impact on the SIC model field in March and September (not shown), with small changes in concentration seen mostly around the ice edge at times of maximum ice extent and within the pack when the minimum extent occurs.

Differences between SIT analyses and forecasts generated by the SIT assimilation experiment and the control are assessed in detail in the following sections, using validation statistics from the assimilation system and comparisons with independent in situ observations.

4 Validation using assimilation statistics

The performance of FOAM SIT and SIC 1-day model forecasts (also referred to as the backgrounds onto which the assimilation increments are added) has been assessed using CryoSat-2 SIT and SSMIS SIC (Tonboe et al., 2017) observations, prior to them being assimilated. The performance of the FOAM 5-day forecasts of SIT and SIC has also been assessed using the same observations, during the months for which these longer forecasts were produced. Matchups between the forecasts and observations were obtained by interpolating the daily-mean model forecast fields to the observation locations. Although this assessment does not use independent data, it aims to demonstrate that the assimilation is working as expected, and that the model is brought into line with high-quality observations, which have been independently validated.

Figure 7 shows timeseries of daily SIT observation-minus-background mean difference and root-mean-square difference (RMSD) for the SIT assimilation experiment and the control, for 1 January 2015 to 31 December 2017. Note that there are no CryoSat-2 SIT observations, and hence statistics, available between May and September of each year. The improvement in both the mean difference and RMSD of the 1-day forecast on assimilation of the CryoSat-2 observations is clear. There are substantial reductions for the SIT assimilation experiment compared to the control of 0.46 m mean difference (0.33 m RMSD)
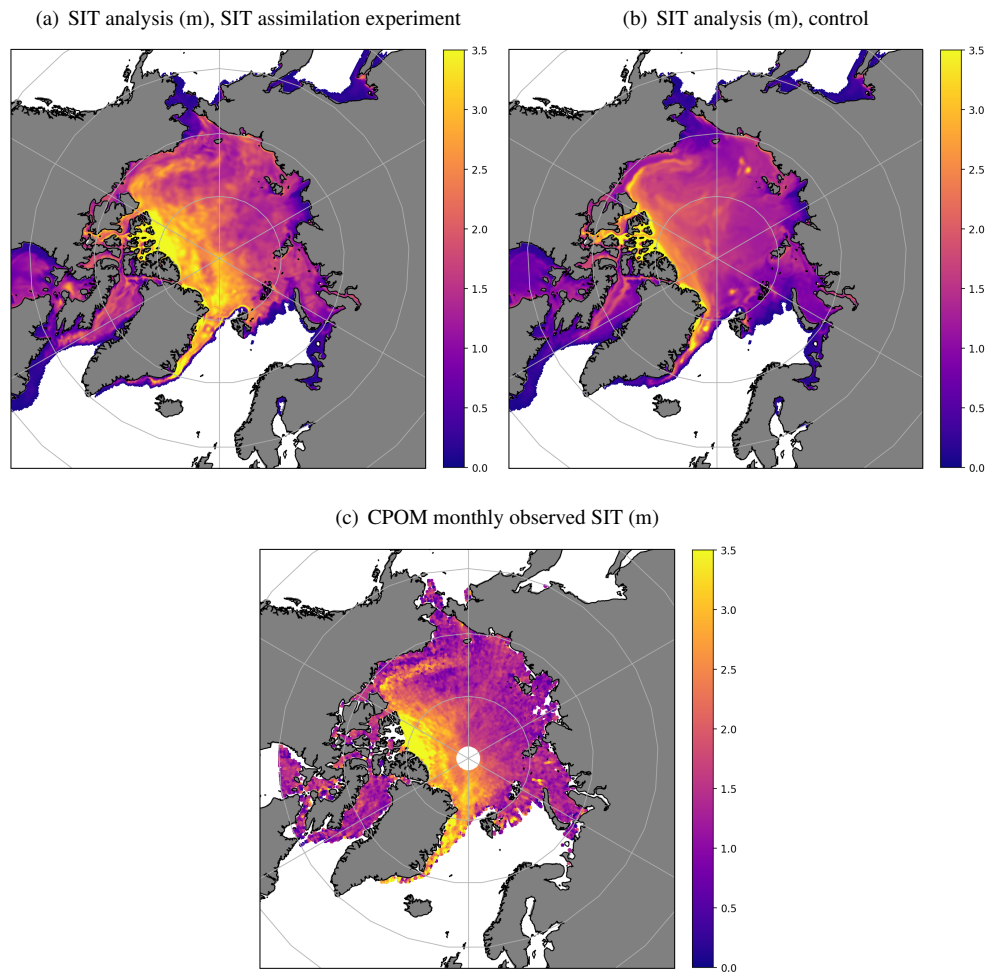
15

(a) SIT analysis (m), SIT assimilation experiment

(b) SIT analysis (m), control

(c) CPOM monthly observed SIT (m)

Figure 5. Modelled and observed SIT (in m). Daily mean FOAM SIT analysis for (a) SIT assimilation experiment and (b) control, 15 January 2015, and (c) 5 km gridded CPOM CryoSat-2 SIT observations, January 2015.

for March-April (ice break-up), and 0.75 m mean difference (0.41 m RMSD) for October-November (ice freeze-up), averaged over the Arctic for 2015 - 2017. Figure 7 also illustrates that the benefit to the model from the SIT assimilation continues throughout the summer months since, despite quite an increase, the differences to the observations still remain lower than in the control when the statistics become available again in October. Additionally, there is a very quick reduction in the model differences to the SIT observations once the assimilation restarts.

Figures 8 and 9 show the same statistics as Fig. 7, but as spatially-binned plots of mean difference and RMSD. March-April and October-November are shown, to examine the regional variation in results for the ice break-up and freeze-up periods respectively, for 2015 as an example year. The reduction (improvement) in the mean difference and RMSD due to SIT assimilation compared to the control is seen over large areas of the Arctic, both at the start of the melt period (Fig. 8, though some differences still remain in the regions of thickest ice, in the Canadian Arctic and along the eastern coast of Greenland), and during the freeze-up (Fig. 9). This effect is also seen in 2016 and 2017 (not shown). The largest regional improvements are in March-April (Fig. 8), with reductions in the mean difference of more than 1.30 m and 1.22 m in the RMSD in the European sector and parts of the Canadian Arctic.

Figure 10 shows spatially-binned plots of SIC observation-minus-background mean difference and RMSD, for July 2015. The difference between the SIC results for the SIT assimilation experiment and control are negligible for the other months. Figure 10 illustrates that both the mean difference and RMSD for SIC are reduced (improved) in the Atlantic sector, specifically in the area immediately north of Svalbard and Franz Josef Land, by around 0.05 sea ice fraction for the SIT assimilation experiment compared to the control. This occurs despite the assimilation of SIT ceasing at the end of April, when CryoSat-2 data production is suspended over the summer months. This effect is likely to be due to the wintertime assimilation introducing thicker ice, which melts more slowly over the summer than in the control, leading to improvements in the SIC field. This process ceases to become important by early August and does not appear to affect SIC in the freeze-up period (not shown). The effect is also seen, though to a lesser extent, in 2016 and 2017 (not shown).

Figure 11 shows validation statistics for FOAM 1- to 5-day forecasts of SIT and SIC compared to observations, for the SIT assimilation experiment and the control. The break-up (March-April) and freeze-up (October-November) periods in the Arctic are shown for SIT, and summertime (July) for SIC, for 2015 as an example year. The benefit of the CryoSat-2 assimilation on the forecasts of SIT is clear, with substantial reductions seen in both the mean difference and RMSD compared to the control (Figs. 11(a) and (b)). The improvement is particularly marked for the freeze-up period (Fig. 11(b)), as was illustrated on Fig. 9 for the 1-day forecast (assimilation background). The uncertainty in the SIT forecasts themselves grows by only a few centimetres over the 5-day period, as the thick sea ice observed by CryoSat-2 does not change rapidly on this timescale.

Figure 11(b) also unexpectedly shows that there is a slight improvement in the mean difference (becomes closer to zero) of the observations minus the model over the 5-day forecast in the freeze-up period, for the SIT assimilation experiment. The improvement may be due to smoothing out over the 5-day period of spatial noise introduced by the assimilation. Noise in the assimilation could be improved by further tuning of the error covariances. However, this could also be a result of the observed ice growing faster than in the model, which has been shown to have a systematic error towards thinner ice in the control (Fig. 5; Figs. 11(a),(b)).
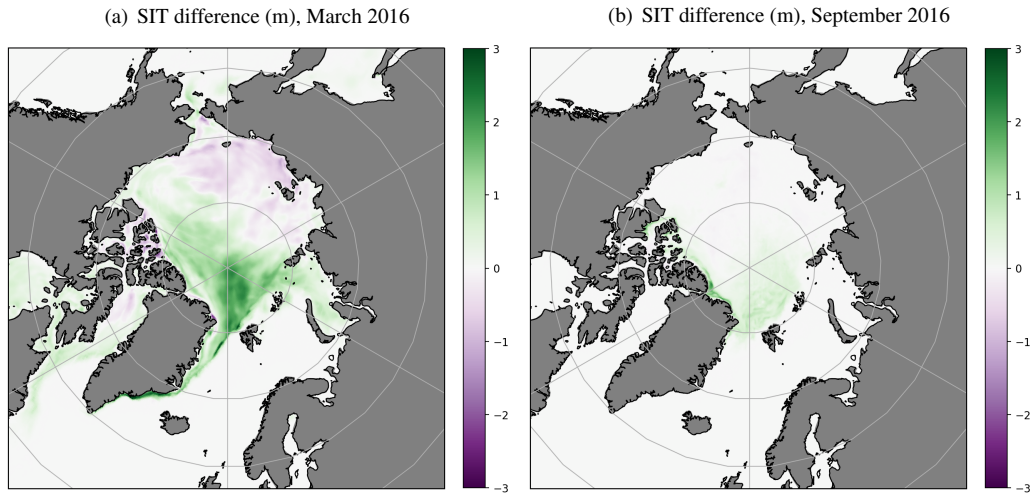
17

Figure 6. Monthly mean SIT differences (in m) for SIT assimilation experiment minus control, March and September 2016. Green (purple) indicates thicker (thinner) ice in the SIT assimilation experiment compared to the control.
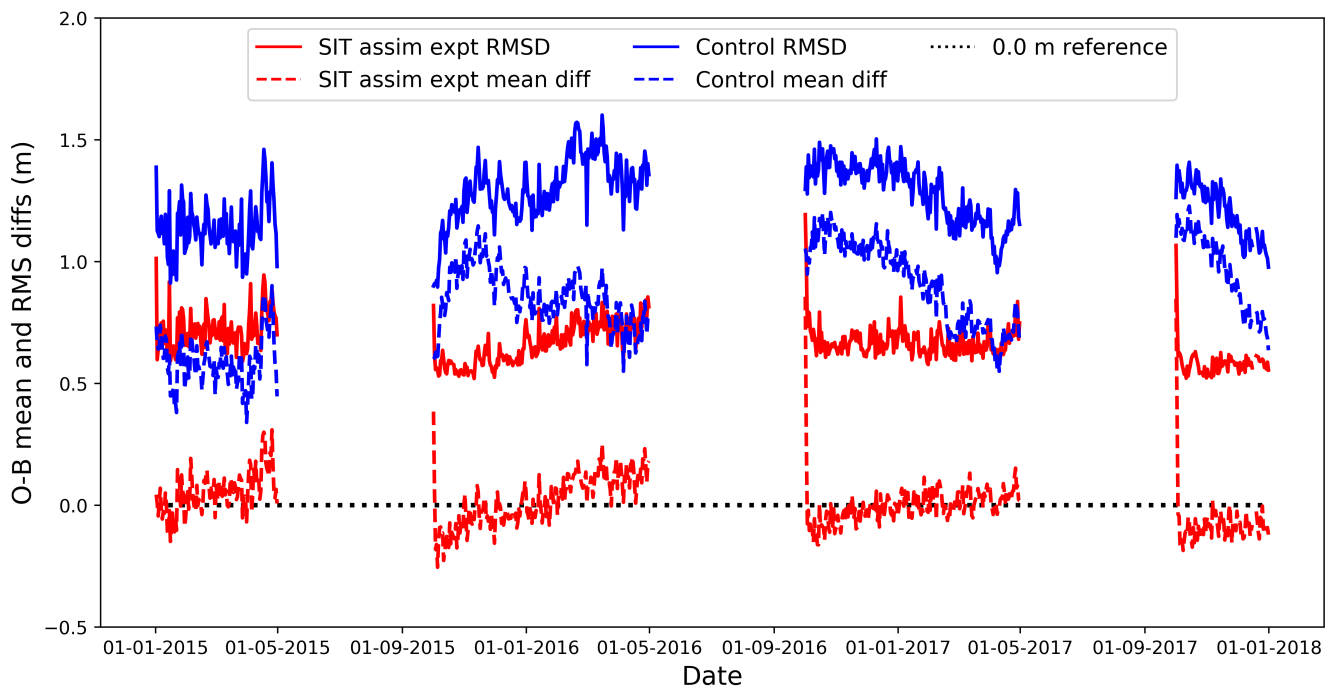

Figure 7. Timeseries of daily SIT observation-minus-background (O-B) mean and RMS differences for 1 January 2015 to 31 December 2017.
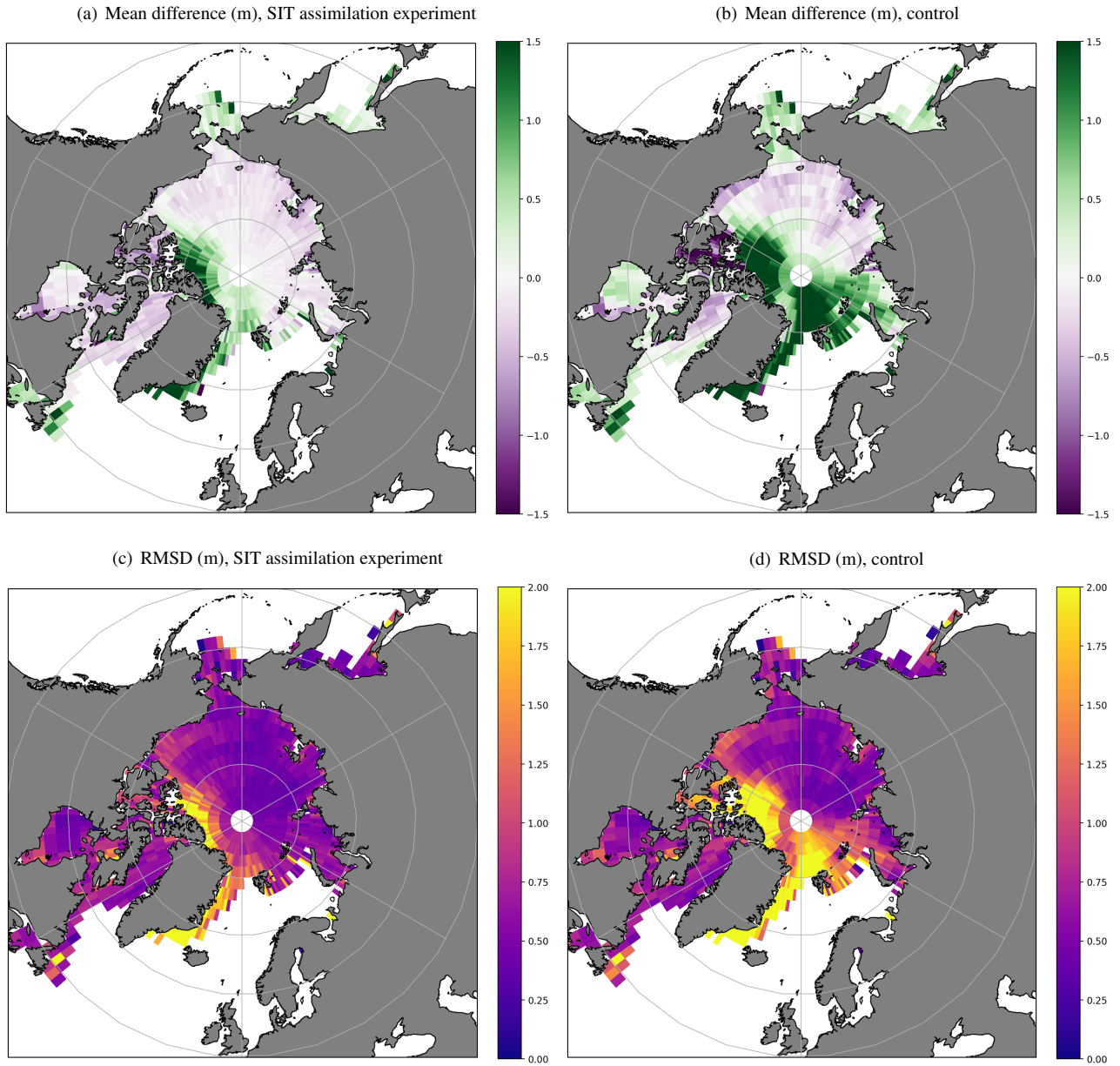
Figure 8. Spatially-binned SIT observation-minus-background statistics (in m) for March-April 2015. For (a), (b): Green (purple) indicates that model ice is thinner (thicker) than the observations. For (c), (d): A reduced RMSD indicates an improvement in the model compared to the observations.
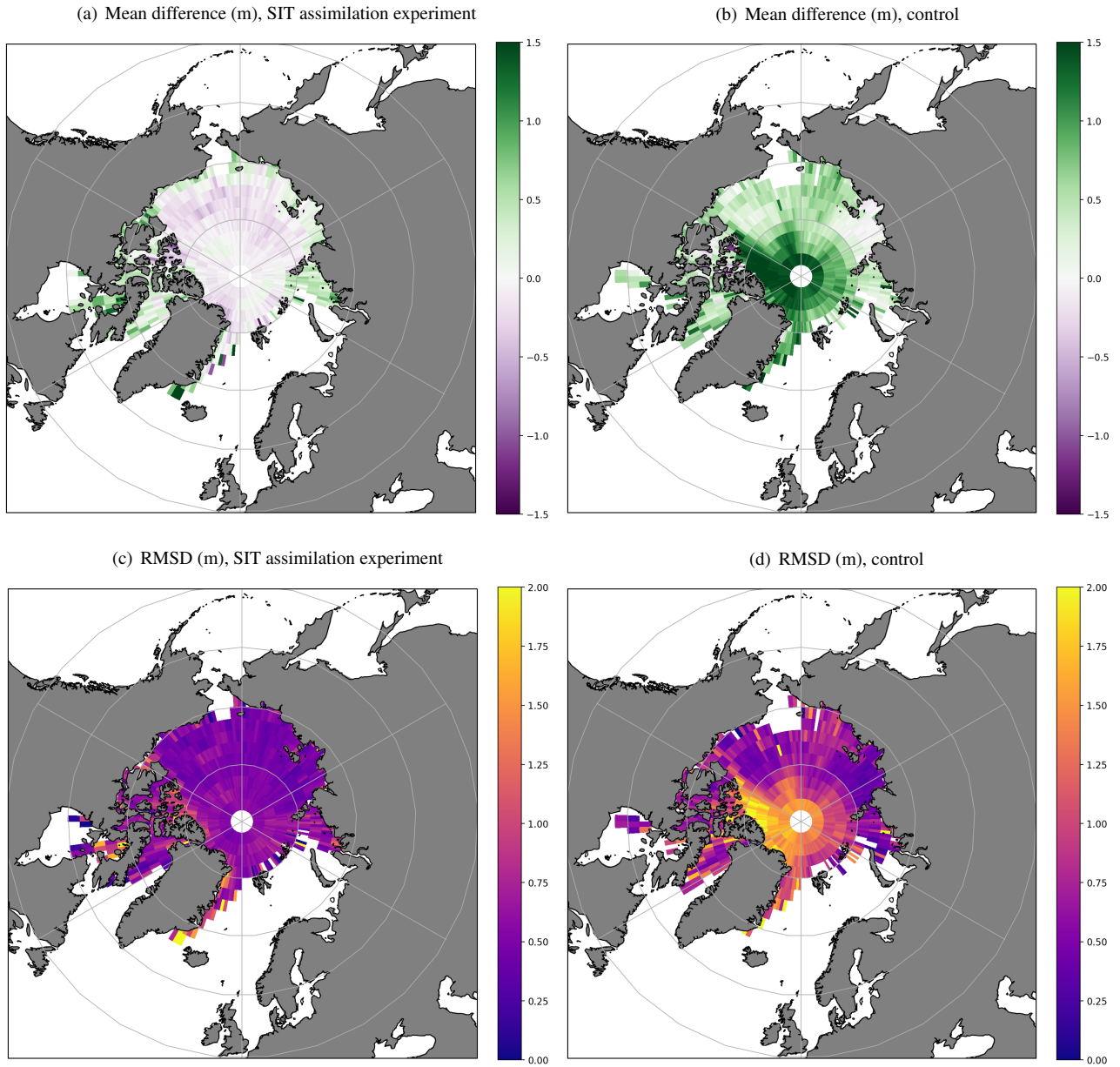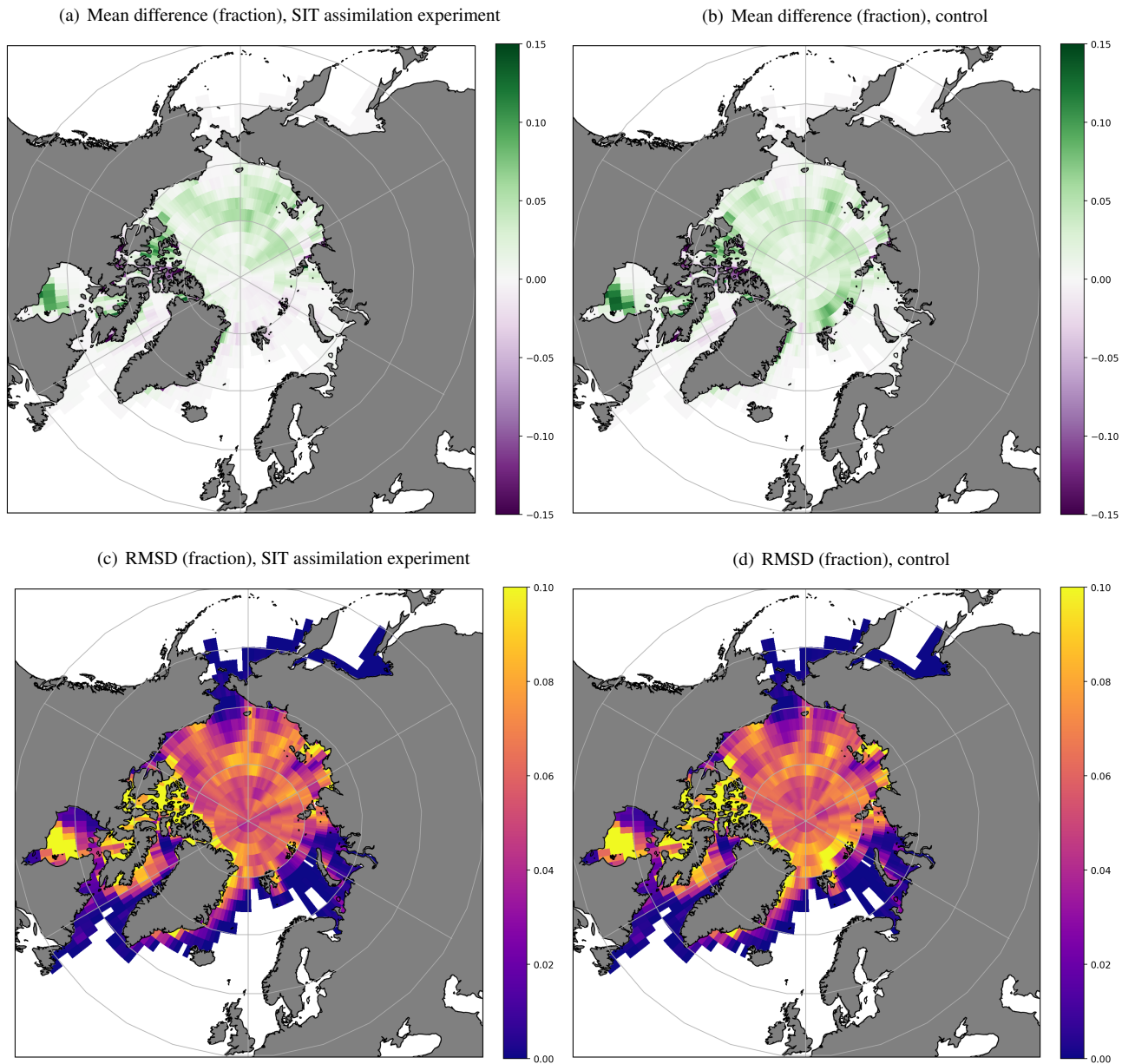
Figure 9. Spatially-binned SIT observation-minus-background statistics (in m) for October-November 2015. For (a), (b): Green (purple) indicates that model ice is thinner (thicker) than the observations. For (c), (d): A reduced RMSD indicates an improvement in the model compared to the observations.

Figure 10. Spatially-binned SIC observation-minus-background statistics (in sea ice fraction) for July 2015. For (a), (b): Green (purple) indicates that model ice concentration is lower (higher) than the observations. For (c), (d): A reduced RMSD indicates an improvement in the model compared to the observations.

Figure 11(c) shows SIC forecast statistics for July 2015 (no SIT observations are available for this time period), and demonstrates improvement in the summertime SIC forecasts for the SIT assimilation experiment compared to the control. This was highlighted above for assimilation background (1-day forecast) assessment results in July (Fig. 10), which demonstrated the improvement is located in the European sector. As discussed, this is likely to be due to the retention of thicker ice introduced by the SIT assimilation earlier in the year, which leads to a reduction in SIC forecast error growth during the summer. Differences between SIC forecasts generated by the SIT assimilation experiment and control at other times of the year are negligible.

5   Validation using independent in situ observations

FOAM model output from the SIT assimilation experiment and the control has been validated using airborne radar and laser altimeter observations of SIT from NASA OIB, moored upward-looking sonar observations of sea ice draft from BGEP, and combined SIT and snow depth observations from PAM-ARCMIP using Air-EM. Details of these datasets and pre-processing applied are given in Sect. 2.3. The CryoSat-2 SIT observations used in the SIT assimilation experiment have also been compared with these in situ datasets, to provide context for the changes in the model due to their assimilation. Tilling et al. (2015) and Tilling et al. (2018) also provide independent validation of CPOM CryoSat-2 SIT observations.

In order to produce matchups between FOAM model fields and the validation observations, the model fields were interpolated to the observation locations. An offline observation operator, part of the NEMOVAR assimilation code (Waters et al., 2015), was used for this purpose. The OIB and Air-EM data are daily means, so the daily mean model field for the date of the observations was used when producing the matchups. For the monthly BGEP data, the monthly mean of the model fields was used.

For matchups of CryoSat-2 SIT observations with the in situ data, 30 days' worth of the quality-controlled, super-obbed, assimilated subset of the full dataset was gridded onto the same 1/4 degree tripolar (ORCA025) grid used by the model. The gridded observation field was then interpolated to the observation locations. The 30-day periods were chosen to centre on the middle of the observation window of each yearly field campaign for the OIB and Air-EM validation datasets, and calendar months were used when producing matchups with the BGEP dataset. The necessary exception to this was for OIB in 2016, with fieldwork dates of 20, 28 April (Table 1). Here, the period of 1-30 April was used in order to acquire 30 days of CryoSat-2 data, since data production ceases on 30 April each year before resuming in the autumn. This means that OIB matchups with CryoSat-2 in 2016 may not be representative of the true relationship between the datasets. Nevertheless, these matchups are not outliers of the OIB matchups group (shown on Figure 12(c)), indicating a comparable level of accuracy. For comparison of CryoSat-2 SIT observations with the combined SIT and snow depth of the Air-EM observations, the 30-day mean of the modelled snow depth was added to the gridded CryoSat-2 SIT observations before producing the matchups. The snow depth from the SIT assimilation experiment rather than the control was used, to maintain consistency with the assimilated SIT observations.
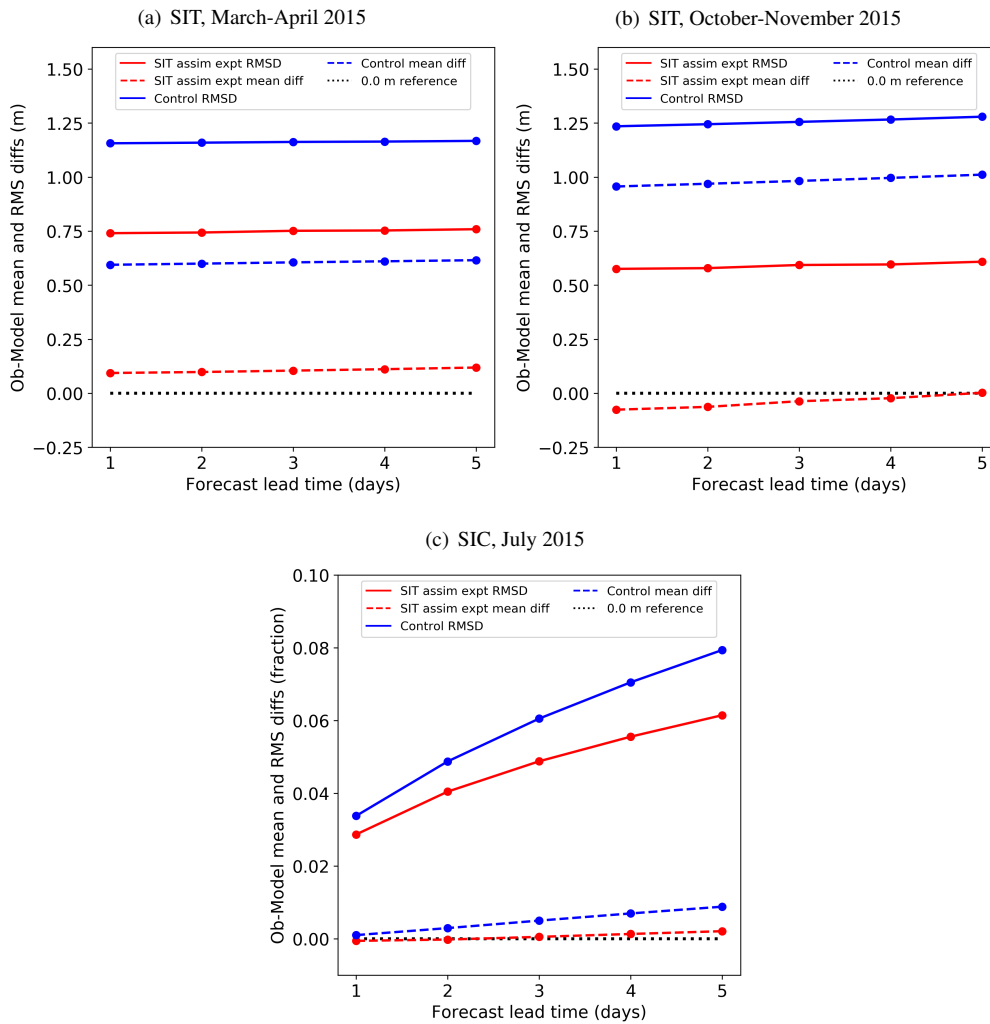
Figure 11. Mean and RMS differences of SIT (in m) and SIC (in fraction) observations minus 1- to 5-day forecasts for (a) SIT in March-April 2015 (break-up period), (b) SIT in October-November 2015 (freeze-up period), and (c) SIC in July 2015.

5.1 Operation IceBridge (OIB)

Figures 12(a) and (b) show OIB cluster average SIT observations from 2015, 2016 and 2017, minus the 5-day SIT forecast
produced by the FOAM system, for the SIT assimilation experiment and the control respectively. The figures demonstrate that,
overall, the SIT assimilation experiment produces a substantial improvement in the SIT forecast field compared to the control.

	SIT assimilation experiment		Control		CryoSat-2 obs
	Analysis	5-day forecast	Analysis	5-day forecast	
Correlation coefficient	0.76	0.80	0.57	0.58	0.78
Mean difference (m)	0.14	0.15	0.75	0.74	0.07
Absolute mean difference (m)	0.51	0.48	0.87	0.86	0.49
RMS of differences (m)	0.69	0.65	1.11	1.11	0.65
Standard deviation of differences (m)	0.68	0.64	0.82	0.82	0.65

Table 2. Statistics for all matchups (total 547) of OIB SIT observations from 2015-2017 with FOAM SIT analysis, FOAM 5-day SIT forecast,
and assimilated CryoSat-2 SIT observations. Differences are in situ observation minus model or satellite observation.

Figure 12(c) shows a scatter plot of matchups between CryoSat-2 SIT observations and OIB cluster average SIT observations
for 2015, 2016 and 2017. Figures 12(d) and (e) show scatter plots of the OIB cluster average SIT observations with the corre-
sponding FOAM 5-day SIT forecast matchup points, for the SIT assimilation experiment and control respectively. Summary
statistics of the relationships are given in Table 2, along with the relationships between the OIB observations and the FOAM
SIT analysis (a 1-day forecast that has been corrected by the assimilated observations, and is used to initialise subsequent
forecasts). These results demonstrate that, using OIB SIT as a reference for validation, the CryoSat-2 SIT observations are
more reliable than the model without SIT assimilation. Assimilating these observations leads to improvements in the model
performance, with fewer outlying points and the best fit line being closer to the ideal 1:1 line. Table 2 illustrates that there are
improvements across all statistics for the SIT assimilation experiment compared to the control. As indicated in Sect. 4 using
assimilation statistics, validation of the SIT assimilation experiment against the independent OIB observations also demon-
strates an unexpected slight improvement in the 5-day forecast compared with the analysis (Table 2). As discussed, this may
indicate that spatial noise is introduced by the assimilation, which is smoothed out over the 5-day forecast period. The good
results for the forecast again demonstrate that, throughout the 5-day forecast period, the model is able to successfully retain
improvements to the SIT field introduced by the initial assimilation. There is very little difference between the statistics for the
SIT analysis and 5-day forecast for the control run, since the control analysis has not assimilated SIT data and the generally
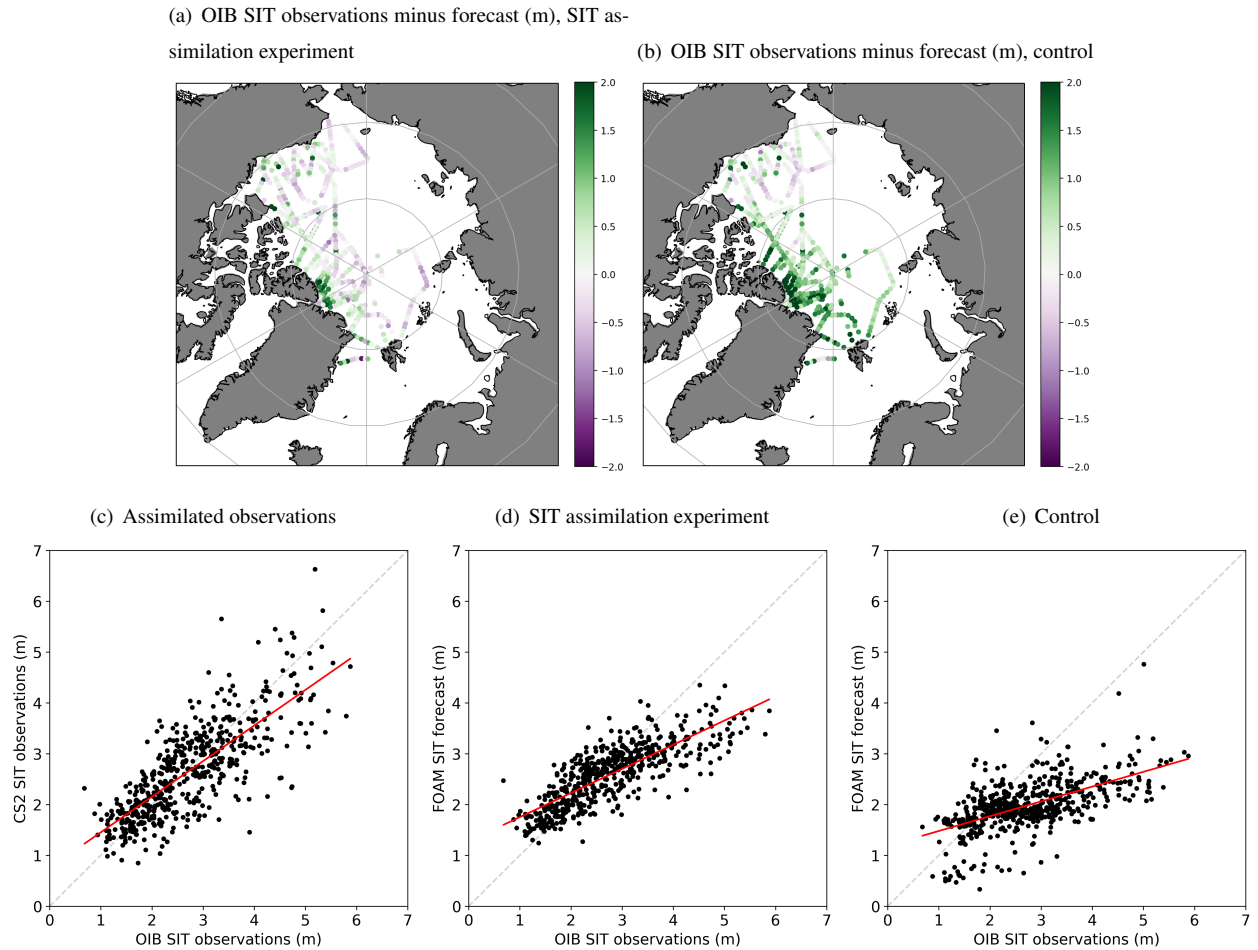thick ice being assessed changes slowly over this timescale.

24

(a) OIB SIT observations minus forecast (m), SIT assimilation experiment

(b) OIB SIT observations minus forecast (m), control

(c) Assimilated observations

(d) SIT assimilation experiment

(e) Control

Figure 12. Validation of CryoSat-2 (CS2) SIT observations and FOAM 5-day SIT forecasts against OIB cluster average SIT observations from 2015-2017. OIB observations minus FOAM forecasts (in m) for (a) SIT assimilation experiment and (b) control. Green (purple) in (a), (b) indicates the model ice is thinner (thicker) than the observations. Scatter plots of OIB observations (in m) with (c) monthly, gridded, quality-controlled CryoSat-2 SIT observations (in m); and FOAM 5-day SIT forecasts (in m) for (d) SIT assimilation experiment and (e) control. Best fit lines for (c), (d) and (e) are shown in red, with a 1:1 reference line in broken grey.

5.2 Airborne electromagnetic combined snow and ice thickness (Air-EM) observations

Figures 13(a) and (b) show Air-EM cluster average observations of combined SIT and snow depth from 2015, minus FOAM
5-day forecasts of the same quantity, for the SIT assimilation experiment and control respectively. Figure 13 also shows scatter
plots of the Air-EM observations, with matchups of CryoSat-2 SIT observations plus modelled snow depth (Fig. 13(c)), and
FOAM 5-day forecasts of SIT plus snow depth for the SIT assimilation experiment and control (Figs. 13(d),(e) respectively).
Summary statistics of the relationships between these datasets and also with the FOAM SIT analysis plus snow depth are given
in Table 3.

For observations in the Canadian Arctic, it can be seen that differences between the Air-EM observations and the FOAM
output are generally reduced for the forecast runs assimilating SIT observations compared to the control. However, there are
some larger negative differences, which indicate that the modelled combined SIT and snow depth for a few points has increased
too much in the SIT assimilation experiment. Figure 13 also illustrates that differences between the observations and the model
in the Beaufort Sea are slightly smaller for the control than for the SIT assimilation experiment, indicating that the assimilation
leads to a degradation in model performance for this region. Validation of the FOAM SIT analysis against Air-EM observations
(Table 3) yields similar results to the forecasts.

Table 3 indicates that the CryoSat-2 SIT observations plus modelled snow depth are substantially thicker than the Air-EM
observations, with a large negative mean difference, and Fig. 13(c) illustrates that this mainly results from observations in
the Canadian Arctic. Since the modelled SIT is too thin, assimilating these SIT observations has the compensating effect of
reducing the mean difference of the model compared to the Air-EM observations, and this also reduces the RMSD (Table 3).
However, the model standard deviation and correlation coefficient are poorer on assimilation of these data. It should be noted
that there are many more OIB observations over both the Canadian Arctic and Beaufort Sea regions (Fig. 12), and these agree
much better with the model output and the CryoSat-2 observations than do the Air-EM observations. This potentially indicates
an uncertainty in the quality of the Air-EM observations. Uncertainty in the snow depth will also be contributing to this issue,
although the difference between the CryoSat-2 and Air-EM observations is greater than the snow depth itself.

Figure 14 shows the difference in snow depth between the SIT assimilation experiment and the control, for 11 April 2015 as
an example date when Air-EM and FOAM matchups are available. This illustrates that the snow is generally ∼5 cm deeper at
the Air-EM matchup locations (shown on e.g. Fig. 13(a)) for the SIT assimilation experiment than for the control. This indicates
that uncertainties in the modelled snow depth will be affecting results, not only for the Air-EM and CryoSat-2 comparisons,
but for assessment of the SIT assimilation experiment and the control using Air-EM validation data. An assessment of the snow
mass budget in the model shows that the deeper modelled snow in the SIT assimilation experiment is due mostly to the thicker
ice from the assimilation reducing the ice surface temperature, which results in reduced evaporation/sublimation compared
to the control. Note that this change, despite increasing the mean difference between the CryoSat-2 and Air-EM matchups,
actually brings the modelled snow depth closer to climatology (see Fig. 1).

Unlike for the validation against OIB observations (Sect. 5.1, Table 2), the Air-EM 5-day forecast statistics do not show an
improvement over the analysis for the SIT assimilation experiment (Table 3). This may be a result of uncertainty due to snow
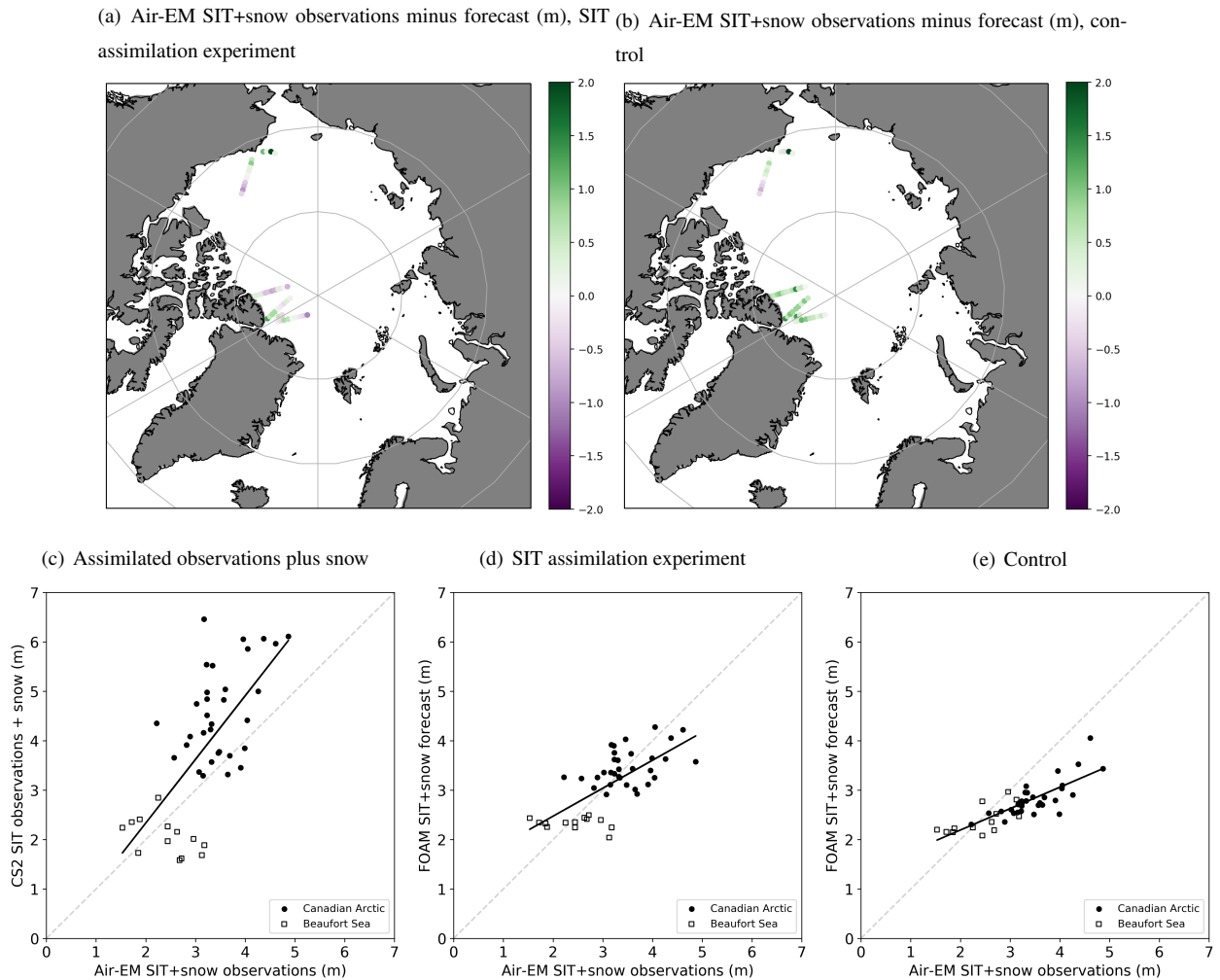
(a) Air-EM SIT+snow observations minus forecast (m), SIT assimilation experiment

(b) Air-EM SIT+snow observations minus forecast (m), control

(c) Assimilated observations plus snow

(d) SIT assimilation experiment

(e) Control

Figure 13. Validation of CryoSat-2 (CS2) SIT observations plus modelled snow depth, and FOAM 5-day SIT plus snow depth forecasts against Air-EM cluster average combined SIT and snow depth observations from 2015. Air-EM observations minus FOAM forecasts (in m) for (a) SIT assimilation experiment and (b) control. Green (purple) in (a), (b) indicates the model ice is thinner (thicker) than the observations. Scatter plots of Air-EM observations (in m) with (c) monthly, gridded, quality-controlled CryoSat-2 SIT observations plus modelled snow depth (in m); and FOAM 5-day forecasts (in m) for (d) SIT assimilation experiment and (e) control. Best fit lines for (c), (d) and (e) are shown in red, with a 1:1 reference line in broken grey.

27
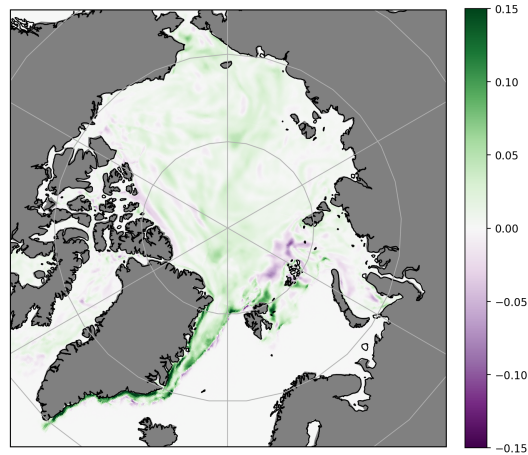
Figure 14. Difference in FOAM daily-mean modelled snow depth (m) for SIT assimilation experiment minus control, 11 April 2015.

	SIT assimilation experiment		Control		CryoSat-2 obs
	Analysis	5-day forecast	Analysis	5-day forecast	
Correlation coefficient	0.68	0.69	0.82	0.82	0.68
Mean difference (m)	-0.03	0.04	0.48	0.48	-0.67
Absolute mean difference (m)	0.46	0.45	0.58	0.57	1.03
RMS of differences (m)	0.55	0.54	0.67	0.67	1.25
Standard deviation of differences (m)	0.55	0.54	0.48	0.47	1.06

Table 3. Statistics for all matchups (total 45) of Air-EM combined SIT and snow depth observations from 2015 with FOAM SIT analysis plus snow depth, FOAM SIT plus snow depth 5-day forecast, and assimilated CryoSat-2 SIT observations plus modelled snow depth. Differences are in situ observation minus model or satellite observation.

depth in the Air-EM validation, a potential sampling error owing to the limited number of Air-EM matchups, or uncertainties in the Air-EM observations themselves.

5.3 Beaufort Gyre Exploration Project (BGEP)

485 Figure 15 shows scatter plots of monthly mean SIT observations derived from BGEP sea ice draft observations for January 2015 to August 2017 (no BGEP data were available after this date), with matchups to CryoSat-2 SIT observations and FOAM SIT analyses. Results for FOAM 5-day forecasts are not shown since these were only produced for selected months, and the limited number of data points would lead to unreliable statistics. Table 4 summarises the statistical relationships between the observation and model or satellite datasets. Results in Fig. 15 and Table 4 are separated into bins of thin ice (below 1 m) and 490 thicker ice (equal to or above 1 m), based on the BGEP SIT. Points on Fig. 15 are colour-coded by seasonal groupings of

months, with May-September together to represent the summer months without SIT assimilation, and when the effect of snow potentially biasing the draft to SIT conversion is negligible.

Figure 15 and Table 4 demonstrate that CryoSat-2 SIT observations over 1 m thick are slightly less reliable than the model without assimilation, using the BGEP data as a validation reference. Consequently, statistics for the SIT assimilation experiment are slightly poorer than for the control. However, overall, for SIT above 1 m, the control, SIT assimilation experiment and CryoSat-2 statistics are good, and the difference is considerably smaller than for observations below 1 m. It is therefore less likely that a dramatic improvement could be achieved for this region by the assimilation of thick SIT observations from CryoSat-2, unlike for other areas such as in the Canadian Arctic where the SIT is much too thin in the control (see Fig. 5).

For matchups where the BGEP SIT observations are under 1 m, the mean difference, standard deviation and RMSD between the model and in situ observations are all poorer for the SIT assimilation experiment than for the control (Figs. 15(b),(c); Table 4), despite the assimilation giving very little weight to CryoSat-2 observations under 1 m. This difference results from the October-November observations (results for the SIT assimilation experiment and control are similar for the other thin ice matchups, in May-September). The negative mean difference between the validation observations and the model indicates that in the Beaufort Sea region, where the BGEP moorings are located, the modelled ice in the SIT assimilation experiment thickens up more quickly in October-November than the BGEP observations. Investigation has shown that this is unrelated to positive SIT increments in the assimilation being spread too far from regions of thicker ice. It is instead likely to be a result of poorly-specified observation uncertainties for the assimilated CryoSat-2 data. At high latitudes there are up to around 100 observations per super-observation due to overlapping orbit tracks of the satellite, but at lower latitudes, towards the ice edge, there can be as few as 2 observations. This means there will be an increased contribution of random error in these observations. Since the measurement uncertainty component of the OBE is based on the magnitude of the SIT observations themselves, this (along with other potential biases in the data) is not taken into account. Mitigating the issue of random error is an avenue for future investigation, and could include, for example, filtering the CryoSat-2 observations by latitude or by a minimum number of observations included in the super-observation. Alternatively, the Bayesian background check discussed in Sect. 2.2.2 could be used to reject observations deviating too far from the model. However, as previously discussed, this relies on the free model (without assimilation) being relatively unbiased. It is also possible that thicker modelled ice in the Canadian Arctic in the SIT assimilation experiment is being advected into this region by the model, following the general pattern of ice circulation in this region. Conversely however, the relationships between SIT and season shown on Fig. 15 illustrate that the summertime (May to September) model SIT is too thin compared to the BGEP observations. The ice mass balance budget of the model for this area could therefore be assessed as an avenue of future investigation. Research into assimilating observations of SIT under 1 m from the SMOS satellite instrument is also underway, which aims to improve the representation of thinner ice in FOAM and will help to mitigate these problems.

The BGEP data for March-April 2015-2017 (Fig. 15(a)) compares better with the CryoSat-2 observations than does the Air-EM data in the Beaufort Sea for April 2015 (Fig. 13(a)). This again suggests that the Air-EM data may be unreliable, as discussed in Sect. 5.2.
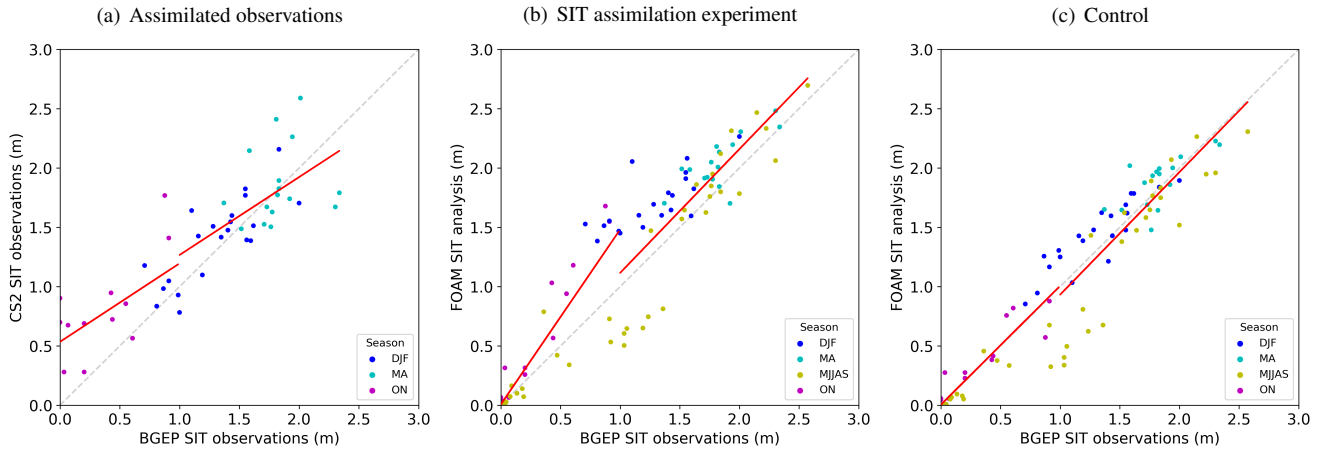
Figure 15. Validation of monthly gridded CryoSat-2 (CS2) observations and monthly mean FOAM SIT analyses against BGEP monthly mean observations of SIT (converted from sea ice draft) from January 2015 to August 2017, colour-coded by season. Scatter plots of BGEP observations (in m) with (a) monthly, gridded, quality-controlled CryoSat-2 SIT observations (in m); and FOAM analyses (in m) for (b) SIT assimilation experiment and (c) control. Note that CryoSat-2 matchups are not available from May to September each year). Best fit lines (in red) are plotted separately for BGEP SIT observations below 1 m, and equal to or above 1 m. 1:1 reference line shown in broken grey.

	SIT assimilation experiment		Control		CryoSat-2 obs	
For SIT observations:	< 1 m	>= 1 m	< 1 m	>= 1 m	< 1 m	>= 1 m
Correlation coefficient	0.89	0.79	0.89	0.82	0.65	0.59
Mean difference (m)	-0.18	-0.14	-0.01	0.05	-0.36	-0.05
Absolute mean difference (m)	0.24	0.28	0.12	0.21	0.38	0.25
RMS of differences (m)	0.36	0.33	0.18	0.27	0.47	0.31
Standard deviation of differences (m)	0.31	0.30	0.18	0.26	0.30	0.30

Table 4. Statistics for all matchups of BGEP SIT observations with FOAM SIT analysis and assimilated CryoSat-2 SIT observations, January 2015 - August 2017, for thickness bins less than 1 m (total matchups 35 for model and 17 for satellite observations) and equal to or above 1 m (total matchups 54 for model and 32 for satellite observations). Differences are observation minus model or satellite observation. Satellite observations not available between May and September each year. CryoSat-2 observations below 1 m are given very little weight in the assimilation.

6 Discussion and conclusions

The feasibility of assimilating sea ice thickness (SIT) derived from CryoSat-2 along-track measurements of sea ice freeboard into a global, coupled ocean-sea ice model, Forecast Ocean Assimilation Model (FOAM), has been demonstrated. This is a novel use of along-track SIT observations, as other centres have previously used only gridded, temporally-averaged SIT mea-

surements. The assimilation results in improvements to the SIT analysis and forecast fields generated by FOAM compared to a

530   control without SIT assimilation, validated using SIT observation-minus-background mean difference and RMSD assimilation statistics. Arctic-wide improvements are 0.75 m mean difference (0.41 m RMSD) in the freeze-up period (October-November) and 0.46 m mean difference (0.33 m RMSD) in the ice break-up period (March-April), calculated using data from 2015-2017. Regional improvements in the Canadian Arctic are particularly notable, where there is a reduction of more than 1.30 m mean difference (1.22 m RMSD) for SIT in the break-up period.

535   Comparison with independent springtime in situ SIT observations from NASA Operation IceBridge (OIB; Kurtz et al., 2019) also indicates that the assimilation results in substantial improvements to the model SIT, of 0.61 m mean difference and 0.42 m RMSD. Validation against monthly sea ice draft measurements from the Beaufort Gyre Exploration Project (BGEP) shows that, for thicknesses above 1 m, the model performance without SIT assimilation is similar to (and slightly better than) that of CryoSat-2 SIT observations. Therefore, the SIT assimilation experiment does not show an improvement compared to the

540   control against this validation dataset, although the statistics for both model runs are good. The SIT assimilation experiment experiences a degradation for thicknesses below 1 m compared to the control, validated using the BGEP observations, despite giving very little weight to SIT observations below 1 m in the analysis. The degradation may be due to poorly-specified observation uncertainties in the assimilated observations. Validation against springtime airborne electromagnetic induction (Air-EM) combined SIT and snow depth observations (Haas et al., 2009) yields poorer results than for the OIB and BGEP

545   datasets, despite covering similar locations. This may be evidence of uncertainty in the Air-EM observations, or sampling error owing to the limited number of matchups available from this dataset. It may also be a result of noise in the SIT analysis, uncertainty in the modelled snow depth, or uncertainty in the assimilated observations.

Despite the lack of CryoSat-2 SIT observations over the summer months due to the presence of meltponds affecting re-trievals, the model has been shown to retain improvements to the SIT field throughout the summer due to previous SIT assim-

550   ilation. Sea ice concentration (SIC) assessment results for the SIT assimilation experiment show a regional improvement in July in the European sector, compared to a control. This is likely due to the slower melting of thicker ice introduced by the SIT assimilation in previous months. Overall, it is concluded that CryoSat-2 along-track measurements, rather than gridded and temporally-averaged observations, can be successfully assimilated.

The heterogeneous nature of sea ice means that SIT observations will always contain sub-gridscale noise, as well as sizeable

555   uncertainties in the data. The mitigating effect of using a model and a well-specified data assimilation scheme avoids the need for gridding or temporal averaging of the noisy CryoSat-2 data. A spatial averaging of sorts is carried out on the freeboard observations through the process of super-obbing, which takes the median observation within a 10 km radius. This aims to reduce the random uncertainty in the observations, while preserving spatial variability at the model grid scale. Remaining noise in the SIT analysis could be reduced by increasing the observation error variance (OBE) for, or filtering out, super-

560   observations with a number of input values below a specified threshold. This issue mainly affects lower latitudes close to the ice edge, and the planned assimilation of Soil Moisture and Ocean Salinity (SMOS) thin ice observations (below 1 m) into FOAM will help to improve the SIT analysis and forecasts in these regions. Measurement uncertainties generated as part of

the processing chain for each CPOM freeboard observation, and not determined solely from the magnitude of the observation itself, would be very useful since the current method is reliant on the observations being unbiased.

565   There are a number of options for further development of SIT assimilation in FOAM. Well-specified OBEs and background error variances (BGEs), and importantly the balance between them, are vital for producing high-quality analyses. Owing to the reduction in Arctic sea ice cover in recent decades (e.g. Meredith et al., 2019), a climatological BGE as used in this study may not be the optimal method for SIT assimilation. Instead, a daily-varying BGE could be obtained for SIT from the latest ensemble system under development at the Met Office. This would be combined with the climatological BGE in a hybrid
570   ensemble/variational framework, as is planned for the ocean variables in FOAM. Use of the NEMOVAR dual length scale capability for background error covariances could also be investigated for SIT. Additional future plans include the assimilation of satellite snow depth observations into the FOAM system, which would reduce this source of uncertainty in the conversion of sea ice freeboard to SIT.

In order to introduce SIT assimilation into operational Met Office forecasting systems, improvements to the delivery time-
575   liness of the near-real-time observations would be required, as well as operational support of their dissemination. The CPOM near-real-time product used in this study currently has a latency of 72 hours and, at the present time, FOAM uses observations which are made available within 48 hours. The coupled numerical weather prediction (NWP) system being implemented in 2021 will require observations to be available as close to real-time as possible, up to a maximum delay of 24 hours. Operational implementation of SIT assimilation in FOAM will directly influence the Met Office's GloSea seasonal forecasting
580   system (MacLachlan et al., 2014), as this is initialised using FOAM analysis fields. Using improved FOAM SIT fields for initialisation, particularly for regions of thicker ice, is expected to be of substantial benefit to the quality of seasonal forecasts of sea ice extent, concentration and thickness produced by GloSea, as was demonstrated by Blockley and Peterson (2018).

CryoSat-2 is currently operating well beyond its 3.5-year nominal lifespan, having been launched in 2010. Freeboard observations from NASA's ICESat-2 (launched 2018, also with a design life of 3 years) are available, though not currently in
585   near-real-time. Although planning is underway for a high-priority ESA candidate satellite mission to observe the polar regions, Copernicus Polar Ice and Snow Topography Altimeter (CRISTAL), there is a real risk of an observation data gap occurring, should CryoSat-2 and ICESat-2 both fail before the CRISTAL mission is realised. The work in this study highlights the significance of the continuation of dedicated satellite missions for monitoring SIT, and demonstrates the suitability of near-real-time, along-track SIT observations for use in operational ocean-sea ice modelling and forecasting.

References

615 Aaboe, S., Breivik, L.-A., Sørensen, A., Eastwood, S., and Lavergne, T.: Global sea ice edge and type product user's manual OSI-402-c & OSI-403-c, Product User Manual SAF/OSI/CDOP2/MET-Norway/TEC/MA/205, Ocean and Sea Ice Satellite Application Facility, Norwegian Meteorological Institute, 2020.

Alexandrov, V., Sandven, S., Wahlin, J., and Johannessen, O. M.: The relation between sea ice thickness and freeboard in the Arctic, The Cryosphere, 4, 373–380, https://doi.org/10.5194/tc-4-373-2010, 2010.

620 Allard, R., Metzger, E., Barton, N., Li, L., Kurtz, N., Phelps, M., Franklin, D., Smedstad, O. M., Crout, J., and Posey, P.: Analyzing the impact of CryoSat-2 ice thickness initialization on seasonal Arctic sea ice prediction, Ann. Glaciol., 61(82), 78–85, https://doi.org/10.1017/aog.2020.15, 2020.

Allard, R. A., Farrell, S. L., Hebert, D. A., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A., and Wallcraft, A. J.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system, Adv. Space Res., 62, 1265–1280,
625 https://doi.org/10.1016/j.asr.2017.12.030, 2018.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Q. J. R. M. S., 134, 1951–1970, https://doi.org/10.1002/qj.339, 2008.

Barbosa Aguiar, A., Waters, J., Price, M., Inverarity, G., Pequignet, C., Maksymczuk, J., Smout-Day, K., Martin, M., Bell, M., King, R., While, J., and Siddorn, J.: The new Met Office global ocean forecast system at $1/12^{th}$ degree resolution, in prep., 2021.

630 Bertino, L. and Lisaeter, K. A.: The TOPAZ monitoring and prediction system for the Atlantic and Arctic Oceans, Journal of Operational Oceanography, 1:2, 15–18, https://doi.org/10.1080/1755876X.2008.11020098, 2008.

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, https://doi.org/doi/10.5194/tc-12-3419-2018, 2018.

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A., and Storkey,
635 D.: Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new Global FOAM forecasts, Geosci. Model Dev., 7, 2613–2638, https://doi.org/10.5194/gmd-7-2613-2014, 2014.

Bloom, S. C., Takacs, L. L., da Silva, A. M., and Ledvina, D.: Data assimilation using incremental analysis updates, Mon. Weather Rev., 124, 1256–1271, 1996.

Carse, F., Martin, M. J., Sellar, A., and Blockley, E. W.: Impact of assimilating temperature and salinity measurements by animal-borne
640 sensors on FOAM ocean model fields, Q. J. R. M. S., 141(693), 2934–2943, https://doi.org/10.1002/qj.2613, 2015.

Collow, T. W., Wang, W., Kumar, A., and Zhang, J.: Improving Arctic sea ice prediction using PIOMAS initial sea ice thickness in a coupled ocean-atmosphere model, Mon. Weather Rev., 143(11), 4618–4630, https://doi.org/10.1175/MWR-D-15-0097.1, 2015.

Comiso, J. C., Parkinson, C. L., Gersten, R., and Stock, L.: Accelerated decline in the Arctic sea ice cover, Geophys. Res. Lett., 35, L01 703, https://doi.org/10.1029/2007GL031972, 2008.

645 Day, J. J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, Geophys. Res. Lett., 41, 7566–7575, https://doi.org/10.1002/2014GL061694, 2014.

Dirkson, A., Merryfield, W. J., and Monahan, A.: Impacts of sea ice thickness initialization on seasonal Arctic sea ice predictions, J. Climate, 30(3), 1001–1017, https://doi.org/10.1175/JCLI-D-16-0437.1, 2017.

Farrell, S. L., Kurtz, N., Connor, L. N., Elder, B. C., Leuschen, C., Markus, T., McAdoo, D. C., Panzer, B., Richter-Menge, J., and Sonntag, J. G.: A first assessment of IceBridge snow and ice thickness data over Arctic sea ice, IEEE Transactions on Geoscience and Remote Sensing, 50(6), 2098–2111, https://doi.org/10.1109/TGRS.2011.2170843, 2012.

Fiedler, E. K., Mao, C., Good, S. A., Waters, J., and Martin, M. J.: Improvements to feature resolution in the OSTIA sea surface temperature analysis using the NEMOVAR assimilation scheme, Q. J. R. M. S., 145, 3609–3625, https://doi.org/doi/10.1002/qj.3644, 2019.

Giles, K. A., Laxon, S. W., Wingham, D. J., Wallis, D. W., Krabill, W. B., Leuschen, C. J., McAdoo, D., Manizade, S. S., and Raney, R. K.: Combined airborne laser and radar altimeter measurements over the Fram Strait in May 2002, Remote Sens. Environ., 111, 182–194, 2007.

Guiavarc'h, C., Roberts-Jones, J., Harris, C., Lea, D. J., Ryan, A., and Ascione, I.: Assessment of ocean analysis and forecast from an atmosphere-ocean coupled data assimilation operational system, Ocean Sci., 15, 1307–1326, https://doi.org/10.5194/os-15-1307-2019, 2019.

Haas, C., Lobach, J., Hendricks, S., Rabenstein, L., and Pfaffling, A.: Helicopter-borne measurements of sea ice thickness, using a small and lightweight, digital EM system, Journal of Applied Geophysics, 67, 234–241, 2009.

Hollingsworth, A. and Lönnberg, P.: The statistical structure of short-range forecast errors as determined from radiosonde data. Part i: The wind field, Tellus, 38A, 111–136, https://doi.org/10.1111/j.1600-0870.1986.tb00460.x, 1986.

Houtekamer, P. L., Lefaivre, L., Derome, J., Ritchie, H., and Mitchell, H. L.: A system simulation approach to ensemble prediction, Mon. Weather Rev., 124, 1225–1242, 1996.

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., and Elliott, S.: CICE: The Los Alamos sea ice model documentation and software user's manual version 5.1, User Manual LA-CC-06-012, Los Alamos National Laboratory, NM, 2015.

Ingleby, N. B. and Huddleston, M.: Quality control of ocean temperature and salinity profiles - Historical and real-time data, J. Marine Sci., 65, 158–175, 2005.

Janjic, T., Bormann, N., Bocquet, M., Carton, J. A., Cohn, S. E., Dance, S. L., Losa, S. N., Nichols, N. K., Potthast, R., Waller, J. A., and Weston, P.: On the representation error in data assimilation, Q. J. R. M. S., 144, 1257–1278, https://doi.org/10.1002/qj.3130, 2017.

Kaleschke, L., Maass, N., Haas, C., Hendricks, S., Heygster, G., and Tonboe, R. T.: A sea-ice thickness retrieval model for 1.4 GHz radiometry and application to airborne measurements over low salinity sea-ice, The Cryosphere, 4, 583–592, https://doi.org/10.5194/tc-4-583-2010, 2010.

Kern, S., Khvorostovsky, K., Skourup, H., Rinne, E., Parsakhoo, Z. S., Djepa, V., Wadhams, P., and Sandven, S.: The impact of snow depth, snow density and ice density on sea ice thickness retrieval from satellite radar altimetry: Results from the ESA-CCI Sea Ice ECV Project Round Robin Exercise, The Cryosphere, 9, 37–52, https://doi.org/10.5194/tc-9-37-2015, 2015.

Koenigk, T., Caian, M., Nikulin, G., and Schimanke, S.: Regional Arctic sea ice variations as predictor for winter climate conditions, Clim. Dyn., 46, 317–337, https://doi.org/10.1007/s00382-015-2586-1, 2016.

Krishfield, R. and Proshutinsky, A.: BGOS ULS Data Processing Procedure, Technical Report, Woods Hole Oceanographic Institution, http://www.whoi.edu/fileserver.do?id=85684&pt=2&p=100409, 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, J. Geophys. Res. Oceans, 119, 1271–1305, https://doi.org/10.1002/2013JC008999, 2014.

Kurtz, N., Studinger, M., Harbeck, J., Onana, V.-D.-P., and Farrell, S.: IceBridge Sea Ice Freeboard, Snow Depth, and Thickness, 2015-2017, Digital media, http://nsidc.org/data/idcsi2.html, 2019.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data, The Cryosphere, 7, 1035–1056, https://doi.org/10.5194/tc-7-1035-2013, 2013.

690 Kurtz, N. T., Galin, N., and Studinger, M.: An improved CryoSat-2 sea ice freeboard retrieval algorithm through the use of waveform fitting, The Cryosphere, 8, 1217–1237, https://doi.org/10.5194/tc-8-1217-2014, 2014.

Kwok, R. and Cunningham, G. F.: Variability of Arctic sea ice thickness and volume from CryoSat-2, Phil. Trans. R. Soc. A, 373, 20140 157, https://doi.org/10.1098/rsta.2014.0157, 2015.

Kwok, R. and Rothrock, D.: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958-2008, Geophys. Res. Lett., 36, 695 L15 501, https://doi.org/10.1029/2009GL039035, 2009.

Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003-2008, J. Geophys. Res., 114, C07 005, https://doi.org/10.1029/2009JC005312, 2009.

Landy, J. C., Petty, A. A., Tsamados, M., and Stroeve, J. C.: Sea ice roughness overlooked as a key source of uncertainty in CryoSat-2 ice freeboard retrievals, J. G. R. Oceans, 125, e2019JC015 820, https://doi.org/10.1029/2019JC015820, 2020.

700 Lavergne, T., Sorensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjaer, G., Eastwood, S., Gabarro, C., Heygster, G., Killie, M. A., Kreiner, M. B., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L. T.: Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records, The Cryosphere, 13, 49–78, https://doi.org/10.5194/tc-13-49-2019, 2019.

Laxon, S., Peacock, N., and Smith, D.: High interannual variability of sea ice thickness in the Arctic region, Nature, 425, 947–950, https://doi.org/10.1038/nature02050, 2003.

705 Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., Kwok, R., Schweiger, A., Zhang, J., Haas, C., Hendricks, S., Krishfield, R., Kurtz, N., Farrell, S., and Davidson, M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, Geophys. Res. Lett., 40, 732–737, https://doi.org/10.1002/grl.50193, 2013.

Lemieux, J.-F., Beaudoin, C., Dupont, F., Roy, F., Smith, G. C., Shlyaeva, A., Buehner, M., Caya, A., Chen, J., Carrieres, T., Pogson, L., DeRepentigny, P., Plante, A., Pestieau, P., Pellerin, P., Ritchie, H., Garric, G., and Ferry, N.: The Regional Ice Prediction System (RIPS): 710 Verification of forecast sea ice concentration, Q.J.R.M.S., 142, 632–643, https://doi.org/10.1002/qj.2526, 2016.

Liang, X., Zhao, F., Li, C., Zhang, L., and Li, B.: Evaluation of ArcIOPS sea ice forecasting products during the ninth CHINARE-Arctic in summer 2018, Advances in Polar Science, 31(1), 14–25, https://doi.org/10.13679/j.advps.2019.0019, 2020.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, https://doi.org/10.5194/tc-9-269-2015, 2015.

715 MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A., Comer, R. E., Camp, J., Xavier, P., and Madec, G.: Global Seasonal Forecast System version 5 (GloSea5): A high resolution seasonal forecast system, Q. J. R. M. S., 141, 1072–1084, https://doi.org/10.1002/qj.2396, 2014.

Madec, G.: NEMO Ocean Engine, Tech. Rep. 27, Pole de modelisation de l'Institut Pierre-Simon Laplace, http://hdl.handle.net/2122/13309, 2017.

720 Mallett, R. D. C., Stroeve, J. C., Tsamados, M., Landy, J. C., Willatt, R., Nandan, V., and Liston, G. E.: Faster decline and higher variability in the sea ice thickness of the marginal Arctic seas when accounting for dynamic snow cover, The Cryosphere, 15, 2429–2450, https://doi.org/10.5194/tc-15-2429-2021, 2021.

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.

725   Meier, W. N., Hovelsrud, G. K., van Oort, B. E. H., Key, J. R., Kovacs, K. M., Michel, C., Haas, C., Granskog, M. A., Gerland, S., Perovich, D. K., Makshtas, A., and Reist, J. D.: Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity, Rev. Geophys., 52, 185–217, https://doi.org/10.1002/2013RG000431, 2014.

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelbert, M., Ottersen, G., Pritchard, H., and Schuur, E.: Polar Regions, in: IPCC Special Report on the Ocean and Cryosphere in a

730   Changing Climate, edited by Pörtner, H.-O., Roberts, D., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N., In press, 2019.

Mirouze, I., Blockley, E. W., Lea, D. J., Martin, M. J., and Bell, M. J.: A multiple length scale correlation operator for ocean data assimilation, Tellus A: Dynamic Meteorology and Oceanography, 68, https://doi.org/10.3402/tellusa.v68.29744, 2016.

Mu, L., Yang, Q., Losch, M., Losa, S. N., Ricker, R., Nerger, L., and Liang, X.: Improving sea ice thickness estimates by assimilating

735   CryoSat-2 and SMOS sea ice thickness data simultaneously, Q. J. R. M. S., 144, 529–538, https://doi.org/10.1002/qj.3225, 2018.

Nandan, V., Geldsetzer, T., Yackel, J., Mahmud, M., Scharien, R., Howell, S., King, J., Ricker, R., and Else, B.: Effect of snow salinity on CryoSat-2 Arctic first-year sea ice freeboard measurements, Geophys. Res. Lett., 44, 10 419–10 426, https://doi.org/10.1002/2017GL074506, 2017.

Parrish, D. and Derber, J. C.: The National Meteorological Center's spectral statistical interpolation analysis system, Mon. Weather Rev.,

740   120, 1747–1763, 1992.

Pfaffling, A., Haas, C., and Reid, J. E.: A direct helicopter EM sea ice thickness inversion, assessed with synthetic and field data, Geophysics, 72, F127–F137, 2007.

Polavarapu, S., Ren, S., Rochon, Y., Sankey, D., Ek, N., Koshyk, J., and Tarasick, D.: Data assimilation with the Canadian middle atmosphere model, Atmosphere-Ocean, 43, 77–100, https://doi.org/10.3137/ao.430105, 2005.

745   Posey, P. G., Metzger, E. J., Wallcraft, A. J., Hebert, D. A., Allard, R. A., Smedstad, O. M., Phelps, M. W., Fetterer, F., Stewart, J. S., Meier, W. N., and Helfrich, S. R.: Improving Arctic sea ice edge forecasts by assimilating high horizontal resolution sea ice concentration data into the US Navy's ice forecast systems, The Cryosphere, 9, 1735–1745, https://doi.org/10.5194/tc-9-1735-2015, 2015.

Ricker, R., Hendricks, S., Helm, V., Skourup, H., and Davidson, M.: Sensitivity of CryoSat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, The Cryosphere, 8, 1607–1622, https://doi.org/10.5194/tc-8-1607-2014, 2014.

750   Ricker, R., Hendricks, S., Perovich, D. K., Helm, V., and Gerdes, R.: Impact of snow accumulation on CryoSat-2 range retrievals over Arctic sea ice: An observational approach with buoy data, Geophys. Res. Lett., 42, 4447–4455, https://doi.org/10.1002/2015GL064081, 2015.

Ricker, R., Hendricks, S., and Beckers, J. F.: The impact of geophysical corrections on sea-ice freeboard retrieved from satellite altimetry, Remote Sensing, 8, 317, https://doi.org/doi.org/10.3390/rs8040317, 2016.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., and Haas, C.: A weekly Arctic sea-ice thickness data record from merged

755   CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607–1623, https://doi.org/10.5194/tc-11-1607-2017, 2017.

Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G. L., West, A. E., and Schroeder, D.: The sea ice model component of HadGEM3-GC3.1, Geosci. Model Dev., 11, 713–723, https://doi.org//10.5194/gmd-11-713-2018, 2018.

Rothrock, D. A., Zhang, J., and Yu, Y.: The Arctic ice thickness anomaly of the 1990s: A consistent view from observations and models, J. Geophys. Res., 108, 3083, https://doi.org/10.1029/2001JC001208, 2003.

760   Schweiger, A. J.: Unified Sea Ice Thickness Climate Data Record, digital media, http://psc.apl.uw.edu/sea_ice_cdr, 2017.

Screen, J. A.: Simulated atmospheric response to regional and pan-Arctic sea ice loss, J. Clim., 30, 3945–3962, https://doi.org/10.1175/JCLI-D-16-0197.1, 2017.

Smith, G. C., Roy, F., Reszka, M., Surcel Colan, D., He, Z., Deacu, D., Belanger, J.-M., Skachko, S., Liu, Y., Dupont, F., Lemieux, J.-F., Beaudoin, C., Tranchant, B., Drevillon, M., Garric, G., Testut, C.-E., Lellouche, J.-M., Pellerin, P., Ritchie, H., Lu, Y., Davidson, F., Buehner, M., Caya, A., and Lajoieb, M.: Sea ice forecast verification in the Canadian Global Ice Ocean Prediction System, Q. J. R. M. S., 142, 659–671, https://doi.org/10.1002/qj.2555, 2015.

Smith, L. C. and Stephenson, S. R.: New Trans-Arctic shipping routes navigable by midcentury, P. Natl. Acad. Sci. USA, 110, E1191–E1195, https://doi.org/10.1073/pnas.1214212110, 2013.

Stonebridge, G., Scott, K. A., and Buehner, M.: Impacts on sea ice analyses from the assumption of uncorrelated ice thickness observation errors: Experiments using a 1D toy model, Tellus A: Dynamic Meteorology and Oceanography, 70:1, 1445 379, https://doi.org/10.1080/16000870.2018.1445379, 2018.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., Kuhlbrodt, T., Rae, J. G. L., and Sinha, B.: UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions, Geosci. Model Dev., 11, 3187–3213, https://doi.org/10.5194/gmd-11-3187-2018, 2018.

Stroeve, J. and Notz, D.: Changing state of Arctic sea ice across all seasons, Environ. Res. Lett., 13, https://doi.org/10.1088/1748-9326/aade56, 2018.

Tian-Kunze, X., Kaleschke, L., Maass, N., Makynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: Algorithm baseline, product specifications and initial verification, The Cryosphere, 8, 997–1018, https://doi.org/10.5194/tc-8-997-2014, 2014.

Tilling, R., Ridout, A., Shepherd, A., and Wingham, D.: Increased Arctic sea ice volume after anomalously low melting in 2013, Nat. Geosci., 8, 643–648, https://doi.org/10.1038/NGEO2489, 2015.

Tilling, R. L., Ridout, A., and Shepherd, A.: Near-real-time Arctic sea ice thickness and volume from CryoSat-2, The Cryosphere, 10, 2003–2012, https://doi.org/10.5194/tc-10-2003-2016, 2016.

Tilling, R. L., Ridout, A., and Shepherd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data, Adv. Space Res., 62, 1203–1225, https://doi.org/doi.org/10.1016/j.asr.2017.10.051, 2018.

Tonboe, R., Lavelle, J., Pfeiffer, R.-H., and Howe, E.: Product user manual for OSI SAF global sea ice concentration product OSI-401-b, Product User Manual SAF/OSI/CDOP3/DMI_MET/TEC/MA/204, Danish Meteorological Institute, 2017.

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R.: Snow depth on Arctic sea ice, J. Climate, 12, 1814–1829, 1999.

Waters, J., Lea, D., Martin, M. J., Mirouze, I., Weaver, A., and While, J.: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, Q. J. R. M. S., 141, 333–349, 2015.

Wei, T., Yan, Q., Qi, W., Ding, M., and Wang, C.: Projections of Arctic sea ice conditions and shipping routes in the twenty-first century using CMIP6 forcing scenarios, Environ. Res. Lett., 15, https://doi.org/10.1088/1748-9326/abb2c8, 2020.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., de Chateau-Thierry, P., Laxon, S. W., Mallow, U., Phalippou, C. M. L., Ratier, G., Rey, L., Rostan, F., Viau, P., and Wallis, D. W.: CryoSat: A mission to determine the fluctuations in Earth's land and marine ice fields, Adv. Space Res., 37, 841–871, https://doi.org/10.1016/j.asr.2005.07.027, 2006.

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, https://doi.org/10.5194/tc-12-3671-2018, 2018.

Yang, Q., Losa, S. N., Losch, M., Tian-Kunze, X., Nerger, L., Liu, J., Kaleschke, L., and Zhang, Z.: Assimilating SMOS sea ice thickness into a coupled ice-ocean model using a local SEIK filter, J. Geophys. Res., 119(10), 6680–6692, https://doi.org/10.1002/2014JC009963, 2014.

Yang, Q., Mu, L., Wu, X., Liu, J., Zheng, F., Zhang, J., and Li, C.: Improving Arctic sea ice seasonal outlook by ensemble prediction using an ice-ocean model, Atmos. Res., 227, 14–23, https://doi.org/10.1016/j.atmosres.2019.04.021, 2019.

Zeng, Q., Lu, T., Lin, K.-C., Yuen, K. F., and Li, K. X.: The competitiveness of Arctic shipping over Suez Canal and China-Europe railway, Transport Policy, 86, 34–43, https://doi.org/10.1016/j.tranpol.2019.11.005., 2020.