Impact of assimilating a merged sea ice thickness from CryoSat-2 and SMOS in the Arctic reanalysis

Jiping Xie[1], François Counillon[1, 2], and Laurent Bertino[1, 2]

1. *Nansen Environmental and Remote Sensing Center, Bergen N5006, Norway*
2. *Bjerknes Center for Climate Research, Bergen, Norway*

*Corrsponding author: Jiping Xie, E-mail: jiping.xie@nersc.no

Abstract

Accurate forecast of Sea Ice Thickness (SIT) represents a major challenge for Arctic forecasting systems. The new CS2SMOS SIT measurements merges measurements from the CryoSat-2 and SMOS satellites and are available weekly during the winter months since October 2010. The impact of assimilating CS2SMOS is tested for the TOPAZ4 system - the Arctic component of the Copernicus Marine Environment Monitoring Service (CMEMS). TOPAZ4 currently assimilates a large set of ocean and sea ice observations with the Deterministic Ensemble Kalman Filter (DEnKF).

Two parallel reanalyses are conducted without (Official run) and with (Test run) assimilation of the previously weekly CS2SMOS for the period from 19th March 2014 to 31st March 2015. The raw observation error is underestimated. An additional term was added to compensate for the underestimation, but it was found a posteriori too large in our analysis. The SIT bias (too thin) is reduced from 16 cm to 5 cm and the RMSD decreases from 53 cm to 38 cm (reduction by 28%) when compared to the simultaneous SIT from CS2SMOS. When compared to independent SIT observations, the errors are reduced by 24% against the Ice Mass Balance (IMB) buoy 2013F and by 12.5% against SIT data from the IceBridge campaigns. When compared to the satellite ice drift product, the RMSDs around the North pole are reduced by about 8-9% in December 2014 and February 2015 relative to that in the Official. There is good improvement for the sea ice volume that extends outside of the assimilation period. Finally, using the Degrees of Freedom for Signal (DFS), we find that CS2SMOS is the main source of observations in the central Arctic and in the Kara Sea. These results suggest that C2SMOS observations should be included in Arctic reanalyses in order to improve the ice thickness and the ice drift.

Keywords: Sea ice thickness; Arctic reanalysis; CS2SMOS; EnKF; Innovation; Impact evaluation;

1. Introduction

Sea ice plays an important role in the Arctic climate system because it prevents the rapid exchange of heat flux between ocean and atmosphere. A decline and a thinning of the sea ice cover has occurred in the past decades (e.g. Johannessen et al., 1999; Comiso et al., 2008; Stroeve et al., 2012). It is expected that this change will have significant impacts on the Arctic Ocean Circulation (e.g. Levermann et al., 2007; Budikova, 2009; Kinnard et al., 2011) and on the future human living environment (Schofield et al., 2011; Bathiany et al., 2016). The interpretation of such changes is severely hampered by the sparseness of the observations and the use of reanalyses that can provide continuous spatio-temporal reconstruction by assimilating existing observations into dynamical models has become increasingly popular tools.

Satellite observation for sea ice concentration (SIC) is available since the 1980s, and has allowed an accurate monitoring of sea ice extent (SIE) during that period. Data assimilation of SIC has been used to improve the evolutions about the sea ice edge (Lisæter et al., 2003; Stark et al., 2008; Posey et al., 2015), but large uncertainty (e.g., Uotila et al, 2018) remains in the estimation of sea ice volume as observations of sea ice thickness (SIT) are very sparse. In addition, recent studies (Day et al. 2014; Guemas et al., 2014; Melia et al. 2015) have shown that SIT anomalies play an important role for the Arctic predictability up to seasonal time scale.

Up to the 1990s, the availability of SIT measurement was limited to sparse in situ measurements and submarines data. With the emergence of satellite, continuous estimates of SIT on basin scale have been achieved using radar and laser altimeters from the satellites: European Remote Sensing (ERS), Envisat and the NASA Ice, Cloud and land Elevation Satellite (ICESat). These were used to document the rapid thinning of sea ice in Arctic (Laxon et al., 2003; Kwok and Rothrock, 2009).

CryoSat-2 launched in April 2010 has been the first satellite dedicated to measure with high accuracy the sea ice freeboard, from which the sea-ice thickness can be derived (Ricker et al., 2014; Tilling et al., 2016). The retrieved SIT still contains considerable uncertainty because of approximations made for example when estimating the snow depth (using climatology), snow penetration and sea ice density (Kern et al, 2015; Khvorostovsky and Rampal, 2016). These uncertainties are comparatively large for thin ice (<1 m). Satellite measurements derived from passive microwave radiometer have allowed retrieval of thin sea ice thickness (Martin et al., 2004; Heygster et al., 2009). The Soil Moisture and Ocean Salinity (SMOS) satellite, measures the brightness temperature in a L-Band microwave frequency (1.4 GHz) that can be used for estimating very thin sea ice thickness (Kaleschke et al., 2010; Tian-Kunze et al., 2014), typically bellow 0.5 m. Although the consistency between the SMOS and CryoSat-2 estimates is still poor (Wang et al., 2016), a recent initiative has combined the two data sets (e.g. Kaleschke et al., 2015; Ricker et al., 2017). A merged product of weekly SIT measurements in Arctic from the CryoSat-2 altimeter and SMOS radiometer (referred to as CS2SMOS) is now available online at http://www.meereisportal.de (Ricker et al., 2017). There is a need to test assimilation of this data set and assess its potential for reanalysis and operational forecasting.

In this study, the CS2SMOS will be assimilated into the TOPAZ4 forecast system, which is a coupled ocean-sea ice data assimilation system using the Deterministic Ensemble Kalman Filter (DEnKF; Sakov and Oke, 2008). The Ensemble Kalman Filter has previously been demonstrated for assimilation of SIT data (Lisæter et al., 2007) or freeboard data (Mathiot et al., 2012) or CS2SMOS data (Mu et al., 2018*)*. TOPAZ4 is the main Arctic Marine Forecasting system in the Copernicus Marine Environment Monitoring Services (CMEMS, http://marine.copernicus.eu). Every day, it provides a 10-day forecast of the ocean and biogeochemistry in the Arctic region through the CMEMS portal for the public. It also provides a long reanalysis from 1990 to present – currently 2016 - that is extended every year. By default, SIT products are not assimilated into the TOPAZ4 reanalysis. This reanalysis has been widely used and validated (Ferreira et al., 2015; Johannessen et al., 2014; Xie et al., 2017). Although the Arctic SIT distribution in TOPAZ4 shows some degree of spatial coherency with that of ICESat in spring and autumn of 2003-2008, it underestimates SIT (up to 1 m) north of Canadian Arctic Archipelago and Greenland and overestimates it by approximately 0.2 m in the Beaufort Sea (Xie et al., 2017). Even though the SIT from ICESat has been reported too thick by about 0.5 m (Lindsay and Schweiger, 2015), the SIT from TOPAZ4 undoubtedly has spatial biases. Similar biases for SIT have been reported for other Arctic coupled ocean-ice models (Stark et al., 2008; Johnson et al., 2012; Schweiger et al., 2012; Yang et al., 2014; Smith et al., 2015) and even reanalyses (Uotila et al, 2018). Xie et al. (2016) have tested assimilation of thin SIT (<0.4 m) from SMOS, and show that the assimilation slightly reduced SIT overestimation near the sea ice edge. The recent availability of the weekly SIT from CS2SMOS provides an opportunity for the TOPAZ4 to constrain the SIT error in the Arctic. This study aims at identifying a suitable practical implementation for assimilating C2SMOS data set and assess its usefulness for the Arctic reanalysis. Although it is expected that a better initialisation of SIT anomalies will enhance the predictability of the system, this is beyond the scope of this paper. A similar assessment over the same time frame has been carried out in the Arctic Cap Nowcast/Forecast System (ACNFS) by Allard et al. (2018) revealing significant improvements of bias and RMSD but little changes in ice velocity except in marginal seas. The proposed study in somewhat complementary to Allard et al. (2018) because TOPAZ4 prediction system uses comparatively a more rudimentary sea ice thermodynamics (no explicit ice thickness distribution) but a more advanced ensemble-based data assimilation method – TOPAZ4 uses strongly coupled data assimilation of ocean and sea ice - Meaning that sea ice observation will impact also the ocean and vice versa (Penny et al., 2017; Kimmritz et al., 2018) - with a flow dependent assimilation method.

Section 2 describes the TOPAZ4 system: namely the coupled ocean and sea ice model, the implementation of EnKF and the observations used for data assimilation and validation. In section 3, we carry an Observing System Experiment (OSE) comparing the two reanalyses: one using the standard observation types used in operational setting and another assimilating the CS2SMOS in addition. Then the performance of the two runs against assimilated and no-assimilated measurements are presented. Section 4 presents the impacts of assimilating the CS2SMOS on sea ice drift and the integrated quantities for sea ice, and quantifies its relative impacts compared to the other observation variables. A summary and discussion are provided in the last Section.

2. TOPAZ4 system descriptions and observations

2.1 The coupled ocean and sea-ice model

TOPAZ4 is a forecasting ocean and sea-ice system developed for the Arctic, having been operational since early of the 2000s (Bertino and Lisæter, 2008). It uses the Hybrid Coordinate Ocean Model (HYCOM: version 2.2) developed initially at University of Miami, which has been successfully applied in global and regional oceans (Chassignet et al., 2003; Counillon and Bertino, 2009; Metzger et al 2014; Xie et al., 2018). The model grids are constructed using conformal mapping (Bentsen et al., 1999; Bertino and Lisæter, 2008) with a 12-16 km resolution shown in Fig. 1 (left). The model uses 28 hybrid layers with reference potential densities selected specifically for the North Atlantic and the Arctic regions (Sakov et al. 2012). A barotropic inflow of Pacific Water is imposed through the Bering Strait, which is balanced by outflowing through the southern model boundary. It has an averaged transport of 0.8 Sv, and seasonally varies with a minimum (0.4 Sv) in January and a maximum (1.3 Sv) in June consistent with the observations proposed in Woodgate et al. (2005). The model account for river discharge for which the seasonal climatology is estimated by feeding the run off from ERA-interim (Dee et al., 2011) to the Total Runoff Integrating Pathways (TRIP, Oki and Sud, 1998) over the period 1989–2009.

A simple sea ice model using a one thickness category has been integrated at NERSC into HYCOM. As such, the sea ice and the ocean are coupled every 3 hours and exchange momentum, salt and heat on the ocean's Arakawa C-grid. The sea ice thermodynamics described in Drange and Simonsen (1996) treat precipitations on ice as snow whenever surface air temperature is below zero. The ice dynamics uses the elastic-viscous-plastic rheology (Hunke and Dukowicz, 1997) with the modification suggested by Bouillon et al. (2013). There is a 0.1 m limit in the model for the minimum thickness of both new ice and melting ice.

2.2 Implementation of the EnKF in the TOPAZ4 system

The TOPAZ4 system uses a deterministic Ensemble Kalman Filter (DEnKF, Sakov and Oke, 2008), which solves the analysis without the need to perturb the observations and is regarded as a square-root filter implementation of EnKF. In the DEnKF, if the model state is represented by $\mathbf{x}$, the ensemble mean is updated by equation:

$$\bar{\mathbf{x}}^{\mathrm{a}} = \bar{\mathbf{x}}^{\mathrm{f}} + \mathbf{K}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^{\mathrm{f}}), \tag{1}$$

where the superscripts “f” and “a” respectively refer to the forecast and the analysis. Following Xie et al. (2017), the model state vector **x** contains 3-dimensional ocean variables in the native hybrid coordinates (u- and v-components of the current velocities, temperature, salinity and model layer thickness), the 2-dimentional ocean variables (u- and v-components of the barotropic velocities, barotropic pressure, and mixed layer depth) and two sea ice variables ice concentration and ice thickness. The assimilated observations are represented by the vector of **y** without perturbation, and the observation operator **H** projects the model variables on the observation space. The misfit between the model and the observation - the bracket term in Eq. (1), is named as innovation. The Kalman gain **K** is calculated by:

$$\mathbf{K} = \mathbf{P^f}\mathbf{H}^{\mathrm{T}}[\mathbf{H}\mathbf{P^f}\mathbf{H}^{\mathrm{T}} + \mathbf{R}]^{-\mathbf{1}} \quad (2).$$

Where $\mathbf{P^f}$ is the matrix of background error covariance, **R** is the matrix of observation error covariance, and the superscript “T” denotes a matrix transpose. The background error covariance is approximated from the ensemble anomalies **A** (where $\mathbf{A} = \mathbf{X} - \bar{\mathbf{x}}\mathrm{I}_N$, $\mathrm{I}_N = [1, \ldots, 1]$, N being the ensemble size) as follows $\mathbf{P} = \frac{\mathbf{A}\mathbf{A}^{\mathrm{T}}}{N-1}$. Here, $\mathbf{X}$ denotes the ensemble of model states, the observation errors are assumed being uncorrelated (i.e. the matrix **R** is diagonal). While this assumption is not always corrected for some types of observations, it requires the sufficient knowledge about the covariance structure for the observation errors if considering the correlations in **R**. Otherwise, an approximation of the correlated observation error can yield a poor analysis so a diagonal approximation combined with an inflation of the observation error is a reasonable approximation (Stonebridge 2018).

To ensure that the sampling error remains small, a localization is used (local framework analysis) with a radius of 300 km and Gaussian tapering. More details about the practical implementation of the model and perturbations can be found in Sakov et al. (2012). The model errors include joint perturbations of winds, heat fluxes as originally recommended by Lisæter et al. (2007). The precipitation perturbation was increased from 30% to 100%, following a log-normal probability distribution of errors (Finck et al. 2013).

2.3 Observations for assimilation and validation

The following observations are assimilated sequentially every week in the TOPAZ4 system (Xie et al. 2017): along-track Sea Level Anomaly; in situ profiles of temperature and salinity; gridded Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) SST; Ocean and Sea Ice Satellite Application Facility (OSI-SAF) sea ice concentration and sea ice drift from satellite observation (Lavergne et al., 2010). All measurements are retrieved from http://marine.copernicus.eu, and are quality controlled and superobed – i.e. all observations falling within the same grid cell are averaged and the observation uncertainty is reduced accordingly (Sakov et al., 2012). For SST and ice concentration, we only retain the analysis at the last day of the assimilation cycle. Similarly, the sea ice drifts during the last 2 days of the assimilation cycle are assimilated from OSI-SAF.

The weekly SITs of CS2SMOS were retrieved from http://data.meereisportal.de/maps/cs2smos/version3.0/n for the period from March 2014 to March 2015. This product is gridded with a resolution of approximate 25 km. The provider uses optimal interpolation to blend the measurements of CryoSat-2 and SMOS based on the best estimate, their uncertainties and their spatial covariance. An estimate of the observation error is provided with the data set but it only accounts for the errors related to the merging and interpolation (Ricker et al., 2017). As such, we expect that this observation error is only accounting for a part of the real error and misses both the sensor errors and the model-related representation errors. In particular the mapping error is based on a no-bias assumption and does not account for inconsistencies between the two satellites, like those reported by Ricker et al. (2017). With an EnKF assimilation system, underestimating the observation error leads to an underestimation of the ensemble spread and makes the system suboptimal. In the worst case, the ensemble spread collapses and the system diverges. Underestimating the errors of one data type also lessens the impact of the other assimilated observations since they compete for the control of a finite number of degrees of freedom. This issue will be addressed in Section 4.3. On the other hand, Oke and Sakov (2008) showed that the performance of the EnKF does not degrade much when observation error is overestimated. It is therefore necessary to increase the observation error to a level at least as high as the optimal value for the performance of the filter (Desroziers et al., 2005; Karspeck, 2016).

In order to estimate the representation error for the SIT observation, we have performed a preliminary sensitivity assimilation experiment for November 2014. We used the diagnostics by Desroziers et al. (2005) as an indicative lower limit for the observation error in the TOPAZ4 system based on the misfits to the CS2SMOS data. Desroziers et al. (2005) estimate the optimal observation error as the following matrix:

$$\tilde{\sigma}^{o}_{SIT} = \sqrt{\frac{1}{\mathbf{p}}\sum_{\mathbf{j=1}}^{\mathbf{p}}(\mathbf{y_j} - \mathbf{H}\overline{\mathbf{x}}^{\mathbf{a}})(\mathbf{y_j} - \mathbf{H}\overline{\mathbf{x}}^{\mathbf{f}})} \quad (3)$$

where p is number of data assimilation steps in the sensitivity run (here 4), and $\mathbf{y}_j$ represents the observed SIT from CS2SMOS at the jth assimilation time. Here, the terms $\bar{\mathbf{x}}^a$ and $\bar{\mathbf{x}}^f$ represent the ensemble mean of the analysis and forecast states. In Fig. 2, the diagnosed observation errors from Desroziers et al. (2005) are larger than the mapping error included in CS2SMOS, but still do not account for biases in the CryoSAT2 and SMOS observations. The CS2SMOS mapping error is particularly low for sea ice below 0.5 m: about 4 times lower than the uncertainties obtained by error propagation in the SMOS processing chain (used in Xie et al. 2016), which would make the assimilation of SMOS SIT too strong. The Desroziers diagnosed errors gradually increase with ice thickness, although they vary unrealistically for SITs above 3 m, possibly due to low counts of either modelled or observed ice thickness in certain thickness ranges. In view of the above considerations, we have added a cautious correction term to the CS2SMOS mapping error estimate, which simply increases linearly with the observed SIT.

$$\boldsymbol{\varepsilon}_{\mathbf{Offset}} = \min(0.5, 0.1 + 0.15 * \mathbf{d}_{\mathbf{SIT}}) \quad (4),$$

where $\mathbf{d}_{SIT}$ is the observed sea ice thickness. At low SIT, the resulting values are slightly higher than those used in Xie et al. (2016) and comparable to the Desroziers diagnostics. At SITs of 1.5 m, for which SMOS and CS2SMOS overlap, the added correction is comparable to reported differences between the two satellites: about 20 cm in the Beaufort Sea and 1 meter in the Barents Sea, see Table 3 in Ricker et al. (2017). Tilling et al., (2018) show that the standard deviations between the CryoSat-2 and independent measurements are between 30 and 70 cm depending of the source of observation and increase with ice thickness (their Figure 16). It should be noted however that the processing of CryoSat2 data differs in CPOM and AWI's algorithms. The total observation error including the added term is shown with blue-squared line in Fig. 2. In the following, we will only use the corrected observation error for the CS2SMOS SIT.

3. Observing system experiment runs and validations

3.1 Experiment and independent observations for validation

A parallel OSE is conducted from 19th March 2014 until end of March 2015. The two assimilation runs cover two special time periods: at the onset of ice melting in March-April 2014 following by a free data period of CS2MSOS, and a whole cold season from October 2014 to March 2015. Both runs are forced by atmosphere forcing from ERA-Interim. The control run named the **Official run** uses the standard observational network in the TOPAZ4 system (Xie et al. 2017), which assimilates on a weekly cycle the SLA, SST, in situ profiles of temperature and salinity, SIC and sea ice drift (SID) data. Another assimilation run named the **Test run** involves the SIT from CS2SMOS as a type of additional observation into the system.

The CS2SMOS ice thickness data are weekly averages and provided on a grid with a 25 km resolution. We discard the SIT closer than 30 km from the coast to account for different coastlines between the model and observations. The innovation of SIT in Eq. (1) is calculated in terms of sea ice volume:

$$\Delta\mathbf{SIT} = \mathbf{d}_{\mathrm{SIT}} - \mathbf{H}(\bar{\mathbf{h}}_{\mathrm{m}} \times \bar{\mathbf{f}}_{\mathrm{m}}), \qquad (5)$$

where $\mathbf{d}_{\mathrm{SIT}}$ is the observed SIT from CS2SMOS as in Eq. (4), $\bar{\mathbf{f}}_{\mathrm{m}}$ is the ensemble mean SIC, and $\bar{\mathbf{h}}_{\mathrm{m}}$ is the ensemble mean ice thickness within the grid cell. We assume the observation error to be uncorrelated (**R** in Eq. (2) is diagonal). While it is clear that this approximation is incorrect, it was shown in Stonebridge et al. (2018) that when the structure of the correlation is unknown, it was best to assume **R** diagonal and to tune the inflation. Although the minimal thickness in the model is set to 0.1 m, the ensemble mean from 100 model members can be as thin as 1 mm, so that we reject the observed SIT for

CS2SMOS only if equal to 0. Every week, the SITs from CS2SMOS are considered to be at the analysis time, neglecting the time delay. However, the associated errors due to the sea ice motions or thermodynamic growth/melt of sea ice remain small within one week compared to the large SIT biases targeted in the present exercise.

In the following, we will investigate the misfits of the forecasted model states by evaluating the bias and the root mean square difference (RMSD):

$$\mathrm{Bias} = \frac{\mathbf{1}}{\mathbf{L}}\sum_{\mathbf{i=1}}^{\mathbf{L}}(\mathbf{H_i}\bar{\mathbf{x}}_{\mathbf{i}}^{\mathbf{f}} - \mathbf{y_i}) \tag{6}$$

$$\mathrm{RMSD} = \sqrt{\frac{\mathbf{1}}{\mathbf{L}}\sum_{\mathbf{i=1}}^{\mathbf{L}}(\mathbf{H_i}\bar{\mathbf{x}}_{\mathbf{i}}^{\mathbf{f}} - \mathbf{y_i})^{\mathbf{2}}} \tag{7}.$$

Where L is the total number of assimilation cycle over the study period, $\bar{\mathbf{x}}_i^{\mathrm{f}}$ is the mean of the model state at the *i*th time, which is comparable to the observations $\mathbf{y}_i$.

Three types of independent observations for SIT are involved for validation. First, the SIT measurements from drifting Ice Mass Balance (IMB: http://imb-crrel-dartmouth.org/imb.crrel/buoysum.htm) buoys (Perovich and Richter-Menge, 2006). Four IMB buoys (2013F, 2014B, 2014C, and 2014F) are available during the experimental time period and their trajectories are shown in Fig.1 (left). Second, three upward looking sonar (ULS) buoys funded by the Beaufort Gyre Exploration Project (BGEP, see http://www.whoi.edu/beaufortgyre) have been moored in the Beaufort Sea. Their locations are shown with the red squares in Fig. 1 (left). They estimate the sea ice drafts since October 2014. Third, the NASA IceBridge Sea Ice Thickness Quick Look data (https://nsidc.org/data/icebridge) collected in aerial campaigns estimates the sea ice thickness in spring (Kurtz et al., 2013) with a better spatial coverage. The locations of the quality-controlled observations of SIT from IceBridge for March and April of 2014 and 2015, are shown with the yellow squares in Fig. 1 (left).

3.2 Validation against CS2SMOS and innovation diagnostics

The first assimilation time is on the 19th March 2014 and the last is on the 25th March 2015. The monthly SITs for the two OSE runs are compared to CS2SMOS in Fig. 3. The SITs in April 2014 are presented for comparison in the upper panels of Fig. 3. In the Official run, the thick sea ice to the north of the CAA is underestimated but thickens slightly in the Test run: the 3 m SIT isoline covers a wider area, in better agreement with the observations. The areas of thinner sea ice north of the Barents Sea, west of the Kara Sea, and the coast of the Beaufort Sea, which were too thick in the Official run, have all been improved also shown by reduced area delimited by the isolines of 1 m or 2 m SIT in the Test run.

After summer of 2014, measurements of SIT from CS2SMOS restart at the end of October. Results are presented for November 2014 in Fig. 3: the thick sea ice in the central Arctic has been further improved in the Test run. The thickest sea ice (> 3 m) is located near the northern coast of Canada instead of north of Greenland in the Official run. The averaged SIT in the Test run around the North pole (>80°N), is increased from 1.3 m in the Official run to 1.6 m, which is closer to CS2SMOS by 43%. In the marginal zones of the East Siberian Sea, the Laptev Sea, and the Kara Sea, the SITs in the Official run is too thin, but is thickened in the Test run. Improvements in marginal seas are due to the contribution of SMOS, while improvements in the ice pack are mainly due to CryoSat-2.

In the last month of the experimental period (March 2015), the thick sea ice pattern in the Test run, shown as the 2 m isoline, is more similar to that of CS2SMOS. The maximal SIT denoted by the 4 m isoline is located north of the CAA in the Test run and in CS2SMOS, while the Official run spreads it out from the northern coast of Canada to north of Greenland. In addition, the SIT north of the Fram Strait is thicker than in the Official run. The SIT is similarly improved near the coast of the Beaufort Sea and to the northwest of Svalbard. As expected with data assimilation, the Test run improves clearly the agreement with the assimilated product. Those improvements are largest in the ice pack and in the marginal Seas, where the model has a considerable deviation compared to the CS2SMOS SITs. On the contrary, the thickness near the sea ice edge is not strongly impacted by the assimilation.

The continuous agreement is confirmed quantitatively: misfits of weekly SIT from the two runs are compared with the corresponding CS2SMOS observations. Time series of bias and RMSD (calculated weekly as in Eq. (6-7) are shown in the top panel of Fig. 4. At the beginning of the period, the SIT

RMSD in the Test run decreases quickly from 0.6 m to 0.4 m before the observations are interrupted. The bias of the two runs are similarly reduced. After the observations resume in the end of October 2014, the SIT RMSD is comparable between the two runs but the bias is slightly lower in the Test run. There is large spike in the bias and RMSD for both systems that relates to an inaccuracy of the CS2SMOS observations (see Section 4.2). After the spike, the RMSD and bias in the Test run are lower than in the Official run. The bias in the Test run converges to 0 and fluctuates around that level but this is likely not the influence from the assimilation as the bias in the Official run also converges to 0 during that time. This is rather due to the compensation of seasonal and regional errors. On average, the bias of SIT (too thin) is decreased from 15 cm to 5 cm by the assimilation of CS2SMOS. The RMSD of SIT is 38 cm in the Test run, which corresponds to a reduction of 28.3% relative to the error in the Official run.

The innovation statistics taken at each assimilation time are used to evaluate how well our data assimilation system is calibrated. In the reliability budget of Rodwell et al. (2016), the total uncertainty of an ensemble data assimilation system is calculated as follow:

$$\sigma_{diag} = \sqrt{Bias^2 + \sigma_{en}^2 + \sigma_o^2} \quad , \qquad (8).$$

where the *Bias* term – i.e. the innovation mean (shown as blue-circled lines) - is calculated as in Eq. (6) at a given assimilation time step, and σ_{en} and σ_o represent respectively the ensemble spread and the standard deviation of the observation errors at the same assimilation time. If the data assimilation system is reliable, the diagnosed total uncertainty should be close to the RMSD, formulated in Eq. (7). In Fig. 4 we can see that the pink and red lines are evolving reasonably in phase but that the diagnosed error σ_{diag} is much larger than the RMSD, meaning that our system is overdispersive. The error budget shows that the observation error (σ_{o}) is too large, suggesting that the offset term in Eq. (4) is overestimated, which we do not expect as a serious problem as explained above.

The innovation statistics for SIC are mostly identical in the two runs (not shown), the mean misfits for SIC vary around ±4% and are most of the time lower than 12%, which is consistent with the evaluation of the TOPAZ4 reanalysis in Xie et al. (2017). It is somewhat disappointing that improvements of ice thickness do not yield visible benefit to ice concentration, but on the other hand a degradation could also have been possible if the thermodynamical model had been over-tuned to an incorrect simulated thickness. It should also be noted that the innovation statistics of SST and SLA are also indiscernible in the two runs and not shown either.

3.3 Validation against independent SIT observations

3.3.1 Ice Mass Balance Buoys

Four IMB buoys are available as independent validation of the impact of the assimilation of CS2SMOS. The buoys are drifting in the Canada Basin (Fig. 1), and only one buoy (2013F) lasted during the whole experimental time period shown (upper panel of Fig. 5). This buoy depicts the seasonal variability of SIT: it reaches 1.5 m in spring 2014, decreases down to 1.0 m in September and rises again to 2 m in March 2015. The seasonal SIT cycle of the Official run shows excessive seasonal variability, with a thin bias in summer 2014 and a thick bias during the winters. In the Test run (shown as the red-dashed line) the seasonal cycle is dampened and more consistent with the observations. The bias is still quite large around March-April and that even at the end of the study period. It should be noted that the impact of CS2SMOS seems largest in summer, when no observations are available. This indicates the persistent effects of winter thickness to improve the predictability of the summer Arctic sea ice (as in Mathiot et al. (2012)). When CS2SMOS is assimilated again in the fall 2014, the Test run initially overestimates slightly the SIT measured at the buoy compared to that in the Official run but is slowly improving as data is assimilated. The time-averaged SIT RMSD for 2013F is reduced from 0.33 m in the Official run down to 0.25 m in the Test run, a reduction of 24.2%.

Two other buoys (2014B and 2014C) cover the early months of the experimental period. At the beginning, the two runs are biased with a too thick of 0.5 m and 0.2 m compared to 2014B and 2014C. For 2014B, there is a slight reduction of the error during the assimilation period that continue to reduce beyond the assimilation window as for 2013F. For 2014C although the error is reduced during the analysis period, the error increases beyond the analysis as the error in the official run reduces. For these three buoys the assimilation corrects the mean SIT values and the amplitude of the seasonal cycle but have little influence on the phase of the seasonal cycle.

The buoy 2014F covers the last 6 months of the experimental period. For that buoy, the assimilation seems to be increasing the error. Initially and as for 2013F at the same time, the initial value of SIT is too large in Test while it is quite reasonable in the Official run. For 2013F it was the consequence of curing the too low bias in September and having a too vigorous SIT increase November. At the start of assimilation, Test shows a clear – albeit too weak – decrease and a slower growth of the ice thickness compared to the Official Run. It should be noted that the SIT growth in 2014F is unlikely weak the area and very different from the buoy 2013F, with an increase from 1.5 m to only 1.6 m in the whole winter. However, the Test Run shows a pronounced decrease of SIT at the start of assimilation, and afterward shows a slower growth of the SIT compared to the Official Run.

3.3.2 The BGEP mooring buoys

In order to convert the sea ice draft measured by ULS from the BGEP buoys to SIT, we used the equation introduced in Tilling et al. (2018):

$$\mathbf{d}_{SIT} = \frac{d_i \rho_w - h_s \rho_s}{\rho_i} \tag{9}$$

where $\mathbf{d}_{SIT}$ is the sea ice thickness, d_i is sea ice draft, h_s is snow depth, ρ_i is sea ice density, ρ_s is snow density and ρ_w is seawater density. The three densities are constant of 900, 300, and 1000 kg/m^3 used as in the model. d_i is the sea ice draft measured by ULS at the fixed locations (see Fig. 1). The snow depth is estimated by the daily snow depths averaged of the two runs interpolated to the buoy locations.

The SIT time series of the measurement and of the two runs are shown on Fig. 9, from October 2014. The gray error bars depict the daily standard deviation. The data indicates a SIT increasing from around 0.5 m in October 2014 to close to 2 m in March 2015. The observed SIT at 14D shows a very large daily variability from end of October to November 2014, especially compared with that of 14A and 14B.

The weekly SIT from CS2SMOS matches well the data set with a RMSDs of 15, 19 and 39 cm during the 6 months, which is lower than in the two model runs. Still, the SIT from CS2SMOS overestimates SIT from October 2014 to middle

January 2015 compared to that of BGEP for buoy 14B, and between in Oct and Nov of 2014 for buoy of 14A. The SITs in the Official run are overestimated in all three locations. The SIT RMSDs are 41, 23 and 51 cm respectively compared to SIT measurement from BGEP buoys. The SITs in the Test run is closer to the observed mooring estimate, thanks to the data assimilation of the SIT from CS2SMOS. The SID RMSDs in the Test run are respective 25, 33 and 36 cm for Buoys 14A, B, D. Error is nicely reduced for 14A and 14D compared to the Official run but increased for 14B mostly caused by the initial mismatch between CS2SMOS and BGEP initially. Similarly to what was found to IMB measurements, it suggests that error of SIT in the Beaufort Sea is reduced by assimilation of CS2SMOS.

3.3.3 IceBridge Quick Look

Another independent observation of SIT with better spatial coverage is the SIT Quick Look data from airborne instruments during NASA's Operation IceBridge campaign (Kurtz et al., 2013). They are available via the National Snow and Ice Data Center (NSIDC), albeit for months of March and April only. Note that the airborne SITs have been reported to be slightly low-biased by about 5 cm compared to in situ measurements (King et al., 2015). Figure 7 shows all observed SITs (upper-left panel) from IceBridge, collected during March and April of 2014-2015. All observed SITs are located in the Canadian Basin and north of Greenland and covers most of the area where sea ice is thicker than 3 m. Sea ice with a thickness between 1~3 m is measured in the Beaufort Sea. The two simulated SITs in the two model runs show systematic differences of SIT (see upper-right panel of Fig. 7) - SIT in the Test has been thinned in the Beaufort Sea and thicken near the North pole. On average, the SIT in the Test run is increased by 0.1 m and by 0.27 m north of 80°N. Fig. 10b shows that the distributions of SITs at the location of the buoys (shown in right of Fig. 1) from the International Arctic Buoy Program (IABP) have been significantly adjusted between the two runs: The thick sea ice (>2.2 m) becomes more abundant in the Test run and the relatively thin sea ice (0.5-1.7 m) more abundant in the Official run. The averaged SIT thus increases from 1.52 m to 1.62 m in the Test run.

The SIT deviations of the two OSE runs compared to IceBridge data are presented in the bottom panels. The sea ice in the Official run is too thin north of the CAA and north of Greenland, with a deviation larger than 1.5 m. In the Beaufort Sea on the contrary, the model is too thick by 0.5 to 1 m. This bias is consistent with that reported in Xie et al. (2017), where the TOPAZ4 reanalysis (Official run) was compared to ICESat observation for the period of 2003-2008. In the Test run, the biases are slightly reduced by SIT assimilation, mainly in the Beaufort Sea and north of Greenland, but the reduction is smaller than the remaining error. On average, the SIT RMSD is 1.05 m, which corresponds to a reduction of 12.5% compared to that in the Official run.

The regression of the SIT observations from IceBridge to the two OSE runs is shown in Fig. 8. The Test run shows improved linear correlations to the observation. The offset at the origin is reduced (0.52 m instead of 0.93 m) and the slope is closer to 1 m than in the Official run. The linear correlation in the Test run is slightly increased as indicated with the correlation squared R^2. There is still a lot of spread that explains why the correlation is on the low side. However, the model still underestimates the thickest ice observed in IceBridge, with a bias as high as 2 m.

4. Impact of CS2SMOS in the data assimilation system

The above results and assimilation diagnostics confirm that the SIT misfits can be controlled - to some degree - by assimilation of the CS2SMOS data, without visible degradation of other assimilated variables. To better understand the advantages and the limits of assimilating the merged SIT product, we further evaluate the impact of CS2SMOS in the assimilation system: first the repercussions on other sea ice variables and integrated quantities, and then through a quantitative impact analysis of CS2SMOS relatively to other assimilated observation types.

4.1. Impact on the sea ice drift

The EnKF implemented in TOPAZ4 updates all the variables in the model state vector using flow-dependent multivariate covariances from the ensemble members (Eqs. 1 and 2). The direct assimilation update of ice drift is however short-lived: the ice drift vectors quickly readjust to wind forcing after assimilation, so the ice drift changes are mostly caused by dynamical readjustments, related to the updated ice thickness and ice concentrations. By the first order approximation of the two-dimentional momentum equation (e.g., Hibler 1986; Hunke and Dukowicz, 1997), the drift velocity of sea ice is mainly controlled by 1) the interactions of atmosphere-sea ice, 2) the interactions of ocean-sea ice and 3) the internal sea ice forces which can be represented by the stress tensor σ_i. The work of Olason and Notz (2014, thereafter called ON14) shows from observations that ice thickness is the main driver changes of ice drift in winter (December to March), while the concentration is the main driver in summer (June to November) and ice drift may increase independently from concentration of thickness in transition periods due to increasing fracturing.

Following the EVP rheology in Hibler (1979), the stress tensor σ_i is forced by a pressure term Q which takes a function of the sea ice thickness and concentration only.

$$\mathrm{Q} = \mathrm{P}^{*}\mathrm{d}_{\mathrm{SIT}}\exp(-\mathrm{C}_{0}(1-\mathrm{A}_{\mathrm{SIC}})), \tag{10}$$

Where C_0 and P^* are empirical constants, $\mathbf{d}_{SIT}$ is SIT, and A_{SIC} is sea ice concentration. ON14 thus show that this type of rheology is able to reproduce the changes of ice drift whenever they are related to changes of concentration and thickness, although not the changes during the transition periods. The sensitivity of ice drift to ice thickness can be directly adjusted by tuning the value of P* in Eq. (10) (see for example Docquier et al., 2017). In the TOPAZ4 model, the sea ice dynamics assume a viscous-plastic material with an adjustment mechanism at short timescales by elastic waves (called EVP, Hunke and Dukowicz, 1997). The ice thickness does as well have an influence on the ice concentrations in the summer due to melting, but this influence is limited in TOPAZ4 by the assimilation of ice concentrations. The winter months in the seasonal cycle (see Figure 6 in ON14) indicate that a 10% increase of ice thickness can reduce the ice drift by 9%. Areas of thinner ice are much more sensitive (see Figure 5 in ON14) and therefore the above numbers are subject to possible biases of ice thickness. The sensitivity on seasonal time scales may also differ from the sensitivity on a weekly time scale (that of the TOPAZ4 assimilation cycle).

The evaluation in Xie et al. (2017) shows the model drift of sea ice is overestimated by 2 km d^{-1} on average on the Arctic with an uncertainty of 5 km d^{-1}. The thickness of thick ice is also too thin, consistently with the too fast drift (Figures 14 and 17 in Xie et al., 2017). So, the assimilation of ice thickness is expected to improve the ice drift by dynamical model adjustment. Figure 9 shows monthly differences of the 2-day sea ice drift (SID) compared to the OSI-SAF estimates based on passive microwave data in April 2014, December 2014 and February 2015. The SID in the Official run is too fast in the central Arctic where the SIT was found too thin in Fig. 3. Despite of the relatively small assimilation impact of CS2SMOS on the SID, there are improvements across the Arctic in all winter months.

The RMSD of sea ice drift speed in two-days trajectories is reduced by about 0.1-0.2 km in April 2014 and February 2015 for the whole Arctic, which corresponds to a reduction of less than 5% of the RMSD. However, near the North Pole (north of 80°N), the reduction of drift RMSDs is more important, by about 0.4-0.5 km. In December 2014 and February 2015 it is about 8-9% of the error in the Official run. Near the North Pole the averaged SIT in March 2015 (Fig. 3) is about 10% thicker in the Test run than in the Official run. The impact is more important there than in the rest of the Arctic and well in line with the sensitivity found in ON14. Additionally, there is a small reduction of the fast SID bias but in the case of TOPAZ4, such biases are dependent on the tuning of the drag coefficients between sea ice and the air or the ocean, which has been optimized for the SIT distribution of the TOPAZ free run. The tuning of the drag coefficient adopted by Rampal et al. (2016) is independent from SIT values since it only uses free-drifting ice for tuning.

To evaluate the potential impact of assimilating the SIT from CS2SMOS on the sea ice motion, we further utilize the data set from the IABP buoys which began in 1990s to monitor ice motion throughout the Arctic Ocean. Only trajectories longer than 30 days and reporting more than 5 times per day are used to estimate the daily drift speed of sea ice. To avoid buoys in open water, the observations are selected based on sea ice concentration (>0.15) and ice thickness (>5 cm) at the nearest model grid cell in both runs. Furthermore, the dataset is restricted in the central Arctic, (delimited by a red line in Fig. 1), where water is deeper than 30 m and further away from the coast than 50 km. A total of 151 buoys are left from this selection, which provide 21,793 daily estimates of drift speed.

The speed distribution for daily drift of sea ice from IABP is shown by a histogram in Fig. 10a. In the central Arctic, the averaged drift speed is about 10.6 km d^{-1} (consistently with Allard et al., 2018) and most speeds (95%) are slower than 24 km d^{-1}. The difference of drift distributions between the two runs is minor compared to the difference to the IABP data. Restricting the analysis to the area North of 80 degrees, the two runs show larger differences in SIT with a Test run about 30 cm thicker (Fig. 10d), the resulting difference in SID in that area is small (0.2 km d^{-1}) and tends to degrade slightly the performance by slowing down the drift speed (Fig. 10c). This is somewhat contradictory to the analysis with OSI-SAF data which indicated a too fast model drift and smaller errors in the Test run. This inconsistency may be due to the poor spatial coverage of the IABP buoys. In Fig. 1 we can see that buoys north of 80°N are mainly found in the Eurasian Basin and sample poorly the region between the Transpolar Drift Stream and the Beaufort Gyre (Sumata et al., 2014), where the SID misfits are largest and where the model drift is too fast. This poor coverage of IABP buoys may as well explain why the SID comparisons in Allard et al. (2018) were inconclusive.

4.2 Impact on the sea ice extent and volume in the central Arctic

In Fig. 3, we show that the Arctic SIT has been improved everywhere, the assessment of the sea ice drift is less conclusive but tends to suggest a slight improvement localized in the central Arctic. However, improving the quantitative match with available observations does necessarily warrant the physical consistency of basin-scale integrated quantities. The impact of CS2SMOS on the Arctic-wide sea ice extent (SIE) and the sea ice volume (SIV) are investigated for the two runs and compared with the estimates from CS2SMOS and OSI-SAF respectively. Due to differences of resolution and land mask (especially important in the Canadian Archipelago), we focus on the central Arctic domain shown as the red line in the right panel of Fig. 1, excluding parts of the marginal seas.

Figure 11 shows the time evolutions of SIE and SIV in the two Official and Test runs. Both are calculated by daily averages in the two model runs. The SIE is classically calculated in the area where the SIC is not less than 15% in the Central Arctic. The SIE shows the expected seasonal cycle with the minimum (close to $3x10^6$ km^2) in September 2014 and saturates at a maximum value corresponding to the area of the Central Arctic region (around $6x10^6$ km^2) from January to March. The timing of the minimum and maximum from the two model runs agree very well with the observed in OSI-SAF and CS2SMOS (using the weekly concentration from the CS2SMOS product). We can also notice the impact of the weekly assimilation cycle that causes some "sawtooth" discontinuity and indicates that the model tends to both melt too fast in August and freeze too fast in September-October. Overall the SIE differences between the two runs (about 8,000 km^2) are indiscernible during the experimental time period.

The time evolutions of the SIV in the two runs show larger differences in the lower panel of Fig. 11. The maximum in the Test run is close to $12x10^3$ km^3 in April-May of 2014 and again end of March 2015, and the minimum is close to $5x10^3$ km^3 in September 2014. On average, the SIV difference in the two OSE runs is about 1,000 km^3, with lower volume in the Official run. Assimilation of the CS2SMOS data yields an annual increase of the SIV by about 8% relative to that in the Official run. The signature of the assimilation cycle is generally less pronounced than on SIE, except in August 2014 due to the SIC updates that are positively correlated to SIT in the summer (as noted in Lisæter et al., 2003). Compared to the observed SIV from the weekly CS2SMOS, the underestimation is significant at beginning of the runs (about $3x10^3$ km^3), but corrected by one third through the first month of assimilation of CS2SMOS. When the CS2SMOS data are missing, the gap between the two runs remains constant throughout the summer due to the long memory of winter ice, as previously noted with the assimilation work of ICESat SIT data in Mathiot et al. (2012). After the end of the summer during which no data of CS2SMOS are available， the SIV from the Test run is in better agreement with the first observed SIV from CS2SMOS. This indicates that the TOPAZ4 Official run has underestimated SIV due to the history of the reanalysis but not as a systematic tendency towards a bias state. The SIV estimates from observations occasionally present sudden discontinuities that seem unrealistic for a large integrated quantity such as the SIV of the central Arctic area. These discontinuities are larger than what the data assimilation system would expect based on the assumed observation error statistics given above. But the time series indicate that the EnKF does, as the name indicates, filter out part of the discontinuities so that only the major spike in early November 2014 causes a discontinuity in the Test run. Fig. 12 shows that the spike corresponds to a large homogeneous increase of SIT in all marginal seas between 26th Oct and 2nd Nov 2014, followed by a large decrease in the subsequent week. The weekly SIT innovation on the 2nd Nov reveals that the increase is largest south of the Eurasian Basin and around the Fram Strait. There, the SIT is thinner than 0.3 m on the 26th Oct which may suggest that the problem comes from the SIT measurement from SMOS. Until such inconsistencies are resolved in the dataset, we would recommend to either discard the first weeks of observations or increase the observation error during that period.

4.3 Quantitative impact for the observational network

The value of the Degrees of Freedom for Signal (DFS) is commonly used to monitor the relative impact of different observations in a data assimilation system (ref. Cardinali et al, 2004; Rodgers 2000; Xie et al, 2018), and is calculated as follows:

$$\mathrm{DFS} = tr\left(\frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{y}}\right) = tr\left\{\frac{\partial[\mathbf{H}(\bar{\mathbf{x}}^{\mathrm{a}})]}{\partial \mathbf{y}}\right\} = tr(\mathbf{KH}) \quad (11).$$

Where $\hat{\mathbf{y}}$ is the analyzed observation vector, the observation operator **H** is same in Eq. (1), and the term *tr* is the trace operator. The DFS is easily calculated and stored while performing the analysis with ensemble data assimilation (see Sakov et al. (2012) for an application to the TOPAZ4 system with the EnKF). It measures the reduction of uncertainty caused by a given observation type expressed as a number of equivalent degrees of freedom. Note that the DFS depend on the observation error statistics but not on the actual observation values (see equation 11). A DFS of 0 indicates that the observation has no impact at all, and a DFS equals to the total number of degrees of freedom indicates that the observation has so much impact that it has collapsed the ensemble to a single value. As the analysis is solved either in observational space or in ensemble space (depending on which is computationally cheapest), the DFS cannot exceed the smaller of the ensemble size and the number of observations used for the local assimilation. The DFS quantity is linear and can be split by observation types and accumulated in time periods. The averaged DFS for the *k*th type of observation can then be noted by $\overline{\mathrm{DFS}_{\mathrm{k}}}$, and thus a corresponding Impact Factor (IF) is defined as:

$$\mathrm{IF}_{\mathrm{k}} = \frac{\overline{\mathrm{DFS}_{\mathrm{k}}}}{\sum_{\mathrm{i}=1}^{\mathrm{o}} \overline{\mathrm{DFS}_{\mathrm{i}}}} \times 100\% \qquad (12).$$

Where *o* represents the number of different observation types assimilated in this time period. IF_{k} represents the relative impact of the k^{th} type of observations with respect to the whole observation network.

Figures 13 and 14 show the IF_{k} for different observations assimilated in the Test run averaged in two typical months: in November 2014 and in March 2015. The SIC impacts are dominant close to the sea ice edge and in the CAA region in the November, with an average IF of 22.7% in the whole Arctic. The SIT impact from CS2SMOS is largest in the central Arctic in November 2014. A relatively smaller impact (>20%) is also noticeable in north of the Barents Sea and west of the Kara Sea. In the open ocean, the SST and SLA have the largest impact. Temperature and salinity profiles have locally an important effect in the ice-covered Arctic, where a few of ice-tethered profilers (ITP) are available and the uncertainty is large. Xie et al. (2016) applied the same DFS method to evaluate the impact of thin SIT from SMOS only. The present results reveal, as expected, much larger impacts of CS2SMOS SITs in the central Arctic, with only a few isolated dips where the ITP profiles are available. The IF is higher where the ice is thicker, even though the observation error increases as a function of ice thickness. It indicates that the ensemble background errors increase even more than the observation errors in thick ice by temporal accumulation of model errors. For example, errors in precipitation grow as the snow accumulates in the Fall, and the resulting inter-member variability of snow cover causes inter-member variability of SIT due to the thermal isolation effect of snow.

In March 2015, CS2SMOS has again a large impact in the central Arctic relative to other assimilated observations even though previous literature indicates a lower impact in the midst of winter than when the ice is growing (Mathiot et al., 2012). The relative IF of SIT indeed remains high even though the absolute DFS is decreasing, due to the lower impact of other assimilated observations, in particular SIC (Lisæter et al., 2003). On average, the IF value of CS2SMOS is about 40%. The high values (>40%) are clearly separated into two areas: one is to the north of the CAA and Greenland; another following the inner side of the sea-ice edge in marginal ice zones. The former is primarily a CryoSat-2 contribution, while the latter corresponds to the thin SITs from SMOS. The high IF in the polar hole is probably undesirable since the observations there are merely extrapolated, so in the future applications we would recommend discarding these data, in order to leave the polar hole filled instead with sea ice advected from areas where trustworthy SIT observations have been assimilated.

5. Conclusions and discussions

CS2SMOS is the first product to monitor the complete pan-Arctic SIT in a systematic way, although only for the winter months. It is a combination of two very different, yet very advanced, technologies onboard the SMOS and CryoSat-2 satellites, calibrated against very few in-situ observations of SIT, freeboard and snow depths. Altogether, the issue of measurements uncertainties is particularly delicate for the assimilation of CS2SMOS data. On the other hand, defining proper model background errors for SIT is just as delicate, when considering that the simulated SIT accumulates errors both in the sea ice dynamics (in particular the rheological model) and in the thermodynamics. The Bayesian approach to confront these two uncertainties is by Monte Carlo propagation of uncertainties, which is what is practiced in the present study for the model background error, although not for the observation error.

This study assesses the impact of assimilating the new SIT product from 19th March 2014 to 31st March 2015. Compared to the assimilated SIT CS2SMOS, the thin bias is reduced from 15 cm to 5 cm, and the RMSD also decreased from 58 cm to 38 cm, a reduction by 28.3%. Other innovation diagnostics show no degradation towards other assimilated variables –namely SIC, SSH, SST and TS profiles.

The SIT is also improved when compared to four independent drifting IMB buoys and three BGEP mooring buoys. The benefits persist throughout the summer although no SIT observations are available then, consistently with the experiments from Mathiot et al. (2012). This is important because it suggests that the model is not attracted to his bias solution. The assimilation reduces the low SIT biases north of the CAA and north of Greenland and the high bias in the Beaufort Sea compared to independent observations from Operation IceBridge. Both the thick pack ice in central Arctic and the thin ice in marginal seas are corrected. On average, the SIT errors in March- April of 2014 and 2015 are reduced by 15 cm, a reduction by 12.5% compared to the Official run. The dynamical adjustment following the assimilation of SIT has partially improved the sea ice drift speeds in the Test run where the SIT has thickened: the monthly averaged drift speed errors north of 80°N are reduced by 0.4-0.5 km per two days in December 2014 and February 2015 (8-9% reduction of the error). This has been revealed by satellite products but not IABP in situ buoys for which the spatial coverage is very poor. However, it should also be reminded that the drag coefficient used in the Test run were tuned for the Official run which has a biased SIT. One would expect some improvement with a retuned drag coefficient value. At term, we consider doing an online parameter estimation of key parameter such as the drag coefficient as tested in Massonnet et al. (2014).

In this study, the DFS information in the ensemble data assimilation system has been applied to quantitatively evaluate the relative contributions of all assimilated observation types. CS2SMOS has the highest impact near the northern coast of Canada, north of Greenland, and on the inner side of the sea ice edge, where the contributions from CryoSat-2 and SMOS SIT were expected. The results, compared to assimilating SMOS only in Xie et al. (2016), show the importance of CryoSat-2, particularly in the winter months to constrain the SIT offsets (also shown by Mu et al. 2018, in a coupled MITgcm model system) and motivate the assimilation of CS2SMOS in the following reanalysis of TOPAZ4. However, the impact of SIT observations may vary with the evaluation of the modelling and observing system. Firstly, the SIC may have been underestimated in central Arctic due to the simplicity of the present sea ice model. Further planned developments of TOPAZ include a new model rheology that is able to resolve the scaling laws of deformation of sea ice (Rampal et al., 2016) and should therefore improve the background errors of ice concentration in winter months and sea ice drift, increase the impact of SIC and SID within the ice pack and reduce the estimated SIT impact accordingly. Other planned changes such as the simulation of melt ponds are not expected to influence these results directly since there are no melt ponds when the SIT data is available. Lastly, if a large number of in situ profiles were available below the sea ice, they would also compete with the SIT observations.

The above OSE results, like others, are necessarily contingent on adequate specifications of observation errors. Those are very much simplified in the case of CS2SMOS, which is not an uncommon case for remote sensing observations: due to the complexity of the physics involved, the specified observation errors are reflecting interpolation errors rather than a nonlinear propagation of errors from their sources (Ricker et al., 2017). In the present study, an offset has been added to account for this difference in Eq. (4), which results in a conservative error estimate with respect to the classical Desroziers optimality criterion and a suboptimal performance in the reliability budget analysis. In the one hand, reducing the observation would have accelerate the convergence to observed SIT and converge to a more accurate solution. On the other hand, this would have made the EnKF less robust to the sudden inconsistencies in the observations as seen in Fig. 11. Further versions of the CS2SMOS data will hopefully improve their temporal continuity and the impact of the data can be increased accordingly.

An alternative to using the scheme CS2SMOS data would have been to assimilate the two data sets CryoSat-2 and SMOS SIT separately and let the EnKF merge them together rather than relying on optimal interpolation, as successfully demonstrated by Mu et al (2018). This would for instance avoid assimilating observations in places where they are the pure result of interpolation/extrapolation but would not resolve the offset between the two satellites, which is arguably the most worrying issue as of the present state of the SMOS and CryoSat-2 data. The assimilation of the separate datasets will be attempted in the future when their consistency is further improved.

The current TOPAZ reanalysis is currently reaching 2016 and extended by one year every year. The current study clearly shows the added value of assimilating SIT. In 2020, a new TOPAZ reanalysis will be provided with the upgraded version of TOPAZ5 which will include SIT assimilation from 2010 onwards.

Acknowledgements

Thanks to the three anonymous reviewers for constructive comments. Thanks to Dr. J. A. Johannessen for nice suggestions and to Dr. S. Hendricks and Dr. R. Ricker for sharing the CS2SMOS data on meereisportal.de. The authors acknowledge the support of CMEMS for the Arctic MFC. Grants of computing time (nn2993k and nn9481k) and storage (ns2993k) from the Norwegian Sigma2 infrastructures are gratefully acknowledged.

Reference:

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. Wallcraft, A. J.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system. *Advances in Space Research*, 62/6), 1265-1280, http://doi.org/10.1016/j.asr.2017.12.030, 2018.

Bathiany, S., Notz, D., Mauritsen, T., Raedel, G., and Brovkin, V.: On the potential for abrupt Arctic winter sea ice loss. *J. Climate*, **29**, 2703–2719, https://doi.org/10.1175/JCLI-D-15-0466.1, 2016.

Bertino, L., and Lisæter, K. A.: The TOPAZ monitoring and prediction system for the Atlantic and Arctic Oceans, Journal of Operational Oceanography, 1(2), 15–19, doi: 10.1080/1755876X.2008.11020098, 2008

Bentsen, M., Evensen, G., Drange, H., and Jenkins, A. D.: Coordinate transformation on a sphere using conformal mapping, Mon. Weather Rev., 127, 2733-2740, doi:http://dx.doi.org/10.1175/1520-0493(1999)127<2733:CTOASU>2.0.CO:2, 1999.

Bouillon, S., Fichefet, T., Legat, V., and Madec, G.: The elastic-viscous-plastic method revised. Ocean Modell., 7, 2-12, doi:10.1016/j.ocemod.2013.05.013, 2013.

Budikova, D.: Role of Arctic sea ice in global atmospheric circulation: A review. Global and Planetary Change, 68,149-163, doi:10.1016/j.gloplacha.2009.04.001, 2009.

Cardinali, C., Pezzulli, S., and Andersson, E.: Influence-matrix diagnostic of a data assimilation system, Q. J. R. Meteorol. Soc., 130, 2767-2786, doi:10.1256/qj.03.205, 2004.

Chassignet, E. P., Smith, L. T., and Halliwell, G. R.: North Atlantic Simulations with the Hybrid Coordinate Ocean Model (HYCOM): Impact of the vertical coordinate choice, reference pressure, and thermobaricity, J. Phys. Oceanogr., 33, 2504-2526. Doi: http://dx.doi.org/10.1175/1520-0485(2003)033<2504:NASWTH>2.0.CO:2, 2003.

Comiso, J. C., Parkinson, C. L., Gersten, R., and Stock, L.: Accelerated decline in the Arctic sea ice cover. *Geophys. Res. Lett.*, **35**, L01703, doi:https://doi.org/10.1029/2007GL031972, 2008.

Counillon, F. and Bertino, L.: High-resolution ensemble forecasting for the Gulf of Mexico eddies and fronts, Ocean Dynam., 59, 83–95, doi:10.1007/s10236-008-0167-0, 2009.

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

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., et al.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quart. J. Roy. Meteor. Soc., 137, 553-597, doi:10.1002/qj.828, 2011

Desroziers, G., Berre, L., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space. Q. J. R. Meteorol. Soc., 131(613), 3385-3396, https://doi.org/10.1256/qj.05.108, 2005.

Docquier, D., François Massonnet, F., Barthélemy, A., Tandon, N. F., Olivier Lecomte, O., and Fichefet, T.: Relationships between Arctic sea ice drift and strength modelled by NEMO-LIM3.6. The Cryosphere, 11, 2829-2846, https://doi.org/10.5194/tc-11-2829-2017, 2017

Drange, H., and Simonsen, K.: Formulation of air-sea fluxes in the ESOP2 version of MICOM, Technical Report No. 125 of Nansen Environmental and Remote Sensing Center, 1996.

Ferreira, A. S. A., Hátún, H., Counillon, F., Payne, M. R., and Visser, A. W.: Synoptic-scale analysis of mechanisms driving surface chlorophyll dynamics in the North Atlantic, Biogeosciences, 12, 3641-3653, https://doi.org/10.5194/bg-12-3641-2015, 2015.

Finck, N., Counillon, F., Bertino, L., Bouillon, S. and Rampal, P.: Validation of sea ice quantities of TOPAZ for the period 1990-2010, Technical Report No. 332 of Nansen Environmental and Remote Sensing Center, 2013.

Guemas, V., Wrigglesworth, E. B., Chevallier, M., et al.: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time scales. Q. J. R. Meteorolog. Soc., 142(695), 546-561, https://doi.org/10.1002/qj.2401, 2014.

Heygster, G., Hendricks, S., Kaleschke, L., Maass, N., et al.: L-Band Radiometry for Sea-Ice Applications, Final Report for ESA ESTEC Contract 21130/08/NL/EL. Institute of Environmental Physics, University of Bremen, 219 pages, 2009.

Hibler, W. D., III: A dynamic thermodynamic sea ice model. *J. Phys. Oceanogr.,* **9,** 817–846, *https://doi.org/10.1175/1520-0485(1979)009<0815:ADTSIM>2.0.CO;2*, 1979.

Hibler, W. D., III: Ice dynamics. chap. 9, *The Geophysics of Sea Ice,* N. Untersteiner, Ed., *NATO ASI Series B: Physics*, Plenum Press, 577–640, 1986.

Hunke, E. C., and Dukowicz, J. K.: An elastic-viscous-plastic model for sea ice dynamics, J. Phys. Oceanogr., 27, 1849-1867, https://doi.org/10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2, 1997.

Johannessen, O. M., Shalina, E. V., and Miles, M. W.: Satellite evidence for an Arctic Sea ice cover in transformation, Science, 286, 1937—1939. Doi:10.1126/science.286.5446.1937, 1999.

Johannessen, J. A., et al.: Toward improved estimation of the dynamic topography and ocean circulation in the high latitude and Arctic Ocean: The importance of GOCE, Surv. Geophys., 35, 661–679, doi:10.1007/s10712-013-9270-y, 2014.

Johnson, M., Proshutinsky A., Aksenov Y., Nguyen A. T., Lindsay R., Haas C., Zhang J., Diansky N., Kwok R., et al.: Evaluation of Arctic sea ice thickness simulated by Arctic Ocean Model Intercomparison Project models. J. Geophys. Res., 117(C8), C00D31, doi:10.1029/2011JC007257, 2012.

Kaleschke, L., Maaß, N., Haas, C., Hendricks, S., Heygster, G., and Tonbøe, R.: 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. Doi: 10.5194/tc-4-583-2010, 2010.

Kaleschke, L., Tian-Kunze, X., Maaß, N., Ricker, R., Hendricks, S., and Drusch, M.: Improved retrieval of sea ice thickness from SMOS and Cryosat-2. Proceedings of 2015 International Geoscience and Remote Sensing Symposium IGARSS, doi: 10.1109/IGARSS.2015.7327014, 2015.

Karspeck, A. R.: An ensemble approach for the estimation of observational error illustrated for a nominal 1 global ocean model. Monthly Weather Review, 144, 1713-1728, DOI: 10.1175/MWR-D-14-00336.1, 2016.

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, doi:10.5194/tc-9-37-2015, 2015.

Khvorostovsky, K., and Rampal, P.: On retrieving sea ice freeboard from ICESat laser altimeter. The Cryosphere, 10, 2329-2346, doi:10.5194/tc-10-2329-2016, 2016.

Kimmritz, M., Counillon, F., Bitz, C.M., Massonnet, F., Bethke, I. and Gao, Y., 2018. Optimising assimilation of sea ice concentration in an Earth system model with a multicategory sea ice model. *Tellus A: Dynamic Meteorology and Oceanography*, *70*(1), 1435945, https://doi.org/10.1080/1600870.2018.1435945, 2018.

King, J., Howell, S., Derksen, C., Rutter, N., Toose, P., Beckers, J. F., Haas, C., Kurtz, N., and Richter-Menge, J.: Evaluation of Operation IceBridge quick-look snow depth estimates on sea ice, Geophys. Res. Lett., 42, 9302–9310, doi:10.1002/2015GL066389, 2015.

King, J., Skourup, H., Hvidegaard, S. M., Rösel, A., Gerland, S., Spreen, G., . . . Liston, G. E. (2018). Comparison of freeboard retrieval and ice thickness calculation from ALS, ASIRAS, and CryoSat-2 in the Norwegian Arctic to field measurements made during the N-ICE2015 expedition. Journal of Geophysical Research: Oceans, 123, 1123–1141. https://doi.org/10.1002/ 2017JC013233

Kinnard, C., Zdanowicz, C. M., Fisher, D. A., Isaksson, E., Vernal, A., and Thompson, L. G.: Reconstructed changes in Arctic sea ice over the past 1,450 years. Nature, 479, 509–512. doi:10.1038/nature10581, 2011.

Kwok, R., and Rothrock, D.: Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008, Geophys. Res. Lett., 36, L15501, doi:10.1029/2009GL039035, 2009.

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, doi:10.5194/tc-7-1035-2013, 2013.

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

Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H., and Breivik, L. -A.: Sea ice motion from low resolution satellite sensors: an alternative method and its validation in the Arctic. *Journal of Geophysical Research*, 115, C10032, 2010. doi: 10.1029/2009JC005958, 2010.

Levermann, A., Mignot, J., Nawrath, S., Rahmstorf, S.: The role of Northern sea ice cover for the weakening of the thermohaline circulation under global warming. J. Climate, 20, 4160-4171, *https://doi.org/10.1175/JCLI4232.1*, 2007.

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

Lisæter, K. A., Rosanova, J. J., and Evensen, G.: Assimilation of ice concentration in a coupled ice ocean model, using the Ensemble Kalman filter. Ocean Dynamics, 53, 368-388, doi:10.1007/s10236-003-0049-4, 2003.

Lisæter, K. A., Evensen, G., and Laxon, S.: Assimilating synthetic CryoSat sea ice thickness in a coupled ice-ocean model, J. Geophys. Res., 112, C07023, doi:10.1029/2006JC003786, 2007.

Martin, S., Drucker, R., Kwok, R., and Holt, B.: Estimation of the thin ice thickness and heat flux for the Chukchi Sea Alaskan coast polynya from Special Sensor Microwave/Imager data, 1990-2001, J. Geophys. Res., 109, C10012, https://doi.org/10.1029/2004JC002428, 2004.

Massonnet, F., Goosse, H., Fichefet, T., and Counillon, F.: Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter, J. Geophys. Res., 119(7), 4168-4184, https://doi.org/10.1002/2013JC009705, 2014.

Mathiot, P., König Beatty, C., Fichefet, T., Goosse, H., Massonnet, F., and Vancoppenolle, M.: Better constraints on the sea-ice state using global sea-ice data assimilation, Geosci. Model Dev., 5, 1501-1515, https://doi.org/10.5194/gmd-5-1501-2012, 2012.

Melia, N., Haines, K., and Hawkins, E.: Improved Arctic sea ice thickness projections using bias-corrected CMIP5 simulations, The Cryosphere, 9, 2237-2251, https://doi.org/10.5194/tc-9-2237-2015, 2015.

Metzger, E. J., Smedstad, O. M., Thoppil, P. G., Hurlburt, H. E., Cummings, J. A., Wallcraft, A. J., Zamudio, L., Franklin, D. S., Posey, P. G., Phelps, M. W. and Hogan, P. J.: US Navy operational global ocean and Arctic ice prediction systems. *Oceanography*, *27*(3), 32-43, https://doi.org/10.5670/oceanog.2014.66, 2014.

Mu, L., Yang, Q., Losch, M., Losa, S. N., Ricker, R., Nerger, L., and Liang, X.: Improving sea ice thickness estimates by assimilating CryoSat-2 and SMOS sea ice thickness data simultaneously. Q. J. R. Meteorol. Soc., 144(711), 529-538, DOI:10.1002/qj.3225, 2018.

Oke, P. R., and Sakov, P.: Representation error of oceanic observations for data assimilation. J. Atmos. Oceanic Technol., 25, 1004–1017, doi:10.1175/2007JTECHO558.1, 2008.

Oki, T., and Sud, Y. C.: Design of Total Runoff Integrating Pathways (TRIP)—A Global River Channel Network. *Earth Interact.,* **2**, 1–37, https://doi.org/10.1175/1087-3562(1998)002<0001:DOTRIP>2.3.CO;2, 1998.

Olason, E., and Notz, D.: Drivers of variability in Arctic sea-ice drift speed, J. Geophys. Res. Oceans, 119,5755–5775, doi:10.1002/2014JC009897, 2014.

Penny, G., Akella, S.R., Frolov, S., Fujii, Y., Karspeck, A., Peña, M., Subramanian, A., Tardif, R., Wu, X., Anderson, J., Kalnay, E., Kleist, D.T., and Todling, R.: Coupled Data Assimilation for Integrated Earth System Analysis and Prediction : Goals , Challenges , and Recommendations. Technical Report, https://ntrs.nasa.gov/search.jsp?R=20170007430, 2017.

Perovich, D. K., and Richter-Menge, J. A.: From points to Poles: extrapolating point measurements of sea-ice mass balance. *Ann. Glaciol.,* **44**, 188–192, doi:10.3189/172756406781811204, 2006.

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, dos:10.5194/tc-9-1735-2015, 2015.

Rampal, P., Bouillon, S., Ólason, E., and Morlighem, M.: neXtSIM: a new Lagrangian sea ice model. *The Cryosphere*, *10*(3), 1055–1073, 2016.

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, doi:10.5194/tc-8-1607-2014, 2014.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J. and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607-1623, doi:10.5194/tc-11-1607-2017, 2017.

Rodgers, C.: Inverse methods for atmospheres: theory and practice, World Scientific, 2000.

Rodwell, M. J., Lang, S. T. K., Ingleby, N. B., Bormann, N., Hólm, E., Rabier, F., Richardson, D. S. and Yamaguchi, M.: Reliability in ensemble data assimilation. Quart. J. Roy. Meteor. Soc., 142, 443–454, doi: 10.1002/qj.2663, 2016.

Sakov, P., and Oke, P. R.: A deterministic formulation of the ensemble Kalman Filter: an alternative to ensemble square root filters. Tellus A, 60(2), 361-371, doi:10.1111/j.1600-0870.2007.00299.x, 2008.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., and Korablev, A.: TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. *Ocean Science*, 8(4), 633–656. http://doi.org/10.5194/os-8-633-2012, 2012.

Schofield, O., Ducklow, H. W., Martinson, D. G., Meredith, M. P., Moline, M. A., and Fraser, W. R.: How Do Polar Marine Ecosystems Respond to Rapid Climate Change? Science (328), 5985, 1520–1523, DOI: 10.1126/science.1185779, 2011.

Schweiger, A., Lindsay, R., Zhang, J., Steels, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic sea ice volume, J. Geophys. R., 116, C00D06, doi:10.1029/2011JC007084, 2012.

Smith, G. C., Roy, F., Reszka, M., Colan, D. S., He, Z., Deacu, D., et al.: Sea ice forecast verification in the Canadian Global Ice Ocean Prediction System. Quart. J. Roy. Meteor. Soc., doi:10.1003/qj.2555, 2015.

Stark, J. D., J. Ridley, M. Martin, M., and Hines, A.: Sea ice concentration and motion assimilation in a sea ice−ocean model, J. Geophys. Res., 113, C05S91, doi:10.1029/2007JC004224, 2008.

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), 1445379, DOI: 10.1080/16000870.2018.1445379, 2018.

Stroeve, J. C., Serreze, M. C., Holland, M. M. et al.: The Arctic's rapidly shrinking sea ice cover: a research synthesis. Climatic change, 10 (3), 1005-1027, doi:10.1007/s10584-011-0101-1, 2012.

Sumata, H., Lavergne, T., Girard-Ardhuin, F., Kimura, N., Tschudi, M. A., Kauker, F., Karcher, M., and Gerdes, R.: An intercomparison of Arctic ice drift products to deduce uncertainty estimates, J. Geophys. Res. Oceans, 119, 4887–4921, doi:10.1002/ 2013JC009724, 2014.

Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M., and Krumpen, T.: SMOS-derived sea ice thickness: algorithm baseline, product specifications and initial verification, The Cryosphere, 8, 997-1018, doi:10.5194/tc-8-997-2014, 2014.

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, doi:10.5194/tc-10-2003-2016, 2016.

Tilling, R. L., Ridout, A., and Sheperd, A.: Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data. *Advances in Space Research*, 62(6), 1203-1225, http://doi.org/10.1016/j.asr.2017.10.051, 2018.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V. S., Marnela, M., Massonnet, F., Mignac, D., Peterson, A., Sadikn, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J., Zhang, Z.: An assessment of ten ocean reanalyses in the polar regions, Climate Dynamics, https://doi.org/10.1007/s00382-018-4242-z, 2018.

Wang, X., Key, J., Kwok, R., and Zhang, J.: Comparison of Arctic sea ice thickness from satellites, aircraft, and PIOMAS data. *Remote Sensing*, *8*(9), 1–17, http://doi.org/10.3390/rs8090713, 2016.

Woodgate, R., Aagaard, K. and Weingartner, T.: Monthly temperature, salinity, and transport variability of the Bering Strait through flow. *Geophys. Res. Lett.*, 32, L04601, DOI: 10.1029/2004GL021880, 2005.

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013. *Ocean Science*, 13(1), 123–144. http://doi.org/10.5194/os-13-123-2017, 2017.

Xie, J., Bertino, L., Cardellach, E., Semmling, M., and Wickert, J.: An OSSE evaluation of the GNSS-R altimetery data for the GEROS-ISS mission as a complement to the existing observational networks, *Remote Sens. Environ.*, 209, 152-165, doi:10.1016/j.rse.2018.02.053, 2018.

Xie, J., Counillon, F., Bertino, L., Tian-Kunze, X., and Kaleschke, L.: Benefits of assimilating thin sea-ice thickness from SMOS into the TOPAZ system. *The Cryosphere*, 10, 2745–2761. http://doi.org/10.5194/tc-10-2745-2016, 2016.

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, 6680–6692, doi:10.1002/2014JC009963, 2014.

Figures:

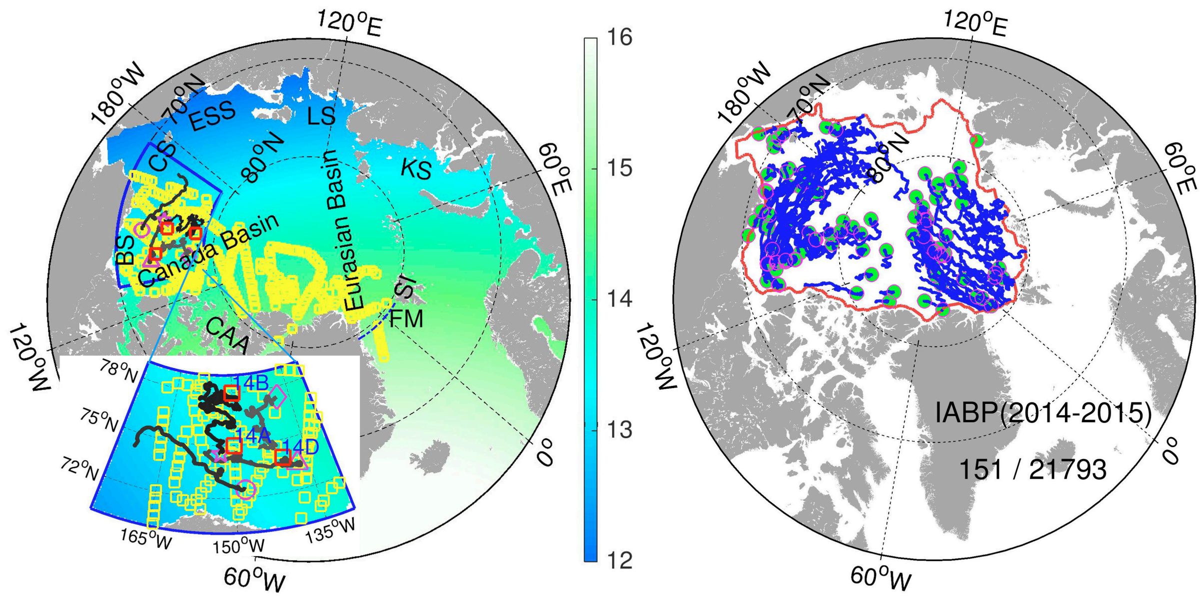

Fig. 1 Left: Horizontal resolution (km) of the model grid in the Arctic (>60°N). The small yellow squares are the locations of IceBridge campaigns during the experimental period. The marginal seas are: Beaufort Sea (BS, ; also shown with the blue line), Chukchi Sea (CS), East Siberian Sea (ESS), Laptev Sea (LS), Kara Sea (KS) and the other regions: Canadian Arctic Archipelago (CAA), Svalbard Island (SI), and Fram Strait (FM). The four purple markers (pentagram, circle, triangle and diamond) are the deployment location of IMB buoys (2013F, 2014B, 2014C, and 2014F respectively) with the following trajectory shown as black solid curves. The three red squares are the fixed locations of the BGEP moorings (14A, 14B, and 14D respectively). **Right**: Trajectories of International Arctic Buoy Program buoys drift during the experimental period. The solid red line delimits the coastal areas excluded in the analysis.

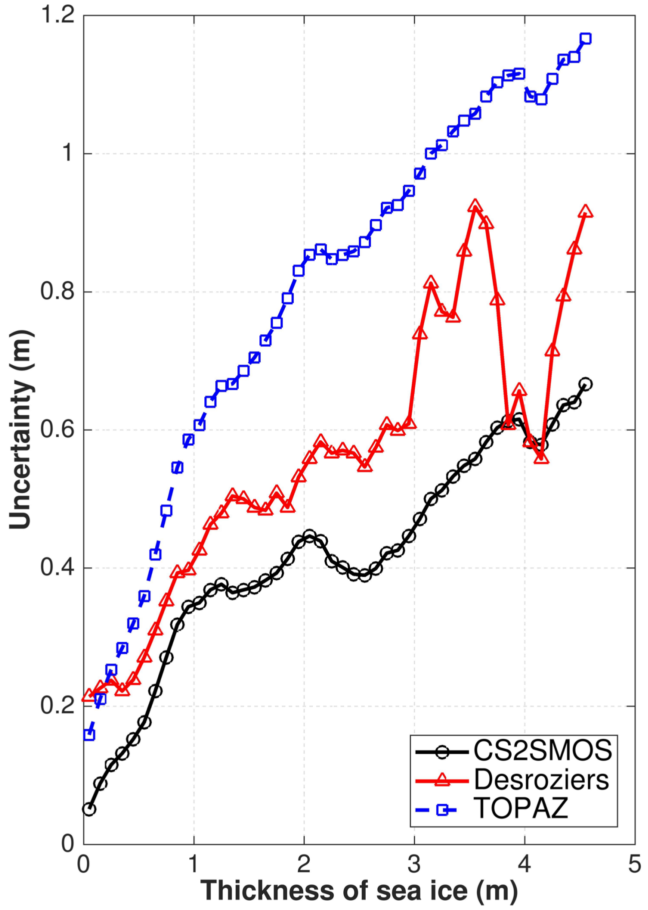

Fig. 2 Observation error uncertainties as a function of sea ice thickness for the original CS2SMOS data set (black line), the estimated observation error using the Desroziers diagnostics with red-triangle line (see Eq. (3)) and the one used in the TOPAZ Test run with blue-square, with an additional error term as Eq. (4) to the original uncertainty.

0

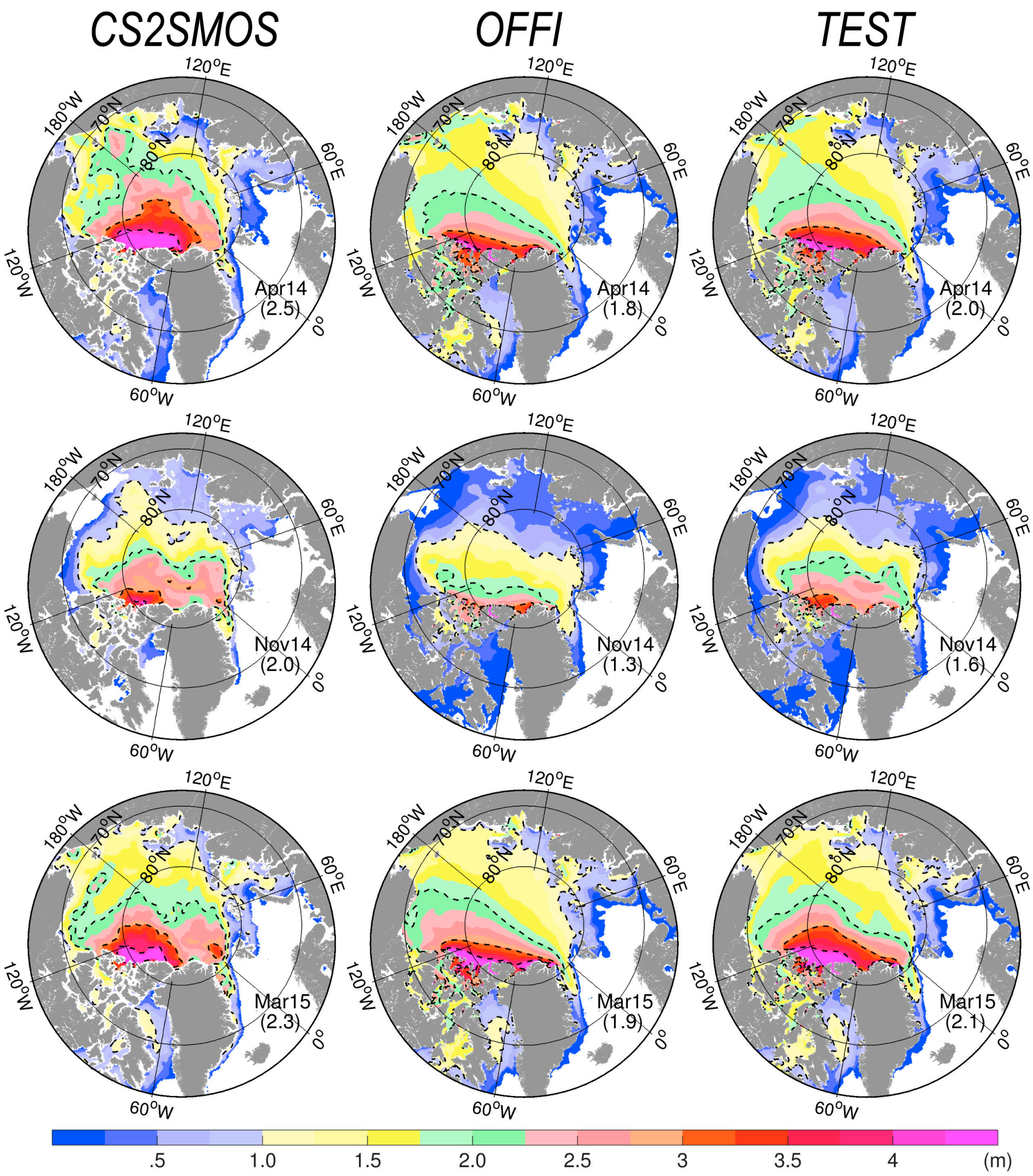

Fig. **3** Monthly SIT from CS2SMOS (left), Official run (middle) and Test run (right) in April 2014, November 2014, and March 2015. The mean SIT estimated for the area north of 80N is indicated in brackets (unit: m). The dashed lines are isolines of 1, 2, 3, and 4 meters SIT respectively.

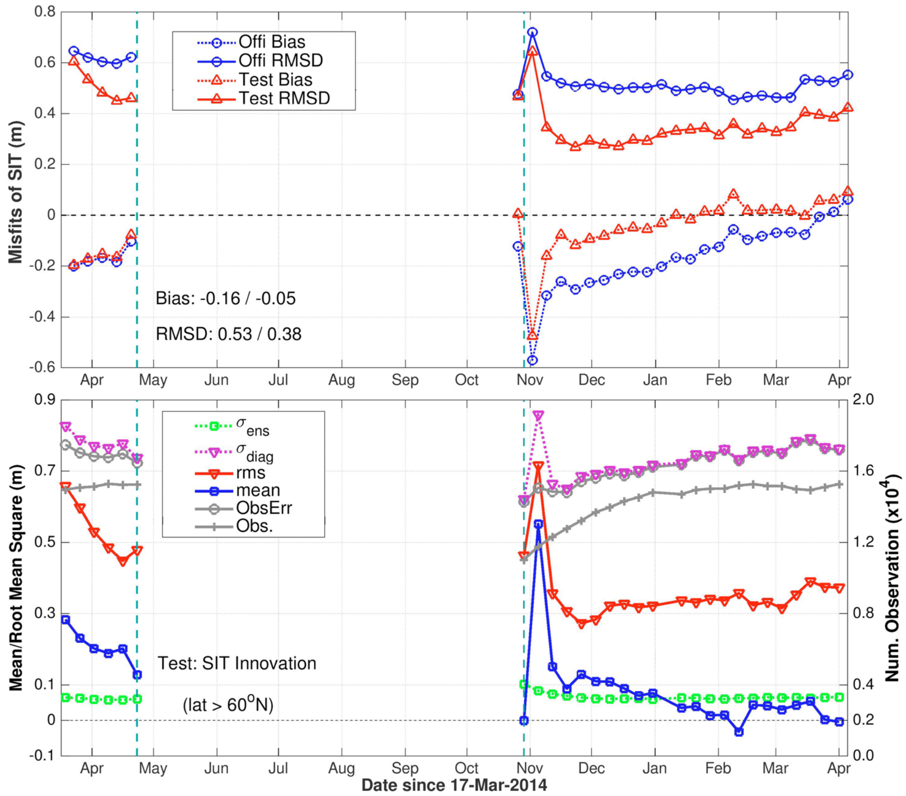

Fig. 4 Top: Bias (dotted line) and RMSD (solid line) of SIT in the two runs - Official (blue) and Test (red) – based on weekly averaged reanalysis and CS2SMOS observations. The time-averaged bias and RMSD are indicated (Official/Test). **Bottom**: SIT innovation statistics in the Test run in the Arctic region (>60°N) from 19th March 2014 to end of March 2015. The blue-squared (resp. red reverted-triangle) line represents the mean (RMSD) of the innovation. The green squared line represents the ensemble spread and the purple reverted-triangle line is the diagnosed total uncertainty (see Eq. (8)). The gray-crossed (gray-circled) line is the number (RMSD observation error) of assimilated observations.

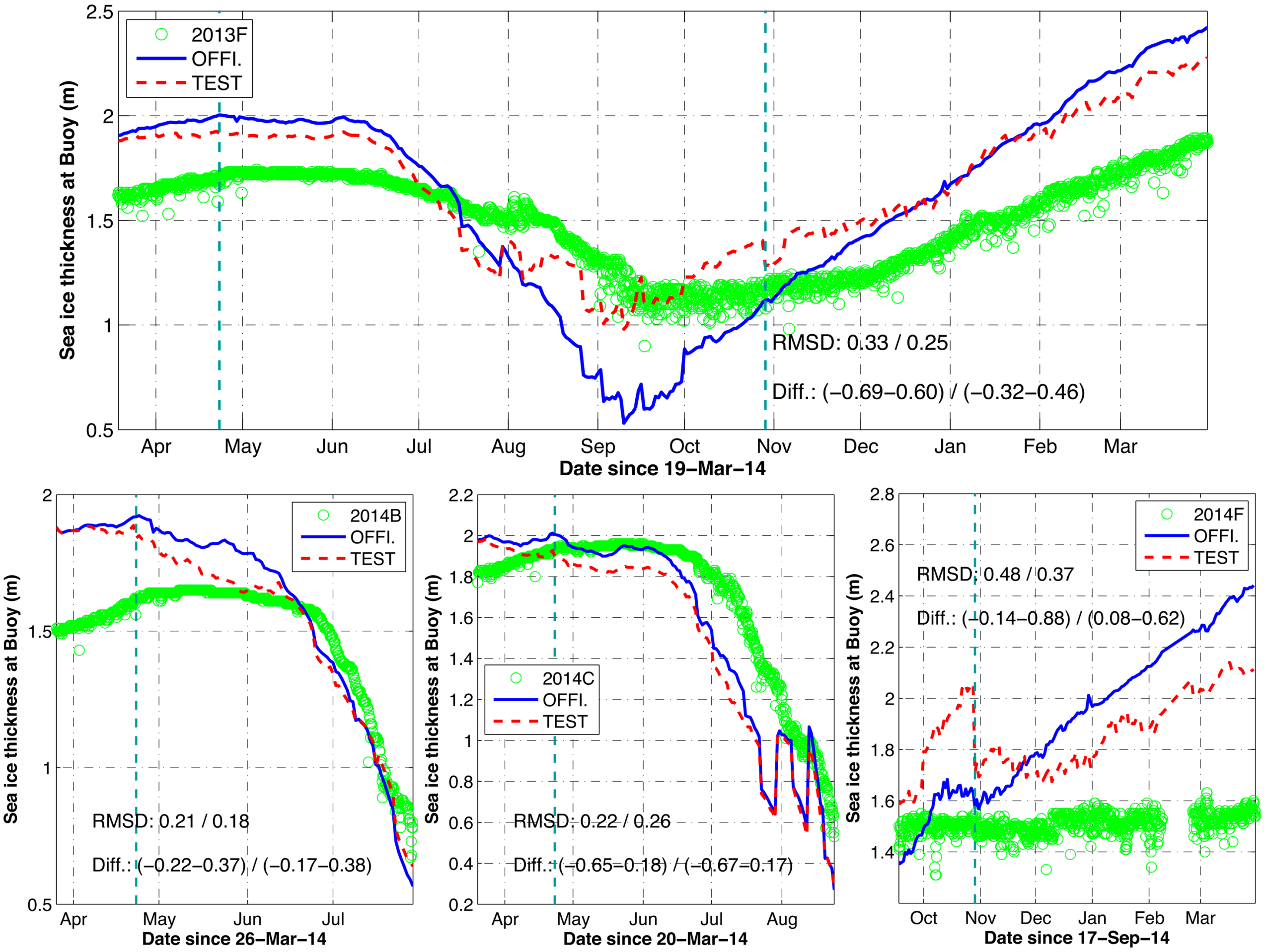

Fig. 5 Time series of SIT along the trajectories of IMB buoys (upper: 2013F; bottom: 2014B, 2014C, and 2014F). Measured SIT (green), daily averages from the Official run (blue line) and the Test run (red line). The vertical cyan-dashed lines indicate the winter period when C2SMOS is assimilated in the Test run.

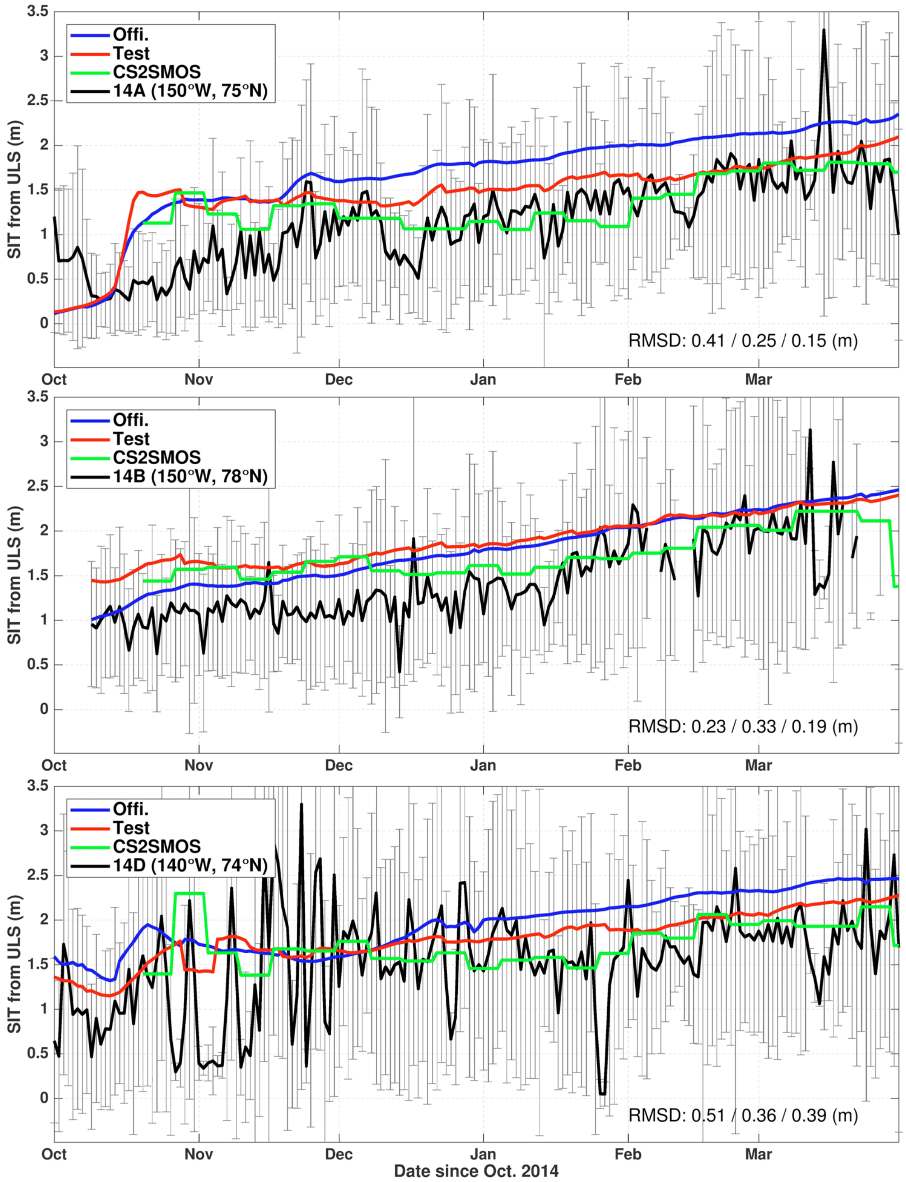

Fig. 6 Daily series of SIT (black line) at the BGEP mooring (14A, 14B, and 14D) compared with the two model runs - Official (blue line) and Test (red line) - and the weekly observed by CS2SMOS (green line). The black line represents the daily average at the mooring location with the standard deviation shown as the error bar. The RMSDs of the Official run, Test run and CS2SMOS are respectively indicated on the bottom of each panels.

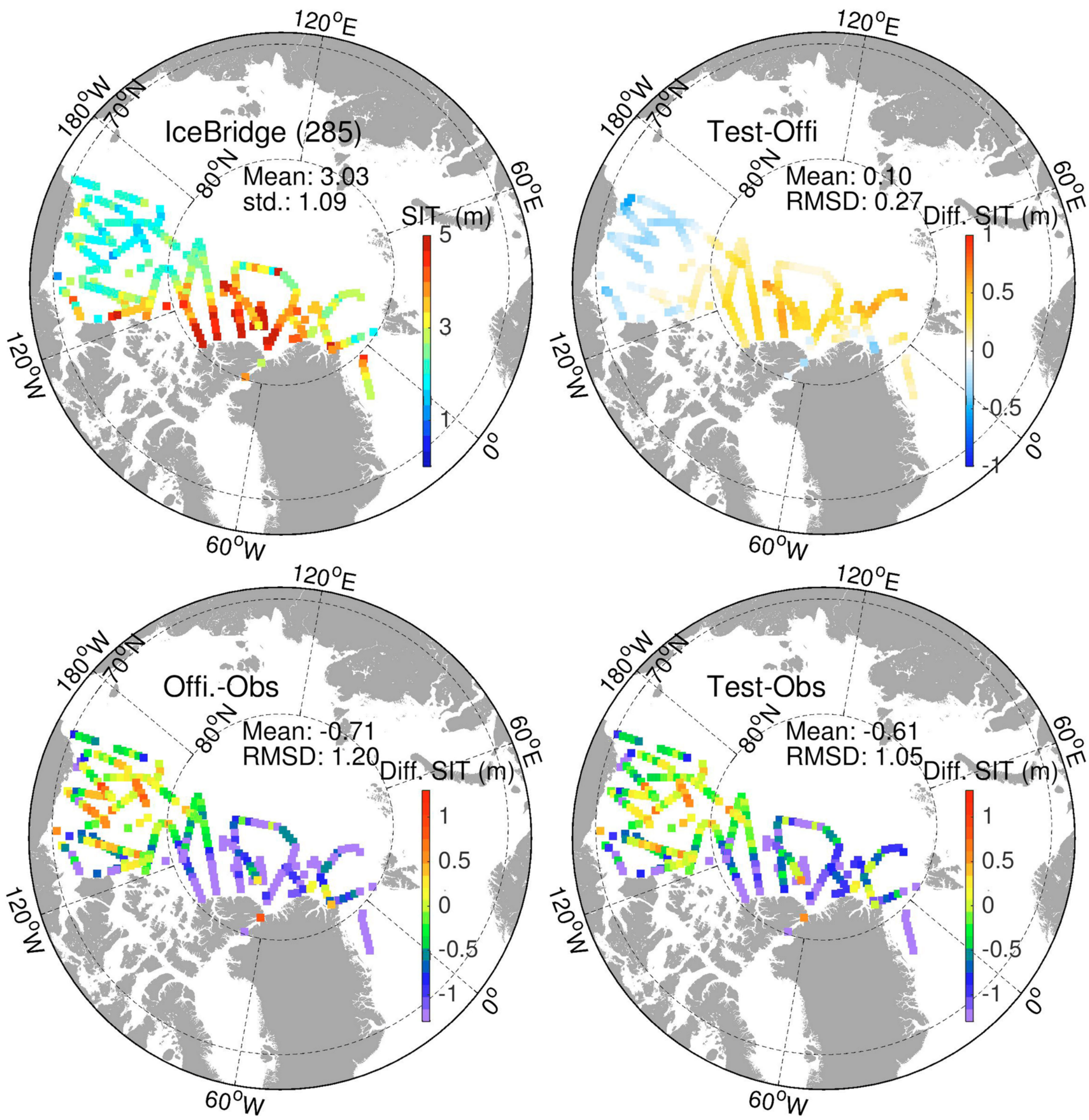

Fig. 7 Top: IceBridge SIT in 2014 and 2015 (left) and the SIT differences in the two model runs according to the observational locations and times (right). **Bottom**: SIT deviations from the Official run (left) and Test run (right) using model daily average at observations time.

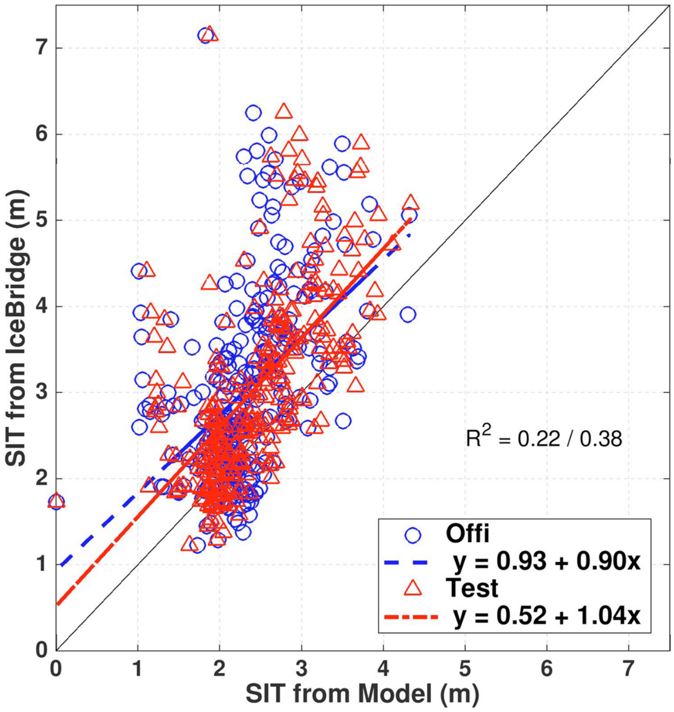

Fig. 8 Scatterplots of SIT daily averaged of Official (blue) and Test (red) runs compared to IceBridge data. The dashed lines are the respective linear regression, the coefficient R^2 is the squared correlation to represent how strong of the linear relationship in Official/Test run. The black line is y=x.

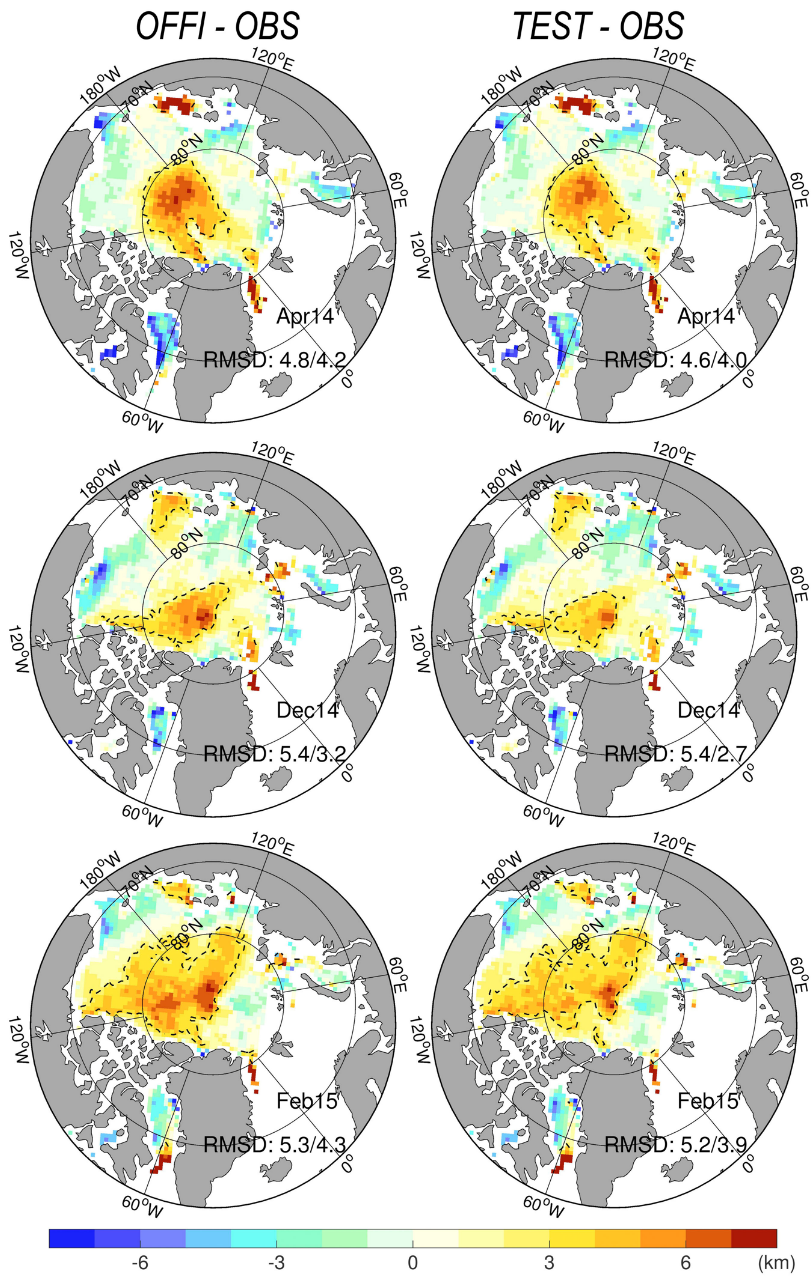

Fig. 9 Sea ice drift misfits (model minus observation, in km per two days) in the Official run (left column) and Test run (right column) compared against the OSI-SAF sea ice drift in April 2014 (top line), December 2014 (middle line), and February 2015 (bottom line). The black dashed delimits the area of fastest drift (drift > 3km per 2 days), and the RMSD relative to the monthly observations is indicated when calculated for the whole domain and at for the region north of 80°N.

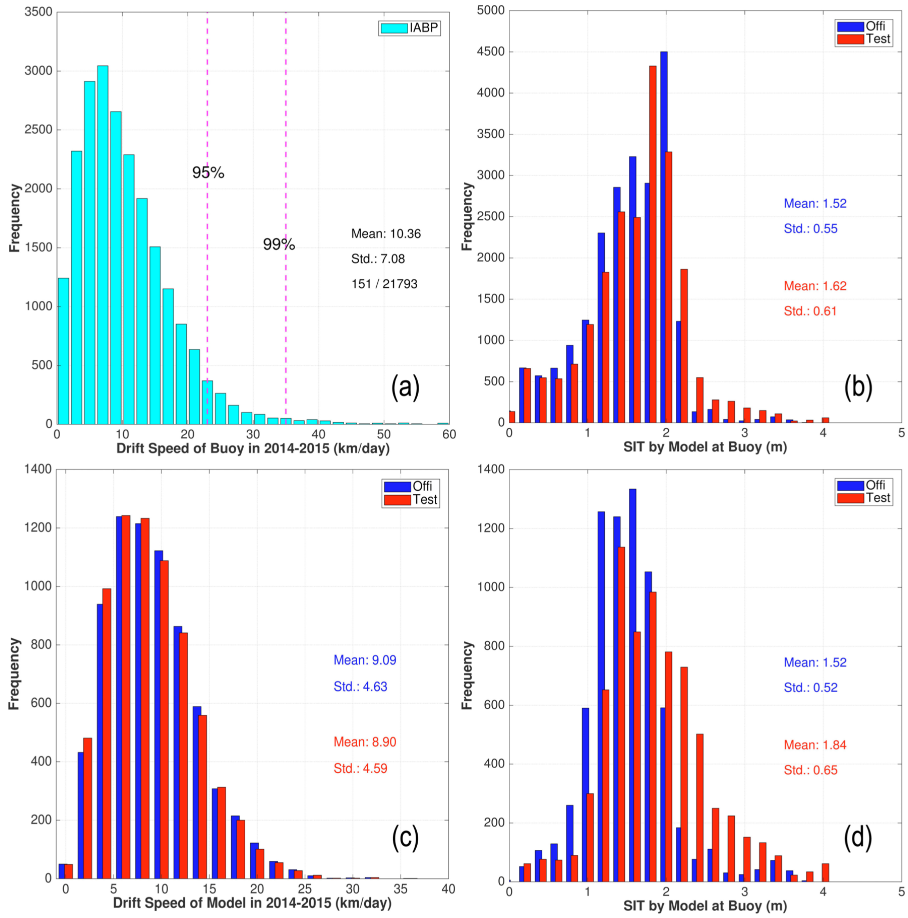

Fig. 10 (a) Histogram of sea ice drift speeds calculated from IABP buoys in the central Arctic for the period 2014-2015. (b) histogram of the simulated SIT at buoys locations in the central Arctic from the two runs. (c) histogram of the drift speed restricted near the North pole (>80N) in the Official (blue) and Test (red) runs; the mean speed and the standard deviation are indicated; (d) histogram of the simulated SIT near the North pole from the two runs;

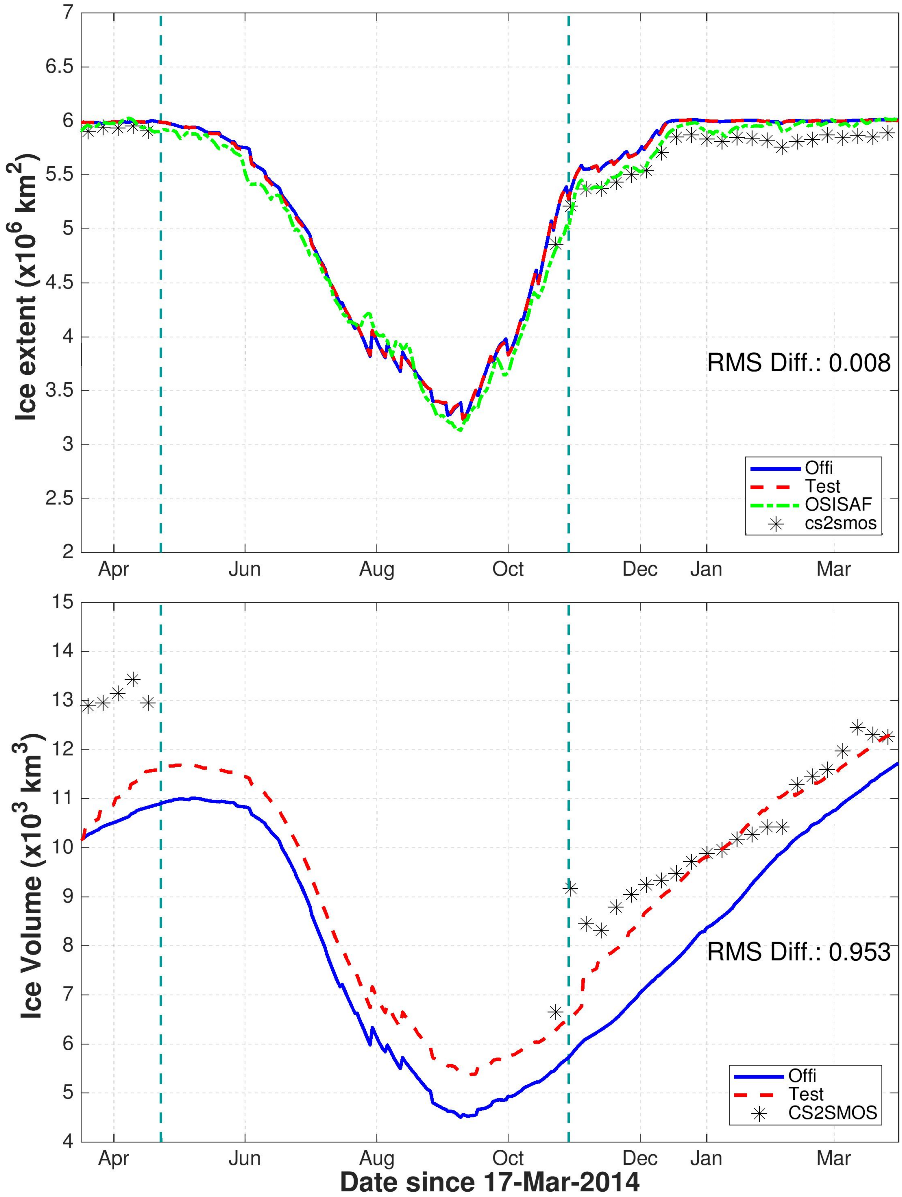

Fig. 11 SIE and SIV in the official run (blue) and the test run (red) in the Central Arctic. The black stars are the corresponding weekly SIE (or SIV) estimated from CS2SMOS. The green dash-dotted line is the daily SIE from OSI-SAF. The averaged differences of the two runs (Offi-Test) are reported. The vertical cyan-dashes delimits the periods when C2SMOS data is assimilated.

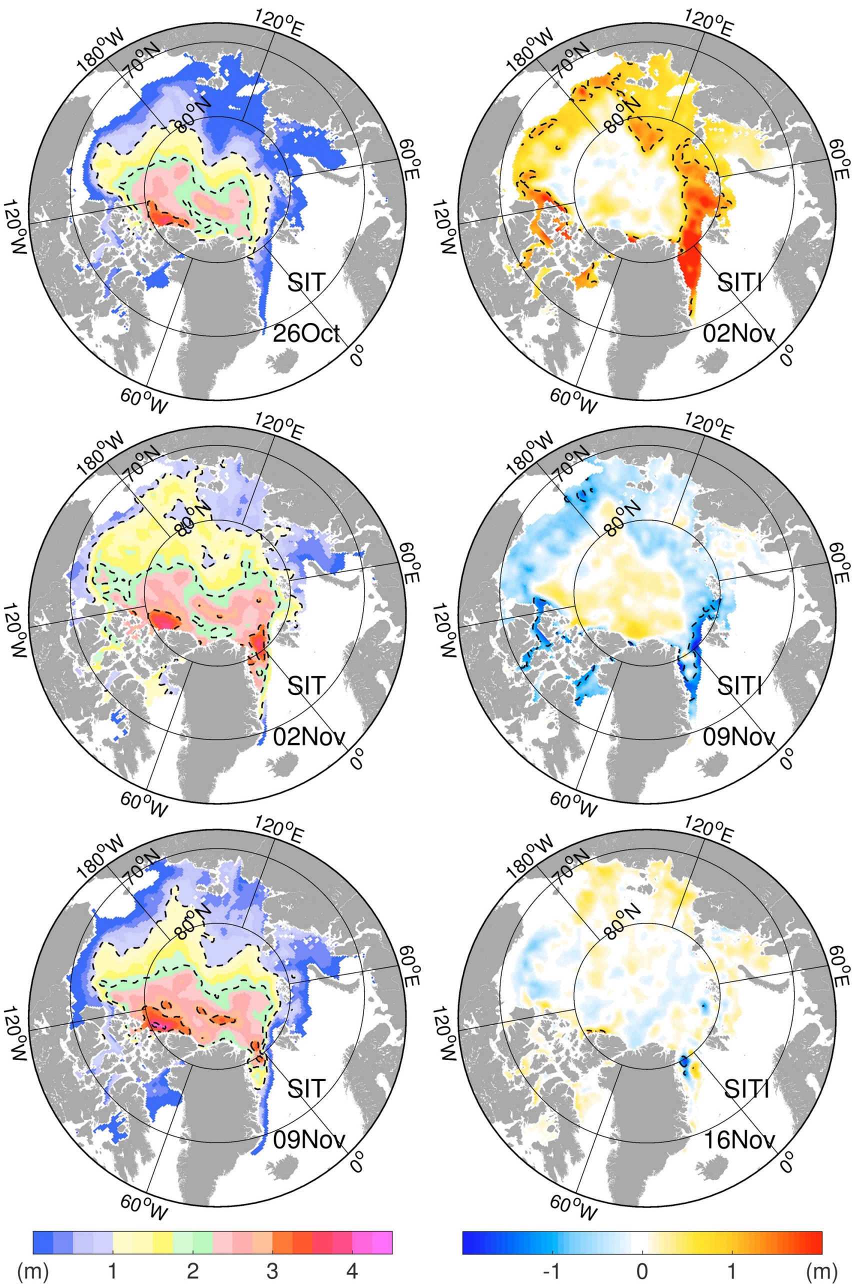

Fig. 12 Left: First three weekly SITs (20^{th}-26^{th} Oct; 27^{th} Oct-2^{nd} Nov; 3^{rd}-9^{th} Nov) from CS2SMOS in the beginning of fall 2014. The dashed white lines denote the 1, 2, 3, and 4 m isolines. **Right**: The associated time increments of SIT relative to the last weekly SIT. The dashed lines denote the -1 and 1 m isolines.

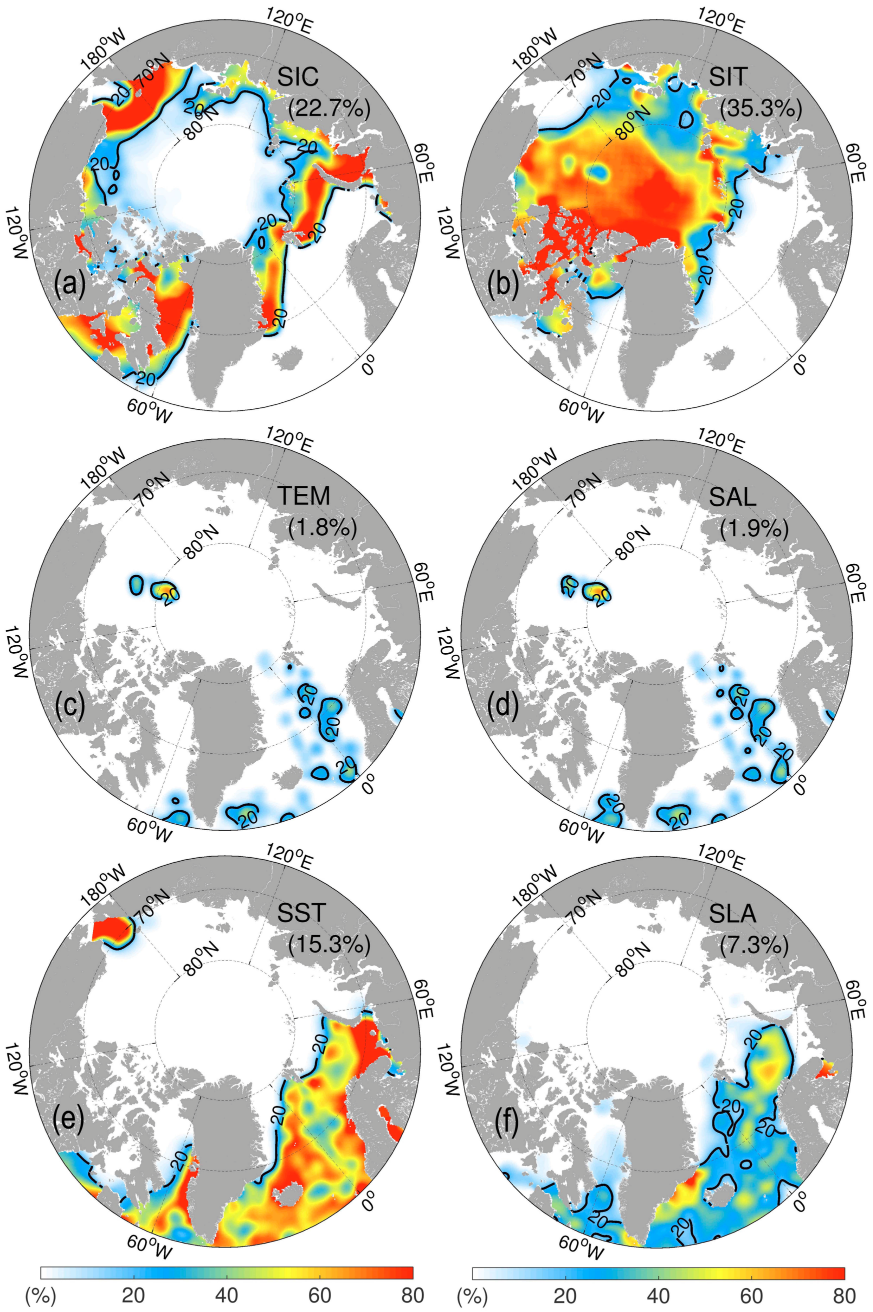

Fig. 13 Relative DFS contributions (IF) of each observation data types in November 2014. (a) SIC from OSI-SAF; (b) SIT from CS2SMOS; (c) temperature profiles; (d) salinity profiles; (e) SST; (f) along-track sea level anomaly (SLA). The black line is the 20% isoline, and the monthly IF (see Eq. 15) is reported between parenthesis.

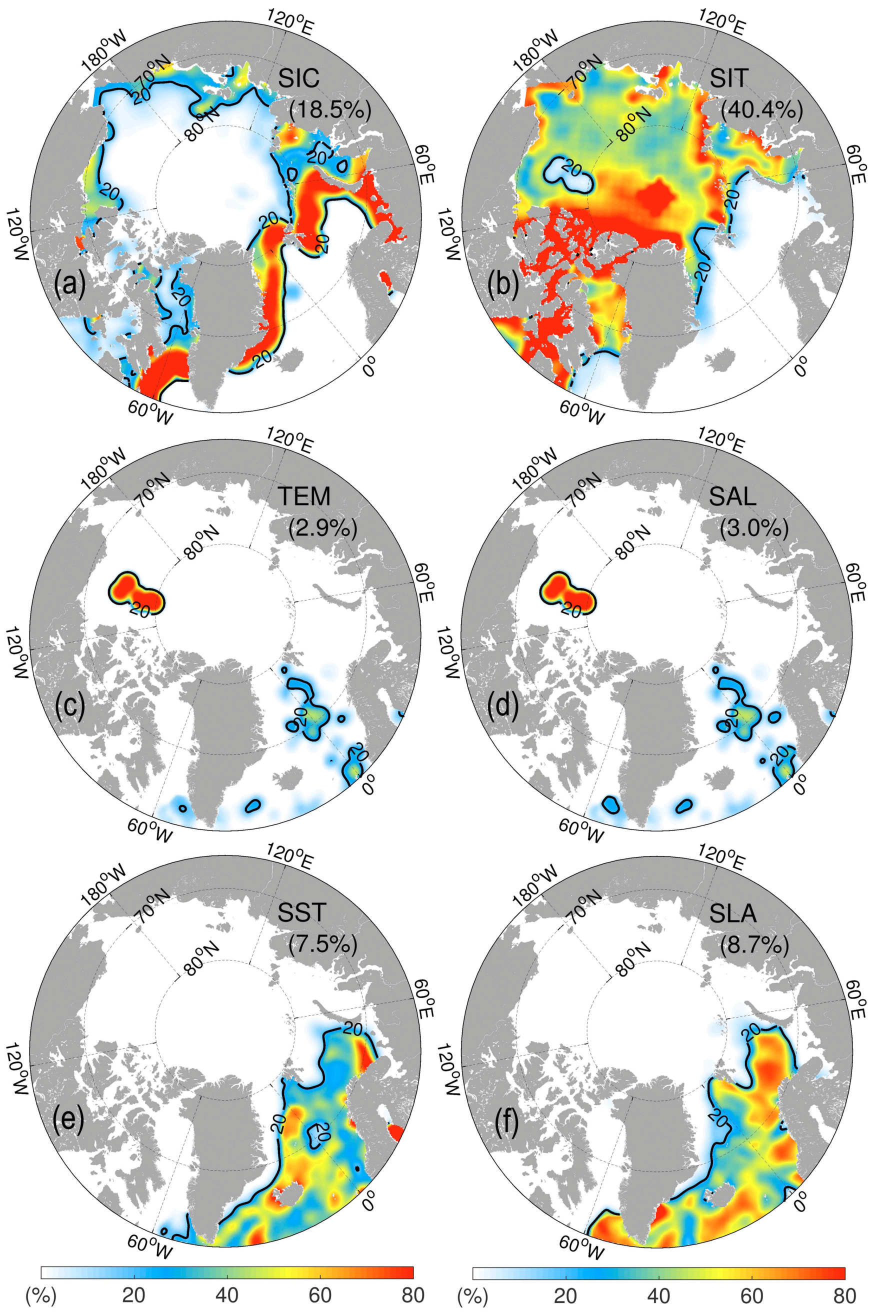

Fig. **14** Same as the above but for March 2015.