TECHNICAL MEMORANDUM

844

Towards Improved Sea Ice Initialization and Forecasting with the IFS

B. Balan Sarojini, S. Tietsche, M. Mayer, M. A. Balmaseda, and H. Zuo

Research Department

March 2019

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/research/publications

Contact: library@ecmwf.int

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Contents

List of Figures

Abstract

The ECMWF analysis system currently assimilates Level-4 sea ice concentration (SIC) from OSTIA (the Operational SST and Sea Ice Analysis produced by the UK Met Office). Here, we evaluate the impact of assimilating Level-3 SIC observations in the ECMWF ocean-sea ice analysis system. Furthermore, we make use of the availability of Arctic-wide sea ice thickness (SIT) observations in the recent years to constrain the modelled sea ice thickness. Coupled forecasts of the ocean-sea-ice-wave-land-atmosphere are then initialized using the improved sea-ice initial conditions from the above assimilation experiments, and the predictive skill of Arctic sea ice up to lead times of 7 months is investigated in a low-resolution analogue of the currently operational ECMWF seasonal forecasting system SEAS5. Results show that the system successfully assimilates Level-3 SIC observations from the OSISAF (EUMETSAT Ocean and Sea Ice Satellite Applications Facility) product OSI-401-b. Differences in the analysis are small and within the observational uncertainties, but the assimilation of Level-3 SIC will result in increased operational reliability. The impact on coupled forecasts is generally positive for SIC at lead month 1 and neutral for longer lead times. Statistically significant improvements are found over the ice edge and coastal seas in the Arctic mostly in the first 2 weeks for forecasts initialized in most calendar months, except for January starts, when the impact is neutral. The positive impact persists up to week 4 for March, May, August, November and December start months. For SIT and sea ice volume, the forecast impact of Level-3 SIC assimilation is neutral in all lead months.

Using SIT information from CS2-SMOS (CryoSat2-Soil Moisture and Ocean Salinity) as an additional constraint results in substantial changes of sea ice volume and thickness in the ocean-sea ice analysis. Forecasts started from these sea-ice initial conditions show a reduction of the positive sea ice bias and an overall reduction of summer-time forecast errors compared to SEAS5. A slight degradation in skill is found in the autumn sea ice forecasts initialized in July and August. While there is improvement in the skill of autumn 2m-temperature forecast initialized in spring, a degradation in skill is found for the October forecasts initialized in August. We conclude that the strong thinning by CS2-SMOS initialization mitigates or enhances seasonally dependent forecast model errors in sea ice and near surface temperatures. Hence, changes in root-mean-square errors are predominantly due to changes in biases. Using a novel metric, the Integrated Ice Edge Error (IIEE), we find significant improvement of up to 28% in the September sea ice extent forecast started from April. Our results demonstrate the usefulness of new sea ice observational products in both data assimilation and forecast verification, and strongly suggest that better initialization of SIT is crucial for improving seasonal sea-ice forecasts.

1 Introduction

Sea ice is an integral part of the Earth's climate system. Forecasting Arctic sea ice has advanced significantly in this decade, with most forecasting centres using prognostic sea-ice models operationally, allowing us to explore the sea ice forecast skill on long lead times from weeks to months to seasons. Possibilities of economically viable shorter shipping routes across the Arctic in the summer are constantly being explored. Monthly and seasonal outlooks of sea ice products are therefore in great demand especially by the Arctic maritime industry.

Moreover, there is increasing scientific evidence that sea-ice loss and warming in the Arctic due to climate change affect the European weather and climate (Balmaseda et al. (2010), Mori et al. (2014), Overland et al. (2016), Ruggieri et al. (2016)). Since reliable estimates of sea-ice extent and volume are needed for understanding climate change and for initializing numerical weather forecasts, there is growing interest in using improved and new types of sea-ice observations in data assimilation systems (Blanchard-Wrigglesworth et al. (2011), Tietsche et al. (2013)). Earlier studies propose that long-term memory in winter sea ice thickness can potentially improve summer sea ice extent forecasts (Guemas

et al. (2016), Tietsche et al. (2014), Day et al. (2014)). While assimilation of sea ice concentration (SIC) is routine practice in operational sea-ice forecasting, assimilation of sea ice thickness (SIT) is at its early stage (Allard et al. (2018)).

ECMWF has been using a prognostic sea ice model for operational coupled ensemble forecasts in the medium and extended-range since November 2016, for seasonal forecasts since November 2017, and for the high-resolution HRES forecast since June 2018. The ocean–sea-ice initial conditions are provided by the OCEAN5 assimilation system (Zuo et al. (2018), Zuo et al. (2019)), consisting of a real-time and a reanalysis component. The latter is called ORAS5, and it is used to initialize the reforecasts needed for products like the Extreme Forecast Index, and to calibrate the extended-range and seasonal forecasts. ORAS5 also contributes to the monitoring efforts of Copernicus Marine Monitoring Services (CMEMS). OCEAN5 uses variational data assimilation. Currently, sea-ice concentration is the only sea-ice parameter which is assimilated. Although the ECMWF sea ice reanalysis and reforecasts compare well with other systems (Chevallier et al. (2017), Uotila et al. (2018), Zampieri et al. (2018)), they are affected by noticeable errors (Tietsche et al. (2018)). There are large biases in sea-ice forecasts from months to seasons, pointing to uncertainties in both the models and observations used in the assimilation and forecasting systems. We explore the pathway to improve the initialization using lower level observations of sea ice concentration (Level-3) and novel observations of sea ice thickness. We then assess the impact of the changed sea ice initial conditions on the forecast skill on long lead times of months to seasons. This work has been motivated and supported by the EU H2020 project SPICES (Space-borne observations for detecting and forecasting sea ice cover extremes), which aimed at improved observations and forecasts of sea-ice cover extremes.

The rest of the report is organised in two sections. The first section describes the ocean-sea ice models, observations and assimilation experiments used for the sea ice initialization and presents the impact of new SIC and SIT observations on the analysed sea ice state. The second section explains the features of the forecasting system we use and assesses the impact of the changed sea-ice initialization on medium-, extended and long-range sea-ice forecast skill using newly available observations and sea-ice specific metrics.

2 Sea Ice Assimilation and Initialization

The sea-ice and ocean initial conditions used in our reforecast experiments are produced by a set of multi-year ocean–sea-ice assimilation experiments or ocean reanalyses. Ocean reanalyses (Carton et al. (2000), Weaver et al. (2003), Balmaseda et al. (2008), Balmaseda et al. (2013), Zuo et al. (2017), Tietsche et al. (2015), Zuo et al. (2018)) are historical reconstructions of the ocean state using an Ocean General Circulation Model (with or without a sea-ice model) forced by atmospheric fluxes and constrained by satellite and in-situ observations of the ocean using data assimilation methods. These spatially and temporally continuous, gridded observationally constrained estimates of the ocean state are highly useful as a) initial boundary conditions for numerical weather prediction systems, b) verification datasets for forecast evaluation, and c) climate monitoring datasets. Here we conduct ocean reanalysis experiments assimilating new sea-ice observations. The impact of these new observations is evaluated by assessing the quality of the resulting reanalyses (fit-to observations), and by assessing the impact on forecast skill when these new reanalyses are used as initial conditions of seasonal re-forecasts. Some results of forecast at shorter lead times are also presented.

2.1 Models and Methodology

We use the NEMOVAR ocean–sea-ice data assimilation system in a 3D-Var FGAT configuration (Mogensen et al. (2012)) to produce the reference and improved initial conditions. Assimilated observations comprise of temperature and salinity profiles, altimeter-derived sea level anomalies and sea ice concentration. Sea-surface temperature is constrained to observations by adjusting the surface heat fluxes. A global freshwater correction is added to reproduce the observed global-mean sea-level change.

In the 3D-VAR FGAT data assimilation approach, a first guess of the model is first calculated and compared to the observational state at appropriate time within the assimilation window. These misfits between model and observations are quantified in a cost function which is minimized using conjugate gradient method. To reduce potential shocks, the resulting assimilation increments are added continuosly to the model state with constant weights during a second model run (Incremental Analysis Updates, see Bloom et al. (1996)). The length of the assimilation window is 10 days in our experiments. Unlike the other ocean variables, SIC assimilation is univariate: the increments are derived by minimising a separate cost function as explained in Tietsche et al. (2015) and Zuo et al. (2017).

The ocean GCM used in these experiments is NEMO version 3.4 (Nucleus for European Modelling of the Ocean) (Madec (2008)) with an approximate horizontal resolution of approximately 1° and 42 vertical layers. The grid is tripolar, with the poles over Northern Canada, Central Asia and Antarctica enabling higher resolution across the Arctic than at the equator. The first model layer is 10 m thick, and the upper 25 levels represent approximately the top 880 m. Both the horizontal and vertical resolution in our setup is lower than that of the operational system, which has a horizontal resolution of approximately 0.25° and 75 vertical levels. The time step is one hour.

The prognostic thermodynamic-dynamic sea-ice model used is LIM2 (Louvain-la-Neuve Sea Ice Model) (Fichefet and Maqueda (1997)). The vertical growth and decay of ice due to thermodynamic processes is modelled according to the three-layer (one layer for snow and two layers for ice) Semtner scheme (Semtner (1976)). The ice velocity is calculated from a momentum balance considering sea-ice as a two-dimensional continuum in dynamical interaction with the atmosphere and ocean. Internal stress within the ice for different states of deformation is computed following the viscous-plastic (VP) rheology proposed by Hibler III (1979). LIM2 has a single sea-ice category to represent sub-grid scale ice thickness distribution, and open water areas like leads and polynyas are represented using ice concentration. Melt ponds are not modelled. LIM2 has a time step of 1 hour and is coupled to the ocean every time step.

2.2 Current operational Level-4 SIC issues and Level-3 SIC

Operational forecasts at ECMWF currently rely on ocean and sea-ice initial conditions from the ocean analysis OCEAN5 (Zuo et al. (2018)), which for years from 1985 onwards assimilates OSTIA SIC observations using the ocean–sea-ice data assimilation system as explained above. OSTIA (Operational SST and Sea Ice Analysis) system (Donlon et al. (2012)) is a daily SST-SIC observational product (available as a near-real time product after 2008 and as a re-processed product before 2008) produced by the UK Met Office. It is a global Level-4 product with ca. 6 km resolution. Level-4 is the last level of different stages of satellite data processing. The observations in the previous level, Level-3, have missing data, for instance if the satellite has a temporary malfunction, or if certain areas like the North Pole are not observed. The OSTIA Level-4 SIC analysis is derived from the EUMETSAT Ocean and Sea Ice Satellite Applications Facility (OSISAF) Level-3 observations.

The data infilling procedure used by OSTIA to convert Level-3 to Level-4 data can lead to erroneous

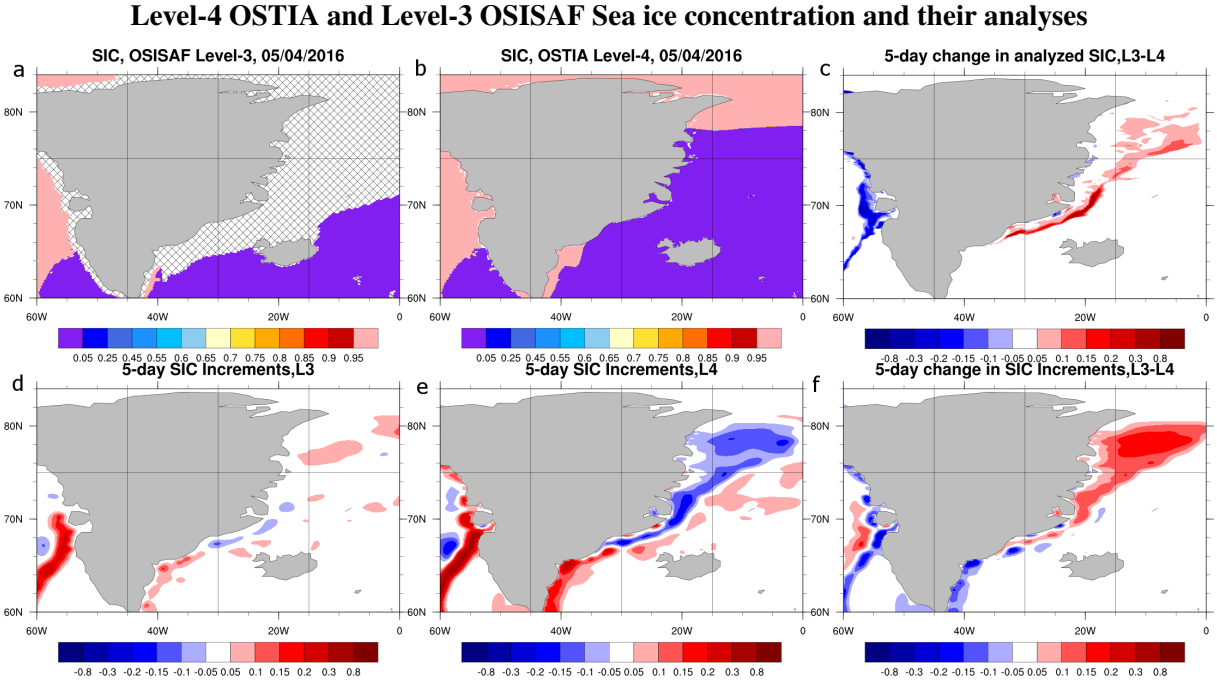

Figure 1: Sea ice concentration (SIC) on 5th April 2016: a) Level-3 OSISAF with hatching where data is missing, b) Level-4 OSTIA. Panel (c) shows the difference in analysed 5-day SIC mean between two analysis experiments assimilating Level-3 and Level-4 SIC respectively. SIC increments over a 5-day window spanning 5–9 April 2016: d) Level-3 assimilation, e) Level-4 assimilation (see text for details), and f) difference in 5-day SIC increments between Level-3 and Level-4 assimilation experiments. The impact of the erroneous Level-4 data on the SIC analysis is apparent in panels (c) and (f).

forecasts for days/areas when/where Level 3 data is not available. To illustrate this point, OSISAF SIC for the 5th of April 2016 is shown in Figure 1a. The swath data was corrupt or missing in a large area east of Greenland. The SIC data in OSTIA over the same region have infilled most of the area of missing data with zeros (Figure 1b). The SIC increments over the 5-day assimilation window spanning the period 5–9 April 2016 in assimilation experiments with either Level-3 (Figure 1d) or Level-4 (Figure 1e) SIC observations show increments in opposite signs north-east of Greenland, resulting in erroneous Level-4 SIC conditions (Figures 1c and 1f). Up to 20% more SIC (Figure 1c) is seen when assimilating Level-3 data compared to assimilating Level-4 data. The positive SIC increments (red) in 1e come from the SIC information in other days within the assimilation window (not shown). The area of SIC increase (red) is smaller in the difference in 5-day mean (Figure 1c) than that in the difference in 5-day increments (Figure 1f). This suggests that the model tries to compensate for the negative increments (Figure 1d) in the Level-4 SIC assimilation by producing more ice in the region, as has been found by Tietsche et al. (2013). This is a good example of where the assimilation increments of SIC directly impact on another model variable – sea-ice growth rates in this case.

There is continuous feedback between ECMWF and the OSTIA providers regarding the quality of the product, and most likely developments in the OSTIA algorithm will be put in place to mitigate the effect of missing data. However, now that ECMWF has a sea-ice model and data assimilation, the use of a Level-4 product is no longer justified. Beyond mitigating the impact of occasionally missing observations better, assimilating Level-3 data from the OSISAF product has several technical advantages compared to assimilating Level-4 data from OSTIA product: first, OSTIA has zero values for SIC instead of missing values over regions where sea ice is never formed such as over tropics, which increases the size of the SIC

control vector and hence introduces unnecessary computations when minimising the cost function. The OSISAF Level-3 product obviously has missing values outside the domain for which the observational retrievals were run. Second, the current implementation of the OSTIA SIC assimilation introduces an additional spatial interpolation onto the model grid, which is not needed with the Level-3 assimilation of OSISAF data, and introduces unnecessary interpolation errors. Third, it is a major advantage for future developments that the spatially and temporarily variable observational uncertainty information available with the OSISAF product can be used directly in the data assimilation. Last but not least, for operational applications Level-3 data can be made more timely available than Level-4 from OSTIA (about 1 day). Considering all of the above, ECMWF is planning to replace the Level-4 OSTIA SIC with the Level-3 OSISAF SIC for future operational updates of the ocean data assimilation system.

2.3 Observations and Ocean-Sea-Ice Assimilation experiments

Our reanalysis experiments are forced by near-surface air temperature, humidity and winds as well as surface radiative fluxes from the atmospheric reanalysis ERA-Interim (Dee et al. (2011)) until 2015 and from the ECMWF operational NWP from 2015 to 2016. We use the same set-up of NEMOVAR used in ORAS5 (Zuo et al. (2019)) except for the following changes: a) a coarser model resolution as described above, b) different assimilated sea-ice observations and, c) a longer assimilation window of 10 days instead of 5 days. The experiments presented here use the sea ice concentration product of the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF, `www.osi-saf.org` ; product identifier OSI-401-b (Tonboe et al. (2017))) and CryoSat2-SMOS (Ricker et al. (2017)) merged sea ice thickness observations.

Level-3 OSISAF SIC observations (OSI-401-b) are a daily real-time product produced by the Danish and Norwegian Meteorological Institutes under the EUMETSAT OSI SAF programme. The brightness temperatures are measured by the SSMIS instrument, which is a polar orbiting conically scanning microwave radiometer with constant incidence angle around 50° and a swath width of about 1700 km. The swath data is then processed by the OSISAF sea ice concentration algorithm to produce filtered and unfiltered SIC observations, which is the percentage of an ocean grid point covered by sea-ice and expressed in decimals. OSI-401-b SIC observations and its uncertainty estimates come in a polar stereographic grid of 10km horizontal resolution with varying pole hole size.

In order to assimilate the OSISAF Level-3 sea ice concentration with the ECMWF ocean-sea-ice data assimilation system, the following steps are necessary: 1) Pre-processing of the OSISAF Level-3 SIC data to convert gridded data into point observations that is compatible to the NEMOVAR feedback format. During this step, observations can be rejected if they do not pass quality control. 2) A stratified random sampling method has been applied to SIC data for observation thinning and perturbation, with pre-defined reduced grids. This is similar to what is done when assimilating sea-level anomalies as implemented in ORAS5 (Zuo et al. 2017), and is essential to ensure that a reasonable number of Level-3 observations (comparable to the number of Level-4 observation) will be assimilated. Other settings remain unchanged to ORAS5 (Zuo et al. (2017), Tietsche et al. (2015)), in particular the constant SIC observation error of 0.2.

While SIC has been monitored for several decades, satellite-remote sensed sea-ice *thickness* observations have only been available for less than a decade (Kwok and Rothrock (2009), Laxon et al. (2013), Kaleschke et al. (2012)). A recent initiative led by the Alfred Wegener Institute provides a merged Level-4 product of Arctic-wide winter ice thickness that combines thick-ice retrievals by CryoSat2 (CS2) satellite and thin-ice retrievals by the Soil Moisture and Ocean Salinity (SMOS) satellite. It is the first ever multi-sensor ice thickness product for the Arctic. CS2 (Hendricks et al. (2016)) measures freeboard (the

height of the ice or snow surface above the water level) using altimetry, whereas SMOS (Tian-Kunze et al. (2014)) measures microwave brightness temperatures in the L-band. Both measurements are converted to ice thickness in metres. Low relative uncertainties are present for ice thinner than 1 m in SMOS retrievals and for ice thicker than 1 m in CS2 retrievals, making them represent the entire thickness range covering the whole Arctic (Ricker et al. (2017)). CS2 and SMOS are merged using an optimal interpolation scheme to produce the CS2-SMOS product, which is available on a weekly basis on an EASE-2 grid with 25km horizontal resolution. It is only available for 5 full months from November to March every year during 2011–2017. Both the CS2 and SMOS retrievals are not possible in the melt season due to the presence of melt ponds, and wet and warm snow and ice.

The above SIT observational information has been used to constrain the simulated SIT after interpolating the weekly CS2-SMOS observations in EASE2 grid to daily data in the ORCA 1° grid. The weekly to daily interpolation is done by appropriately weighting two adjacent weekly-mean fields. Instead of using the variational data assimilation, the constraint is imposed via a linear nudging technique, which is easy to implement for gridded products. The relationship between the modelled and observed sea ice thickness is described by the following equation:

$$SIT^n = SIT^m - [\frac{\Delta t}{\tau}(SIT^m - SIT^o)] \tag{1}$$

where SIT^n is the nudged thickness, SIT^m is the modelled thickness, SIT^o is the observed thickness, Δt is the sea-ice model time step of 1 hour, and τ is the nudging coefficient corresponding to a relaxation time scale of 10 days. The choice of a 10-day time scale makes sense as a first trial, since it is consistent with the length of the assimilation window.

In order to assess the impact of new sea-ice observations on the assimilation, we carry out three data assimilation experiments (reanalyses) as shown in Table 1. They are 1) a reference experiment with current Level-4 SIC assimilation (L4), 2) an experiment with Level-3 SIC assimilation (L3), and 3) an experiment with Level-3 SIC assimilation and sea ice thickness constraint (L3T). Experiments L4 and L3 are run for the years 2005 to 2016. Experiment L3T is run for the years 2011 to 2016, because these are the years for which CS2-SMOS observations were available at the time of experimentation.

Reanalysis	SIC observations assimilated	SIT constraint	Time period	Description
L4	OSTIA SIC	No	2005-2016	Level-4 SIC assimilation
L3	OSISAF SIC	No	2005-2016	Level-3 SIC assimilation
L3T	OSISAF SIC	Yes	2011-2016	Level-3 SIC assimilation and SIT nudging

Table 1: Specifications of the ocean-sea ice assimilation experiments

We have also tested the sensitivity of different nudging strengths by running variants of L3T with a relaxation time scale of 20, 30 and 60 days, but chose the experiment with the strongest constraint (10-day relaxation time) for initializing the ensemble reforecasts. For the L4 and L3 experiments, sensitivity to high resolution has also been tested by running experiments at the higher resolution of the operational data assimilation system ($1/4°$ with 75 vertical levels). Test results indicate that large-scale results are similar between high- and low-resolution experiments. As explained in the next section, the largest impact is due to the thickness nudging, i.e. between L4 and L3T, rather than between L4 and L3. Therefore, the main focus of our study will be on the difference between the L4 and the L3T experiments.

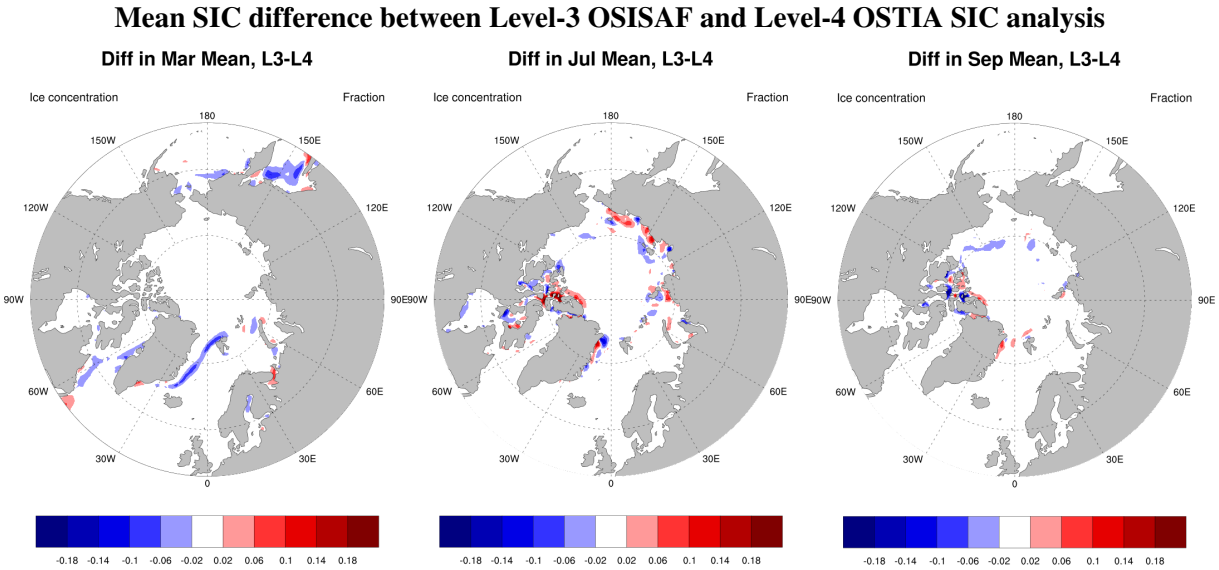

Figure 2: Difference in monthly-mean sea ice concentration between L3 and L4 for March (left), July (middle) and September (right), averaged over the years 2005–2015.

2.4 Impact of improved observations on the sea ice state

2.4.1 Level-3 sea ice cover assimilation versus Level-4 sea ice cover assimilation

Assimilating Level-3 SIC observations has only a small impact on the analysed SIC. Figure 2 shows the mean difference in SIC between L3 and L4 (L3 minus L4), for March, July and September averaged over 2005–2015. The differences are mostly within the observational uncertainty. Regional patterns of reduced sea ice concentration of up to 8 to 10% is seen over the Sea of Okhotsk and north east of Greenland along the sea-ice edge in March. Reduced and increased SIC patterns over Canadian Archipelago, Siberian shelf seas and other coastal seas is seen in the melt season (July). During the annual sea-ice minimum in September, smaller patterns of changes in the Canadian Archipelago region are seen. The pattern of change originates from the SIC increments (not shown) confirming the impact of Level-3 SIC assimilation.

The indirect effect of Level-3 SIC assimilation on the analysed ice thickness suggests robust pattern of change, but this is a small effect and mostly within the observational uncertainty of 0.5m (not shown). Similar results are also obtained for analysed Sea Ice Volume (SIV) - changes are within the observational uncertainty of 1500 km^3 (Tilling et al. (2015) (see Figure 3). Time evolution of SIV shows identical results for both the assimilation experiments with notable interannual variations in the melt season.

In conclusion, the Level-3 assimilated sea ice state is very similar to the Level-4 state, which indicates the assimilation system is robust. Since the differences are often within the observational uncertainties, it is difficult to ascertain from the reanalyses themselves whether L3 is improved with respect to L4. However, we will see later that forecasts from L3 are detectably better than forecasts started from L4. Furthermore, L3 offers many technical advantages, as discussed in Section 2.2

Time evolution of November Sea Ice Volume

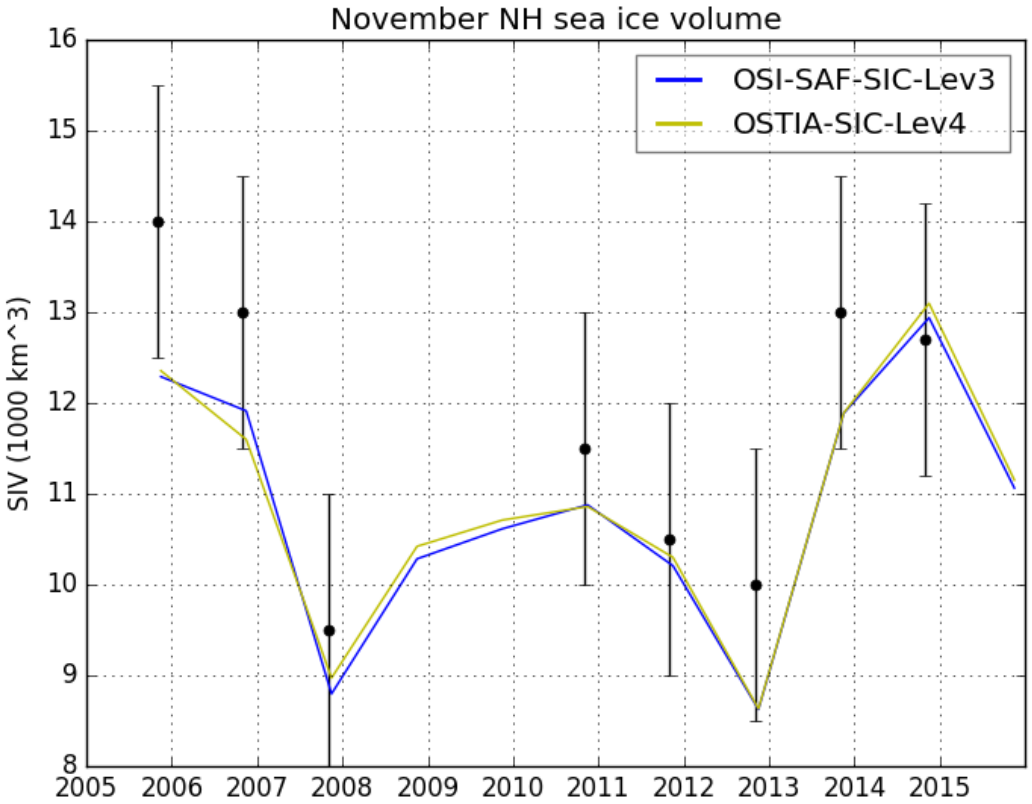

Figure 3: Time series of sea ice volume ($\times 10^3 km^3$) in November, L3 reanalysis (blue) and L4 reanalysis (yellow) evaluated against IceSat (Kwok and Rothrock (2009)) and CryoSat2 (Tilling et al. (2015)) observations.

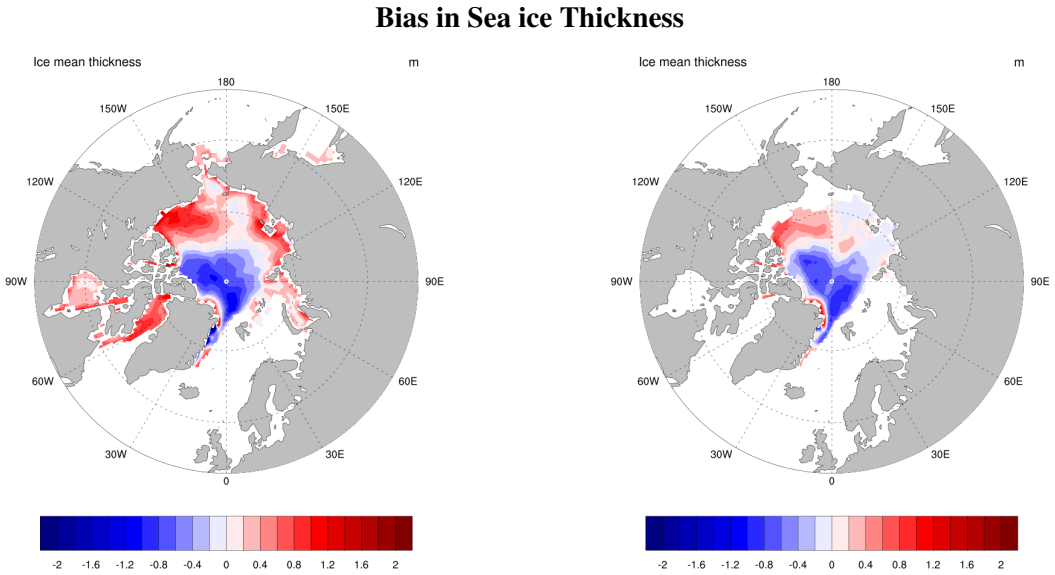

Figure 4: Difference in March (left) and November (right) sea ice thickness (m) between L3 and CS2-SMOS observations averaged over 2011-2016.

2.4.2 Sea ice thickness constraint against Level-4 sea ice cover assimilation

Both the L4 and L3 reanalyses suffer from large ice thickness biases of up to 1.4 m when validated against the CS2-SMOS observations. Predominant patterns for all available months are an underestimation of ice thickness ($>$ 1 m) in the central Arctic, and an overestimation in the Beaufort Gyre of the order of 1 m (Figure 4 for November and March multi-year means). In March, widespread overestimation in the coastal Arctic seas is present.

The impact of SIT nudging is shown in Figure 5. An example month of January in multi-year mean shows similar bias patterns (left), and how much the thickness is constrained with the strongest nudging strength of 10 days (right): the large-scale pattern of underestimation and overestimation of sea ice in L4 is not present in L3T. Time series of March SIV in L3 and in the experiments with different thickness nudging strengths (not shown) suggests that the experiment with 10-day nudging strength has the least sea ice volume and is closest to the CS2-SMOS observations. As noted earlier, out of our 4 nudging experiments the initial conditions of 10-day relaxation time scale is chosen to initialize our seasonal reforecasts.

3 Sea Ice Forecasting

Currently, ECMWF seasonal forecasts of sea-ice parameters are made operationally every month using the seasonal forecasting system SEAS5 (https://www.ecmwf.int/sites/default/files/medialibrary/2017-10/System5_guide.pdf). SEAS5 (Stockdale et al. (2018), Johnson et al. (2018)) consists of the same ocean and sea-ice model (NEMO3.4/LIM2) used for our reanalyses experiments, and is coupled to the ECMWF atmospheric model IFS (Integrated Forecast System). IFS is run with a horizontal resolution of 36 km, corresponding to a cubic octahedral reduced Gaussian grid at truncation TCo319. It has 91 vertical levels, and the top of the atmosphere is around 80 km. SEAS5 also includes the land surface model CHTESSEL (Carbon-Hydrology Tiled ECMWF Scheme for Surface

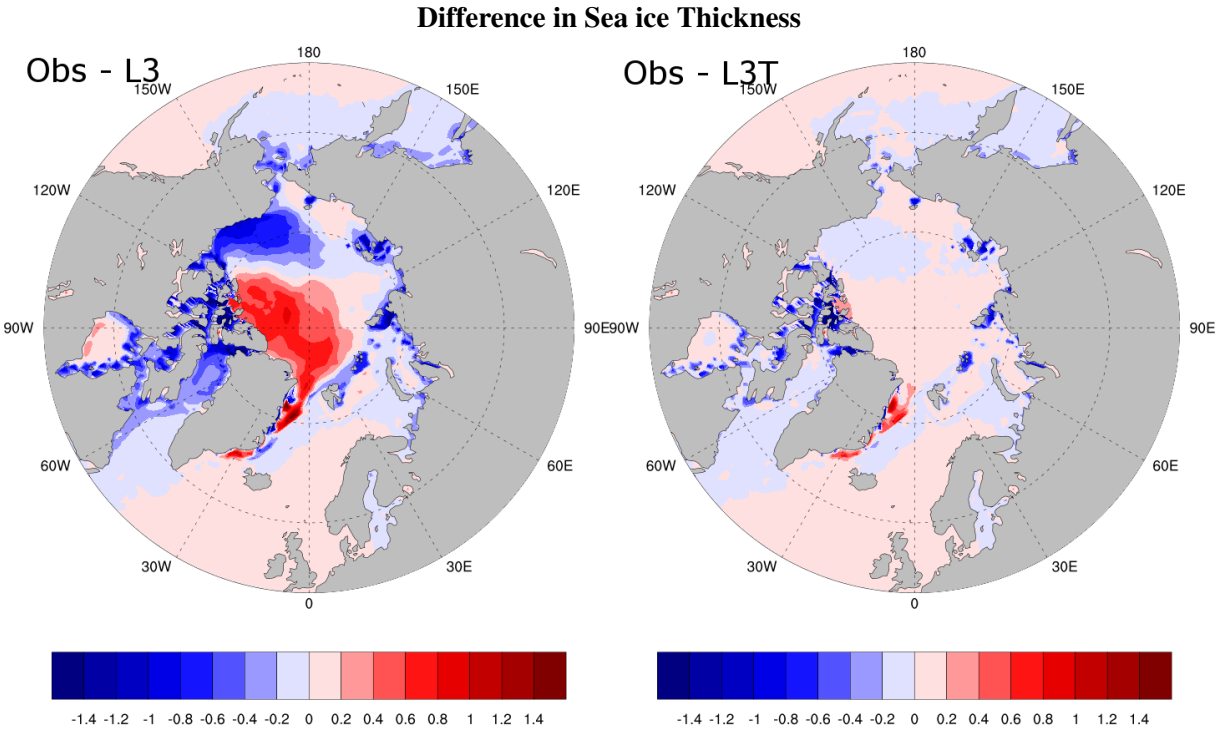

Figure 5: Difference in January mean (2011-2016) sea ice thickness (m) between (left) L3 and CS2-SMOS and (right) L3T and CS2-SMOS. L3T is the assimilation experiment with a thickness relaxation time scale of 10 days.

Exchanges over Land) and the ocean surface wave model WAM. The coupling of the atmosphere and ocean is done using a Gaussian interpolation method, and the coupling frequency is 1 hour.

The SEAS5 forecasts have 51 ensemble members. They are started from the 1st of each calendar month and run for 7 months ahead. The ocean and sea-ice initial conditions are provided by the real-time equivalent of ORAS5 (Zuo et al., 2018). As the model is run freely for many months to produce the forecast in the seasonal range, model errors play an important role in the outcome. One of the operational ways to tackle model biases is to calibrate the ensemble forecast output at the post processing stage. For this, retrospective hindcasts (the so-called reforecasts) are run for the past 36 years (1981 - 2016). These reforecasts are also used to assess the skill of the seasonal forecasting system.

3.1 Sea-ice related SEAS5 biases

To put our results into context, some aspects of the SEAS5 biases are first discussed. Figure 6 gives an overview of bias in sea ice area w.r.t. ORAS5 reanalysis as a function of forecast initial and target month (units are $10^6\,\mathrm{km}^2$). Error at lead month 1 are generally small throughout the year. However, for longer lead times, there is a strong over-prediction of sea ice area in summer months. Also, a moderate under-prediction of autumn sea-ice conditions can be seen. Delayed refreeze of sea ice in autumn could be the reason. The forecast biases are generally small in winter months. These biases are discussed in more detail in Stockdale et al. (2018)

The exact cause of these biases is still under investigation. However, it seems that surface radiation errors caused by a known model bias of excessive low-level clouds in the Arctic are a main contributor. This is illustrated in Figure 7, which for the month of July shows the differences between the downwelling

Current Bias in SEAS5 sea-ice area forecasts

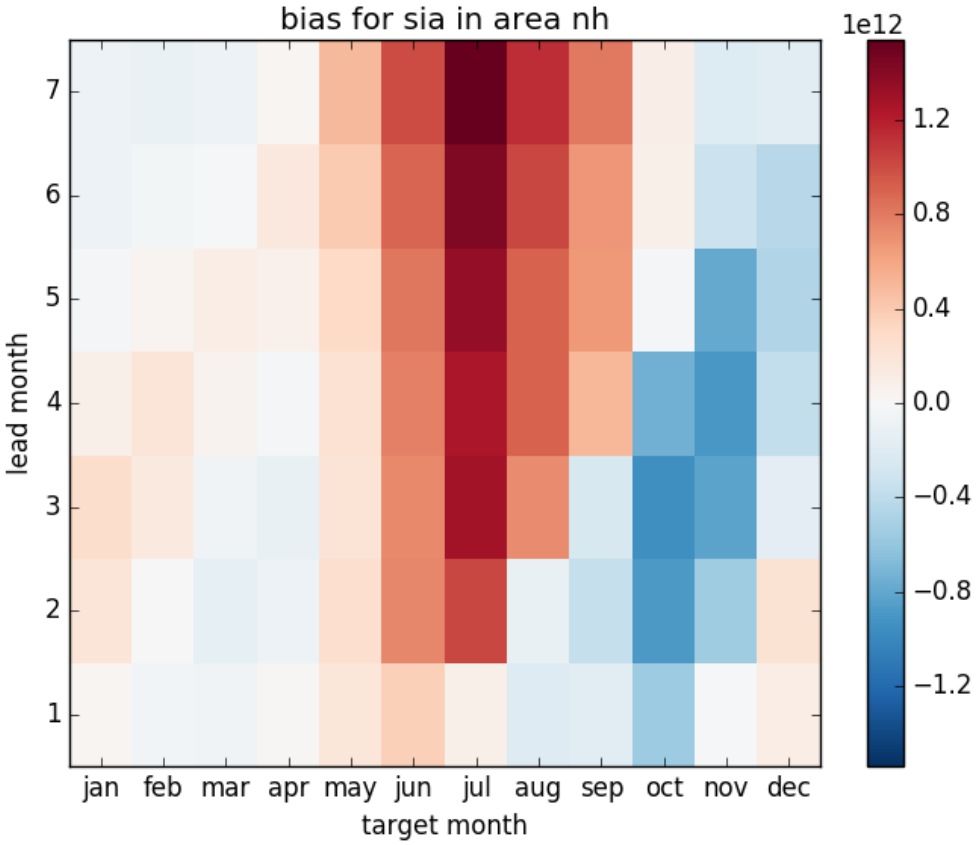

Figure 6: Bias in forecast sea-ice area ($\times 10^{12} m^2$) in SEAS5 w.r.t. ORAS5 as a function of initial and target month. Red colour denotes over-prediction of sea-ice area, and blue colour denotes under-prediction.

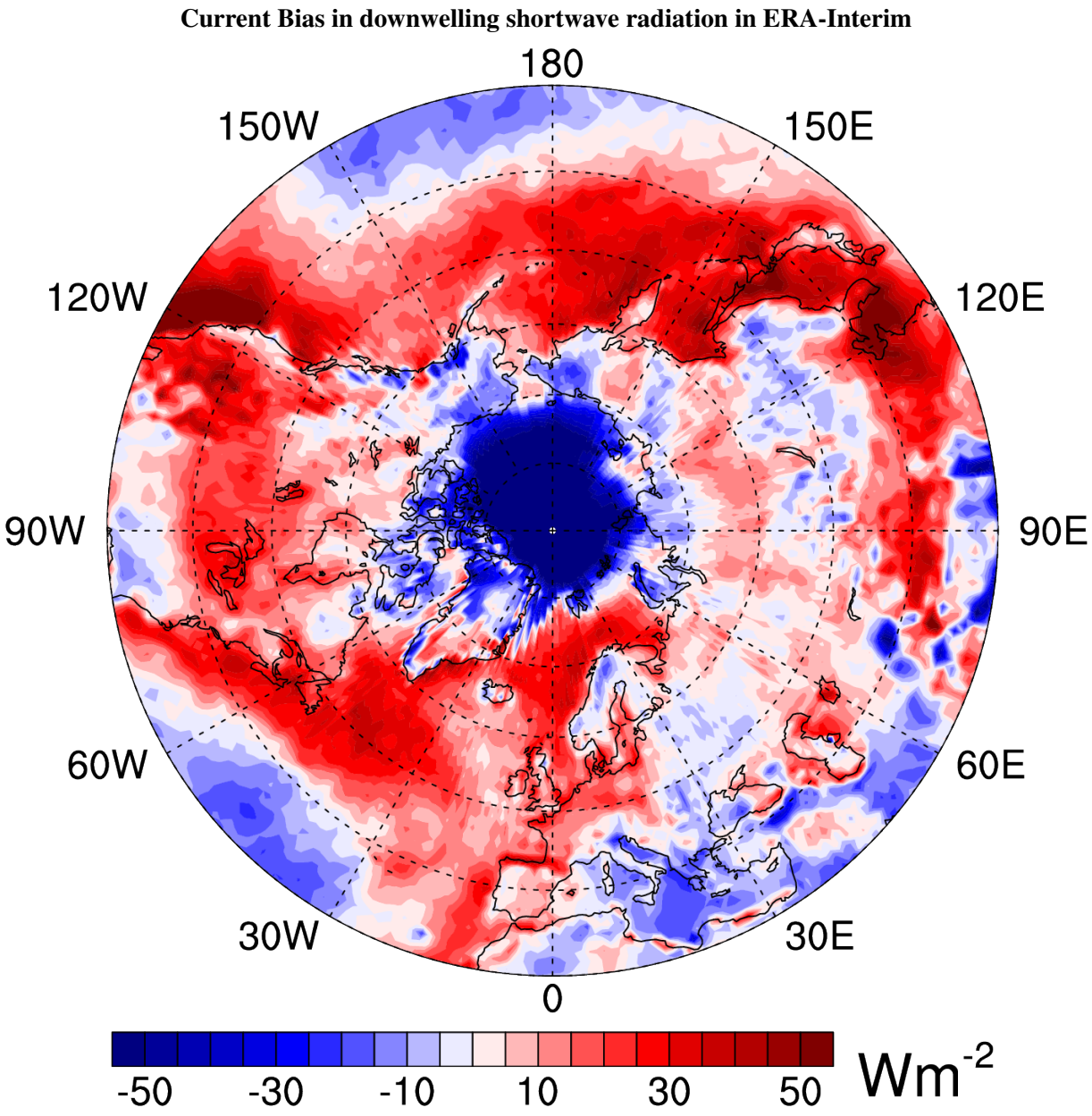

Figure 7: Biases in downwelling short wave radiation at the surface ($Wm^{-}2$) in ERA-Interim reanalysis w.r.t CERES observational estimates, July, 2010-2015. Blue colour denotes underestimation in downwelling short wave and red colour denotes overestimation.

shortwave radiation at the surface in ERA-Interim against observational estimates based on CERES (Clouds and the Earth's Radiant Energy System) (Kato et al. (2018)). The dark blue colours over the Arctic indicate a strong underestimation of downwelling shortwave radiation of about 50 W/m^2. This lack of downwelling shortwave radiation at the surface could be causing the positive sea ice area bias in summer (Hogan et al. (2017)).

3.2 Coupled reforecasts with improved initialization in SEAS5

In order to assess the impact of Level-3 sea-ice cover initialization and the CS2-SMOS sea ice thickness initialization on sea-ice forecast, we performed 3 sets of reforecast experiments as shown in Table 3.2, which are started from the corresponding data assimilation experiments shown in Table 1. They are carried out using a low-resolution version of SEAS5. The only differences between the reforecast experiments is their ocean-sea ice initial conditions, which are produced using the reanalyses discussed in Section 2: 1) as in current operations, a reference reforecast with Level-4 OSTIA SIC initialization (FC-L4), 2) a reforecast with Level-3 OSISAF SIC initialization (FC-L3), and 3) a reforecast with both Cryosat2-SMOS thickness and Level-3 OSISAF SIC initialization (FC-L3T). For 1) and 2) the reforecasts are performed from 12 start dates, i.e., on the 1st of each month of each year for 2005-2016 period. In order to get a robust estimate of the impact of thickness initialization, the reforecasts for 3) are performed for all start months of each year in the 2011-2016 period when the CryoSat2-SMOS winter observations are available. The forecast initialized from each start date has 25 ensemble members for all the 3 sets of reforecasts.

Experiment	Start years	Lead months	Ens. size	Init. cond.	Description
FC-L4	2005–2016	7	25	L4	Level-4 SIC initialization
FC-L3	2005–2016	7	25	L3	Level-3 SIC initialization
FC-L3T	2011–2016	7	25	L3T	Level-3 SIC and SIT nudging initialization

Table 2: Overview of the reforecast experiments. For each of the start years, forecasts are started on the 1st of every calendar month.

3.3 Sea-ice forecast impact of Level-3 sea-ice concentration assimilation

Seasonal forecasts of ice edge are in great demand for exploring economically viable Arctic shipping routes. Integrated Ice Edge Error (IIEE) is one of the new user-relevant sea ice metrics on ice extent or ice edge (Goessling et al. (2016), Bunzel et al. (2017)). Since it can be decomposed into over- and under-prediction, it is more useful than the traditional basin-wide sea ice extent error.

As expected, the difference between SIC forecasts started from L3 initialization and L4 initialization is small overall, especially beyond the first forecast month. Figure 8 shows the IIEE of FC-L3 and FC-L4 when verified against OSI-401-b observations. For monthly means in a long-range forecast (Figure 8b), there is a small reduction in seasonal ice edge error in FC-L3, which is not statistically significant. However, when considering daily forecast samples over all start months and years for lead month 1, improvements in FC-L3 forecast skill are statistically significant (black triangles) at the 5% level (Del-Sole and Tippett (2016)) for most lead days in the medium-range (Figure 8a). The few days when skill

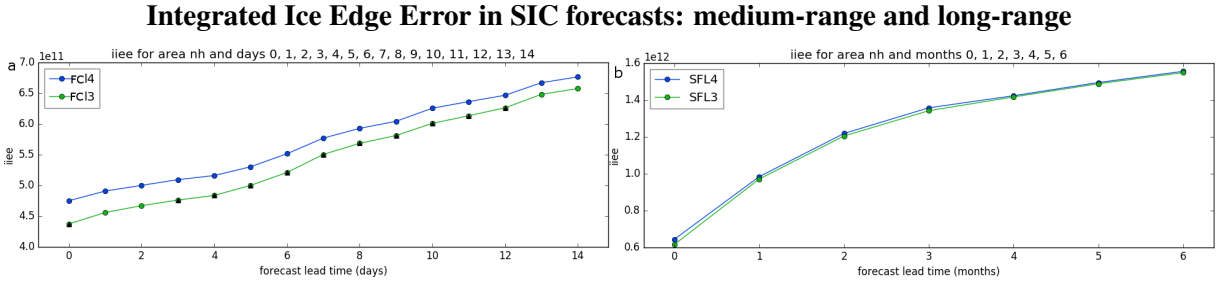

Figure 8: Integrated ice Edge Error in SIC forecasts averaged over the Northern hemisphere over all start months in years 2006–2015 in FC-L3 and FC-L4 verified against OSI-401-b observations. Panel (a) shows daily means the medium-range (units are $10^{11} m^2$), and panel (b) shows monthly means for the seasonal range (units are 10^{12} m^2). Improvement in forecast skill in FC-L3 is significant at the 5% level for most days (marked with black triangles) when considering a large sample (N=120) of daily forecasts of the first 2 weeks of all start months. The non-significant forecast days are border-line cases.

improvements are not significant are border-line cases where test statistic is very close but below the threshold for rejecting the null hypothesis of same-skill forecasts.

Daily forecasts in the short-range show an ice-edge error reduction of about 0.04×10^6 km^2 with L3 initialization (Figure 8a). Similar significant improvements are obtained when verifying the reforecasts against their own analysis (not shown). Errors of 0.47 and 0.43 $\times 10^6$ km^{2}are seen at day 1 in FC-L4 and FC-L3, respectively. When verified against the L3 analysis, the error in FC-L3 is slightly reduced to 0.4×10^6 km^2. There is a clear initialization shock for sea-ice cover in our experimental setup: The ice-edge error increases an order of magnitude faster during the first day of the forecast than it does during subsequent days. This is not the case in operational SEAS5 and ENS forecasts: using the exact same diagnostic methods as for FC-L3 and FC-L4, we find that SEAS5 and ENS forecasts have a day-1 IIEE of about 0.08×10^6 km^2. This initialization shock should be kept in mind when interpreting the results of the reforecast experiments discussed here.

When grouping daily forecasts for each start month separately (Figure 9), we find an overall improvement in SIC forecast skill in FC-L3 for all start months except for January. Forecasts started in February, March, April, May, August, September, October, November and December are statistically significant (black triangles) throughout the medium range. For forecasts started in March, May, August, November and December, skill improvements are statistically significant also in the extended range. January forecasts consistently show neutral impact in all the years. A non-significant degradation (light red) seen in week 4 for October start months occurs mainly over Kara Sea and north-west Laptev Sea. This becomes neutral in the area means of bias-corrected mean absolute error (not shown). Overall, the significance test rejects the null hypothesis that the forecast skill of FC-L3 and FC-L4 forecasts are the same with 95% confidence for most start months in the medium-range, and for March, May, August, November and December start months in the extended-range.

Maps of changes in root-mean-square error (RMSE) for monthly and daily means of forecasts started in May (when significant improvements in forecast skill are found) show that the improvements are seen over the ice edge and coastal seas in the Arctic when verified against both OSI-401-b observations (Figure 10) and ORAS5 (not shown). Similar patterns of improvement over large regions and degradation over isolated regions are found for lead month 1 and lead days (e.g. day 3 in Figure 10a). Improvements in SIC RMSE of up to 18% are seen close to the ice edge in both the Atlantic and Pacific sectors of the Arctic.

Forecasts over-predict sea ice volume (SIV) for May start dates (not shown), which is consistent with

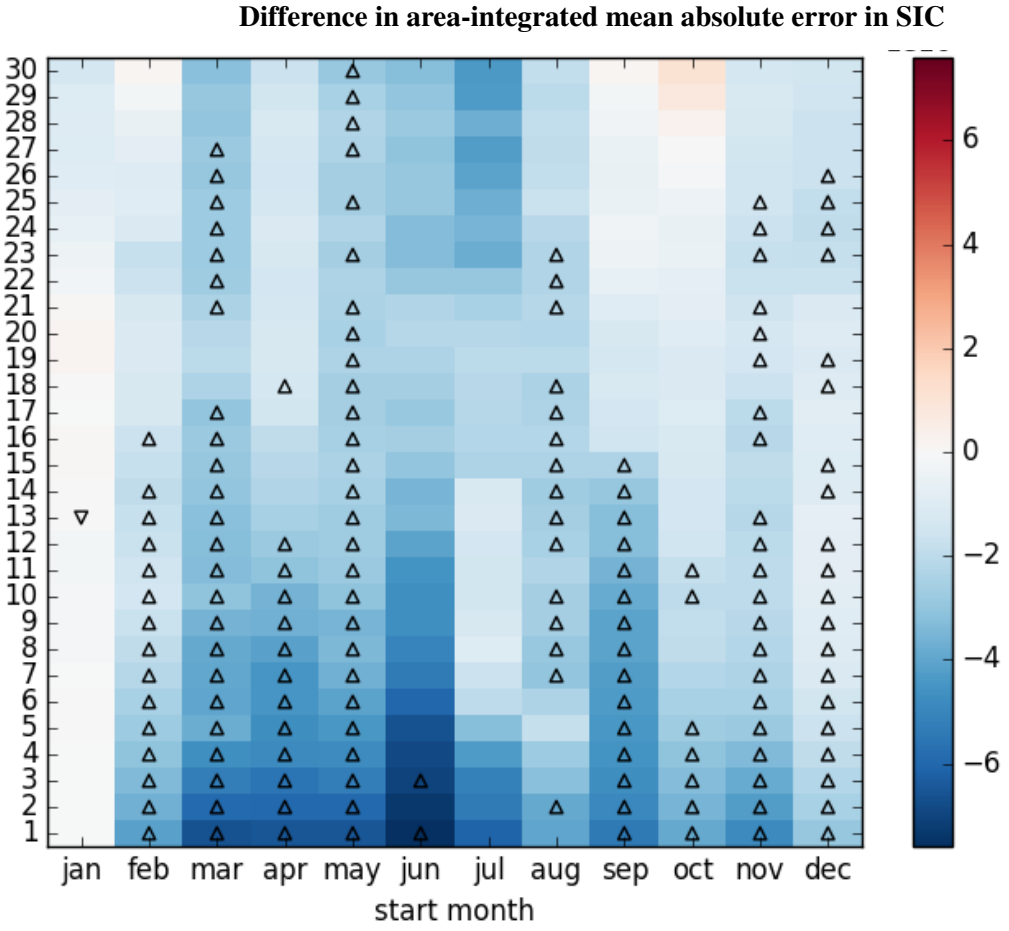

Figure 9: Difference in NH area-integrated mean absolute error in 10^{10} m^2 *between FC-L3 and FC-L4 SIC reforecasts for the first 30 forecast days w.r.t. OSI-401-b observations (2006-2015). Blue colour denotes reduced error in sea ice cover in FC-L3, and red colour denotes increased error in FC-L3. Black triangles represent statistical significance at the 95% confidence level from the test described in DelSole and Tippett (2016).*

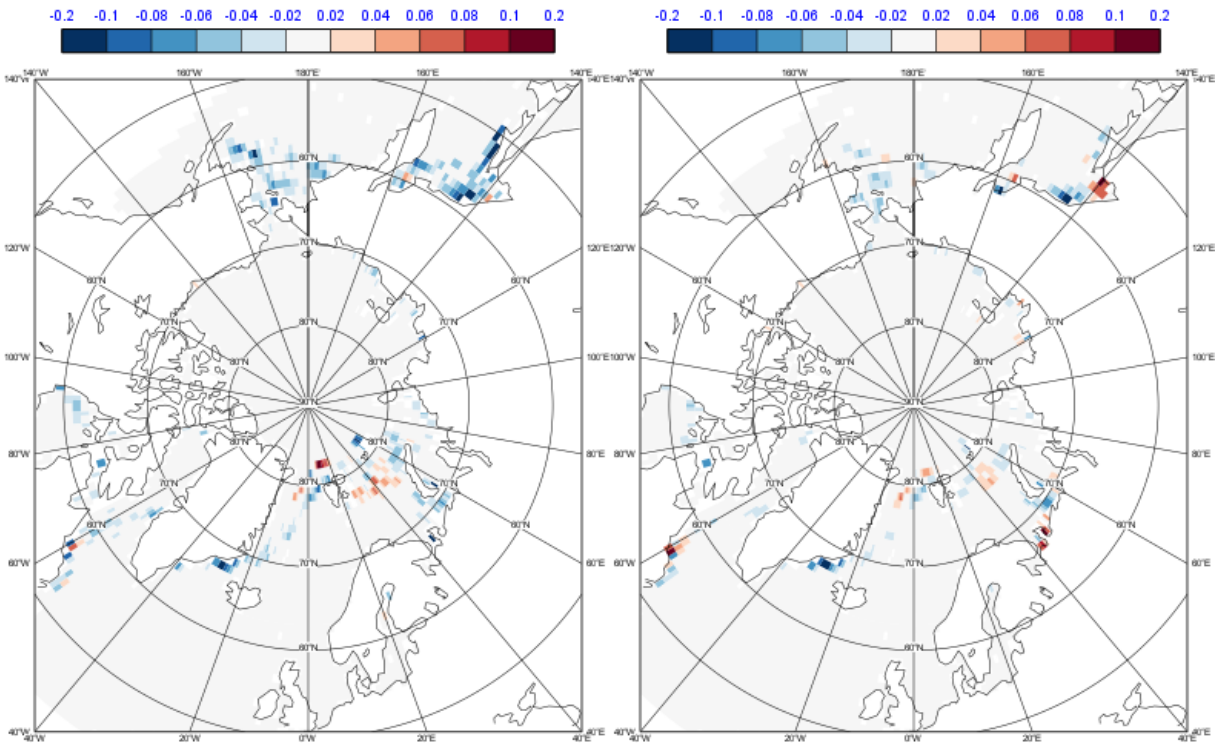

Figure 10: Difference in root mean square error between FC-L3 and reference FC-L4 SIC May reforecasts (2006-2015) w.r.t OSI-401-b, at lead month 1 (right), and at lead day 3 (left). Blue colour denotes reduced RMSE in FC-L3 and red colour denotes increased error in FC-L3.

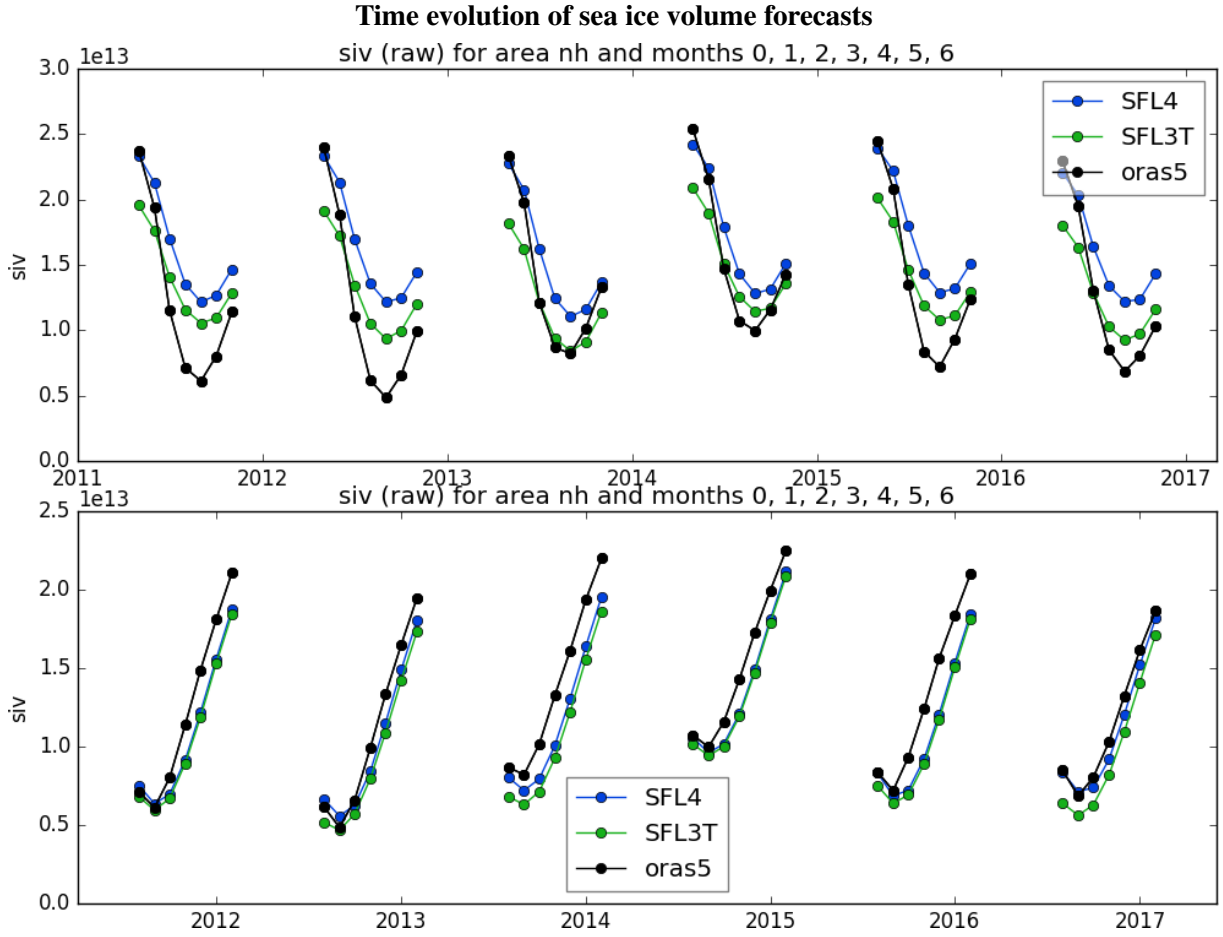

Figure 11: Time series of sea ice volume (units are 10^{12} m^3) forecast for May start date (top) and August start date (bottom) in reference Level-4 (FC-L4, blue) and Level-3 with thickness initialization (FC-L3T, green) reforecasts compared to ORAS5 (black). Each dot represents sea ice volume for each lead month.

the summer-time sea-ice area bias in SEAS5 discussed in Section 3.1. The only small difference seen in SIV forecasts between FC-L3 and FC-L4 around 2007 could be the impact of large SIV differences between L3 and L4 assimilations around the same time, which are possibly due to an algorithm change in the reprocessing of L4 data. In short, no degradation in seasonal forecasts is found for all start months. Hence, forecast impact of initializing with Level-3 SIC instead of Level-4 SIC in our experiments is neutral for long-range forecasts and positive for medium to extended-range.

3.4 Sea-ice forecast impact of ice thickness initialization

The CS2-SMOS thickness initialization has a much larger impact on seasonal sea-ice forecasts than the change from Level-4 to Level-3 SIC assimilation. Figure 11 shows ensemble-mean sea-ice volume for each lead month for the forecasts started in May (top) and August (bottom) in the years 2011-2016, together with the ORAS5 reanalysis. It is remarkable that the shape of the seasonal cycle is preserved between FC-L4 and FC-L3T, and only the starting point is changed. FC-L3T starts from a thinner ice state. For the May forecasts, the distance between each corresponding black and blue dots suggests that the melting is slower in the forecast than it is in the reanalysis. Similarly, for the October and November forecasts, the slope of the last 2 black dots is in general steeper than the corresponding 2 green dots

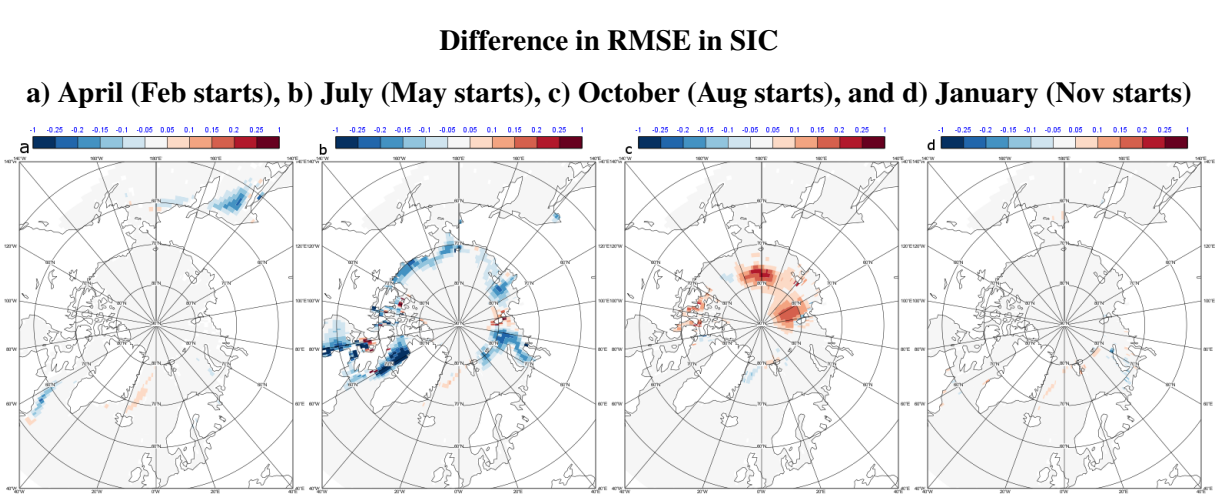

Figure 12: Difference in SIC RMSE between FC-L3T and FC-L4 re-forecasts for a) April forecasts started in February, b) July forecasts started in May, c) October forecasts started in August, and d) January forecasts started in November. Verification is for the years 2011-2016 w.r.t. OSI-401-b SIC observations.

indicating even slower autumn refreeze in FC-L3T compared to that in ORAS5. This could be due to the indirect effect of a thinner starting point in FC-L3T combined with the already existing slow refreeze bias of the model (cp. Section 3.1). However, it should also be kept in mind that ORAS5 probably overestimates the sea ice annual cycle, i.e. both melt rates in summer and growth rates in winter are too high (see Uotila et al. (2018) or Mayer et al. (in prep)). Similar to the autumn forecasts from May start dates, L3T initialization deteriorates autumn forecasts started in August for certain years in FC-L3T.

The thickness initialization has also changed the sea ice cover forecast skill. Figure 12 shows the change in RMSE between reference FC-L4 and FC-L3T verified against OSI-401-b for four start months representative of the seasonal cycle. July forecasts from May start dates show more than 25% RMSE reduction over large parts of the Arctic, and October forecasts from August start dates show larger RMSE of up to 20% over the Siberian side of the Central Arctic. April forecasts from February start dates also show improvement in forecast skill, and January forecasts from November start dates show neutral impact.

Impact of thickness initialization has not only improved the summer SIV and SIC forecasts, but it has also improved the sea ice extent forecasts as measured by the integrated ice-edge error IIEE (Figure 13). IIEE for all lead months and start months verified against OSI-401-b suggests reduced error in sea ice edge (blue colours) in FC-L3T overall. Figure 13 can be compared to Figure 9 to appreciate that the SIT constraint has a much larger impact than the Level-3 SIC assimilation: the units shown are $10^{11}\,m^2$ in Figure 13 but $10^{10}\,m^2$ in Figure 9, indicating that the impact is an order of magnitude larger. The most striking feature is the improvement in summer forecast error for long lead times in FC-L3T compared to FC-L4. The main contribution to the error reduction of up to 30% in summer forecasts comes from the reduction of the model bias leading to consistent over-prediction. For the traditional September sea ice extent forecast started in April, an improvement of 28% is found. For October and November forecasts started in July and August, a slight degradation is found, which is caused by under-prediction. This could again be due to the indirect effect of a thinner starting point in FC-L3T and a lower SIC in L3T, combined with the already existing slow refreeze nature of the model. Changes in sea ice edge forecast skill is consistent with the changes in the sea ice cover forecast skill.

Our results demonstrate the positive impact of a) constraining thickness in the assimilation system to improve model realism and b) exploiting the longer memory contained in winter sea-ice thickness to improve summer forecast skill of sea ice variables. These findings also confirm earlier studies using

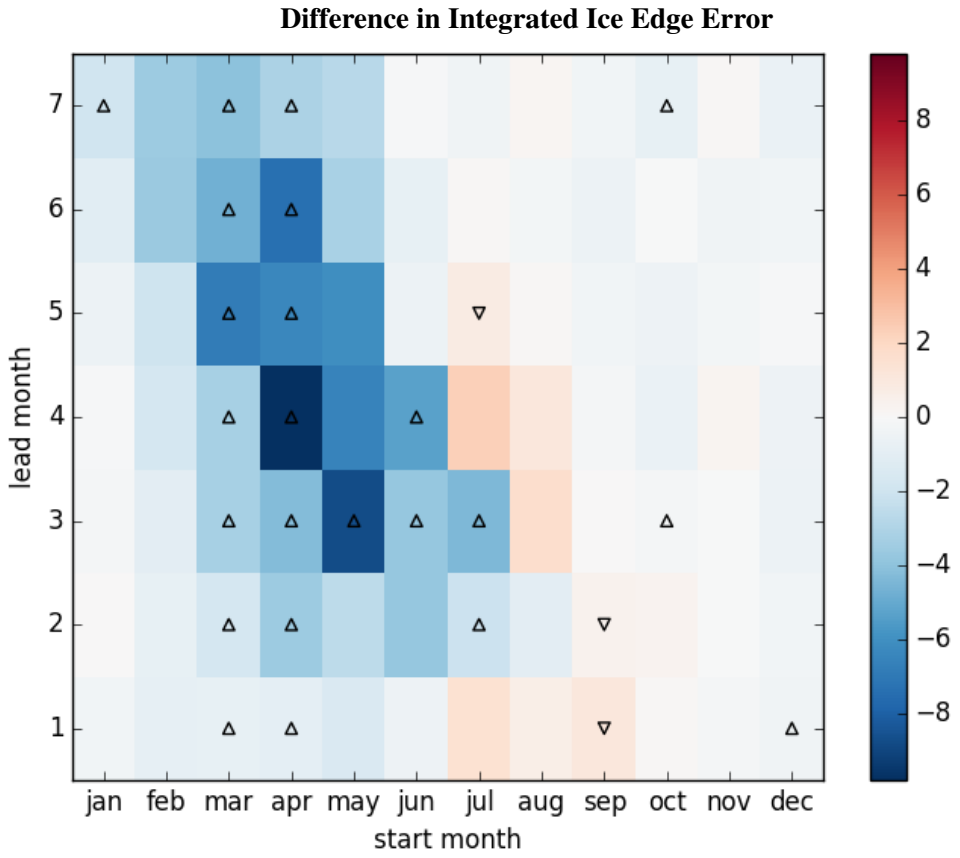

Figure 13: Difference in Integrated Ice Edge Error in 10^{11} m^2 *between FC-L3T and FC-L4 re-forecasts 2011–2016 w.r.t. OSI-401-b observations. Blue colour denotes reduced error in sea ice edge in FC-L3T and red colour denotes increased error in FC-L3T. Black triangles represent statistically significant (DelSole and Tippett (2016)) changes.*

thickness initialization through perfect model approach (Day et al. (2014)), and they agree with recent results using the UK Met Office operational seasonal forecasting system (Blockley and Peterson (2018)).

3.5 Impact of ice thickness initialization on atmospheric variables

The changes in sea ice cover forecasts associated with ice thickness initialization have profound impacts on near-surface temperature. The focus here is on reforecasts for October since the impacts on the atmosphere have been found to be strongest in the late summer and early freezing season.

Figure 14a) shows ensemble mean differences between 2m temperature (t2m) reforecasts for October initialized on 1st of May from L3T and L4 sea ice initial conditions. It is evident that the reduced end of summer sea ice concentration in FC-L3T w.r.t. FC-L4 (not shown) leads to substantially higher t2m everywhere over the Arctic Ocean. The warming is strongest (>2K) in the areas of the largest sea ice concentration differences, e.g. the Siberian Sea, but it spreads also southward onto the continents. For example, reforecasts for northern Siberia and northern America are significantly warmer in FC-L3T. Only Greenland and the western Greenland Sea show negative t2m differences, which could be related to increased sea ice cover east of Greenland in FC-L3T w.r.t FC-L4. The generally positive differences can also be seen for shorter lead times (not shown) but are strongest during the early freezing season, when the ocean releases the extra heat absorbed during summer due to the reduced sea ice cover (ice-albedo feedback).

Figure 14b) also shows October ensemble mean t2m differences between FC-L3T and FC-L4 re-forecasts, but initialized in August. The difference pattern in the Arctic is similar compared to the May start dates but less pronounced, probably because the time for the ocean to absorb additional heat during melt season is much shorter compared to the May start dates. However, comparison of Figures 14 (a) and (b) reveals that several signals in the mid-latitudes are quite similar for both start dates, e.g. the warmer temperatures over the Atlantic, North America, and Scandinavia. This suggests that at least some of the mid-latitude impacts are robust and large enough to be detectable even from the small re-forecast sample available to us.

The mid-latitude changes in 2m-temperature can also be found at higher atmospheric levels, e.g. at 850 hPa, but with smaller magnitude (not shown). To put the results from Figure 14a) and Figure 14b) into context, Figure 14c) and Figure 14d) show October ensemble mean t2m differences between FC-L3 and FC-L4 reforecasts initialized in May and August, respectively. Differences are relatively small and generally insignificant, consistent with only small changes in sea ice concentration.

Changes in RMSE of 2m-temperature re-forecasts are different for May and August start dates. There are substantial RMSE reductions for the May start dates (Figure 15a) in the regions where warming is seen in the ensemble mean (compared to Figure 14a). Conversely, RMSE increases for the August start dates (Figure 15b) in the regions where warming is seen in the ensemble mean. This opposite behaviour can be understood by considering the existing sea ice biases in the FC-L4 reforecasts (see Figure 6). May start dates of the FC-L4 reforecasts exhibit only small positive sea ice concentration bias in October, but the large positive summer sea ice bias leads to too little ocean heat absorption in summer and consequently too little ocean heat release in fall. The latter causes a negative 2m-temperature bias of the FC-L4 reforecasts for October started in May. Hence, the warming seen in the FC-L3T reforecasts helps to reduce this negative bias. In contrast, FC-L4 August start dates have a negative sea ice bias in October due to too slow re-freezing (see Figure 6) and hence exhibit a positive 2m-temperature bias. Thus, the temperature increases resulting from thickness initialization move the ensemble mean in the wrong direction in this case. It should be noted that largest part of the RMSE changes are due to the

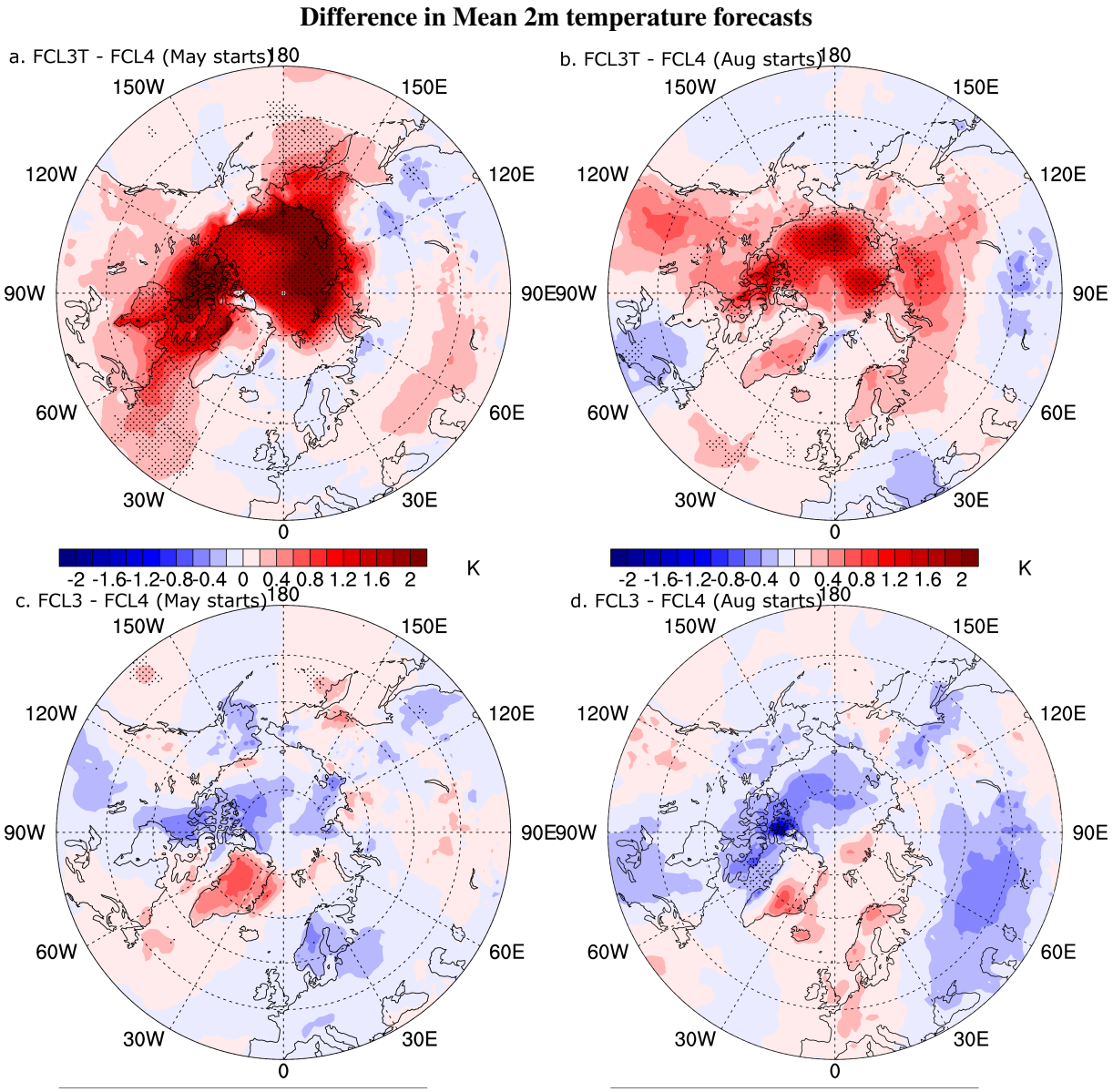

Figure 14: Differences in ensemble means of 2m temperature re-forecasts valid in October: a) FC-L3T minus FC-L4 (May starts), b) FC-L3T minus FC-L4 (August starts), c) FC-L3 minus FC-L4 (May starts), and d) FC-L3 minus FC-L4 (Aug starts). All diagnostics represent 2011-2016 averages. Stippling in a-d) denotes regions where the distributions of t2m are significantly different on the 95% level according to the Komolgorov-Smirnov test.

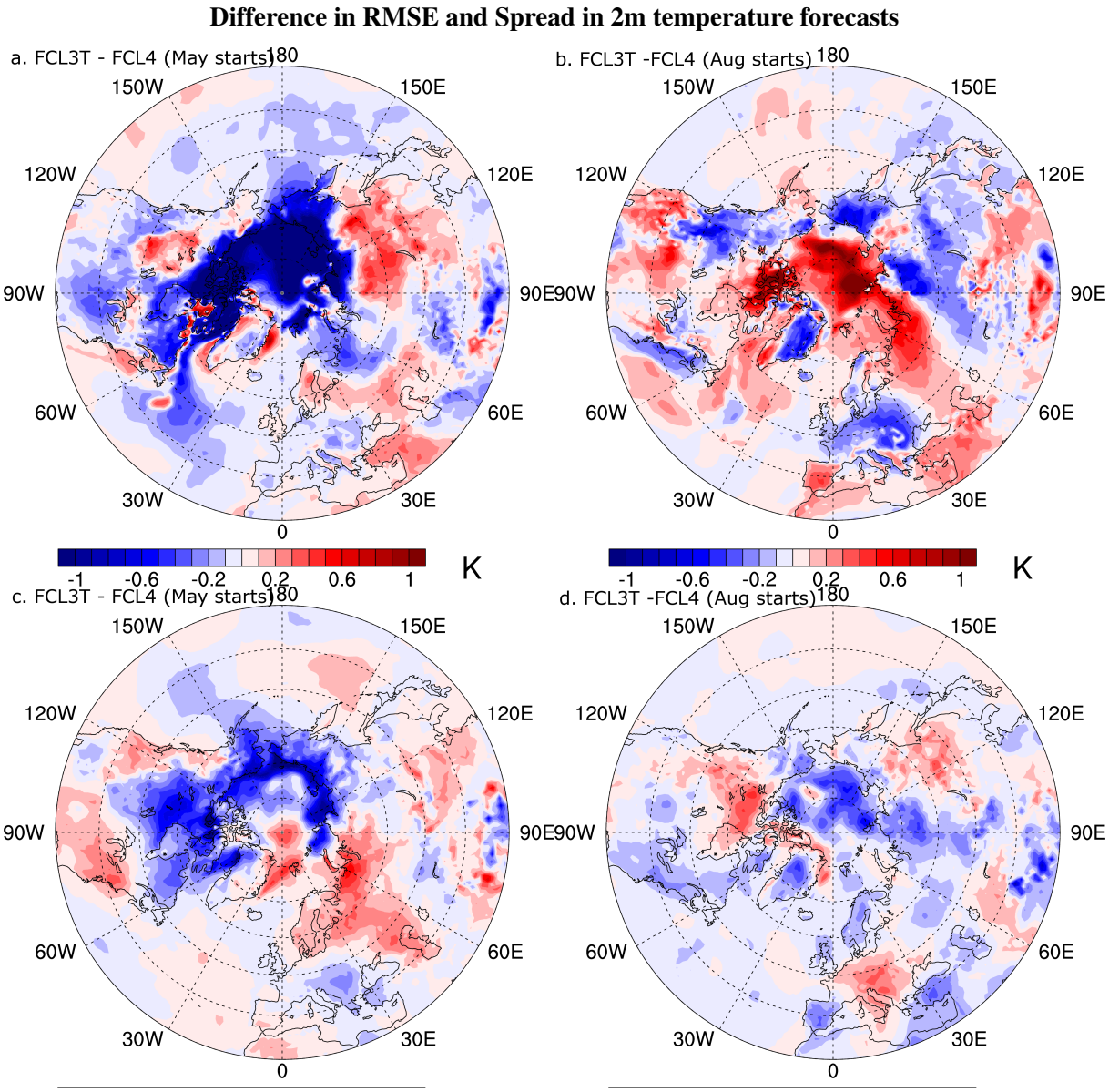

Figure 15: Differences in RMSE of 2m-temperature re-forecasts valid in October: a) FC-L3T minus FC-L4 (May starts) and b) FC-L3T minus FC-L4 (August starts); Differences in ensemble spread of 2m-temperature re-forecasts valid in October between c) FC-L3T minus FC-L4 (May starts), and d)FC-L3T minus FC-L4 (Aug starts); All diagnostics represent 2011-2016 averages.

changes in bias. Changes in bias-corrected 2m-temperature RMSE are generally small (not shown).

Changes in ensemble spread of 2m-temperature reforecasts for October between L3T and L4 initialization in May and August are presented in Figure 15c) and Figure 15d). Ensemble spread is reduced where sea ice concentration is reduced in L3T w.r.t. L4 initialization. The reason for this is that 2m-temperature is quite strongly coupled to SSTs over open sea via air-sea exchanges, while a frozen surface behaves more like land and provides a weaker constraint for 2m-temperature. It should be noted that the statistical significances indicated in Figure 14a) and (b) are also affected by changes in ensemble spread, since the employed statistical test (Komolgorov-Smirnov; e.g. Chakravarti and Roy (1967)) checks for differences in distributions rather than only differences in means.

Impacts on other atmospheric fields have been explored as well. The picture for circulation-related fields such as mean sea-level pressure and 500 hPa geopotential height is less clear compared to 2m-temperature, indicating that much of the changes found at the surface in the Arctic are due to local thermodynamic effects. However, the 2m-temperature changes in the mid-latitudes as discussed above are likely also related to changes in circulation, which needs to be investigated further.

4 Summary and Conclusions

We have assessed the impact of assimilating sea-ice concentration from the Level-3 OSI-401-b product instead of the Level-4 OSTIA product as in current operations at ECMWF. Additionally, we have used Arctic-wide sea ice thickness observations to constrain the modelled sea ice thickness, and have studied its impact on the ocean-sea ice reanalysis. We have initialized coupled forecasts of the ocean-sea-ice-wave-land-atmosphere using the above assimilation experiments, and have investigated the forecast skill of Arctic sea ice up to lead times of 7 months.

We find that the new Level-3 OSISAF SIC assimilation is robust compared to the existing Level-4 OSTIA SIC assimilation. The forecast impact is generally positive for SIC at lead month 1 and neutral for long lead times. Statistically significant improvements are found over the ice edge and coastal seas in the Arctic mostly in the first 2 weeks for forecasts initialized in most calendar months, except for January starts, when the impact is neutral. The positive impact persists up to week 4 for March, May, August, November and December start months.

Our findings are drawn from initial conditions produced from low resolution (ORCA1, 43r3 cycle) NEMOVAR experiments with ERA-Interim forcing and no ocean EDA. The assessment of Level-3 SIC initialization should be repeated in more recent model versions and with higher resolution before eventual operational implementation. Though the change towards Level-3 SIC assimilation is of technical nature, it is encouraging to see statistically significant improvements in the medium and extended range. For sea ice thickness and volume, the forecast impact is small and within the observational uncertainties.

CryoSat2-SMOS ice thickness nudging results in a substantial reduction of sea-ice volume and thickness in the ocean–sea-ice analysis. Constraining the analysis to observed sea-ice thickness observations and using the constrained state for the initialization of the forecasts has reduced some of the forecast biases and improved forecast skill in general. CS2-SMOS initialization of sea-ice forecasts has led to drastic thinning of the initial conditions. This acts to reduce the existing forecast bias during summer, but in turn increases the errors during autumn.

The impact of sea-ice thickness initialization on seasonal forecast skill for Arctic sea ice variables, namely sea ice cover, sea ice thickness, sea ice volume and sea ice edge is positive for seasonal forecasts started from February and May start dates. Neutral forecast impact for November start dates is found.

However, a slight degradation is seen in autumn forecasts started from July and August start dates, which warrants further investigation. Thinning of sea ice by CS2-SMOS mitigates or enhances seasonally dependent forecast model error. We find significant improvement of up to 28% in the traditional September sea ice extent forecasts started from April start dates as shown by Integrated Ice Edge Error.

The impact of sea-ice thickness initialization on atmospheric variables has also been investigated. Changes in ensemble mean 2m-temperature over the Arctic region are significant for October forecasts initialized from May start dates. While there is improvement in the skill of October 2m-temperature forecasts initialized from May, a degradation in forecast skill is found for the October forecasts initialized from August start dates. Changes in 2m-temperature RMSE are predominantly due to corresponding changes in biases of the sea ice variables. Our results illustrate that information on sea ice thickness is relevant for identifying model errors and for exploiting the long-term memory present in ice thickness for seasonal forecasts of sea ice and near surface temperatures.

These findings demonstrate that making use of newly available sea-ice thickness observations from satellites has the potential to significantly improve the sea-ice state in currently operational ocean–sea-ice reanalyses. Estimation of uncertainty of the available gridded Arctic-wide thickness observational datasets and their latency need to be addressed in the future, in order to make progress in sea ice forecasting.

Acknowledgements

We acknowledge the European Union Horizon 2020 SPICES project (640161) for funding this work.

References

Allard, R. A., Farrell, S. L., Hebert, D. A., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A. et al. (2018), 'Utilizing cryosat-2 sea ice thickness to initialize a coupled ice-ocean modeling system', *Advances in Space Research* **62**(6), 1265–1280.

Balmaseda, M. A., Ferranti, L., Molteni, F. and Palmer, T. N. (2010), 'Impact of 2007 and 2008 Arctic ice anomalies on the atmospheric circulation: Implications for long-range predictions', *Quarterly Journal of the Royal Meteorological Society* **136**(652), 1655–1664.
URL: *http://doi.wiley.com/10.1002/qj.661*

Balmaseda, M. A., Mogensen, K. and Weaver, A. T. (2013), 'Evaluation of the ECMWF ocean reanalysis system ORAS4', *Quarterly Journal of the Royal Meteorological Society* **139**(674), 1132–1161.
URL: *http://doi.wiley.com/10.1002/qj.2063*

Balmaseda, M. A., Vidard, A. and Anderson, D. L. T. (2008), 'The ECMWF Ocean Analysis System: ORA-S3', *Monthly Weather Review* **136**(8), 3018–3034.
URL: *http://journals.ametsoc.org/doi/abs/10.1175/2008MWR2433.1*

Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M. and DeWeaver, E. (2011), 'Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations', *J. Climate* **24**(1), 231–250.
URL: *http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3775.1*

Blockley, E. W. and Peterson, K. A. (2018), 'Improving met office seasonal predictions of arctic sea ice using assimilation of cryosat-2 thickness', *The Cryosphere* **12**(11), 3419–3438.

Bloom, S., Takacs, L., Da Silva, A. and Ledvina, D. (1996), 'Data assimilation using incremental analysis updates', *Monthly Weather Review* **124**(6), 1256–1271.

Bunzel, F., Notz, D. and Tietsche, S. (2017), 'Definition of a new set of observation based metrics relevant for regional applications. spices deliverable, d8.2 report'.

Carton, J. A., Chepurin, G., Cao, X. and Giese, B. (2000), 'A simple ocean data assimilation analysis of the global upper ocean 1950–95. part i: Methodology', *Journal of Physical Oceanography* **30**(2), 294–309.

Chakravarti, L. and Roy (1967), *Handbook of methods of applied statistics* **1**, 392–394.

Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A. et al. (2017), 'Intercomparison of the arctic sea ice cover in global ocean–sea ice reanalyses from the ora-ip project', *Climate Dynamics* **49**(3), 1107–1136.

Day, J. J., Hawkins, E. and Tietsche, S. (2014), 'Will Arctic sea ice thickness initialization improve seasonal forecast skill?', *Geophysical Research Letters* **41**(21), 7566–7575.
URL: *http://doi.wiley.com/10.1002/2014GL061694*

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F. (2011), 'The ERA-Interim reanalysis: configuration and performance of the data assimilation system', *Quarterly Journal of the Royal Meteorological Society* **137**(656), 553–597.
URL: *http://doi.wiley.com/10.1002/qj.828*

DelSole, T. and Tippett, M. K. (2016), 'Forecast Comparison Based on Random Walks', *Monthly Weather Review* **144**(2), 615–626.
URL: *http://journals.ametsoc.org/doi/10.1175/MWR-D-15-0218.1*

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012), 'The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system', *Remote Sensing of Environment* **116**, 140–158.
URL: *http://www.sciencedirect.com/science/article/pii/S0034425711002197*

Fichefet, T. and Maqueda, M. A. M. (1997), 'Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics', *Journal of Geophysical Research* **102**(C6), 12609–12646.
URL: *http://doi.wiley.com/10.1029/97JC00480*

Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E. and Jung, T. (2016), 'Predictability of the Arctic sea ice edge', *Geophysical Research Letters* **43**(4), 1642–1650.
URL: *http://doi.wiley.com/10.1002/2015GL067232*

Guemas, V., Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., Fučkar, N. S., Germe, A., Hawkins, E., Keeley, S. et al. (2016), 'A review on arctic sea-ice predictability and prediction on seasonal to decadal time-scales', *Quarterly Journal of the Royal Meteorological Society* **142**(695), 546–561.

Hendricks, S., Ricker, R. and Helm, V. (2016), 'User guide-awi cryosat-2 sea ice thickness data product (v1. 2)'.

Hibler III, W. D. (1979), 'A Dynamic Thermodynamic Sea Ice Model', *J. Phys. Oceanogr.* **9**(4), 815–846.

Hogan, R., Ahlgrimm, M., Balsamo, G., Beljaars, A., Berrisford, P., Bozzo, A., pe, F. D. G., Forbes, R., Haiden, T., Lang, S., Mayer, M., Polichtchouk, I., Sandu, I., Vitart, F. and Wedi, N. (2017), 'Radiation in numerical weather prediction', *ECMWF Technical Memorandum* (816).
URL: *https://www.ecmwf.int/node/17771*

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H. and Monge-Sanz, B. (2018), 'Seas5: The new ecmwf seasonal forecast system', *Geoscientific Model Development Discussions* **2018**, 1–44.
URL: *https://www.geosci-model-dev-discuss.net/gmd-2018-228/*

Kaleschke, L., Tian-Kunze, X., Maaß, N., Mäkynen, M. and Drusch, M. (2012), 'Sea ice thickness retrieval from SMOS brightness temperatures during the Arctic freeze-up period', *Geophysical Research Letters* **39**(5), L05501.
URL: *http://doi.wiley.com/10.1029/2012GL050916*

Kato, S., Rose, F. G., Rutan, D. A., Thorsen, T. J., Loeb, N. G., Doelling, D. R., Huang, X., Smith, W. L., Su, W. and Ham, S.-H. (2018), 'Surface irradiances of edition 4.0 clouds and the earths radiant energy system (ceres) energy balanced and filled (ebaf) data product', *Journal of Climate* **31**(11), 4501–4527.

Kwok, R. and Rothrock, D. A. (2009), 'Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008', *Geophys. Res. Lett.* **36**, L15501.

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. (2013), 'CryoSat-2 estimates of Arctic sea ice thickness and volume', *Geophysical Research Letters* **40**(4), 732–737.
URL: *http://doi.wiley.com/10.1002/grl.50193*

Madec, G. (2008), NEMO ocean engine, Technical report, Institut Pierre-Simon Laplace (IPSL).
URL: *http://www.nemo-ocean.eu/About-NEMO/Reference-manuals*

Mayer, M., Tietsche, S., Haimberger, L., Tsubouchi, T., Mayer, J. and Zuo, H. (in prep), 'An improved estimate of the coupled arctic energy budget'.

Mogensen, K., Balmaseda, M. A. and Weaver, A. (2012), The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4, Technical Report 668, European Centre for Medium-Range Weather Forecasts.

Mori, M., Watanabe, M., Shiogama, H., Inoue, J. and Kimoto, M. (2014), 'Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades', *Nature Geoscience* **7**(12), 869–873.
URL: *http://dx.doi.org/10.1038/ngeo2277*

Overland, J. E., Dethloff, K., Francis, J. A., Hall, R. J., Hanna, E., Kim, S.-J., Screen, J. A., Shepherd, T. G. and Vihma, T. (2016), 'Nonlinear response of mid-latitude weather to the changing arctic', *Nature Climate Change* **6**(11), 992.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J. and Haas, C. (2017), 'A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data', *The Cryosphere*

11(4), 1607–1623.
URL: *https://www.the-cryosphere.net/11/1607/2017/*

Ruggieri, P., Buizza, R. and Visconti, G. (2016), 'On the link between barents-kara sea ice variability and european blocking', *Journal of Geophysical Research: Atmospheres* **121**(10), 5664–5679.

Semtner, A. J. (1976), 'A Model for the Thermodynamic Growth of Sea Ice in Numerical Investigations of Climate', *J. Phys. Oceanogr.* **6**, 379–389.

Stockdale, T., Alonso-Balmaseda, M., Johnson, S., Ferranti, L., Molteni, F., Magnusson, L., Tietsche, S., Vitart, F., Decremer, D., Weisheimer, A., Roberts, C. D., Balsamo, G., Keeley, S., Mogensen, K., Zuo, H., Mayer, M. and Monge-Sanz, B. (2018), 'Seas5 and the future evolution of the long-range forecast system', *ECMWF Technical Memorandum* (835).
URL: *https://www.ecmwf.int/node/18750*

Tian-Kunze, X., Kaleschke, L., Maaß, N., Mäkynen, M., Serra, N., Drusch, M. and Krumpen, T. (2014), 'SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification', *The Cryosphere* **8**(3), 997–1018.
URL: *http://www.the-cryosphere.net/8/997/2014/ http://www.the-cryosphere.net/8/997/2014/tc-8-997-2014.html*

Tietsche, S., Balmaseda, M. a., Zuo, H. and Mogensen, K. (2015), 'Arctic sea ice in the global eddy-permitting ocean reanalysis ORAP5', *Climate Dynamics* .
URL: *http://link.springer.com/10.1007/s00382-015-2673-3*

Tietsche, S., Balmaseda, M., Rosnay, P., Zuo, H., Tian-Kunze, X. and Kaleschke, L. (2018), 'Thin Arctic sea ice in L-band observations and an ocean reanalysis', *The Cryosphere* **12**(6), 2051–2072.
URL: *https://www.the-cryosphere.net/12/2051/2018/*

Tietsche, S., Day, J. J., Guemas, V., Hurlin, W. J., Keeley, S. P. E., Matei, D., Msadek, R., Collins, M. and Hawkins, E. (2014), 'Seasonal to interannual Arctic sea-ice predictability in current global climate models', *Geophysical Research Letters* **41**, 1035–1043.
URL: *http://onlinelibrary.wiley.com/doi/10.1002/2013GL058755/abstract*

Tietsche, S., Notz, D., Jungclaus, J. H. and Marotzke, J. (2013), 'Assimilation of sea-ice concentration in a global climate model – physical and statistical aspects', *Ocean Science* **9**, 19–36.
URL: *http://www.ocean-sci.net/9/19/2013/os-9-19-2013.html*

Tilling, R. L., Ridout, A., Shepherd, A. and Wingham, D. J. (2015), 'Increased Arctic sea ice volume after anomalously low melting in 2013', *Nature Geoscience* **8**(643-646).
URL: *http://dx.doi.org/10.1038/ngeo2489*

Tonboe, R., Lavelle, J., Pfeiffer, R. and Howe, E. (2017), 'Product user manual for osi saf global sea ice concentration (product osi-401-b)'.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthelemy, A., Bricaud, C., Carton, J., Fuckar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V. S., Marnela, M., Massonnet, F., Mignac, D., Peterson, K. A., Sadikni, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J. and Zhang, Z. (2018), 'An assessment of ten ocean reanalyses in the polar regions', *Climate Dynamics* **(in press)**, 1–64.

Weaver, A., Vialard, J. and Anderson, D. (2003), 'Three-and four-dimensional variational assimilation with a general circulation model of the tropical pacific ocean. part i: Formulation, internal diagnostics, and consistency checks', *Monthly Weather Review* **131**(7), 1360–1378.

Zampieri, L., Goessling, H. F. and Jung, T. (2018), 'Bright prospects for arctic sea ice prediction on subseasonal time scales', *Geophysical Research Letters* **45**(18), 9731–9738.

Zuo, H., Balmaseda, M. A. and Mogensen, K. (2017), 'The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals', *Climate Dynamics* **49**(3), 791–811.
URL: *http://link.springer.com/article/10.1007%2Fs00382-015-2675-1*

Zuo, H., Balmaseda, M. A., Mogensen, K. and Tietsche, S. (2018), 0cean5: The ecmwf ocean reanalysis system oras5 and its real-time analysis component, Technical report, ECMWF Tech. Memo 2018. Available online: https://www. ecmwf. int/en/elibrary .

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K. and Mayer, M. (2019), 'The ecmwf operational ensemble reanalysis-analysis system for ocean and sea-ice: a description of the system and assessment', *Ocean Science Discussions* **2019**, 1–44.
URL: *https://www.ocean-sci-discuss.net/os-2018-154/*