

Presentation and evaluation of the Arctic sea ice forecasting system neXtSIM-F

Timothy Williams¹, Anton Korosov¹, Pierre Rampal^{2,1}, and Einar Ólason¹

¹Nansen Environmental and Remote Sensing Center, Thormøhlensgate 47, 5006 Bergen, Norway and the Bjerknes Center for Climate Research, Bergen, Norway

²Université Grenoble Alpes/CNRS/IRD/G-INP, Institut Géophysique de l'Environnement, Grenoble, France

Correspondence: Timothy Williams (timothy.williams@nersc.no)

Abstract. The neXtSIM-F forecasting system consists of a stand-alone sea ice model, neXtSIM, forced by the TOPAZ ocean forecast and the ECMWF atmospheric forecast, combined with daily data assimilation of sea ice concentration. It uses the novel Brittle Bingham-Maxwell (BBM) sea ice rheology, making it the first forecast based on a continuum model not to use the viscoplastic (VP) rheology. It was tested in the Arctic for the time period November 2018– June 2020 and was found to perform well, although there are some shortcomings. Despite drift not being assimilated in our system, the sea ice drift is good throughout the year, being relatively unbiased, even for longer lead times like 5 days. The RMSE in speed and the total RMSE are also good for the first 3 or so days, although they both increase steadily with lead time. The thickness distribution is relatively good, although there are some regions that experience excessive thickening with negative implications for the summer-time sea ice extent, particularly in the Greenland Sea.

The neXtSIM-F forecasting system assimilates OSI-SAF sea ice concentration products (both SSMI and AMSR2) by modifying the initial conditions daily and adding a compensating heat flux to prevent removed ice growing back too quickly. The assimilation greatly improved the sea ice extent for the forecast duration.

1 Introduction

Arctic sea ice has been in great decline in the last number of years (Meier, 2017). Perovich et al. (2018) report that in 2018, the summer extent was the sixth lowest and the winter extent was the second lowest in the satellite record (1979–2018). Moreover, surface air temperatures in the Arctic continued to warm at twice the rate relative to the rest of the globe, and Arctic air temperatures for the past five years (2014–18) have exceeded all previous records since 1900 (Overland et al., 2018), which will also contribute to future sea ice decline if it continues.

With less sea ice comes an increase in summertime accessibility for shipping. Azzara et al. (2015) considered a range of different scenarios and projected an increase in the number of vessels operating in the Bering Strait and the U.S. Arctic of between 100 and 500%. The International Maritime Organization has also recognized that shipping would increase and adopted an International Code for Ships Operating in Polar Waters (Polar Code)¹ on 1 January 2017. This polar code addresses the

¹<http://www.imo.org/en/MediaCentre/HotTopics/polar/Pages/default.aspx>

increased safety and pollution risks of operating in the Arctic. A recent example of the risks and concomitant costs of accidents in the Arctic is the rescue of the fishing vessel Northguider, which ran aground between Spitzbergen and Nordaustlandet (Svalbard) after getting into trouble with sea ice. The crew had to be rescued by the Norwegian Coast Guard icebreaker K.V. Svalbard, who then had to drain 300 kl of diesel from the damaged vessel.²

5 Thus sea ice forecasting is becoming increasingly important. As well as search and rescue/accident prevention, other applications are optimized ship (icebreaker) routing based on forecasts (Kaleschke et al., 2016) and support of research activities – e.g. Schweiger and Zhang (2015) give an example of scheduling of high-resolution SAR images in order to follow the drift of some ice-mass balance (IMB) buoys by using the PIOMAS/MIZMAS forecast from the University of Washington. The year-long drift of the Polarstern from September 2019 (part of the MOSAIC project) also relied heavily on sea ice and weather
10 forecasts.

Tonani et al. (2015) give a good overview of the 2015 status of operational forecasting, while Hunke et al. (2020, Table 1) give a comprehensive list of modelling systems that include sea ice, most of which are used in national forecasting capacities. They vary in resolution, in complexity (with regards to the modelled processes and the coupling between these processes) and in the data assimilation schemes that are used. We note however that they do not vary in their numerical framework (e.g.
15 Eulerian advection schemes in all systems) nor in the rheological framework that they use for sea-ice (e.g.. viscous-plastic). In this review, the authors claim that the trend is towards fully-coupled systems at high resolution. For example, the ECMWF forecast coupled ice and ocean models to their atmospheric model in between the two papers (this system went operation in June 2018) at a resolution of 0.1° . neXtSIM-F uses this latest ECMWF product (IFS: integrated forecast system: Owens and Hewson, 2018) to provide forecast atmospheric forcing, along with ocean forcing from TOPAZ Sakov et al. (2012).

20 Another relevant example is the replacement of RIPS³ (Lemieux et al., 2016a) by RIOPS⁴ in 2016, having NEMO coupled to the system, as well as having other processes like tides added. Like neXtSIM-F, RIPS used a stand-alone sea ice model based on the CICE sea ice model which used 3DVAR assimilation of OSI-SAF SSMI and AMSR2 concentration, as well as ice charts from the Canadian Ice Service.

In this paper, we introduce a new sea ice forecasting system, neXtSIM-F, which is based on a stand-alone version of the
25 sea ice model neXtSIM (Rampal et al., 2016, 2019). neXtSIM is a Lagrangian finite element model, and we are running it with a nominal triangle side length of 10 km, with a distance from one point of a triangle to the opposite edge being about 7.5 km. The name neXtSIM-F refers to the entire platform, including data input/output and assimilation, model initialisation and simulation, export, visualisation and evaluation of results (see Section 3). Due to the relatively recent arrival of the sea ice model, neXtSIM-F is simpler than most other platforms, both in terms of assimilation scheme (data insertion with nudging)
30 and model components (uncoupled to ocean or atmosphere). (See Section 5 for planned improvements.) However, it is the first forecasting system based on a model with brittle sea ice rheology instead of the traditional viscoplastic (VP) rheology.

²<https://www.highnorthnews.com/en/svalbard-preparing-extreme-pumping-operation-using-small-boats>

³Regional Ice Prediction Service: operated by ECCC (Environment and Climate Change Canada) from 2013 to 2016.

⁴Regional Ice Ocean Prediction Service: operated by ECCC. See https://eccccc-msc.github.io/open-data/msc-data/nwp_riops/readme_riops_en/#technical-documentation.

It is also the first system being based on a Lagrangian (adaptive) deforming grid, as opposed to the other ones being based on the standard Eulerian (fixed) grids.

neXtSIM-F entered into operations as part of the Copernicus Marine and Environmental Monitoring Services (CMEMS) on 7 July 2020. It was equipped with neXtSIM v1.0 based on the Maxwell-Elasto-Brittle (MEB) rheology (Dansereau et al., 2016), which had been shown to reproduce Arctic sea ice drift and deformation particularly well Rampal et al. (2016, 2019). MEB consisted of an elastic spring in series with a dashpot, together with two main modifications to improve localisation and to prevent excessive convergence: the damage value was only used to modify the stress when it exceeded a threshold of 0.95, and a kind of viscoplastic stress term was added which only played a role when the ice was very damaged (see Appendix A1). This improves the thickness field somewhat over longer simulations (1–2 years; Rampal et al. (2019)), but not quite enough for longer than that.

In September 2020 the core of the forecasting platform was replaced with a new model: neXtSIM v2.0 based on a preliminary version of the novel Brittle-Bingham-Maxwell (BBM) sea ice rheology (Ólason et al., in prep.). The new version of neXtSIM-F enters into operations in December 2020.

BBM consists of an elastic spring in series with a composite element that contains a dashpot and a frictional sliding element in parallel (Ólason et al., in prep.). (For a summary of the BBM equations, see Appendix A2.) The BBM’s main physical achievement has been to stabilise the thickness for decadal-scale simulations; computationally it is also 5–6 times faster, since it is able to be solved explicitly, unlike our version of the MEB. It also kept the improvements in the model’s representation of the main spatial and temporal characteristics of observed deformation that were gained by MEB: strain localization and scaling (Marsan et al., 2004; Rampal et al., 2008; Stern and Lindsay, 2009), multifractality and intermittency in time (Rampal et al., 2019), meaning that higher deformations are more localised in space and more intermittent in time than smaller ones. These properties have strong implications for things like distribution and size of lead openings and how long they will stay open, which controls the heat and salt fluxes across the ocean–ice–atmosphere coupled system. More pertinently in a forecast context, they are also highly relevant for navigation.

The paper is organised as follows: we begin by introducing the data and methods that we use throughout, and then evaluate the neXtSIM model’s general performance for a free run from November 2018 to June 2020 in terms of concentration/extent, thickness and drift. This free run uses hindcast forcing fields. We then evaluate the neXtSIM-F forecast platform for the same period, when we assimilate concentration but use forecast forcing fields.

2 Data sources

2.1 Forecast ocean forcing from TOPAZ4

The official European forecast for the Arctic is developed and run by the CMEMS ARC-MFC (Arctic Marine Forecasting Centre)⁵. This uses the TOPAZ system (Simonsen et al., 2018; Sakov et al., 2012), which uses version 2.2.37 of the Hybrid

⁵Three institutes contribute to the ARC-MFC: the Nansen Environmental and Remote Sensing Center, the Norwegian Meteorological Institute and the Norwegian Institute for Marine Research.

Coordinate Ocean Model (HYCOM) (Bleck, 2002). In the current version (4) of TOPAZ, HYCOM is coupled to a sea ice model derived from version 4.1 of the Community Ice Code (CICE: Hunke et al., 2010); ice thermodynamics are described in Drange and Simonsen (1996), while the dynamics are based on the visco-plastic (VP) sea ice rheology (implemented with the elastic-viscous-plastic (EVP) solver of Hunke and Dukowicz, 1997). The model's native grid covers the Arctic and North Atlantic Oceans and has a horizontal resolution of between 11 and 16 km. TOPAZ4 uses the Ensemble Kalman filter method (EnKF; Sakov and Oke, 2008) to assimilate remotely sensed sea level anomalies, sea surface temperature, sea ice concentration, sea ice thickness and Lagrangian sea ice velocities (the latter two in winter only), as well as temperature and salinity profiles from Argo floats and ice-tethered profilers. Data assimilation is performed weekly.

To force neXtSIM, we use the following variables from TOPAZ: the near-surface (30 m) ocean velocity, the mixed layer depth (MLD), and the sea surface (3 m) temperature and salinity (SST and SSS, respectively). We give more details of how they are used in Section 3.1 below.

2.2 Forecast atmospheric forcing from ECMWF

For our forecast demonstration, we use the latest version (Cycle 45r1) of the Integrated Forecast System from ECMWF (IFS: Owens and Hewson, 2018) to provide atmospheric forcing fields to neXtSIM. It consists of an atmospheric model coupled to the NEMO 3.4 ocean model (Nucleus for European Modelling of the Ocean), the LIM2 (Louvain-la-neuve Sea Ice Model) sea ice model, the ECWAM (ECMWF WAVE Model) wave model, and a land surface model (HTESSEL). Its spatial resolution is 0.1° .

We use the 10-m wind velocity, the 2-m air and dew point temperatures (the latter is used to determine the specific humidity of air for the latent heat flux calculation), the mean sea level pressure, the long- and short-wave downwelling radiation, and the total precipitation (this becomes snow if the 2-m air temperature is below 0°C).

2.3 Sea ice concentration products from OSI-SAF

OSI-SAF provides estimates of sea ice concentration derived from the Special Sensor Microwave Imager Sounder (SSMIS) radiometer (Tonboe et al., 2016; Tonboe and Lavelle, 2016; Lavelle et al., 2017) and from AMSR2 (Lavelle et al., 2016a, b; Tonboe and Lavelle, 2015). The SSMIS algorithm uses the 19 GHz frequency (vertically polarized, footprint size about 56 km) and the 37 GHz frequency (both vertically and horizontally polarized, footprint size about 33 km). The AMSR2 algorithm uses three frequencies: 18.7, 36.5 and 89 GHz (also in vertical and horizontal polarizations) with footprints from 22 to 5 km). The AMSR2 data are presented on a 10-km grid, and we chose this product over the higher resolution (3.25 km) ASI product as we found it less noisy near the ice edge. These products are available daily within 12 hours after acquisition and processing so it is possible to assimilate this data in operational forecasts. However, the file for the day before the bulletin date (the day the model is run) doesn't arrive early enough to be assimilated in our daily run, which is launched at 03:00 (European western time). Therefore, we use the file from two days before the bulletin date.

As specified in the validation reports cited above the SSMIS has lower resolution ice concentration but has the advantage of higher accuracy, while the AMSR2 algorithm has higher resolution but also higher uncertainties. In order to combine the

advantages of these products we generated a blended product that was used both for assimilation during the forecasts. Blending was performed with a weighted average of the two products (using the errors in the products to calculate the weights):

$$C_B = \frac{C_L \sigma_L^{-2} + C_H \sigma_H^{-2}}{\sigma_H^{-2} + \sigma_L^{-2}} \quad (1)$$

where C denotes sea ice concentration, σ denotes the concentration uncertainty, index H denotes high resolution (AMSR2) and L - low resolution (SSMIS).

Sea ice extent, used as an evaluation metric of the model, was calculated from the concentration product as a sum of areas of all pixels within the model domain with concentration above 15%. Sea ice extent uncertainty was calculated as a difference between the extents calculated from the sum of concentration and uncertainty and concentration alone.

We use both SSMI and AMSR2 products for assimilation by the forecasts, but only SSMI for evaluation of the free run and our forecasts. This was because we found the AMSR2 product somewhat inconvenient due to missing sections of data, which made our evaluation statistics quite noisy. (OSI-SAF SSMI is therefore not an independent validation dataset for the forecasts.)

2.4 Sea ice drift from OSI-SAF

We use this product for evaluation of both our free run and our forecasts. (It is therefore an independent validation dataset for our forecasts.) To produce it, low-resolution ice drift datasets are computed on a daily basis from aggregated maps of passive microwave (e.g. SSM/I, AMSR-E) or scatterometer (e.g. ASCAT) signals (all channels are used) using the continuous maximum cross-correlation method (CMCC: Laverne et al., 2010; Laverne and Eastwood, 2010; Laverne, 2010). Daily 48-hour ice drift vectors can be obtained at a spatial resolution of 62.5 km. As part of our evaluation we sometimes apply a filter on the uncertainty given in the product to concentrate on the more precise observations, only considering those where the 2-day drift uncertainty is less than 2.5 km. This completely excludes the summer period of May to September (RMS uncertainty is about 12 km), since surface melting and a denser atmosphere preclude the retrieval of precise information. From October to April, we can still retain about 75% of the observation vectors after using this threshold, with the removed vectors generally being close to the ice edge, coast or the north pole. The error is higher in these regions as the sub-images on which the CMCC method is applied must be reduced to limit them to being inside the ice mask (Laverne and Eastwood, 2010). In the case of the north pole, there are fewer observations there, while the vectors in the MIZ have especially high uncertainties (sometimes up to 12 km) due to the high velocities in those regions, combined with the relatively long time interval over which the drift is calculated.

In order to compare neXtSIM drift to this product, every day at 12:00 we place Lagrangian drifters at the grid points of the OSI-SAF drift product and advect them for 48 hours at the ice velocity. The total drift is then compared to the OSI-SAF drift product.

2.5 Sea ice thickness from CS2-SMOS

We use this product (version 2.2: Ricker et al., 2017) for long-term evaluation of the free run. This product is a daily hybrid product which combines thickness estimates from the Cryosat-2 (CS2) altimeter (more reliable for thicker ice ($\gtrsim 0.5$ m) and

from SMOS (Soil Moisture and Ocean Salinity: better for thin ice). The CS2 altimeter tracks are somewhat sparse so optimal interpolation (OI) is used to fill the gaps between them and the areas of thin ice. For this reason each file covers a 7-day period. The OI method requires a background field to be created with full coverage, and that is independent of the target week so it is created from the CS2 values for the two weeks ahead and behind the target week, and from the SMOS values for the day
5 before the target week.

The errors from this approach can be particularly high in coastal areas of thick ice (for example north of Greenland and Canada), which are too thick to be measured by SMOS, but may be only covered by altimeter tracks every 2–3 weeks. For these gaps in coverage, the product uses the background field.

There is now an operational version of this product, which only uses CS2 data from one week before and after the target
10 week. A delay of one week for thickness would probably be acceptable for assimilation in a real time forecast in the future.

3 Description of the forecast platform

As shown in Figure 1 the platform consists of pre-processing steps (including generation of the model mesh and initialization of the forecast or assimilation of observations), running of the forecast by the neXtSIM model, and post-processing steps (including forecast evaluation, and visualisation). These steps are explained in detail in the section below.

15 3.1 The neXtSIM model

neXtSIM is a stand-alone sea-ice model which can use winds and currents from a variety of atmospheric and oceanic models (hindcasts or forecasts). This makes it quite flexible and light to run and therefore ideal for a forecasting context. Its dynamical core is the new Brittle-Bingham-Maxwell (BBM) rheology (Ólason et al., in prep.). (Also see Appendix A2 for a summary.) Rampal et al. (2016) presented results using a previous version of neXtSIM including an elasto-brittle (EB) rheology as de-
20 scribed in Bouillon and Rampal (2015), showing good agreement with observed drift and concentration in particular. More recently, Rampal et al. (2019) showed the ability to reproduce the characteristic multi-fractal scalings of deformation when using the previous version of neXtSIM, which included the Maxwell-elasto-brittle (MEB) rheology (Dansereau et al., 2016). Rampal et al. (2019) updated the paper of Rampal et al. (2016), evaluating the model’s improved performance when using the MEB rheology. The key addition of the MEB rheology was the addition of a viscous dissipation of stress in areas where
25 the ice is damaged, allowing it to move more freely. However, in longer term simulations the MEB showed unrealistic pile-up of ice particularly along the north-east coast of Greenland and the north-west coast of Svalbard. This led to the addition of a frictional element to the rheology (Ólason et al., in prep.), which provides some resistance to compression (up to a threshold). This framework, of a spring in series with a composite element made up a dashpot and fraction element in parallel, is called BBM.

30 The dynamical equations are solved with a finite element method on a Lagrangian (moving) triangular mesh. The code is currently a parallelised C++ code, used by Rampal et al. (2019), and presented by Samaké et al. (2017). Momentum input comes from the wind and ocean stresses (a turning angle of 25° is applied to the ocean velocity from the ocean forcing), the

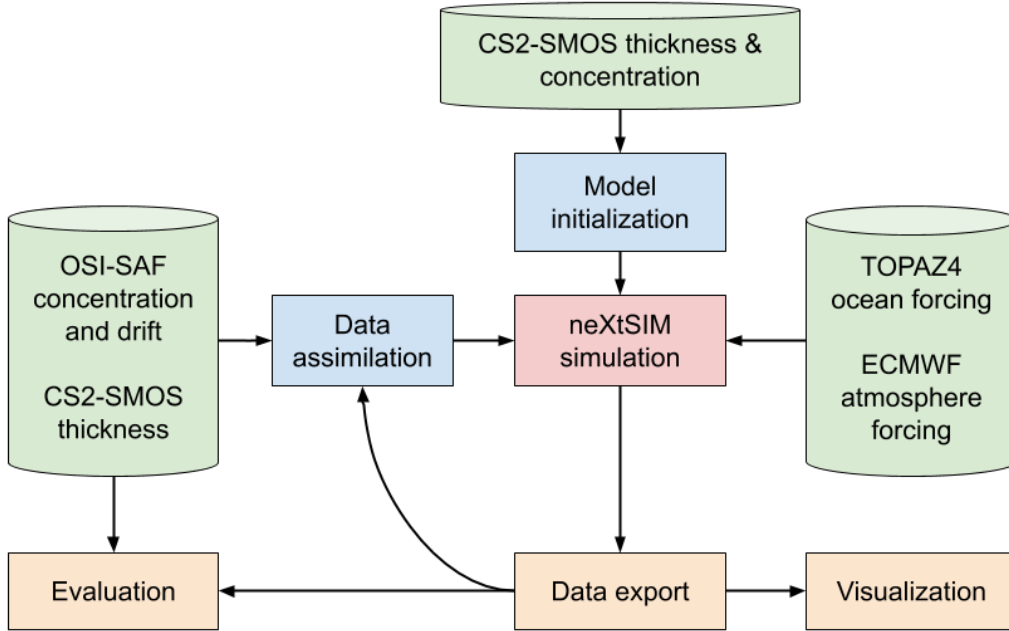


Figure 1. Overall scheme of the neXtSIM-F forecasting platform. Blue blocks: pre-processing steps; red block: running of the model core; yellow blocks: post-processing steps; green blocks: input data. Note that we only assimilate OSI-SAF SSMI and AMSR2 concentrations.

Coriolis force and sea surface slope, and there is also a basal stress applied at the bottom of the ice when it becomes grounded. For this basal stress we follow the scheme of Lemieux et al. (2016b), using the parameters $k_1 = 10$ m, $k_2 = 15$ m, $\alpha_b = 20$, $u_0 = 5 \times 10^{-5} \text{ m s}^{-1}$.

There is also a thermodynamic component of the code and beneath the ice is a slab ocean with three variables: temperature, salinity and thickness. The temperature and salinity are modified by the heat and salinity fluxes determined by the thermodynamical model as ice melts and freezes and as the model interacts with the atmosphere. They are relaxed towards the SST and SSS from TOPAZ over a time scale of about one month, while the thickness of the slab ocean is taken directly to be the MLD of TOPAZ. This varies spatially and evolves with time according to the forecast from TOPAZ. The thermodynamical model is a three-category model (detailed in Rampal et al., 2019, Appendix A): open water, newly-formed ice (treated as one ice layer and one snow layer; Semtner, 1976) and older ice (treated as two ice layers and a snow layer; Winton, 2000).

The older ice is characterised in the model by concentration, c , and thickness averaged over the entire cell (effective thickness or, in other words, volume per unit area), h . The absolute thickness of ice can be computed as the ratio: $H = h/c$; there is also an effective snow thickness, s . The young ice has concentration c_y , effective and absolute thicknesses, h_y and $H_y = h_y/c_y$, and snow thickness, s_y . H_y is constrained so that $H_{\min} \leq H_y \leq H_{\max}$. If $H_y < H_{\min}$, c_y is reduced to $c'_y = (H_y c_y)/H_{\min}$; if $H_y > H_{\max}$, some ice is moved to the older ice category. During this process H_y is reduced to $H'_y = H_{\max}$, while c_y is also reduced to $c'_y = c_y(H_{\max} - H_{\min})/(H_y - H_{\min})$. This concentration reduction is intended to give some lateral decrease in thin ice volume and not just vertical. The corrected values for h_y and s_y are $h'_y = c'_y H_{\max}$ and $s'_y = c'_y (s_y/c_y)$; the corresponding properties for the older ice h , c and s are then increased in a ice-and-snow-volume-conserving manner. The values that we used for the thresholds on the absolute thin ice thickness are $H_{\min} = 0.05$ m and $H_{\max} = 0.275$ m.

The domain we present simulations for is a pan-Arctic one, with a resolution of about 7.5 km. A 7-day forecast for this domain runs in about 30 minutes using 16 processors.

3.2 Initialization of the model fields

Before the model can be run it has to be initialized - the triangular mesh for the destination domain should be generated and initial values of sea ice concentration and thickness and the slab ocean surface temperature and salinity have to be set. The mesh is generated with Unref (a component of Gmsh: Geuzaine and Remacle, 2009) and using the Global Self-consistent, Hierarchical, High-resolution Shoreline Database (GSHHS) (Wessel and Smith, 1996). The model fields are initialized from the observed sea ice concentration and thickness from the CS2SMOS product (sec. 2.5) and simulated sea surface temperature and salinity from TOPAZ4 (sec. 2.1). The ice concentration in the CS2SMOS product comes from the low-resolution OSI-SAF product and has to be adjusted in areas with low thickness (Ivanova et al., 2015). Therefore total sea ice concentration (c_t) is calculated by adjusting the observed SIC using the observational operator \mathcal{O} :

$$c_t = c_O / \mathcal{O}(h_{C2S})$$

$$\mathcal{O}(h_{C2S}) = a_0 \sinh\left(\frac{h_{C2S} - a_1 a_2}{2a_1}\right) \quad (2)$$

where c_O is the OSI-SAF concentration, h_{C2S} is the CS2SMOS (effective) thickness, and $a_0 = 0.9569$, $a_1 = 0.06787$ and $a_2 = 0.4255$ are parameters fitted from the observations (personal communication with Thomas Lavergne).

As mentioned above the model has two ice categories - young ice and older ice with different rheological behaviour. At the initialization step young ice c_y is set to comprise 20% and the older ice c 80% of the total SIC. If the observed absolute thickness h_{C2S}/c_t is below the young ice upper thickness limit H_{\max} , then thickness is distributed between young and older ice in the same proportion, otherwise young ice thickness is calculated as $h_y = c_y H_{\max}$ and for the older ice: $h = h_{C2S} - h_y$. It was identified that the model has little sensitivity to the fraction of young ice and it may vary within 5 - 30%.

The temperature and salinity of the slab ocean is taken to be equal to the TOPAZ4 surface forecast, while the ice velocity and damage are set to zero.

3.3 Assimilation of concentration

The assimilation is performed before each forecast run using the data insertion method — an updated concentration (c_U) is calculated as a function of the forecast variable (c_F) and observations (c_O): $c_U = f(c_F, c_O)$. Other variables (particularly ice and snow thicknesses and the SST of the slab ocean) are adjusted for consistency, and the simulation is then restarted using the updated variables and the model is run for 8 days to provide one day of hindcast and a 7-day forecast. The assimilation is performed on the model mesh — the satellite observations, originally provided on a regular grid, are linearly interpolated onto the centers of the mesh elements so that they can be directly compared with the neXtSIM prognostic variables.

The concentration is done as follows. The observed concentration c_O is taken to be the blended concentration from OSI-SAF (see Section 2.3) and is combined with the model concentration using a weighted average approach:

$$c'_U = \frac{c_F \sigma_F^{-2} + c_O \sigma_O^{-2}}{\sigma_O^{-2} + \sigma_F^{-2}}. \quad (3)$$

The uncertainties of the observed concentration (σ_O) are obtained from the input products, while the value for the uncertainty in the forecast concentration is set to 0.3 to reduce strong discontinuities in the modelled fields. It can also be thought of as introducing a time scale (in days) of $1 + \sigma_O^2 / \sigma_F^2$ for relaxation towards the observations that depends on the relative uncertainties. Assimilation of both concentration is performed in all elements that have valid observations.

In the pack, the total (young and old ice) model concentration c_F is generally about 1, while the OSI-SAF concentration can be around 0.9. We found that using c'_U as the actual update c_U produced too much of a drop in concentration there, lowering the internal stress to near zero and allowing too much drift. This quickly led to rapid build up of very thick ice in unusual places. Therefore we had to take a quite conservative approach to our correction and only change the model to make sure it had the correct ice mask (this is a kind of assimilation of extent):

$$c_U = \begin{cases} 0 & c'_U < 0.15, \\ c'_U & c_F < 0.15 \text{ and } c'_U \geq 0.15, \\ c_F & \text{otherwise.} \end{cases} \quad (4)$$

After calculating the updated variables as specified in (3) the fractions and thicknesses of young and older ice are calculated in the same way as in the initialisation procedure (see sec. 3.2).

Once all the assimilation steps have been performed the model fields are checked for consistency with each other and corrected if necessary. First, the concentration of young ice is reduced in the elements where the total concentration exceeds 100%. Second, the volume of ridged ice, damage and the ice and snow thicknesses are set to zero in the added ice. Lastly, we need to correct the SST of the slab ocean. Where new ice is added during assimilation it is set to the freezing point, but if ice is removed then the situation is more complicated. Then we proceed as in the following section.

3.4 Compensating heat flux

One of the side effects of assimilation is that the heat balance in the model is disturbed: reducing the concentration opens more water and given the temperatures, atmosphere humidity and ocean salinity provided by the forcing, the heat flux out of the ocean increases and the ice freezes up again very fast. A compensating heat flux is added to the total ocean heat flux in order to keep the heat balance and prevent refreezing of ice where it was removed by assimilation, thus prolonging the effect of assimilation. If c_t and c'_t are respectively the total concentrations before and after assimilation, a compensating heat flux (Q_c) is calculated if ice was removed — i.e. if $c'_t < c_t$ — according to the following formula

$$Q_c = Q_o((c'_t/c_t)^n - 1) \quad (5)$$

where Q_o is the total heat flux from the ocean (sum of flux from the ocean to sea ice and to the atmosphere), and n is a parameter controlling the strength of correction. The function was chosen so that Q_c is zero when the concentration update is zero, and $Q_c = -Q_o$ when $c'_t = 0$, i.e. all the ice was removed by assimilation. With $n = 1$ the reduction of Q_c from 0 to $-Q_o$ is linear and with $n > 1$ it becomes steeper for values of c'_t closer to c_t . We use $n = 4$ for the runs presented here, as this gives a suitably strong heat flux for a modest reduction in concentration.

4 Results

We begin with an evaluation of a free run over the 20-month period from 1 Nov 2018 to 30 June 2020. This shows the general strengths and weaknesses of the model. We then evaluate the performance of the forecast system over this period to show the improvements gained by the assimilation of OSI-SAF concentration.

4.1 Evaluation of free model run

In this section we demonstrate that the model is generally able to reproduce the overall drift, concentration, and thickness patterns in observations. For all comparisons we average the model fields in time over an appropriate time window (in practice 1, 2 or 7 days), apply some spatial smoothing (being guided by the size of the satellite footprint), and interpolate onto the observation grid. For scalar variables, we define bias as $\langle V_{\text{mod}} - V_{\text{obs}} \rangle$ (with $\langle \cdot \rangle$ defining the spatial mean over pixels where both model and observation are defined, and where either model or observations have ice.) We also define RMSE as $\langle (V_{\text{mod}} - V_{\text{obs}})^2 \rangle^{1/2}$. For the ice extent, we define bias in extent as $A_{10} - A_{01}$, where A_{10} is the total area of pixels where neXtSIM predicts the presence of ice (total concentration greater than 15 %) but the observation has no ice, while A_{01} is the total area of pixels where neXtSIM predicts no ice (total concentration greater than 15 %) but the observation does have ice. Analogously to the RMSE for other quantities, we define IIEE (Integrated Ice Edge Error, Goessling et al., 2016) as $A_{10} + A_{01}$. Thus the IIEE is always positive, and the bias is positive if neXtSIM is overestimating the total extent and negative if it is underestimating it. For the ice velocity we define the bias and RMSE as the mean and RMS values of the difference in speed respectively — i.e. $\langle |\mathbf{u}_{\text{mod}}| - |\mathbf{u}_{\text{obs}}| \rangle$ and $\langle (|\mathbf{u}_{\text{mod}}| - |\mathbf{u}_{\text{obs}}|)^2 \rangle^{1/2}$. We also define the vector RMSE (VRMSE), as the RMS of the vector difference, $\langle |\mathbf{u}_{\text{mod}} - \mathbf{u}_{\text{obs}}|^2 \rangle^{1/2}$. Note that these drift errors can only be calculated where both model and observation have ice.

Table 1. Accuracy of the free run. Concentration and extent are evaluated against OSI-SAF SSMI; thickness is evaluated against CS2-SMOS; drift against OSI-SAF drift, where observations with reported error up to 10 km/day are considered. Results are 2-monthly-averaged.

	Concentration, %		Extent, 10 ⁶ km ²		Drift, km/day		
	Bias	RMSE	Bias	RMSE	Bias	RMSE	VRMSE
Nov-Dec 2018	4.91	9.97	0.09	0.24	0.00	3.63	5.16
Jan-Feb 2019	2.73	6.94	0.09	0.17	-0.31	2.96	4.00
Mar-Apr 2019	1.41	5.38	-0.03	0.18	-0.40	2.77	3.65
May-Jun 2019	-1.65	17.84	-0.15	0.81	-0.91	3.19	4.19
Jul-Aug 2019	-2.47	18.79	-0.06	1.05	0.23	3.25	4.75
Sep-Oct 2019	1.90	14.43	-0.06	0.47	0.81	2.80	4.50
Nov-Dec 2019	3.47	10.86	-0.11	0.41	0.32	2.96	4.53
Jan-Feb 2020	2.76	6.06	0.05	0.17	0.09	3.15	4.48
Mar-Apr 2020	1.36	4.90	-0.03	0.14	-0.15	3.06	4.34
May-Jun 2020	-0.99	16.14	-0.08	0.71	-0.74	3.30	4.41

Table 1 shows 2-monthly summary statistics for observations that are available year round (although they are much less reliable between May and September). Errors which are particularly high are the RMSE in ice concentration and the IIEE for extent between May and August — we will discuss reasons behind this below. The bias in these quantities is quite low, being highest between November and December, reflected an ice advance which is too fast. The drift on the other hand compares quite well throughout the period, with the VRMSE being generally less than 4.5 km/day and the RMSE in speed less than 3.3 km/day (with the exception of the first two months). The bias is quite low, staying between ± 0.8 km/day. The first two months have a bias of zero, despite their high VRMSE and RMSE, with a too-slow Beaufort Gyre cancelling a too-fast drift in the triangle between the north pole, Svalbard and Franz Josef land.

Table 2. Accuracy of the free run for the 2018–2019 and 2019–2020 winters. Thickness is evaluated against CS2-SMOS; drift against OSI-SAF drift, where only observations with reported error less than 1.25 km/day are considered. Results are 2-monthly-averaged.

	Thickness, m		Drift, km/day		
	Bias	RMSE	Bias	RMSE	VRMSE
Nov-Dec 2018	-0.09	0.26	-0.13	2.62	4.06
Jan-Feb 2019	0.09	0.36	-0.26	2.22	3.27
Mar-Apr 2019	0.07	0.51	-0.24	2.21	3.03
Nov-Dec 2019	0.04	0.40	0.19	2.59	4.11
Jan-Feb 2020	0.23	0.46	0.27	2.41	3.69
Mar-Apr 2020	0.19	0.56	0.01	2.39	3.64

Table 2 shows the effect of restricting the drift evaluation to regions where the observational error is less than 1.25 km/day. As discussed in Section 2.4 this removes the entire period from May to September, as well as the MIZ, the area around the north

pole and areas close to the coast. Without the less accurate observations, the drop in VRMSE is between 0.3 – 1 km/day, with their values now between 3 and 4.1 km/day. However, the less accurate regions are also regions where the model has difficulties with other variables like thickness and extent, so this could also contribute to the reduction in error. Notwithstanding this, the Nov – Dec 2018 period is now much less of an outlier, now being quite close to Nov – Dec 2019. The RMSE in speed is now less than 2.6 km/day while the bias is between ± 0.3 km/day. There is a clear deterioration going from the first to the second winter.

Table 2 also summarises the comparison of model thickness with the CS2-SMOS product. RMSE is lower (0.26 m) in Nov-Dec 2018 since thickness was initialised using CS2-SMOS, and grows to 0.51 m over the first winter. The RMSE grows from 0.4 m to 0.56 m over the second winter. The bias is quite low in the first winter, but about 0.2 m from Jan – Apr 2020.

Figures 2 and 3 compare the mean concentration and extent from neXtSIM and the OSI-SAF concentration product (see 2.3). Since we do not have enough information to create an error model for the concentration observations (for example it cannot be Gaussian due to its restriction to being between 0 and 100%) we could not generate confidence intervals for the mean concentration, but instead plot a shaded area corresponding to $\pm \langle \sigma_{\text{OSISAF}} \rangle^{1/2}$ for reference when considering the bias and RMSE in concentration. The seasonal cycle for mean concentration and extent is captured reasonably well, but the bias is clearly positive in winter and clearly negative in summer. As noted from Table 1 the RMSE in concentration and the IIEE are very high between mid-April and August. The RMSE in concentration is generally comparable to the RMS error level outside these periods.

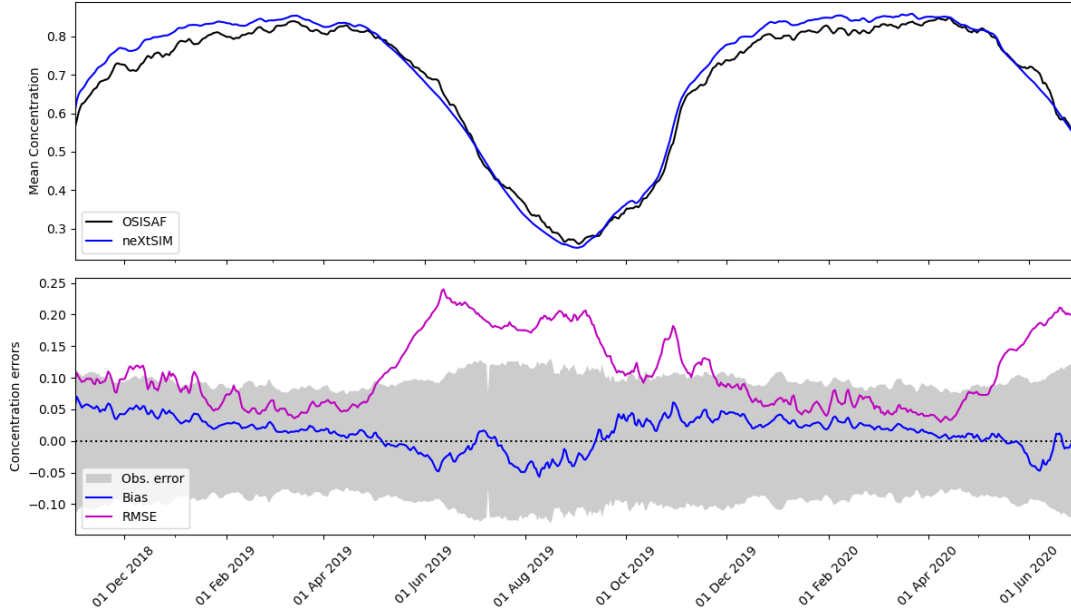


Figure 2. Temporal comparison of model and OSI-SAF SSMI concentrations. The shaded area in the lower plot shows the RMS uncertainty of the concentration for reference when considering the errors. The bias is defined as model – observation.

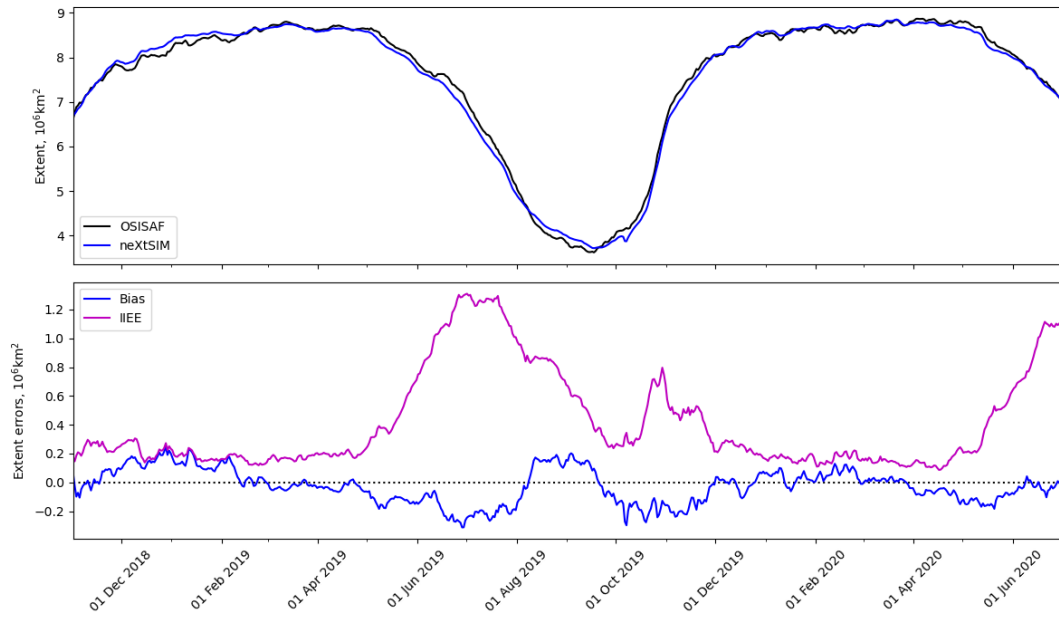


Figure 3. Temporal comparison of extents from model and OSI-SAF SSMI concentrations. The shaded area in the lower plot shows an estimate for the uncertainty in the extent. The bias is defined as model – observation.

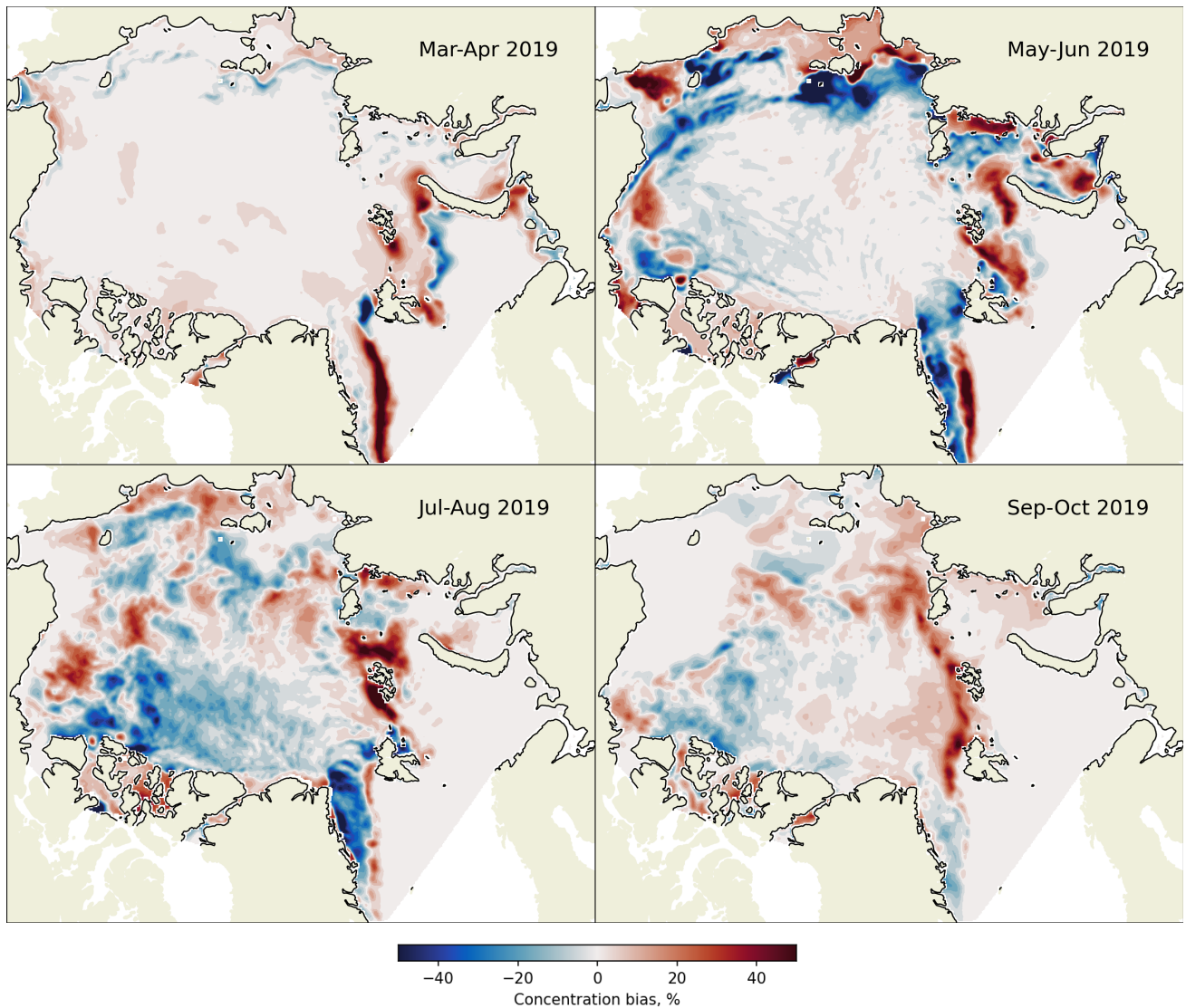


Figure 4. Spatial comparisons of selected 2-monthly-averaged concentrations from the free model run and from OSI-SAF SSMI. The bias is defined as model – observation.

Figure 4 shows some selected maps of the concentration bias. In Jan-Feb there is a general overestimation in the pack as the model is approximately 100% with some leads, while the SSMI field has large regions of lower concentration that contribute to the errors (85 – 90%). There is also a general overestimation in ice extent, which is also the case in Nov-Dec. primarily located in the Greenland Sea, and the Barents Seas. The areas north of Novaya Zemlya (around the Santa Anna Trough), around Franz Josef land, and northwest of Svalbard stand out as places where the ice edge is not being located correctly. This continues into Mar-Apr. The region of underestimation to the northwest of Svalbard is related to the development of an arch between there and the northeast tip of Greenland. (This is also related to relatively large thicknesses in these areas, which increases the

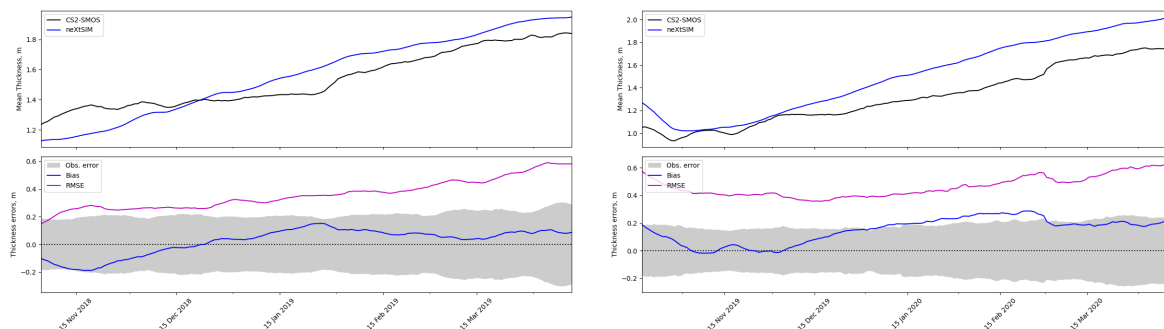


Figure 5. Comparison of model and CS2-SMOS thickness. The shaded regions show the RMS uncertainty in the CS2-SMOS product for reference. The bias is defined as model – observation.

capacity of undamaged ice to stay undamaged.) This arching also reduces the ice export through the Fram Strait, the effect of which is seen in the maps for May–Jun and Jul–Aug. In May–Jun we also see that the ice that is in the Greenland Sea is displaced too far to the east, producing a double penalty effect in the RMSE score. This is also a result of the arching in this area — the ice at the corner of Greenland is too thick to be exported, and so the ice that is exported has detached from the arch away from the coast and has then travelled roughly parallel to the coast without being pushed back towards it. In Jul–Aug the thicker ice has mostly melted and there is less of a dipole situation and more of a clear-cut underestimation.

Another problem area is around the Novo Sibirski Islands and the Laptev Sea. April sees the model opening slightly too far to the north; in May and June there is a strong underestimation to the north of the islands and too much fast ice. The underestimation is partly from too fast melting and that the fast ice has not detached and flowed into this area.

There is a similar problem to the west of Wrangel Island in the Chukchi Sea — an opening that is too large compared to the observations and too much land-fast ice close to the coast which should be flowing into this region. There is also a large region of overestimation to the east of Wrangel Island, which is actually an artefact of the boundary conditions that we are using at the open Bering Strait boundary. If there is inflow at any open boundaries the value of any tracers is taken to be its value in the nearest mesh element, and in this case we are getting too much ice being imported through the Bering Strait, which leads to a build-up at this location.

The last things to mention about the Jul–Aug map are the strong overestimation around Franz Josef land, and the dipole situation in the Beaufort sea, which is the result of a too-slow Beaufort Gyre.

However, by Sep–Oct, the situation has improved substantially. The main disagreement is in the ice edge from Severnaya Zemlya round to the just past Svalbard. This is due to an ice advance that is slightly too fast.

Figure 5 shows time series of mean thicknesses and of thickness errors when compared to CS2-SMOS, which have already been discussed to some extent in the context of Table 2. The mean observed and modelled thicknesses show a steady increase after November, but the modelled value is increasing faster than the observed one. The timing of when the modelled increase starts is close to when the observed increase starts. The lower plots have the RMS uncertainties plotted for reference, and the model bias is generally at a similar level to this uncertainty. The RMSE is about twice this uncertainty in the second winter.

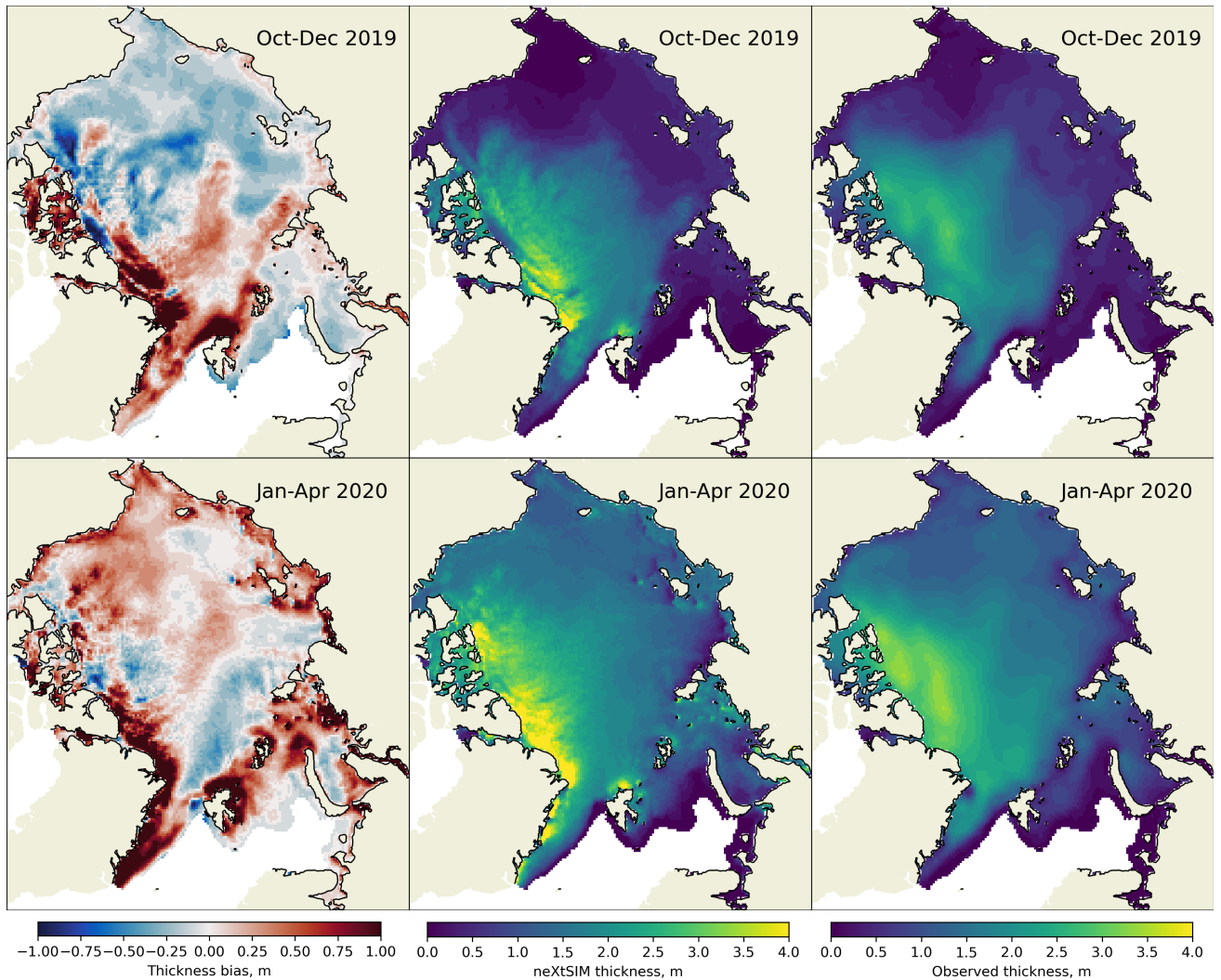


Figure 6. Comparisons of 2-monthly averaged model and CS2-SMOS thickness. Left hand column: bias maps; centre column: neXtSIM thickness; right hand column: CS2-SMOS thickness. The bias is defined as model – observation.

We note here that the error levels in the CS2-SMOS product are only the interpolation error, and are thus a lower bound as they don't include uncertainties in the individual CS2 and SMOS products. CS2 in particular is sensitive to the ice and snow densities used or the snow thickness which affect the conversion from freeboard to thickness (Zygmuntowska et al., 2014).

Figure 6 also shows the spatial distribution of the errors. Throughout the winter there is overestimation off the north coast of Greenland round to Axel Heiberg Island. Further to the west there is a thinning (but not an opening) from Axel Heiberg Island to Ellef Ringnes Island which also seems to be related to the westward drift along this coast being too high. This persists throughout the winter but the affected area reduces with time. In Oct – Dec there is a dipole pattern where the ice is too thin in

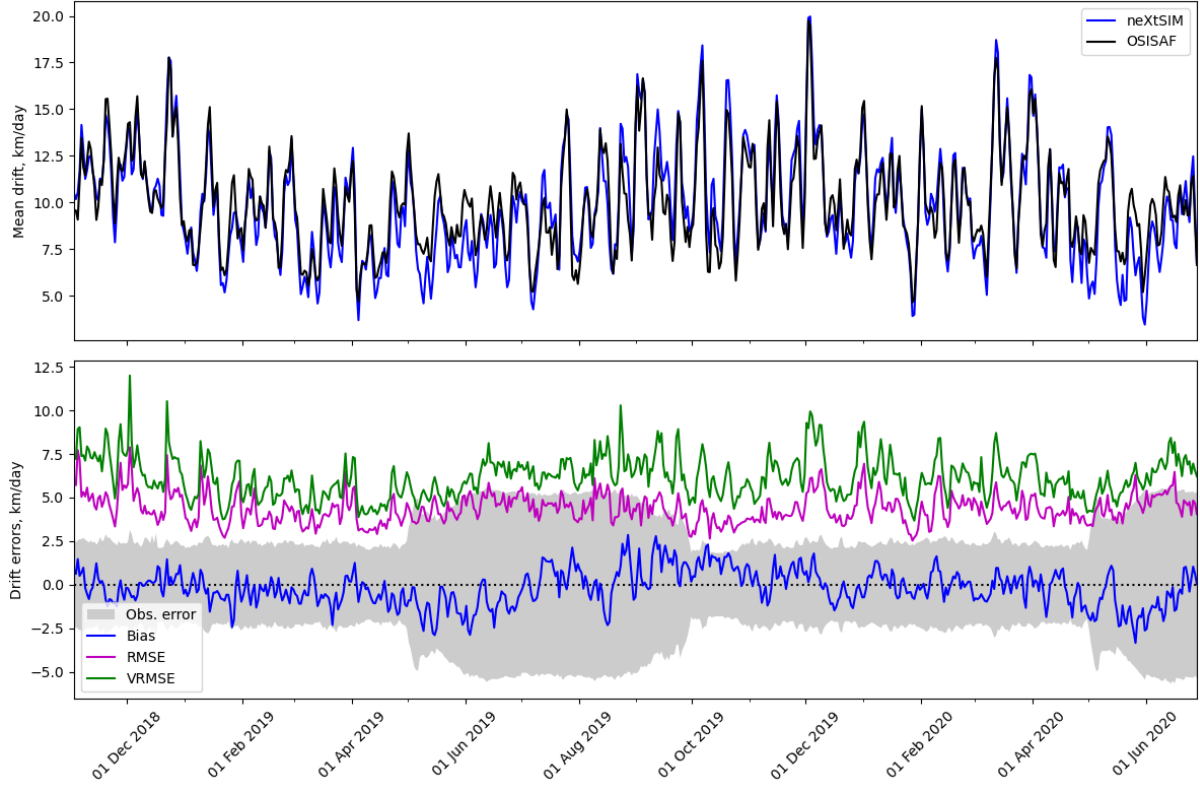


Figure 7. Comparison of model and OSI-SAF drift. The shaded region shows the RMS uncertainty in the OSI-SAF drift product. The bias is defined as model – observation.

the Beaufort Sea and too thick in the triangle between the north pole, Svalbard and Franz Josef Land. However, in Jan – Apr, this dipole reverses.

The modelled thickness is also quite high around the north of Svalbard, in contrast with the CS2-SMOS product, which has quite thin ice. This build-up contributes to the arching discussed in relation to the concentration errors, reducing the ice export through the Fram Strait. This stems from damaged ice not having enough resistance to compression. In the BBM rheology there is a balance between resisting compression enough to stop the build-up of ice at the coasts and resisting it so much that drift becomes too slow.

Figure 7 shows the drift bias and RMSE of neXtSIM when compared to the OSI-SAF drift product. The shaded area corresponds to $\langle 2\sigma_{\text{OSI-SAF}}^2 \rangle^{1/2}$, where the factor of 2 comes from the fact that drift has two components. We also note that if one cell has components (x_i, y_i) with (not-necessarily-Gaussian) noise $(\varepsilon_i, \delta_i)$ added to it (each with zero mean and variance σ_i^2), the mean-squared drift in the presence of noise,

$$\langle \tilde{d}^2 \rangle = \frac{1}{N} \sum_{i=1}^N \langle (x_i + \varepsilon_i)^2 + (y_i + \delta_i)^2 \rangle = \frac{1}{N} \sum_{i=1}^N \langle x_i^2 + y_i^2 + \varepsilon_i^2 + \delta_i^2 \rangle = \frac{1}{N} \sum_{i=1}^N (x_i^2 + y_i^2 + 2\sigma_i^2) = \langle d^2 \rangle + 2\langle \sigma_{\text{OSI-SAF}}^2 \rangle \quad (6)$$

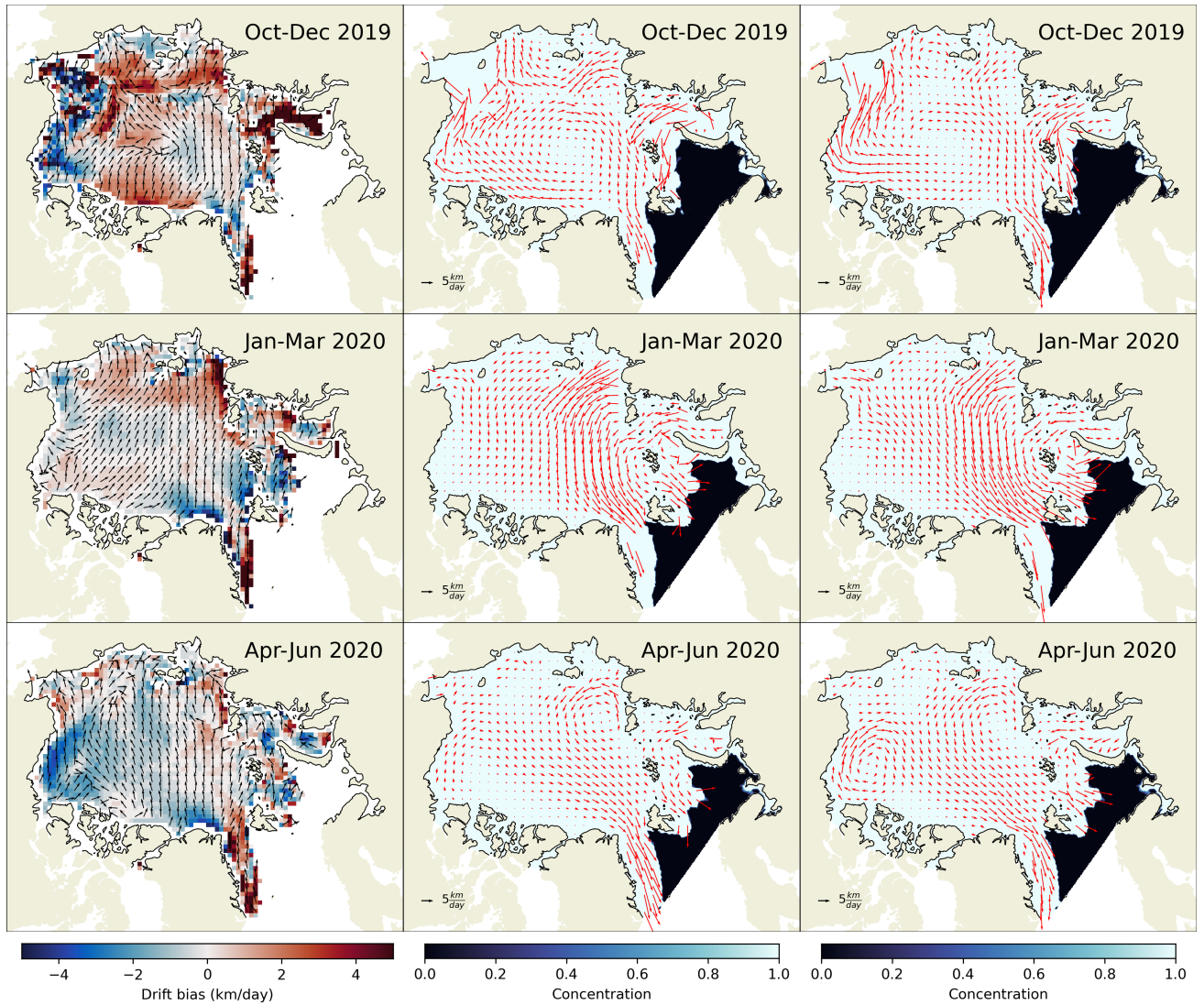


Figure 8. Comparisons of 2-monthly averaged model and OSI-SAF drift. The drift bias colour maps (left column) show the bias in speed, while the directions show the difference between the model and observation directions (arrows pointing up indicate the directions are the same). The central column shows ice velocity vectors from neXtSIM, while the right hand column shows OSI-SAF vectors. The concentration colour maps show the average neXtSIM concentration over the analysis period to indicate the approximate ice mask. The bias is defined as model – observation.

is higher than the noise-free ("true") value $\langle d^2 \rangle$ by $2\langle \sigma_{\text{OSI-SAF}}^2 \rangle$. It is tempting to then just subtract this amount from the speed when comparing with the model, but we decided that the uncertainty in the quoted uncertainty could be too influential and also that the model drift has some unknown uncertainty associated with it so we persisted with directly comparing the drift in the OSI-SAF product with the model drift in order to judge when the model is too fast or too slow, and plotted the error level $\langle 2\sigma_{\text{OSI-SAF}}^2 \rangle^{1/2}$ for reference. The difference in speed (model speed - observed speed) fluctuates somewhat, but stays largely within the error limits. The RMS error for this product is approximately 1.8 km/day from October to May, but increases to about 4 km/day from May to September. However, the model is starting to show signs of being too slow in April and May 2020. The RMSE in speed ranges between about 3–4.5 km/day, which is outside the estimated error for the more accurate colder months. The VRMSE (vector RMSE) is higher than the RMSE by about 1.5 km/day.

Figure 8 shows the general spatial pattern in the drift errors. The maps for Oct–Dec 2019 show that the general circulation of the model and the observations are agreeing well, with the main feature being the gyre in the central Arctic Ocean. However, this gyre is slightly too fast and there is also too much westward drift along the Canadian archipelago. The observations also show a small gyre in the East Siberian sea which is not apparent in the model drift. The strongest positive bias is in the Kara Sea although the errors in the observations are high there, being both close to the coast and the MIZ. The model is also showing strong positive biases in the Laptev and East Siberian seas where the ice is thinner. On the other hand, the drift in the Beaufort and Chukchi Seas is too slow.

In Jan–Mar 2020, the general pattern in the observations is a strong flow from the Laptev over to the Greenland Sea, with a weaker transpolar drift from the Chukchi Sea joining this outflow. The model also produces the flow from the Laptev Sea, but the flow from the Chukchi Sea is more directed towards the Beaufort Sea than across to the Greenland Sea. This may be contributing to the biases seen in the thicknesses — too thick in the Beaufort Sea while too thin in the region between the Laptev and Greenland Seas. The drift is also starting to become too slow north of Greenland and Svalbard, and this is most likely due to the ice build-up in those areas.

In Apr–Jun 2020 there is a stronger Beaufort Gyre and a similarly strong cyclonic gyre in the Laptev Sea. These connect to form quite a wide stream from the central Arctic to the Greenland Sea. While the model captures the gyre in the Laptev Sea, the Beaufort Gyre is less well-defined in this period, possibly due to the ice being too thick in the Beaufort Sea at this time. As in Jan–Mar, Drift is also too slow to the north of Greenland.

4.2 Evaluation of forecasts with assimilation

The performance of the forecasting system with assimilation of concentration was evaluated over the same period as the free run was in Section 4.1 (that is the 20 months from November 2018 to the end of June 2020). In order to reduce computation time, the 7-day forecasts were launched approximately only every seven days with 1-day forecasts being launched in between so that assimilation was still performed daily.

The model performance was evaluated against satellite observations of the sea ice SSMI concentration and drift from OSI-SAF (see Sections 2.3 and 2.4). The metrics we used in our assessment were: bias and RMSE in concentration; bias and IIEE in extent; bias, RMSE and VRMSE in drift. A persistence forecast of concentration (the initial concentration, defined for

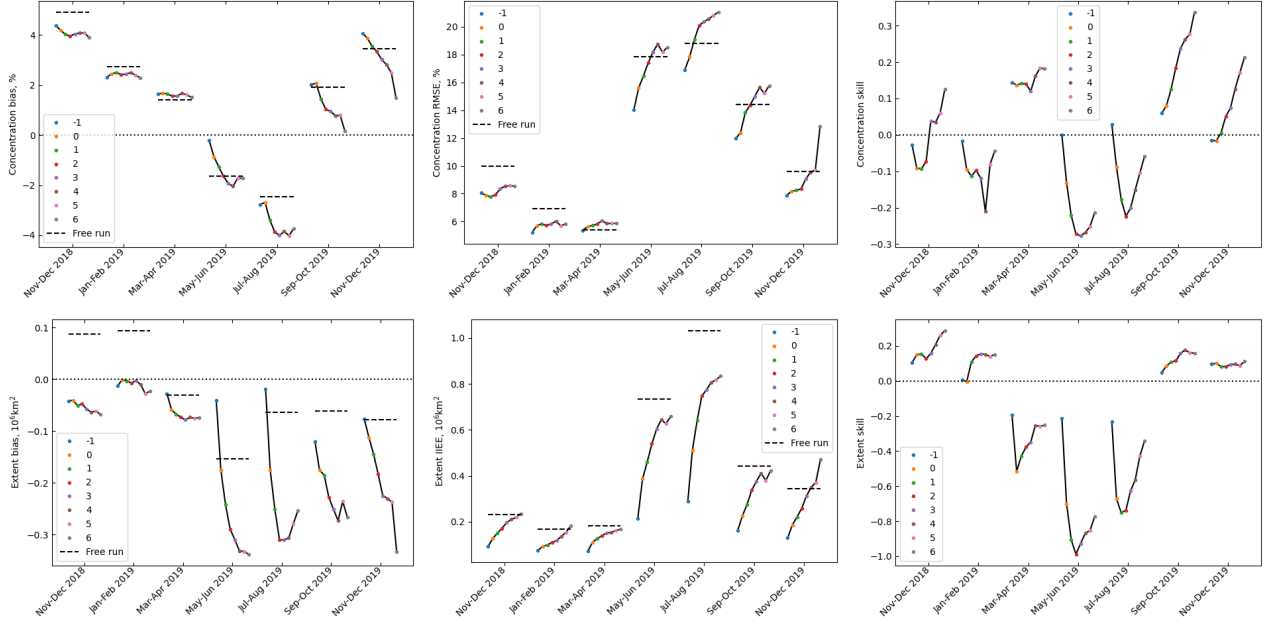


Figure 9. 2-monthly-averaged forecast errors grouped by lead time (lead times are indicated in the figure legends). The free run errors for the corresponding periods are plotted as dotted lines for reference on the bias and RMSE/IIEE plots. Upper row: 2-monthly-averaged bias, RMSE and forecast skill for concentration grouped by lead time. Middle row: 2-monthly-averaged bias, IIEE and forecast skill for extent grouped by lead time. Lower row: 2-monthly-averaged bias, IIEE and forecast skill for extent grouped by lead time.

technical convenience as the average of the first hour) was used as a benchmark for concentration and extent, and forecast skill was defined as:

$$\text{Concentration skill} = 1 - \left(\frac{\text{RMSE}}{\text{RMSE}_{\text{persistence}}} \right)^2, \quad (7a)$$

$$\text{Extent skill} = 1 - \frac{\text{IIEE}}{\text{IIEE}_{\text{persistence}}}. \quad (7b)$$

- 5 Thus, a model that is agreeing perfectly with the observations has skill equal to 1, while it is negative (with no lower bound) if the persistence error is lower than the forecast error. We decided that given the strong dependence of drift on wind, its rapid variability in time would render a persistence forecast for drift a straw man, so for this variable we just present the forecast errors. We have rough benchmarks from some other models however. Although they use different products for their evaluation so we can't make a direct comparison, we note that the drift from the TOPAZ forecast generally has a bias in speed of about
- 10 2 km/day and a VRMSE of about 5–8 km/day Melsom et al. (2018), while Metzger et al. (2017) report an RMS drift speed error of about 5–8 km/day in the Arctic for the GOFS 3.1 system.

Figure 9 shows the 2-monthly averages for the different metrics applied to the concentration and extent. We only plot results for the first year and two months, since the error statistics from the 2020 results were nearly identical to the previous years. As with the free run, the concentration from the forecast is generally higher than OSI-SAF in the winter months. As mentioned

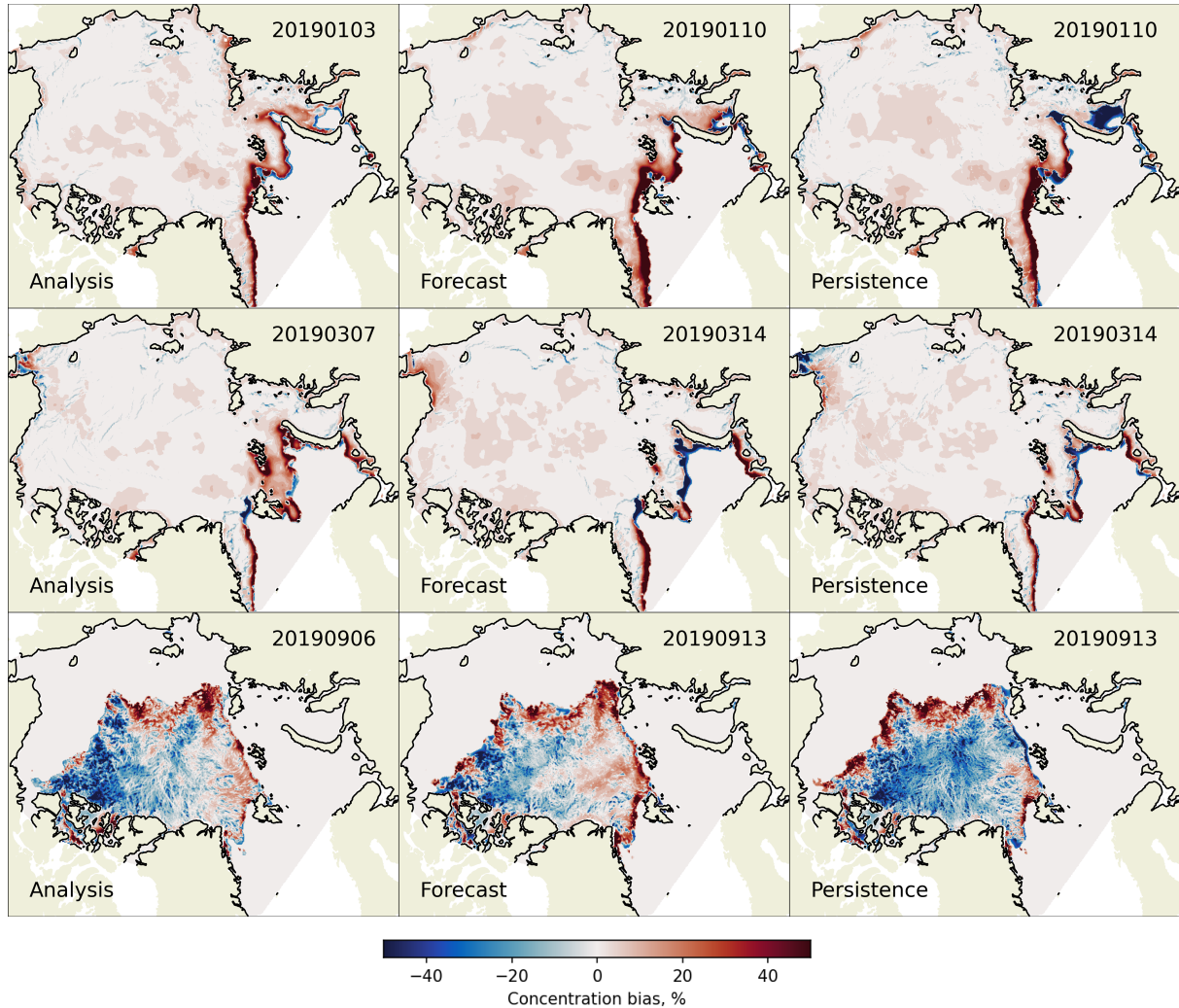


Figure 10. Maps of daily-averaged biases in concentration compared to OSI-SAF SSMI for three examples of forecasts, starting on 3 January 2019 (upper row, positive skill) and the other on 7 March 2019 (middle row, negative skill) and 6 September 2019 (lower row, positive). The left hand columns show the evaluation for the first day of simulation (lead time -1: analysis or hindcast) while the right hand columns show the evaluation for the eighth day of simulation (lead time 7). The right hand columns show the bias in the persistence forecast for lead time 7. The bias is defined as model – observation.

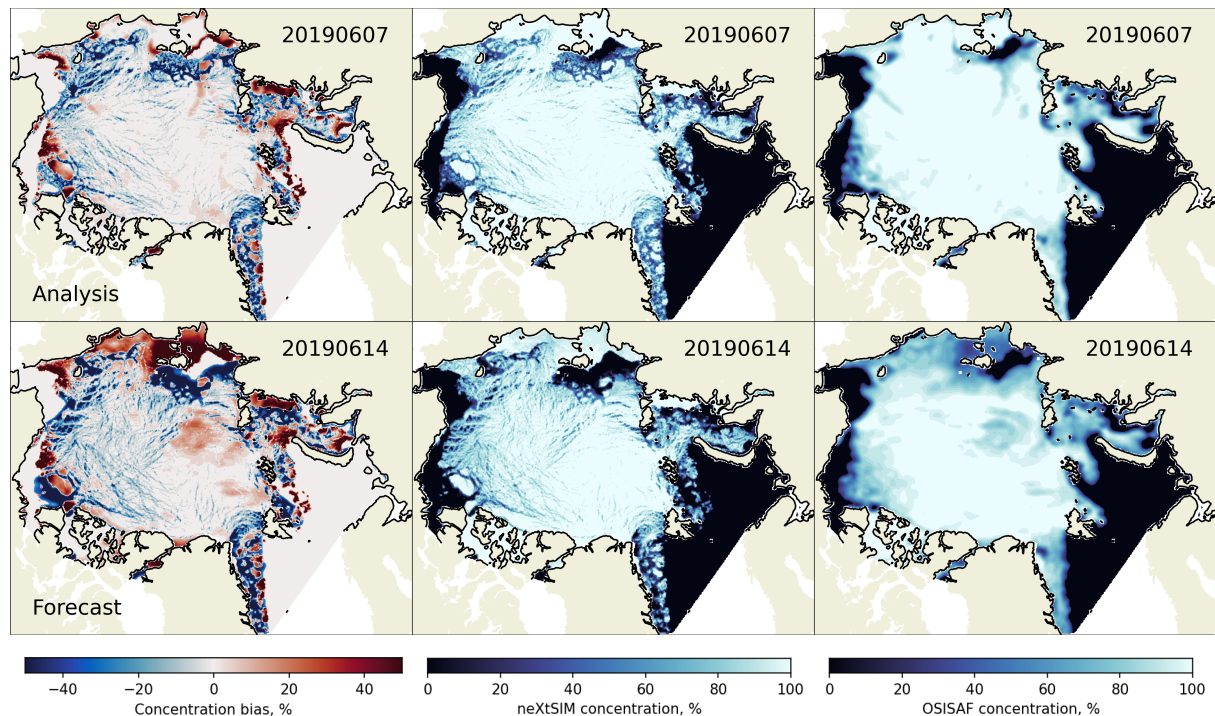


Figure 11. Maps of daily-averaged biases in concentration compared to OSI-SAF SSMI, and modelled and observed concentrations, for an example forecast starting on 6 June 2019. The upper row shows the comparison for the first day of simulation while the lower row shows the comparison for the seventh day of simulation. The bias is defined as model – observation.

earlier, we don't try to correct this as reducing the concentration in the pack causes serious problem with the drift and thickness. The extent, which is adjusted during assimilation is also lower than the observed one for the whole period, by differing amounts. Figure 10 show some example forecasts from winter and autumn, one from January, March and September 2019 respectively. The first and third examples are forecasts with positive skill in extent (model does better than persistence), while the second has negative skill. However, they all have relatively low IIEE and RMSE in concentration. The January example shows that the rapid advance in the Kara Sea is captured relatively well by the model, and while the extent in the Greenland and Barents Seas is too high, the ice edge is quite variable and the model is doing better than the persistence at this time. The March example shows an example where there is quite a persistent negative bias to the northwest of Svalbard and in the Barents Sea. These errors probably originate in the ocean forcing and which we are unable to overcome, even with our flux compensation. The persistence forecast does well in this example since the ice conditions are not changing very much. In the September example, the bias is similar to the free run but it has been reduced by the assimilation. There is still a too-fast ice advance in the European Arctic but it is closer to the observations than the persistence forecast.

In the summer, we are systematically lower in concentration, partly due to differences in the pack, which we don't correct for, and in the extent. We are not too concerned by differences in the pack, as the differences are still comparable to observation

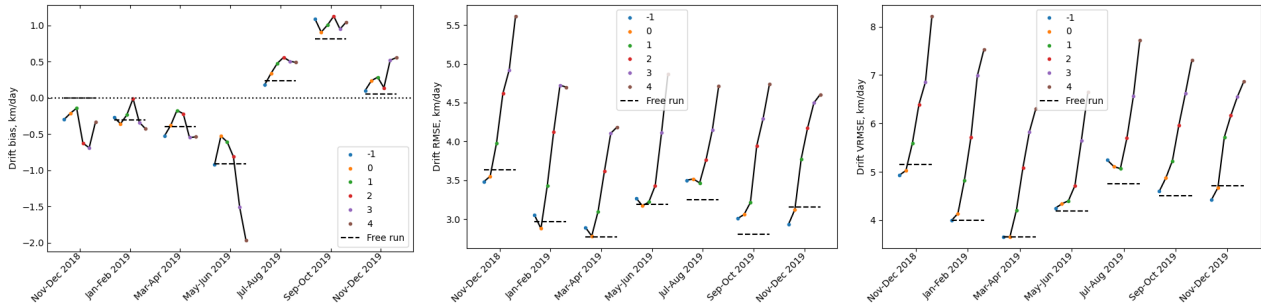


Figure 12. 2-monthly-averaged bias, RMSE and VRMSE for drift grouped by lead time (lead times are indicated in the figure legends). The lead times refer to the start of each 2-day drift evaluation period e.g. “0” covers the period from day 0 (12:00) to day 2 (12:00). The free run errors for the corresponding periods are plotted as dotted lines for reference. The bias is defined as model – observation.

error. Differences in the extent however are more concerning. In the summer, we score particularly badly with this metric, largely due to the dynamical problems also seen in the free run: arching above the Fram Strait, and fast ice in the Laptev and East Siberian seas that is too slow to detach. Figure 11 shows a typical example from this time period. Despite the correction in extent during assimilation the ice that is added in the Fram Strait is quite thin and it melts quite quickly.

- 5 Figure 12 shows the bias, RMSE and VRMSE for the forecast drift. It has similar quality to the free run, which is to be expected since we are not modifying fields in the pack, with minor differences that are probably just due to differences in extent determining which observations are included or not. There is a small negative bias in speed in the winter months, and a slight positive bias in the summer. The bias is not showing a large variability with forecast lead time, although there are some exceptions. With the exception of the first two months, which is slightly higher than the others, the RMSE (RMSE in
- 10 speed) is quite consistent throughout the whole evaluation period with a steady increase with lead time. The RMSE for the final day is still mostly less than 5 km/day. The VRMSE, which evaluates the drift direction as well as its speed, shows a larger dependence on lead time, starting between 4.5 – 5 km/day but getting up to between 7 – 8 km/day by the final day (occasionally being under 7 km/day or over 8 km/day). This dependence on lead time is almost certainly due to increasing inaccuracy in the forecast winds.

15 5 Discussion and conclusions

neXtSIM-F became operational in July 2020 as part of CMEMS, and this paper presents an upgraded version that will be included in December 2020. Here we have evaluated the neXtSIM model itself in a free run for the period 1 November 2018 to 30 June 2020, as well as the neXtSIM-F forecast platform which corrects the initial conditions daily by assimilating sea ice concentration.

- 20 The free run, which used hindcast winds, had good drift, being relatively unbiased and having a low RMSE in speed of 3 – 4 km/day. The RMSE was closer to 2.5 km/day when less accurate observations (uncertainty less than 2.5 km for the 2-day

drift) were filtered out, although the observations so removed would also be near the coast and in the MIZ where we do have problems with thickness and extent respectively. Considering the VRMSE, the RMSE in the final position of each trajectory, added about 1 km/day.

The model thickness was biased slightly too thick in general, and there was also a tendency to have a dipole behaviour with one or too areas being too thick, and one or two being too thin. The location of these too-thin regions varied over the winter (the only time when the CS2-SMOS observations were available), but the thickness off northeast Greenland and north Svalbard was invariably too high. This caused the formation of an arch across the Fram Strait, which limited the export of ice through this strait in the summer, especially close to the coast.

When we evaluated the concentration and extent we found that the sea ice extent in the summer was greatly affected by the arching mentioned above, in that the total extent was too small and that the ice was also displaced from where it should be — there was no ice close to the east coast of Greenland (unlike in the observations) and it extended too far to the east.

There were also less severe differences in the ice edge which were probably partly due to errors in the ocean and/or atmospheric forcings, but the other big issue was the extent of the landfast ice in May and June in the Laptev and East Siberian Seas — this was too slow to detach and also contributed to there being too little ice further away from the coast at this time.

Owing to the fairly conservative approach that we had to take to the assimilation the forecasts generally had the same properties as the free run with regards to thickness and drift. The bias in drift remains close to zero for the whole 8-day simulation, although the RMSE in speed and the VRMSE begin to deteriorate after about 4–5 days of simulation (up to a lead time of 4) — this is almost certainly due to less accurate winds at these lead times. The forecasts also had greatly improved IIEE scores compared to the free run. However, the more serious problems with the summer extent — in the Greenland Sea (due to arching) and to a lesser degree the Laptev and East Siberian seas (land fast ice) — were still present in the forecasts. The land-fast ice could possibly be improved with further tuning of the basal grounding scheme, while the arching problem is more difficult to solve. However, recent experiments indicate that changing the role of concentration in the rheology could give some improvements here — note that the viscous relaxation time in (A3) and thus the entire left-hand side of (A6) only depend on damage and not on concentration so larger stresses do not drop very quickly if the ice is undamaged and the concentration drops (this makes the right-hand side of (A6) drop to near zero). There may be additional modifications to the rheology that could be made to help reduce arching in this region.

Assimilation of different data might also help:

1. Sea ice thickness: we could use either the hybrid CS2-SMOS product or the CS2 trajectories themselves to limit the build up of ice on either side of the Fram Strait. However, this data stops in mid-April so build-up that occurs later may still be enough to cause the arching.
2. Sea ice deformation: Korosov and Rampal (2017) were able to derive sea ice drift from SAR data using a combination of feature tracking and pattern matching. Deformation can then be calculated to modify the damage variable and thus induce break-up at the right place. Unfortunately, this option is also time-limited as surface melt makes the ice smoother and stops the feature tracking algorithm from working as effectively.

Another possibility could be to take the approach of Ying (2019) and to morph the modelled ice mask onto the observed one so that, for example, ice in the Greenland sea could be moved over towards the coast, instead of removing thicker ice at the ice edge and adding thinner ice towards the coast. Possibly a better source of ice extent than OSI-SAF concentration could also be used in conjunction with such a morphing approach. For example, the United States Naval Ice Center⁶ produces daily, pan-Arctic ice charts. Other alternatives for assimilating extent are MASIE (Multisensor Analyzed Sea Ice Extent: Fetterer et al., 2010) and IMS (Interactive Multisensor Snow and Ice Mapping System: Helfrich et al., 2007), as done by the GOFS 3.1 system for example. Automatic ice type classification from SAR is also possible now e.g. the algorithm developed by Park et al. (2019) and upgraded by Boulze et al. (2020) will become operationally distributed by CMEMS in 2021.

A framework to produce an ensemble forecast with neXtSIM-F is also being developed (Cheng et al., submitted) (as a follow-up of the work of Rabatel et al., 2018), with the ultimate aim of using the Ensemble Kalman Filter (EnKF) assimilation method. Work on using EnKF with models running on adaptive meshes (like neXtSIM) is being developed in parallel at NERSC (Aydoğdu et al., 2019). This may also be effective at addressing some of the errors in the forecast.

Finally, in order to provide more consistent ocean inputs to the sea ice model (as well as to provide more realistic stresses and fluxes to the ocean model), work is ongoing to couple neXtSIM with the ocean models HYCOM and NEMO, and to add an atmospheric boundary layer model to mediate between the atmospheric model and the ice and/or ocean models. Also, NeXtSIM is already coupled to the WAVEWATCH 3 wave model (Boutin et al., 2020), so there is scope for neXtSIM-F to include more components like wave and ocean models.

Appendix A: Summary of rheological equations

For completeness we present the rheological equations for the current operational version of neXtSIM-F (see Appendix A1), and for the version used in this paper, which will become operational in December 2020 (see Appendix A2).

A1 Modified Maxwell-Elasto-Brittle (MEB) rheology

We refer readers to Rampal et al. (2019) for most of the details but give some important modifications below. The momentum balance is

$$\rho_i(h + h_y) \frac{D\mathbf{u}}{Dt} = \nabla \cdot (h(\boldsymbol{\sigma} + \boldsymbol{\sigma}_p)) + \tau - \rho_i(h + h_y)(f\mathbf{k} \times \mathbf{u} + g\nabla\eta), \quad (\text{A1})$$

where $\mathbf{u} = (u, v)$ is the ice velocity, ρ_i is the density of ice, h_y is the mean young ice thickness, h is the mean thickness of older ice, $\boldsymbol{\sigma}$ is the internal stress, $\boldsymbol{\sigma}_p$ is an additional plastic stress that is most active when damaged ice is under convergence, τ is the sum of stresses applied by the wind and ocean currents and by keels grounding on the sea floor, f is the Coriolis parameter, $\mathbf{k} \times \mathbf{u} = (v, -u)$, g is the acceleration due to gravity and η is the sea surface height.

⁶https://www.natice.noaa.gov/Main_Products.htm

There is an evolution equation for the damage, which is unchanged from (Rampal et al., 2019, see equations A8 – A15). The constitutive relations are

$$\dot{\boldsymbol{\sigma}} + \frac{\boldsymbol{\sigma}}{\lambda} = E\mathbf{K} : \dot{\boldsymbol{\varepsilon}}, \quad (\text{A2a})$$

$$\boldsymbol{\sigma}_P = \frac{h^2 e^{-20(1-c)} P_*}{|\nabla \cdot \mathbf{u}| + \delta_P} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \dot{\varepsilon}_{11} \\ \dot{\varepsilon}_{22} \\ \dot{\varepsilon}_{12} \end{pmatrix}, \quad (\text{A2b})$$

5 where \mathbf{K} is the stiffness tensor (see equation A4 of Rampal et al., 2019), and

$$\lambda = \lambda_0 (1 - d)^{\alpha-1} \quad (\text{A3})$$

is the viscous relaxation time, λ_0 is its (large) value for undamaged ice, d is damage, and $\alpha > 1$ is an exponent which can be tuned to control how fast λ drops due to damaging. Note we don't consider young ice in the constitutive relations since it is assumed to be weak.

10 Also, P_* is the resistance to compression and δ_P is a small parameter needed for numerical stability, c is the concentration of older ice and E is the Young's modulus:

$$E = e^{-20(1-c)} (1 - d_E) E_0, \quad (\text{A4a})$$

$$d_E = \begin{cases} d & \text{if } d \geq 0.95, \\ 0 & \text{otherwise,} \end{cases} \quad (\text{A4b})$$

where E_0 is the undamaged Young's modulus at 100% concentration, and d_E is a kind of "effective damage". Changes to the

15 MEB rheology as presented by Rampal et al. (2019) are the extra stress $\boldsymbol{\sigma}_P$ in the momentum equation and the use of d_E instead of just d , which helped to improve localisation of deformation.

A2 Brittle-Bingham-Maxwell (BBM) rheology

In this update to the rheology (for full details and further validation, see Ólason et al., in prep.), we could simplify things and make the model much faster (since we could use an explicit solver) by setting $\boldsymbol{\sigma}_P = 0$. Also we no longer needed an "effective

20 damage" d_E , so E is simply

$$E = e^{-20(1-c)} (1 - d) E_0. \quad (\text{A5})$$

The new constitutive relation is

$$\dot{\boldsymbol{\sigma}} + \frac{\boldsymbol{\sigma}}{\lambda} \left(\tilde{P} + \frac{\lambda \dot{d}}{1 - d} \right) = E\mathbf{K} : \dot{\boldsymbol{\varepsilon}}, \quad (\text{A6})$$

where

$$\tilde{P} = \begin{cases} 1 & \text{for } \sigma_n \leq 0 \text{ (diverging),} \\ 0 & \text{for } 0 < \sigma_n < P_{\max} \text{ (elastic),} \\ 1 - \frac{P_{\max}}{\sigma_n} & \text{for } \sigma_n > P_{\max} \text{ (ridging),} \end{cases} \quad (\text{A7})$$

where $\sigma_n = \frac{1}{2}(\sigma_{11} + \sigma_{22})$ is the normal stress (defined to be positive under convergent conditions), and

$$P_{\max} = P_* h^2 e^{-20(1-c)} \quad (\text{A8})$$

- 5 is the concentration- and thickness-dependent threshold for when the Bingham element starts moving (P_* is a constant tuning parameter).

Author contributions. TW and AK wrote the paper and did the experiments. PR and EO contributed to writing and planning the paper and experiments. EO was also behind many background developments in the model code.

Competing interests. No competing interests.

- 10 *Acknowledgements.* We would like to acknowledge support from the Research Council of Norway (RCN) projects neXtWIM #244001 and Nansen Legacy #276730. The projects ImpSIM 18CP10 funded by the French Navy (SHOM) and Cryosphere Virtual Laboratory (CVL AO/1-9465/18/I-NS) funded by the European Space Agency also supported the last phase of this manuscript preparation. We also benefitted from work done during the project SWARP funded by the European Commission FP7 programme (grant number 607476); from interesting discussions with Véronique Dansereau on the MEB and BBM rheology; from the work of Abdoulaye Samaké who parallelised the neXtSIM
- 15 code; and from the contribution of Mika Malila to the evaluation scripts. Finally, we would like to thank Sylvain Bouillon and Philip Griewank for their original contribution to building a forecasting system around the neXtSIM sea ice model.

References

- Aydoğdu, A., Carrassi, A., Guider, C. T., Jones, C. K. R. T., and Rampal, P.: Data assimilation using adaptive, non-conservative, moving mesh models, *Nonlinear Processes in Geophysics Discussions*, pp. 1–32, 2019.
- Azzara, A. J., Wang, H., Rutherford, D., Hurley, B. J., and Stephenson, S. R.: A 10-year projection of maritime activity in the US Arctic region, Tech. rep., 2015.
- Bleck, R.: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates, *Ocean Modelling*, 4, 55–88, 2002.
- Bouillon, S. and Rampal, P.: Presentation of the dynamical core of neXtSIM, a new sea ice model, *Ocean Modelling*, 91, 23–37, <https://doi.org/10.1016/j.ocemod.2015.04.005>, 2015.
- Boulze, H., Korosov, A., and Brajard, J.: Classification of sea ice types in Sentinel-1 SAR data using convolutional neural networks, *Remote Sensing*, 12, 2165, 2020.
- Boutin, G., Williams, T., Rampal, P., Olason, E., and Lique, C.: Wave–sea-ice interactions in a brittle rheological framework, *The Cryosphere Discussions*, 2020, 1–39, <https://doi.org/10.5194/tc-2020-19>, <https://tc.copernicus.org/preprints/tc-2020-19/>, 2020.
- Cheng, S., Aydoğdu, A., Rampal, R., Carassi, A., and Bertino, L.: Probabilistic forecasts of sea ice trajectories in the Arctic: impact of uncertainties in surface wind and ice cohesion, *Oceans*, submitted.
- Dansereau, V., Weiss, J., Saramito, P., and Lattes, P.: A Maxwell elasto-brittle rheology for sea ice modelling, *The Cryosphere*, 10, 1339–1359, 2016.
- Drange, H. and Simonsen, K.: Formulation of air-sea fluxes in the ESOP2 version of MICOM, Tech. Rep. 125, Nansen Environmental and Remote Sensing Center, Thormøhlensgate 47, Bergen 5006, Norway, 1996.
- Fetterer, F., Savoie, M., Helfrich, S., and Clemente-Colón, P.: Multisensor Analyzed Sea Ice Extent - Northern Hemisphere (MASIE-NH), Version 1, <https://doi.org/doi.org/10.7265/N5GT5K3K>, 2010.
- Geuzaine, C. and Remacle, J.-F.: Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities, *International Journal for Numerical Methods in Engineering*, 79, 1309–1331, <https://doi.org/10.1002/nme.2579>, <https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.2579>, 2009.
- Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E., and Jung, T.: Predictability of the Arctic sea ice edge, *Geophysical Research Letters*, 43, 1642–1650, <https://doi.org/10.1002/2015GL067232>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL067232>, 2016.
- Helfrich, S. R., McNamara, D., Ramsay, B. H., Baldwin, T., and Kasheta, T.: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS), *Hydrological Processes: An International Journal*, 21, 1576–1586, 2007.
- Hunke, E., Allard, R., Blain, P., Blockley, E., Feltham, D., Fichet, T., Garric, G., Grumbine, R., Lemieux, J.-F., Rasmussen, T., et al.: Should Sea-Ice Modeling Tools Designed for Climate Research Be Used for Short-Term Forecasting?, *Current Climate Change reports*, pp. 1–16, 2020.
- Hunke, E. C. and Dukowicz, J. K.: An Elastic–Viscous–Plastic Model for Sea Ice Dynamics, *J. Phys. Oceanogr.*, 27, 1849–1867, 1997.
- Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., and Elliott, S.: CICE: the Los Alamos Sea Ice Model Documentation and Software User’s Manual Version 4.1 LA-CC-06-012, T-3 Fluid Dynamics Group, Los Alamos National Laboratory, 675, 2010.
- Ivanova, N., Pedersen, L. T., Tonboe, R. T., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., and Shokr, M.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, *The Cryosphere*, 9, 1797–1817, <https://doi.org/10.5194/tc-9-1797-2015>, <https://www.the-cryosphere.net/9/1797/2015/>, 2015.

- Kaleschke, L., Tian-Kunze, X., Maaß, N., Beitsch, A., Wernecke, A., Miernecki, M., Müller, G., Fock, B. H., Gierisch, A. M., Schlünzen, K. H., Pohlmann, T., Dobrynin, M., Hendricks, S., Asseng, J., Gerdes, R., Jochmann, P., Reimer, N., Holfort, J., Melsheimer, C., Heygster, G., Spreen, G., Gerland, S., King, J., Skou, N., Søbjaerg, S. S., Haas, C., Richter, F., and Casal, T.: SMOS sea ice product: Operational application and validation in the Barents Sea marginal ice zone, *Remote Sensing of Environment*, 180, 264 – 273, <https://doi.org/https://doi.org/10.1016/j.rse.2016.03.009>, <http://www.sciencedirect.com/science/article/pii/S003442571630102X>, special Issue: ESA's Soil Moisture and Ocean Salinity Mission - Achievements and Applications, 2016.
- Korosov, A. A. and Rampal, P.: A combination of feature tracking and pattern matching with optimal parametrization for sea ice drift retrieval from SAR data, *Remote Sensing*, 9, 258, <https://doi.org/10.3390/rs9030258>, <http://www.mdpi.com/2072-4292/9/3/258>, 2017.
- Lavelle, J., Tonboe, R., Pfeiffer, H., and Howe, E.: Product User Manual for the OSI SAF AMSR-2 Global Sea Ice Concentration, Tech. Rep. SAF/OSI/CDOP2/DMI/TEC/265, Danish Meteorological Institute, http://osisaf.met.no/docs/osisaf_cdop2_ss2_pum_amr2-ice-conc_v1p1.pdf, 2016a.
- Lavelle, J., Tonboe, R., Pfeiffer, H., and Howe, E.: Validation Report for The OSI SAF AMSR-2 Sea Ice Concentration, Tech. Rep. SAF/OSI/CDOP2/DMI/SCI/RP/259, Danish Meteorological Institute, http://osisaf.met.no/docs/osisaf_cdop2_ss2_valrep_amr2-ice-conc_v1p1.pdf, 2016b.
- Lavelle, J., Tonboe, R., Jensen, M., and Howe, E.: Product user manual for osi saf global sea ice concentration, Tech. Rep. SAF/OSI/CDOP2/DMI/SCI/RP/225, Danish Meteorological Institute, 2017.
- Lavergne, T.: Validation and Monitoring of the OSI SAF Low Resolution Sea Ice Drift Product, Tech. Rep. SAF/OSI/CDOP/Met.no/T&V/RP/131, Norwegian Meteorological Institute, 2010.
- Lavergne, T. and Eastwood, S.: Low Resolution Sea Ice Drift Product User's Manual, Tech. Rep. SAF/OSI/CDOP/met.no/TEC/MA/128, Norwegian Meteorological Institute, 2010.
- 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: Oceans*, 115, 2010.
- Lemieux, J.-F., Beaudoin, C., Dupont, F., Roy, F., Smith, G. C., Shlyayeva, A., Buehner, M., Caya, A., Chen, J., Carrieres, T., et al.: The Regional Ice Prediction System (RIPS): verification of forecast sea ice concentration, *Quarterly Journal of the Royal Meteorological Society*, 142, 632–643, 2016a.
- Lemieux, J.-F., Dupont, F., Blain, P., Roy, F., Smith, G. C., and Flato, G. M.: Improving the simulation of landfast ice by combining tensile strength and a parameterization for grounded ridges, *Journal of Geophysical Research: Oceans*, 121, 7354–7368, <https://doi.org/10.1002/2016JC012006>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JC012006>, 2016b.
- Marsan, D., Stern, H., Lindsay, R. W., and Weiss, J.: Scale dependence and localization of the deformation of Arctic sea ice, *Phys. Rev. Lett.*, 93, 178 501, 2004.
- Meier, W. N.: Losing Arctic sea ice: Observations of the recent decline and the long-term context, in: *Sea Ice*, edited by Thomas, D. N., chap. 11, pp. 290–303, John Wiley & Sons, 3 edn., 2017.
- Melsom, A., Simonsen, M., Bertino, L., Hackett, B., Waagbø, G. A., and Raj, R.: Quality Information Document For Arctic Ocean Physical Analysis and Forecast Product ARCTIC_ANALYSIS_FORECAST_PHYS_002_001_A, Tech. Rep. CMEMS-ARC-QUID-002-001a, Norwegian Meteorological Institute, 2018.
- Metzger, E., Helber, R. W., Hogan, P. J., Posey, P. G., Thoppil, P. G., Townsend, T. L., Wallcraft, A. J., Smedstad, O. M., Franklin, D. S., Zamudo-Lopez, L., and Phelps, M. W.: Global Ocean Forecast System 3.1 validation test, Tech. rep., NAVAL RESEARCH LAB STENNIS DETACHMENT STENNIS SPACE CENTER MS STENNIS SPACE ... , 2017.

- Ólason, E., Boutin, G., Korosov, A., Rampal, P., and Kimmritz, M.: A new brittle rheology for large-scale sea-ice models, in prep.
- Overland, J. E., Hanna, E., Hanssen-Bauer, I., Kim, S.-J., Walsh, J. E., Wang, M., Bhatt, U. S., and Thoman, R. L.: Surface air temperature, in: Arctic Report Card 2018, NOAA, <https://www.arctic.noaa.gov/Report-Card/Report-Card-2018>, 2018.
- Owens, R. G. and Hewson, T.: ECMWF Forecast User Guide, Tech. rep., ECMWF, Reading, <https://doi.org/10.21957/m1cs7h>, <https://www.ecmwf.int/node/16559>, <p> Replaces previous editions that were available as PDF documents.</p>, 2018.
- 5 Park, J.-W., Korosov, A. A., Babiker, M., Won, J.-S., Hansen, M. W., and Kim, H.-C.: Classification of Sea Ice Types in Sentinel-1 SAR images, The Cryosphere Discussions, 2019, 1–23, <https://doi.org/10.5194/tc-2019-127>, <https://www.the-cryosphere-discuss.net/tc-2019-127/>, 2019.
- Perovich, D., Meier, W., Tschudi, M., Farrell, S., Hendricks, S., Gerland, S., Haas, C., Krumpen, T., Polashenski, C., Ricker, R., and Webster, M.: Sea ice, in: Arctic Report Card 2018, NOAA, <https://www.arctic.noaa.gov/Report-Card/Report-Card-2018>, 2018.
- 10 Rabatel, M., Rampal, P., Carrassi, A., Bertino, L., and Jones, C. K. R. T.: Impact of rheology on probabilistic forecasts of sea ice trajectories: application for search and rescue operations in the Arctic, The Cryosphere, 12, 935–953, <https://search.proquest.com/docview/2014143795?accountid=8579>, copyright - Copyright Copernicus GmbH 2018; Last updated - 2018-08-21; SubjectsTermNotLitGenreText - Arctic region, 2018.
- 15 Rampal, P., Weiss, J., Marsan, D., Lindsay, R., and Stern, H.: Scaling properties of sea ice deformation from buoy dispersion analysis, J. Geophys. Res., 113, <https://doi.org/10.1029/2007JC004143>, 2008.
- Rampal, P., Bouillon, S., Ólason, E., and Morlighem, M.: neXtSIM: a new Lagrangian sea ice model, The Cryosphere, 10, 1055–1073, <https://doi.org/10.5194/tc-10-1055-2016>, 2016.
- Rampal, P., Dansereau, V., Ólason, E., Bouillon, S., Williams, T. D., and Samaké, A.: On the multi-fractal scaling properties of sea ice deformation, The Cryosphere, pp. 1–45, 2019.
- 20 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, Cryosphere, 11, 1607–1623, <https://www.the-cryosphere.net/11/1607/2017/>, 2017.
- Sakov, P. and Oke, P. R.: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus A: Dynamic Meteorology and Oceanography, 60, 361–371, 2008.
- 25 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, 633–656, 2012.
- Samaké, A., Rampal, P., Bouillon, S., and Ólason, E.: Parallel implementation of a Lagrangian-based model on an adaptive mesh in C++: Application to sea-ice, J. Comp. Phys., 350, 84–96, <https://doi.org/10.1016/j.jcp.2017.08.055>, <http://www.sciencedirect.com/science/article/pii/S0021999117306368>, 2017.
- 30 Schweiger, A. J. and Zhang, J.: Accuracy of short-term sea ice drift forecasts using a coupled ice-ocean model, Journal of Geophysical Research: Oceans, 120, 7827–7841, <https://doi.org/10.1002/2015JC011273>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JC011273>, 2015.
- Semtner, A. J.: A model for the thermodynamic growth of sea ice in numerical investigations of climate, J. Phys. Oceanogr., 6, 379–389, 1976.
- 35 Simonsen, M., Hackett, B., Bertino, L., Røed, L. P., Waagbø, G. A., Drivdal, M., and Sutherland, G.: PRODUCT USER MANUAL For Arctic Ocean Physical and Bio Analysis and Forecasting Products ARCTIC_ANALYSIS_FORECAST_PHYS_002_001_A ARCTIC_ANALYSIS_FORECAST_BIO_002_004 ARCTIC_REANALYSIS_PHYS_002_003 ARCTIC_REANALYSIS_BIO_002_005, Tech. Rep. CMEMS-ARC-PUM-002-ALL, Norwegian Meteorological Institute, 2018.

- Stern, H. and Lindsay, R.: Spatial scaling of Arctic sea ice deformation, *Journal of Geophysical Research: Oceans*, 114, 2009.
- Tonani, M., Balmaseda, M., Bertino, L., Blockley, E., Brassington, G., Davidson, F., Drillet, Y., Hogan, P., Kuragano, T., Lee, T., et al.: Status and future of global and regional ocean prediction systems, *Journal of Operational Oceanography*, 8, s201–s220, 2015.
- 5 Tonboe, R. and Lavelle, J.: The EUMETSAT OSI SAF AMSR-2 Sea Ice Concentration Algorithm Algorithm Theoretical Basis Document, Tech. Rep. SAF/OSI/CDOP2/DMI/SCI/MA/248, Danish Meteorological Institute, http://osisaf.met.no/docs/osisaf_cdop2_ss2_atbd_amsr2-sea-ice-conc_v1p1.pdf, 2015.
- Tonboe, R. and Lavelle, J.: Product user manual for osi saf global sea ice concentration, Tech. Rep. SAF/OSI/CDOP/DMI/SCI/MA/189, Danish Meteorological Institute, 2016.
- 10 Tonboe, R., Lavelle, J., Pfeiffer, R.-H., and Howe, E.: Product user manual for osi saf global sea ice concentration, Tech. Rep. SAF/OSI/CDOP3/DMI_MET/TEC/MA/204, Danish Meteorological Institute, 2016.
- Wessel, P. and Smith, W. H. F.: A global, self-consistent, hierarchical, high-resolution shoreline database, *Journal of Geophysical Research: Solid Earth*, 101, 8741–8743, <https://doi.org/10.1029/96JB00104>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JB00104>, 1996.
- Winton, M.: A Reformulated Three-Layer Sea Ice Model, *J. Atmos. Ocean Tech.*, 17, 525–531, 2000.
- 15 Ying, Y.: A multiscale Alignment method for ensemble filtering with displacement errors, *Monthly Weather Review*, 147, 4553–4565, 2019.
- Zygmuntowska, M., Rampal, P., Ivanova, N., and Smedsrud, L. H.: Uncertainties in Arctic sea ice thickness and volume: new estimates and implications for trends, *The Cryosphere*, 8, 705–720, 2014.