The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

Arctic Mission Benefit Analysis: Impact of Sea Ice Thickness, Freeboard, and Snow Depth Products on Sea Ice Forecast Performance

Thomas Kaminski[1], Frank Kauker[2,5], Leif Toudal Pedersen[3], Michael Voßbeck[1], Helmuth Haak[4], Laura Niederdrenk[4], Stefan Hendricks[5], Robert Ricker[5], Michael Karcher[2,5], Hajo Eicken[6], and Ola Gråbak[7]

[1]The Inversion Lab, Tewessteg 4, 20249 Hamburg, Germany
[2]Ocean Atmosphere Systems, Hamburg, Germany
[3]eolab.dk, København, Denmark
[4]Max Planck Institute for Meteorology, Hamburg, Germany
[5]Alfred Wegener Institute, Bremerhaven, Germany
[6]University of Alaska Fairbanks, Fairbanks, USA
[7]European Space Agency ESRIN, Frascati (Rome), Italy

Correspondence to: Thomas Kaminski (Thomas.Kaminski@Inversion-Lab.com)

Abstract. Assimilation of remote sensing products of sea ice thickness (SIT) into sea ice-ocean models has been shown to improve the quality of sea ice forecasts. Open questions are whether the assimilation of rawer products such as radar freeboard (RFB) can achieve yet a better performance and what performance gain can be achieved by the joint assimilation with a snow depth product. The Arctic Mission Benefit Analysis (ArcMBA) system was developed to address this type of question. Using the quantitative network design (QND) approach, the system can evaluate, in a mathematically rigorous fashion, the observational constraints imposed by individual and groups of data products.

We present assessments of the observation impact (added value) in terms of the uncertainty reduction in a four-week forecast of sea ice volume (SIV) and snow volume (SNV) for three regions along the Northern Sea Route by a coupled model of the sea ice-ocean system. The assessments cover seven satellite products, three real products and four hypothetical products. The real products are monthly SIT, sea ice freeboard (SIFB), and RFB, all derived from CryoSat-2 by the Alfred Wegener Institute. These are complemented by two hypothetical monthly laser freeboard (LFB) products (one with low accuracy and one with high accuracy), as well as two hypothetical monthly snow depth products (again one with low accuracy and one with high accuracy).

On the basis of the per-pixel uncertainty ranges that are provided with the CryoSat-2 SIT, SIFB, and RFB products, the SIT achieves a much better performance for SIV than the SIFB product, while the performance of RFB is more similar to that of SIT. For SNV, the performance of SIT is only low, the performance of SIFB higher and the performance of RFB yet higher. A hypothetical LFB product with low accuracy (20 cm uncertainty) lies in performance between SIFB and RFB for both SIV and SNV. A reduction in the uncertainty of the LFB product to 2 cm yields a significant increase in performance.

Combining either of the SIT/freeboard products with a hypothetical snow depth product achieves a significant performance increase. The uncertainty in the snow product matters: A higher accuracy product achieves an extra performance gain. The pro-

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

vision of spatial and temporal uncertainty correlations with the EO products would be beneficial not only for QND assessments, but also for assimilation of the products.

1 Introduction

Over the last decades the state of the Arctic climate system has undergone a rapid change. Most pronounced are decreases of the summer sea ice extent and of the year-round sea ice volume. This transformation is affecting marine ecosystems and coastal communities in an unprecedented way. Economic activities such as resource extraction, maritime transportation, and tourism may benefit from these changes provided that risks, e.g. of sea ice hazards, can be managed. In this context, the performance of short-term to seasonal forecasts of sea ice conditions is of crucial importance.

Forecasts of the sea ice-ocean dynamics are routinely performed by coupled sea ice-ocean models that are driven by prescribed atmospheric conditions. Such forecasts suffer from uncertainty in the model's initial state, the atmospheric boundary conditions, and the parametrisation of physical processes. Only observations can help to reduce such uncertainties and, thus, improve the forecast quality. Recently Earth observation (EO) products of sea ice thickness (SIT) have been shown to provide particularly valuable constraints (Lisaeter et al., 2007; Yang et al., 2014; Day et al., 2014; Kauker et al., 2015; Xie et al., 2016). The constraint from rawer EO products that are used to derive SIT products may be even stronger, because these rawer products are typically more accurate. In the example of the CryoSat-2 SIT product (Ricker et al., 2014) retrieved at the Alfred Wegener Institute (AWI) the uncertainty in the radar freeboard (RFB) product underlying their SIT retrieval is smaller by about two orders of magnitude (Figure 13). This is a consequence of uncertainty in assumptions in particular on snow and ice density and snow depth, which are used to retrieve SIT from RFB. For direct assimilation of RFB these variables can be taken from the model into which the data are assimilated, but still they are uncertain. Hence, the trade-off between assimilation of SIT or RFB requires a rigorous quantitative assessment. This is even more important, when the products are assimilated jointly with products of further variables such as snow depth (SND) that bring in complementary information.

Such rigorous assessments can be performed in an efficient manner by the quantitative network design (QND) approach. QND allows an objective evaluation of observation impact on a given aspect of a model simulation or forecast. The technique originates from seismology (Hardt and Scherbaum, 1994) and was first applied to the climate system by Rayner et al. (1996), who optimised the spatial distribution of in situ observations of atmospheric carbon dioxide. After an initial QND study that demonstrated the feasibility of the approach for remote sensing of the column-integrated atmospheric carbon dioxide concentration (Rayner and O'Brien, 2001) QND is now routinely applied in the design of CO2 space missions (e.g., Patra et al., 2003; Houweling et al., 2004; Crisp et al., 2004; Feng et al., 2009; Kadygrov et al., 2009; Kaminski et al., 2010; Hungershoefer et al., 2010; Rayner et al., 2014; Bovensmann et al., 2015). For the western Arctic domain, the QND approach was successfully demonstrated through the evaluation of the combination of hypothetical airborne altimeter/radar observations of SIT/SND (Kaminski et al., 2015). The study evaluated two idealised flight transects derived from NASA's Operation IceBridge airborne ice surveys in terms of their potential to improve ten-day to five-month forecasts of sea ice conditions.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

The present study describes the implementation of the QND methodology into a system for Arctic mission benefit analysis (ArcMBA) and then applies the system to investigate the impact of a series of EO products onto forecasts of snow and ice volume over three regions along the Northern Sea Route. It addresses products of SIT, SIFB, RFB, laser freeboard (LFB), and SND. The layout of the remainder of this article is as follows: Section 2 will describe the methodological aspects, including the QND approach, the coupled sea ice-ocean model, and the EO products. Section 3 will present the simulated sensitivities of target quantities and observation equivalents to the model's control vector that is composed of process parameters, initial and boundary conditions. Section 4 will present the QND assessments, followed by their discussion in Section 5. Finally, Section 6 provides a summary and conclusions.

2 Methods

2.1 Quantitative Network Design

The QND methodology is presented by Kaminski and Rayner (2017), partly based on algebra by Tarantola (2005) and Rayner et al. (2016). For the sake of self-containedness we provide a shortened form of the presentation by Kaminski and Rayner (2017). As mentioned the QND formalism performs a rigorous uncertainty propagation from the observations via the control vector to a target quantity of interest through a dedicated modelling chain. Hence, it is worth recalling the four influence factors which produce uncertainty in a model simulation:

1. Uncertainty caused by the formulation of individual process representations and their numerical implementation (structural uncertainty).
2. Uncertainty in constants (process parameters) in the formulation of these processes (parametric uncertainty).
3. Uncertainty in external forcing/boundary values (such as surface winds or precipitation) driving the relevant processes.
4. Uncertainty in the state of the system at the beginning of the simulation (initial state).

The first factor reflects the implementation of the model (code) while the others can be understood as input quantities controlling the behaviour of a simulation using the given model implementation. The QND procedure formalises these input quantities through the definition of a control vector, x. The choice of the control vector is a subjective element in the QND procedure. A good choice covers all input factors with high uncertainty and high impact on simulated observations d_{mod} or target quantities y (Kaminski et al., 2012; Rayner et al., 2016).

The target quantity may be any quantity that can be extracted from a simulation with the underlying model, i.e. any potential model output (in the current study regional integrals of predicted sea ice and snow volumes, see Section 2.2), but also any component of the control vector, for example a process parameter such as the albedo of the snow. In the general case, where the target quantity is not part of the control vector, the QND procedure operates in two steps (Figure 1). The first step (inversion step) uses the observational information to reduce the uncertainty in the control vector, i.e. from a prior to a posterior state of information, and the second step (prognostic step) propagates the posterior uncertainty forward to the simulated target quantity.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

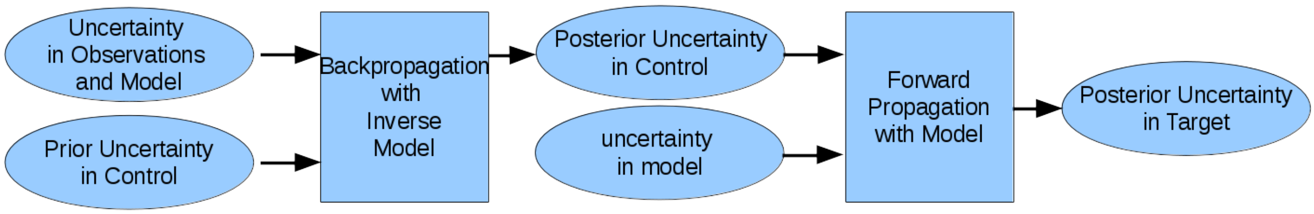

Figure 1. Data flow through two-step procedure of QND formalism. Ovals boxes denote data, rectangular boxes denote processing. Figure taken from Kaminski and Rayner (2017)

In this procedure we take uncertainty into account by representing all variables, i.e. the prior and posterior control vectors as well as the observations, their equivalents simulated by the model, and the simulated target quantity by probability density functions (PDFs). We typically assume a Gaussian form for the prior control vector and the observations, if necessary after a suitable transformation. The Gaussian PDFs' covariance matrices express the uncertainty in the respective quantities, i.e. $\mathbf{C}(x_0)$ and $\mathbf{C}(d_{\text{obs}})$ for the prior control vector and the observations.

For the first QND step we use the model M as a mapping from control variables onto equivalents of the observations. In our notation the observation operators that map the model state onto the individual data streams (see Kaminski and Mathieu (2017) and Section 2.5) are absorbed in M. Let us first consider the case of a linear model, for which we denote by $\mathbf{M}'$ the Jacobian matrix of M, i.e. the derivative of M with respect to x. In this case the posterior control vector is described by a Gaussian PDF with covariance $\mathbf{C}(x)$, i.e. the uncertainty is given by

$$\mathbf{C}(x)^{-1} = \mathbf{M'}^T \mathbf{C}(d)^{-1}\mathbf{M}' + \mathbf{C}(x_0)^{-1} \tag{1}$$

where the data uncertainty $\mathbf{C}(d)$ combines $\mathbf{C}(d_{\text{obs}})$ with the uncertainty $\mathbf{C}(d_{\text{mod}})$ in the simulated equivalents of the observations $M(x)$:

$$\mathbf{C}(d)^2 = \mathbf{C}(d_{\text{obs}})^2 + \mathbf{C}(d_{\text{mod}})^2 \tag{2}$$

The first term in Equation (1) expresses the observational constraint and the second term the prior information content. In the non-linear case we use Equation (1) as an approximation of $\mathbf{C}(x)$.

In the second step, the Jacobian matrix $\mathbf{N}'$ of the model (now used as a mapping from the control vector onto target quantities and denoted by N) is employed to propagate the posterior uncertainty in the control vector $\mathbf{C}(x)$ forward to the uncertainty in a target quantity $\sigma(y)$:

$$\sigma(y)^2 = \mathbf{N}'\mathbf{C}(x)\mathbf{N'}^T + \sigma(y_{\text{mod}})^2 \,. \tag{3}$$

If the model was perfect, $\sigma(y_{\mathrm{mod}})$ would be zero. In contrast, if the control variables were perfectly known, the first term on the right hand side would be zero. The terms $\mathbf{C}(d_{\mathrm{mod}})$ in Equation (2) and $\sigma(y_{\mathrm{mod}})$ in Equation (3) capture the structural uncertainty as well as the uncertainty in those process parameters, boundary and initial values that are not included in the control vector. These two terms typically rely on subjective estimates. When comparing the effect of different data sets in the same setup, $\sigma(y_{\mathrm{mod}})$ acts as an offset (for the respective variance) in Equation (3)). To sharpen the contrast between the products we remove it from the assessment and report two plausible estimates separately.

To conduct a valuable QND assessment, the requirement on the model is not that it simulates the target quantities and observations under investigation realistically, but the requirement is that it provides a realistic *sensitivity* of the target quantities and observations under investigation with respect to a change in the control vector. If these sensitivities, i.e. the Jacobians, are realistic, but the simulation of target quantities and observations incorrect, we can always make a valuable QND assessment with appropriate model uncertainty. The result of the assessment may then be that a particular data stream is not useful in constraining a particular target quantity given current modelling capabilities. In this situation we could operate the QND system with reduced model uncertainty to explore which accuracy of the model is required for a data stream to be a useful constraint on a given target quantity. In particular when it comes to new data streams and target quantities the accuracy of both, the simulation and the sensitivities, are hard to assess. In the case of a model that misses relevant processes we may expect errors in both the simulation and the sensitivities, and consequently also in the QND assessment.

Our performance metric is the (relative) reduction in posterior target uncertainty $\sigma(y)^2$ with respect to a reference. To compare against the case without any observations we use, as the reference, the prior target uncertainty

$$\sigma(y_0)^2 = \mathbf{N}'\mathbf{C}(x_0)\mathbf{N}'^T + \sigma(y_{\mathrm{mod}})^2 \,. \tag{4}$$

The uncertainty reduction with respect to the prior,

$$\frac{\sigma(y_0) - \sigma(y)}{\sigma(y_0)} = 1 - \frac{\sigma(y)}{\sigma(y_0)}, \tag{5}$$

quantifies the impact of the entire network. A schematic illustration of the approach with the prior and posterior uncertainty ranges is shown in Figure 2. The observations d_1 and d_2 render a range of trajectories unlikely, which in the first step leads to a reduction of uncertainty in the control vector (from $C(x_0)$ to $C(x)$) and in the second step to a reduction in the target uncertainty (from $\sigma(y_0)$ to $\sigma(y)$).

We note that (through Equation (1) and Equation (3)) the posterior target uncertainty solely depends on the prior and data uncertainties, the contribution of the model error to the uncertainty in the simulated target variable, $\sigma(y_{\mathrm{mod}})$, as well as the observation and target Jacobians (quantifying the linearised model responses of the simulated observation equivalent and of the target quantities). The QND formalism does not require real observations and can thus be employed to evaluate hypothetical candidate networks. Candidate networks are defined by a set of observations characterised by observational data type, location, time, and data uncertainty. Here, we define a network as the complete set of observations, d, used to constrain the model. The term network is not meant to imply that the observations are of the same type or that their sampling is coordinated. For example, a network can combine different types of in situ and satellite observations.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

The Cryosphere Discussions

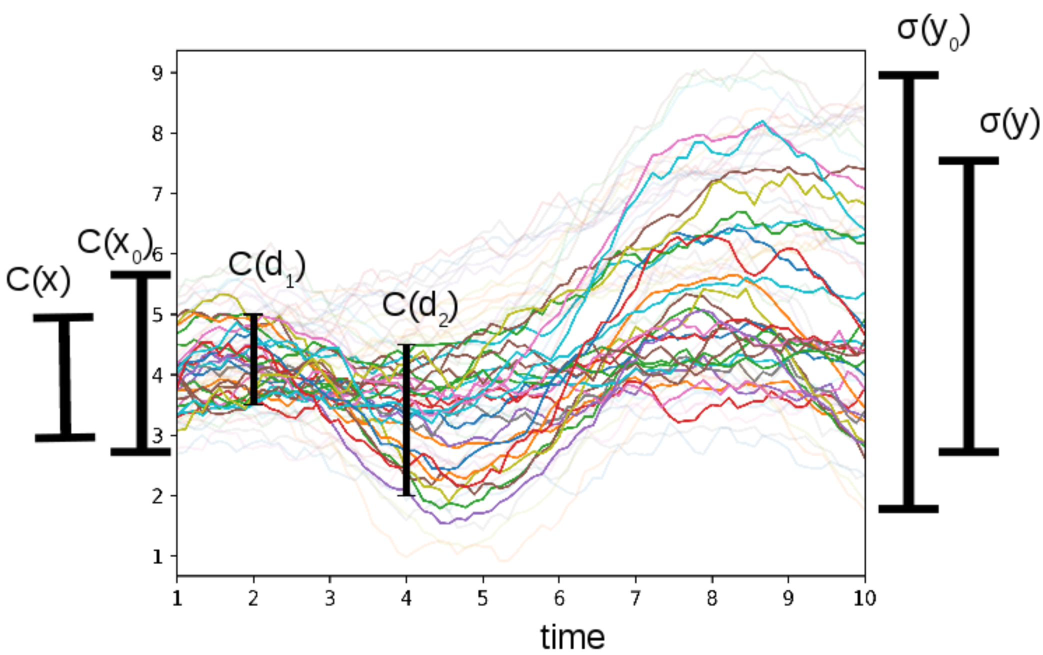

Figure 2. Schematic Presentation of QND procedure.

In practice, for pre-defined target quantities and observations, model responses can be pre-computed and stored. A network composed of these pre-defined observations can then be evaluated in terms of the pre-defined target quantities without any further model runs. Only matrix algebra is required to combine the pre-computed sensitivities with the data uncertainties. This aspect is exploited in our ArcMBA system.

2.2 Target Quantities

For this study we selected target quantities that are particularly relevant for maritime transport, namely predicted sea ice volume (SIV) and snow volume (SNV) over three regions along the Northern Sea Route (NSR). These three regions are displayed in Figure 3 and respectively denoted as “West Laptev Sea” (WLS), “Outer New Siberian Islands” (ONSI), and “East Siberian Sea” (ESS). We perform these predictions for May 28, 2015, a point in time at which there is still sufficient snow cover for our prediction to be relevant. These predictions are started on April 1 and are constrained by observational information until April 30, i.e. we perform a four-week prediction (Figure 4).

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

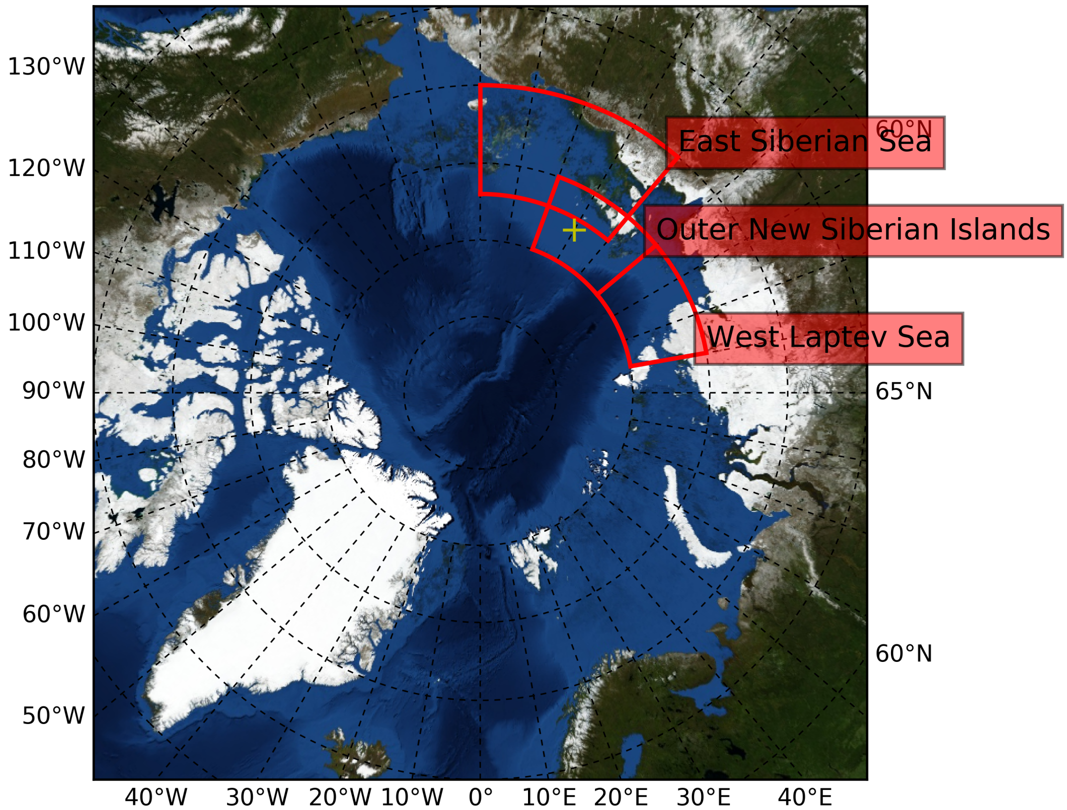

Figure 3. Target regions along the Northern Sea Route.

2.3 Model

The requirement on the dynamical model of the coupled sea ice-ocean system is that it simulates in a realistic manner the sensitivity of the observation equivalents and the target quantities to changes in the control variables. In the present study we use the Max-Planck-Institute Ocean Model (MPIOM) (Jungclaus et al., 2012, 2013; Haak et al., 2003), i.e. the sea ice-ocean component of the Max-Planck-Institute Earth System Model (MPI-ESM) (Giorgetta et al., 2013). MPI-ESM regularly provides climate projections for the Intergovernmental Panel on Climate Change (IPCC) in particular to the IPCC's 5th assessment report (Stocker et al., 2013) and the upcoming 6th assessment report (AR6) and within the seasonal to decadal prediction

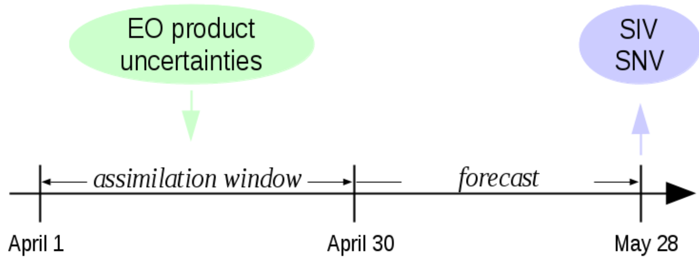

Figure 4. Time line of experimental setup.

system (Müller et al., 2012). In the following we provide a brief description of the model, largely following Jungclaus et al. (2006) and Niederdrenk (2013).

MPIOM is based on the primitive equations, a set of nonlinear differential equations that approximate the oceanic flow and are used in most oceanic models. They consist of three main sets of balance equations: A continuity equation representing the conservation of mass, the Navier-Stokes equations ensuring conservation of momentum, and a thermal energy equation relating the overall temperature of the system to heat sources and sinks. Diagnostic treatment of pressure and density is used to close the momentum balance. Density is taken to be a function of model pressure, temperature and salinity (UNESCO, 1983). Recent development of the model includes the treatment of horizontal discretisation which has undergone a transition from a staggered E-grid to an orthogonal curvilinear C-grid. The treatment of subgridscale mixing has been improved by the inclusion of a new formulation of bottom boundary layer slope convection, an isopycnal diffusion scheme, and a Gent and McWilliams style eddy-induced mixing parameterisation. Along-isopycnic diffusion is formulated following Redi (1982) and Griffies (1998). Isopycnal tracer mixing by unresolved eddies is parameterised following Gent et al. (1995). For the vertical eddy viscosity and diffusion the Richardson number–dependent scheme of Pacanowski and Philander (1981) is used. An additional wind mixing proportional to the cube of the 10-m wind speed (decaying exponentially with depth) compensates for too low turbulent mixing close to the surface. Static instabilities are removed through enhanced vertical diffusion. A viscous–plastic rheology (Hibler, 1979) is used for the sea ice dynamics. The thermodynamics is formulated using a Semtner (1976) zero-layer model relating changes in sea ice thickness to a balance of radiant, turbulent, and oceanic heat fluxes. In the zero-layer model the conductive heat flux within the sea ice/snow layer is assumed to be directly proportional to the temperature gradient across the sea ice/snow layer and inversely proportional to the thickness of that layer, i.e. the sea ice does not store heat. The effect of snow accumulation on sea ice is included, along with snow–ice transformation when the snow/ice interface sinks below the sea level because of snow loading (flooding). The effect of ice formation and melting is accounted for within the model assuming a sea ice salinity of 5 psu. MPIOM allows for an arbitrary placement of the model's poles on an orthogonal curvilinear grid. In the setup used here (taken from Niederdrenk (2013); Mikolajewicz et al. (2015); Niederdrenk et al. (2016)) the poles are located over Russia and North America (as can be seen in Figure 5). Placement over land avoids numerical singularities that for poles over the ocean would be caused by the convergence of the meridians, and the non-diametric placement allows to reach high resolution (average of about 15 km) of the Arctic. In the following we will call the model in that configuration Arctic MPIOM.

As forcing data at the ocean's surface, the model needs heat, freshwater and momentum. These data are taken from ECMWF's ERA-Interim reanalysis (Dee et al., 2011). ERA-Interim is a global atmospheric reanalysis (of the period from 1979 to present) that is produced by a 2006 release of the Integrated Forecasting System (IFS – version Cy31r2) and applies a 4-dimensional variational analysis with a 12-hour analysis window. The spatial resolution of the data set is approximately 80 km (T255 spectral) on 60 vertical levels from the surface up to 0.1 hPa. ERA-interim surface variables to force Arctic MPIOM are 2-meter temperature, 2-meter dew point temperature (surrogate of 2-meter specific humidity – not delivered by ECMWF), 10-meter zonal and meridional wind velocity (to calculate the wind speed), total cloud cover and the following fluxes (delivered in accumulated form over the 12-hourly forecast window): surface downward solar radiation, surface downward thermal radiation, total precipitation, zonal and meridional wind stress. Land runoff into the ocean is taken from the German Ocean Model Intercomparison Project (OMIP, Röske, 2001).

In this study, all model experiments will be started from a restart file generated from a hindcast run of Arctic MPIOM which is initialised on 1.1.1979 (start time of ERA-Interim). This initialisation is based on a set of observations that consists of a topography data set (ETOPO5 5-minute gridded elevation data (NOAA, 1988)), and a hydrographic climatological data set (Polar science center Hydrographic Climatology, PHC3; Steele et al., 2001) containing potential temperature and salinity. The ocean is assumed to be at rest. Sea ice is assumed to be present if the sea surface temperature falls below the freezing temperature of sea water. 100% ice cover and a sea ice thickness of 2m is assumed where sea ice is present. From this initial state the model is integrated with the ERA-Interim surface forcing until 31.3.2015. While a 34 year integration is certainly too short to spin up the deep ocean, it is sufficient for the purpose of this study, because the spinup time of sea ice and the upper ocean (depth above about 500m) is generally assumed to be only a few decades.

Arctic MPIOM has been validated against remotely sensed ice concentration from the reprocessed OSI SAF sea ice concentration product (Eastwood et al., 2015) and against a combination of in-situ and remotely sensed ice thickness observations. In-situ observations of sea ice thickness still have a high uncertainty and each data source has its own strengths and weaknesses. As of today the most reliable information about pan-Arctic sea ice thickness stems from a combination of various sources of in-situ observation and remotely sensed satellite sea ice thickness products by Lindsay and Schweiger (2015).

The reprocessed OSI SAF sea ice concentration product is available daily on a 10 km spatial grid and includes spatially and temporally varying uncertainty estimates. For an assessment of the performance of the Arctic MPIOM, the sea ice concentration has been compared to the long-term means of the March, June and September monthly means for the period 1990 to 2008 (Figure 6). In March (panel d) and June (panel e) only small misfits to the OSI SAF ice concentration are found. The sea ice margin in the Nordic Seas and Barents Sea is captured well. The anomalies apparent in March correspond to the results of a study performed with the MPIOM version of the Max-Planck-Institute's Earth System model MPI-ESM-LR (Notz et al.,

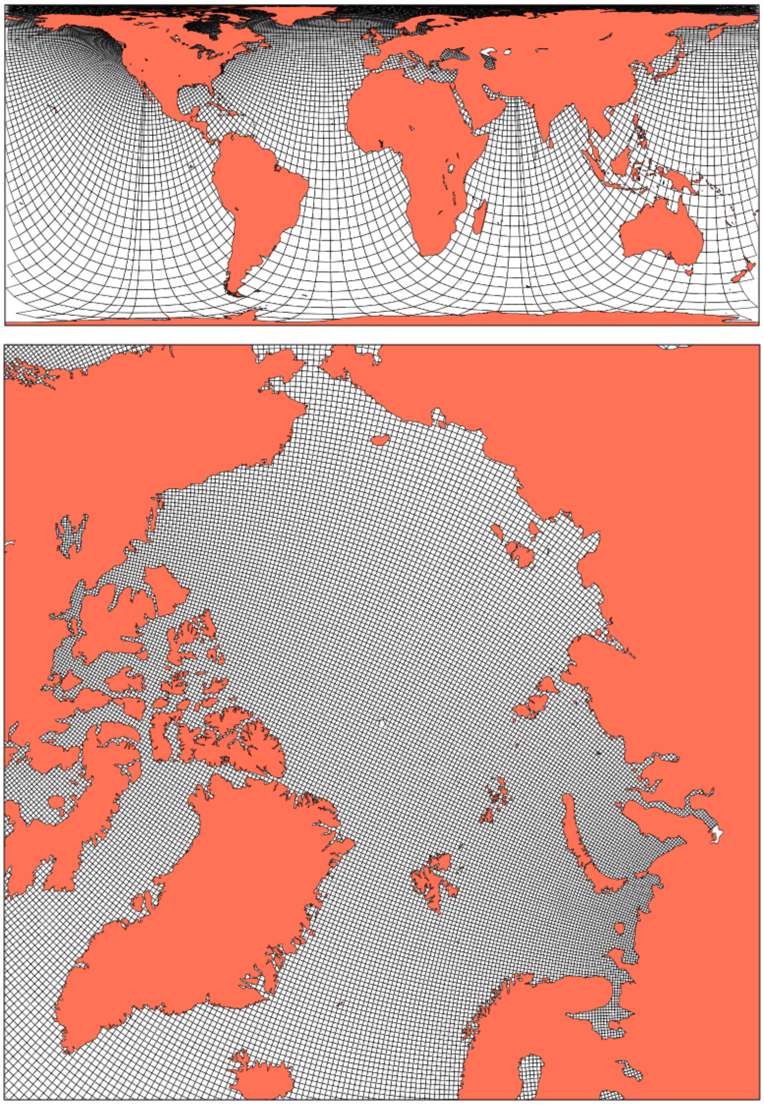

Figure 5. Model grid, global (upper panel) and over the Arctic (lower panel); Upper panel shows groups of 4 by 4 grid cells and lower panel groups of 2 by 2 grid cells.

2013), for which the MPIOM was forced with the same atmospheric forcing data set as used in our study (ERAinterim) (see panel f of their Figure 3).

In September large misfits to the OSI SAF sea ice concentration are obtained (Figure 6 panel f). Especially over the Eurasian basin the model's sea ice margin is located too far north but also over the central Arctic the model is underestimating the sea ice concentration. In our target reions the misfit remains relatively small. The aforementioned analysis by Notz et al. (2013)

shows similar misfits (see panel f of their Figure 4) to a different sea ice concentration data set, namely NSIDC-CDR (National Snow and Ice Data Center Climate Data Record).

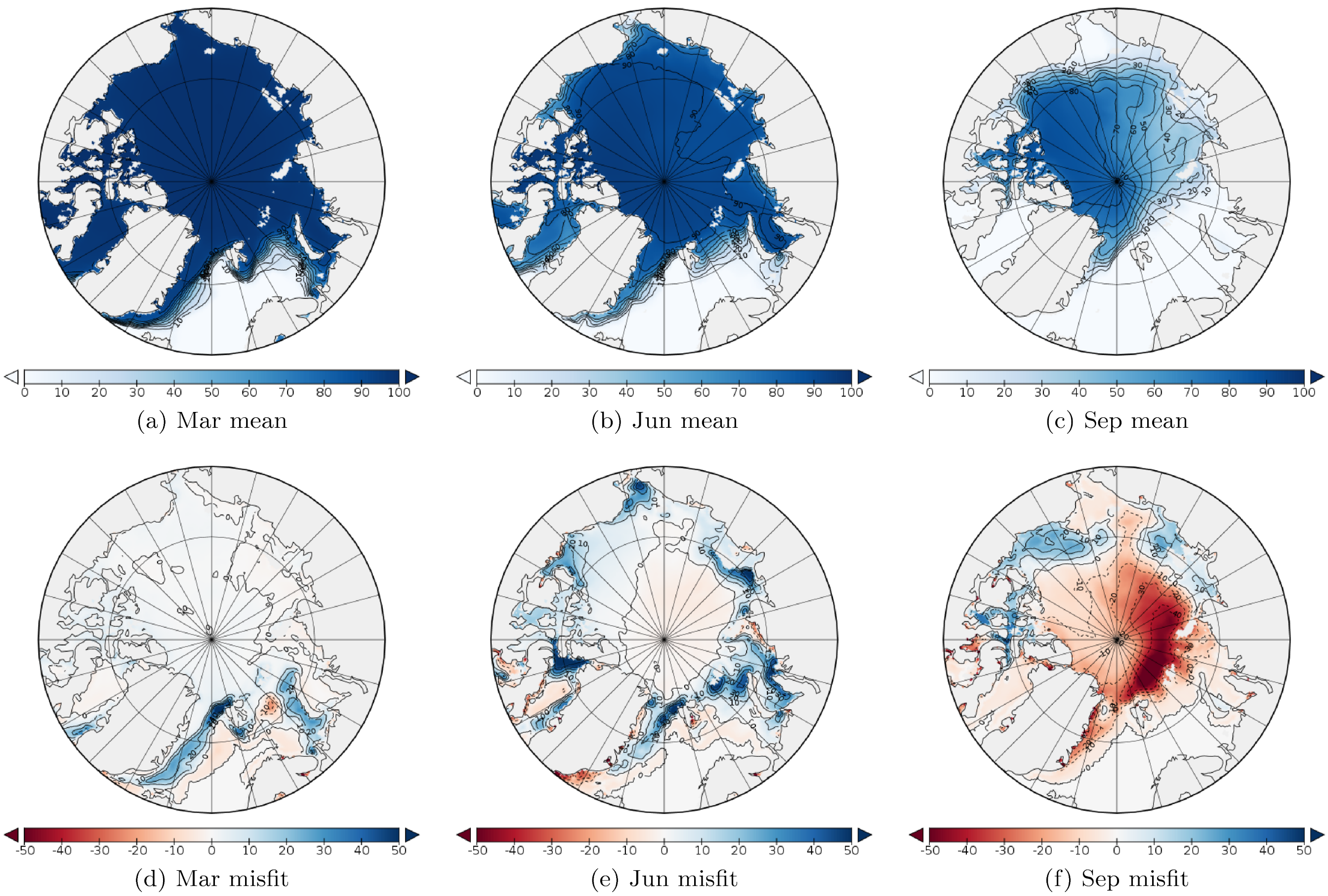

Figure 6. The long-term mean sea ice concentration [%] of the Arctic MPIOM for 1990 to 2008 for March, June and September (panel a to c) and the misfit to the OSI SAF sea ice concentration (panel d to f).

An evaluation of the hindcast simulation with Arctic MPIOM with respect to the modelled SIT is much more difficult, because the observation-based products exhibit large uncertainties reflecting the corrections imposed by the respective measurement principle. For example, Electro-magnetic Air-EM measurements detect the air-snow interface, and not the interface between snow and sea ice, introducing significant errors in the SIT estimates that are corrected by assumptions or measurements of snow depth. Moored and submarine ULS measurements have to be corrected for the first return echo. Differences in the observed and measured spatial scales further complicate the comparison. The aforementioned study of Lindsay and Schweiger (2015) synthesises all available in-situ and remotely sensed satellite SIT products in an ice thickness regression procedure (ITRP) for the time period 2000 to 2012. Low order spatial and temporal polynomials are fitted to the available sea ice thickness measurements. The resulting sea ice thickness regression product describes the evolution in the central Arctic and is linear in time plus a quadratic time-dependent component, i.e. it does not contain year-to-year variability. Uncertainty ranges are deduced from the uncertainty of the individual regression coefficients. The year-to-year variability is reflected in this uncertainty. Lindsay and Schweiger (2015) could show for example that the ICESat ice thickness product from the Jet Propulsion Laboratory (ICESat-JPL, Kwok and Cunningham (2008)), which is widely used for model validation, had a large positive bias.

Here we compare the modelled long-term mean (2000 to 2012) sea ice thickness of the Arctic MPIOM model experiment to the ITRP sea ice thickness for the two-months periods February/March and October/November. We selected these two-month periods, because the availability of the ICESat satellite product ensures a high data coverage in the ITRP. The long-term mean sea ice thickness of the the Arctic MPIOM hindcast simulation for February/March and October/November is depicted in Figure 7 (panel a and panel b) together with the misfit to the ITRP ice thickness (panel c and panel d). A prominent feature is a strong underestimation of the Arctic MPIOM sea ice thickness north and west of Fram Strait and in the strait itself. In the regions of interest of our QND study, in the areas around the Northern Sea Route, the misfit is moderate in February/March (overestimation of about 25%) with the exception around the East Siberian Islands where the misfit can reach more then 1 meter (overestimation of about 50%). In October/November the misfit is very moderate in these areas except for Bering Strait where Arctic MPIOM underestimates the sea ice thickness by more then 50cm.

As we base our QND experiments on simulations from April 1 to May 28, we show the April mean and the May 28 mean of the modelled SIT and the misfit of the April mean thickness to that retrieved from CryoSat-2 (Figure 8). The misfit to the CryoSat-2 ice thickness in April 2015 is similar to the misfit to the ITRP shown in Figure 7: a strong underestimation north of the Canadian Archipelago and North and West of Fram Strait and a moderate overestimation in the area of the target quantities of about or less then $50cm$ (about 25% relative error). Figure 9 depicts the April mean and the May 28 mean of the modelled snow depth and the misfit to the modified Warren climatology (Warren et al., 1999) that is used in the CryoSat-2 retrieval. Note that on May 28 parts of the target regions are almost snow free already. The misfit to the modified Warren climatology in the target area East Siberian Sea is on the order of about $10cm$ (50% relative error) but much less for the other target areas.

Overall the misfits of the Arctic MPIOM are acceptable in particular for our target regions along the Northern Sea Route (Figure 3) and are comparable to misfits found in sea ice-ocean model intercomparison projects (e.g, Chevallier et al. (2017)).

2.4 Control Vector

The definition of the control vector and the specification on prior uncertainty follow Kaminski et al. (2015): The components and their prior uncertainty are listed in Table 1. The largest possible control vector in our modelling system is the superset of initial and surface boundary conditions as well as all parameters in the process formulations, including the observation operators. To keep our ArcMBA system numerically efficient, two and three-dimensional fields are partitioned into regions. More precisely, we divide the Arctic domain into nine regions (shown in Figure 10). In each of these regions we add a scalar perturbation to each of the forcing fields (indicated in Table 1 by “f” in the type column - the perturbation is applied for the entire simulation time). Likewise we add a scalar perturbation to six initial fields (indicated in by “i” in the type column). For

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

The Cryosphere Discussions
Open Access

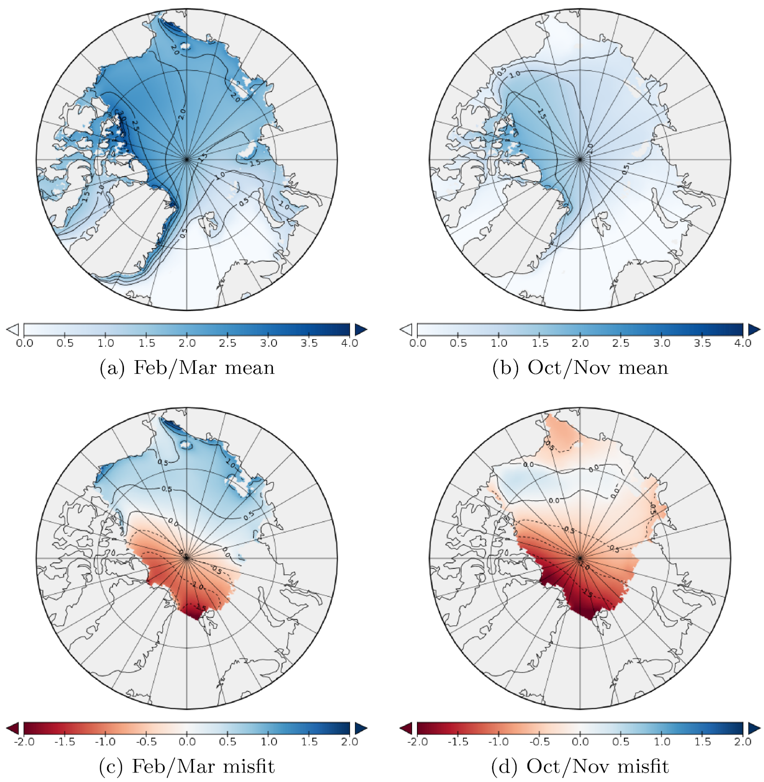

Figure 7. The long-term mean (2000 to 2012) of the simulated ice thickness [m] for the two-month periods February and March and October and November (panel a and b) and the misfit (model – observation) to the ITRP.

the ocean temperature and salinity the size of the perturbation is reduced with increasing depth (and zero below $500m$). Finally we have selected 29 process parameters from the sea ice–ocean model plus two parameters from the observation operators for freeboard products (see Section 2.5 for details). This procedure resulted in a total of 157 control variables. We assume the prior uncertainty to have diagonal form, i.e. there are no correlations among the prior uncertainty relating to different components of the control vector. The diagonal entries are the square of the prior standard deviation. For process parameters this standard deviation is estimated from the range of values typically used within the modelling community. The standard deviation for the components of the initial state is based on a model simulation over the past 37 years and computed for the 37 member ensemble corresponding to all states on the same day of the year. Likewise the standard deviation of the surface boundary conditions is computed for the 37 member ensemble corresponding to the April-October means of the respective year.

2.5 Data Sets and Observation Operators

The study evaluates three data sets retrieved by the AWI (Ricker et al., 2014) from observations provided by the CryoSat-2 mission, two data sets characterising hypothetical LFB products, and two data sets characterising hypothetical SND products.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

Table 1. Control variables. Column 1 lists the quantities in the control vector; column 2 gives the abbreviation for each quantity; column 3 indicates whether the quantity is an atmospheric boundary field (forcing, i.e. f), an initial field (i), or a process parameter (p); column 4 gives the name of each quantity; column 5 indicates (the standard deviation of) the prior uncertainty and the corresponding units and provides the magnitude of the parameter value in parenthesis, where applicable; and column 6 identifies the position of the quantity in the control vector – for initial and boundary values (which are differentiated by region) this position refers to the first region, while the following components of the control vector then cover regions 2 to 9.

Index #	Name	Type	Meaning	Prior uncertainty (value)	Start
1	hiccp	p	(alias pstar) ice strength (devided by density)	$15(20)\,[N\,m^{-2}\,kg^{-1}]$	1
2	hibcc	p	(alias cstar) ice strength depend. on ice conc.	$5.0(20.0)$	2
3	hicce	p	(alias eccen) squared yield curve axis ratio	$0.5(2.0)$	3
4	rleadclose1	p	extra lead closing (Notz et al., 2013)	$0.2(0.25)$	4
5	rleadclose2	p	extra lead closing (Notz et al., 2013)	$1.0(3.0)$	5
6	rleadclose3	p	extra lead closing (Notz et al., 2013)	$1.0(2.0)$	6
7	h_0	p	lead closing	$1.0(0.5)\,[m]$	7
8	hmin	p	mimimal ice thickness	$0.04(0.05)\,[m]$	8
9	armin	p	minimal ice compactness	$0.15(0.15)$	9
10	hsntoice	p	limit on flooding	$0.45(0.45)$	10
11	sice	p	salinity in sea ice	$2.0(5.0)[psu]$	11
12	albi	p	freezing ice albedo	$0.1(0.75)$	12
13	albm	p	melting ice albedo	$0.1(0.70)$	13
14	albsn	p	freezing snow albedo	$0.1(0.85)$	14
15	albsnm	p	melting snow albedo	$0.1(0.70)$	15
16	rhoice	p	density of sea ice	$20(910)[kg/m^3]$	16
17	rhosn	p	density of snow	$20(330)[kg/m^3]$	17
18	cw	p	ocean drag coeff.	$2.0\times10^{-3}(4.5\times10^{-3})$	18
19	av0	p	coeff vertical viscosity	$1.\times10^{-4}(2.\times10^{-4})$	19
20	dv0	p	coeff vertical diffusitiv	$1.\times10^{-4}(2.\times10^{-4})$	20
21	aback	p	background coeff vertical viscosity	$3.\times10^{-5}(5.\times10^{-5})$	21
22	dback	p	background coeff vertical diffusivity	$1.\times10^{-5}(1.05\times10^{-5})$	22
23	cwt	p	vertical wind mixing parameter tracers	$2.0\times10^{-4}(3.5\times10^{-4})$	23
24	cwa	p	vertical wind mixing parameter momentum	$0.4\times10^{-3}(0.75\times10^{-3})$	24
25	cstabeps	p	vertical wind mixing stability parameter	$0.03(0.06)$	25
26	cdvocon	p	coefficient for enhanced vertical diffusivity	$0.1(0.15)$	26
27	bofric	p	linear bottom friction	$2.\times10^{-4}(3.\times10^{-4})$	27
28	rayfric	p	quadratic bottom friction	$0.5\times10^{-3}(1.\times10^{-3})$	28
29	$jerlov_a$	p	jerlov atten - ocean-water types	$0.04(0.08)$	29
30	$jerlov_b$	p	jerlov bluefrac - ocean-water types	$0.20(0.36)$	30
31	albw	p	open water albedo	$0.05(0.1)$	31
32	sit	i	initial ice thickness	$0.5\,[m]$	32
33	siconc	i	initial ice concentration	0.1	41
34	sicsno	i	initial snow thickness	$0.2\,[m]$	50
35	thetao	i	initial ocean temperature	$0.5\,[K]$ (vertically decreasing)	59
36	so	i	initial salinity	$0.5\,[psu]$ (vertically decreasing)	68
37	zos	i	sea level elevation	$0.08\,[m]$	77
38	cloud	f	cloud cover	0.07	86
39	prec	f	total precipitation	$0.4\times10^{-8}\,[m\,s^{-1}]$	95
40	swrad	f	solar downward radiation	$6.\,[W\,m^{-2}]$	104
41	tdew	f	dew pointe temperature	$1.1\,[K]$	113
42	tem	f	2m air temperature	$1.2\,[K]$	122
43	wind10	f	10m scalar wind speed	$0.6\,[m\,s^{-1}]$	131
44	wix	f	zonal wind stress x component	$0.02\,[N\,m^2]$	140
45	wiy	f	meridional wind stress y component	$0.02\,[N\,m^2]$	149

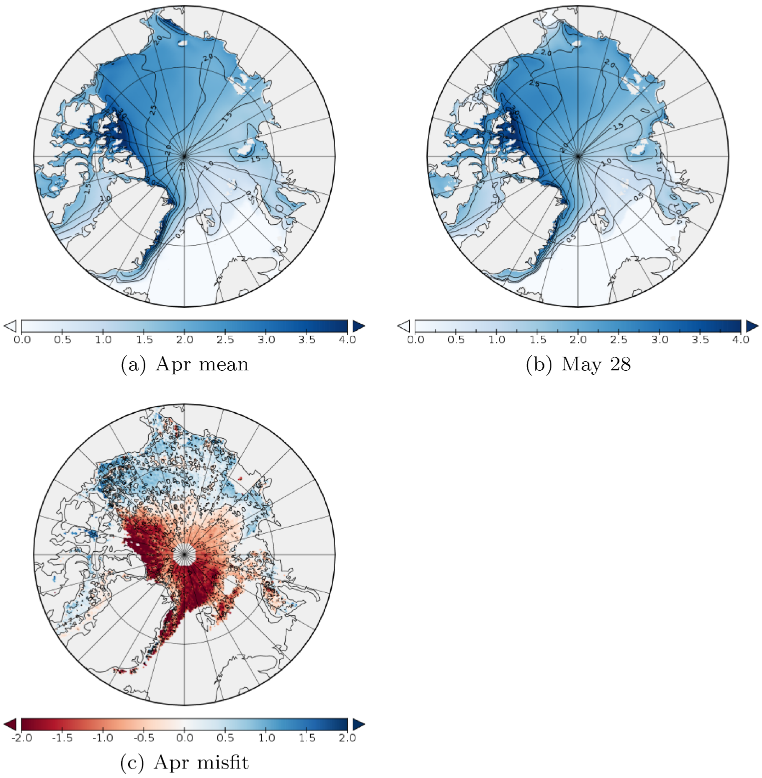

Figure 8. The a) modelled mean April 2015 sea ice thickness [m], b) the modelled sea ice thickness on May 28 2015, and c) the mean April 2015 misfit of the modelled sea ice thickness to the CryoSat-2 sea ice thickness.

In the following we describe these data sets and the simulation of their model equivalents, i.e. the respective observation operators that provide the link from the model's state variables (Kaminski and Mathieu, 2017) .

The three products derived by AWI from CryoSat-2 are SIT (h_i), SIFB (f_i), and RFB (f_r). Their definition is illustrated in Figure 11 together with that of LFB (f_l).

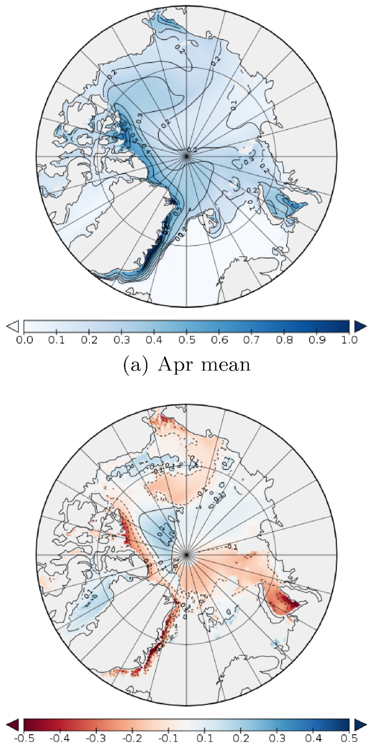

Figure 9. The a) modelled mean April 2015 snow depth [m], b) the modelled snow depth on May 28 2015, and c) the mean April 2015 misfit of the modelled snow depth to the modified Warren climatology used in the CryoSat-2 sea ice thickness retrieval.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

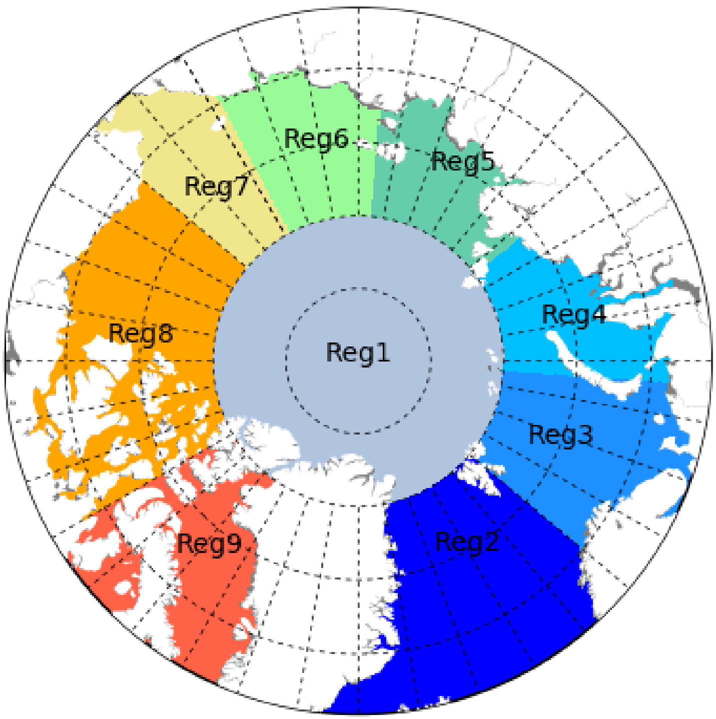

Figure 10. Sub-regions for spatial differentiation of initial and boundary values in the control vector. 1 (light plum): central Arctic; 2 (dark blue): North Atlantic; 3 (blue) Barents Sea; 4 (light blue) Kara Sea; 5 (green) Laptev Sea, 6 (light green) East Siberian Sea; 7 (yellow): Bering Strait/Chukchi Sea; 8 (orange): Beaufort Sea; 9 (red): Baffin Bay.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

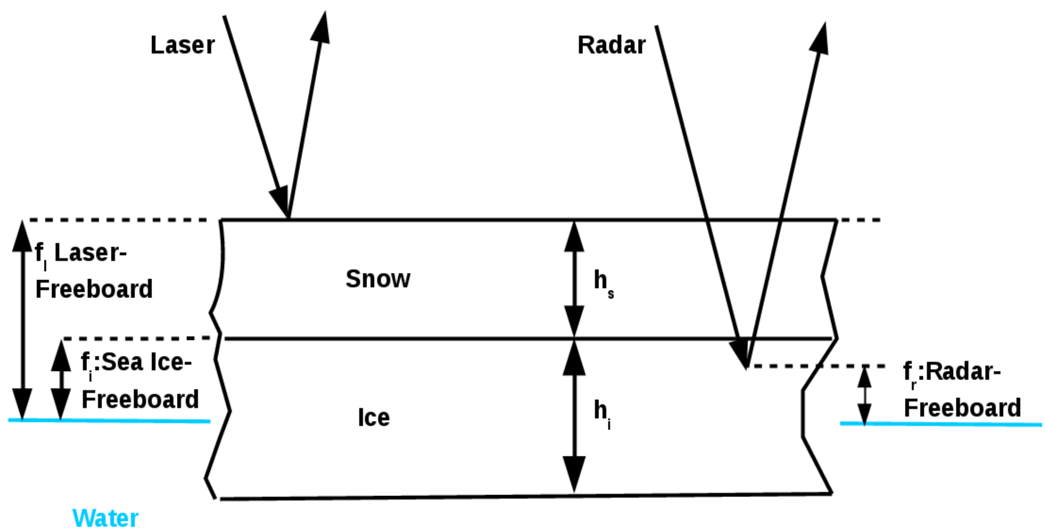

Figure 11. Schematic illustration of sea ice thickness and different freeboard variables.

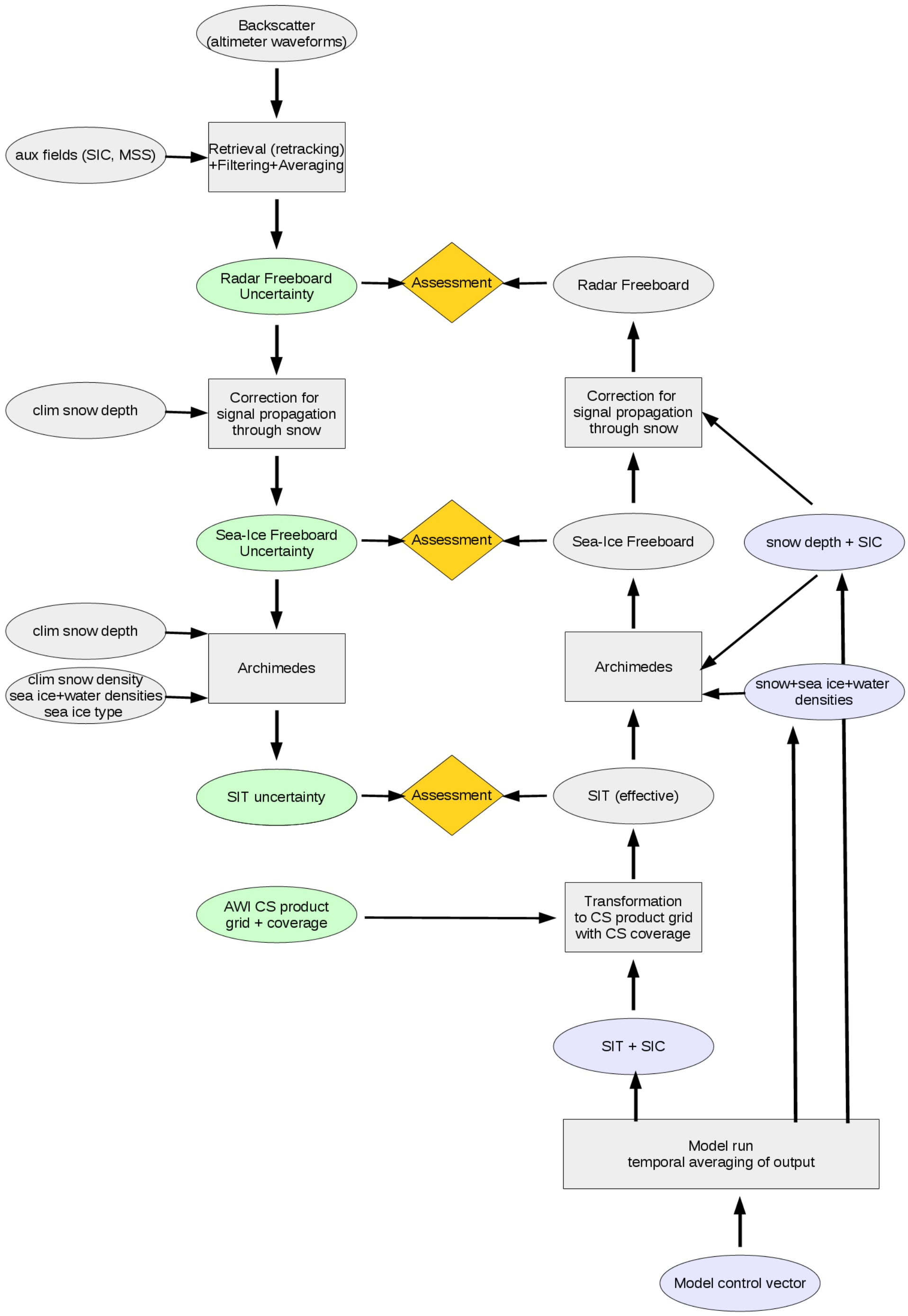

Figure 12. Overview on processing chain for CryoSat-2 product retrievals (left hand side) and chain for modelling product equivalents (right hand side). Oval boxes denote data and rectangular boxes processing steps. Green colour indicates remote sensing products and violet colour model variables.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

The retrieval chain is described in detail by Ricker et al. (2014) and Hendricks et al. (2016). Recall that for each product, in order to run an assessment we need the spatio-temporal coverage as well as the uncertainty ranges. The left-hand side of Figure 12 summarises the main steps in the retrieval chain, starting with the rawest product (RFB) on top. When descending from RFB via SIFB to SIT each step adds further assumptions, which contribute to the product uncertainty. The other element required to evaluate a given product is the observational Jacobian, i.e. the sensitivity of the model simulation to a change in the control vector. The right-hand side of the graph illustrates how this Jacobian is derived from the Jacobians of the model variables, which are denoted in violet colour. On this side of the graph, the complexity increases from bottom to top, i.e. from SIT via SIFB to RFB. For example, in the assessment of the SIT product, the uncertainty in quantities needed to apply the Archimedes' principle (including that of snow depth derived from climatology) is contained in the retrieval product, whereas the observation operator that extracts the product equivalent from the model is relatively simple. We note that, while retrieved SIT is the effective SIT ($h_{i,eff}$), i.e. refers to the average over the ice-covered area of a grid cell, simulated SIT refers to the grid-cell average, i.e. for the Jacobian calculation it has to be divided by the simulated sea ice concentration (SIC denoted by c):

$$h_{i,eff} = h_i/c \tag{6}$$

Likewise for snow depth:

$$h_{s,eff} = h_s/c \tag{7}$$

At the level of RFB, by contrast, it is the observation operator that includes, on the modelling side, the application of Archimedes' principle, for which it requires simulated snow depth and the densities of snow (ρ_s), sea ice (ρ_i), and water (ρ_w), while the retrieval product is relatively raw. In particular the retrieval product is not affected by uncertainties caused by assumptions on snow depth, ρ_s, ρ_i, and ρ_w.

The observation operators for f_i, for f_r, and for f_l are:

$$\begin{aligned} f_i &= h_i/c - (\rho_i h_i/c + \rho_s h_s/c)/\rho_w \\ &= (1-\rho_i/\rho_w)h_i/c - (\rho_s/\rho_w)h_s/c \end{aligned} \tag{8}$$

$$\begin{aligned} f_r &= f_i - 0.22 h_s/c \\ &= (1-\rho_i/\rho_w)h_i/c - (0.22 + \rho_s/\rho_w)h_s/c \end{aligned} \tag{9}$$

$$\begin{aligned} f_l &= f_i + h_s/c \\ &= (1-\rho_i/\rho_w)h_i/c + (1-\rho_s/\rho_w)h_s/c \end{aligned} \tag{10}$$

The term $-0.22h_s/c$ in Equation (9) adds to the simulated f_i the correction for the signal propagation through snow, which is contained in f_r We note that, in these three observation operators, f_i, f_r, and f_l have the same sensitivity to h_i, but sensitivities to h_s and c differ. The sea ice component of the MPIOM uses constant densities of snow, sea ice, and water. As simulated freeboard is relatively sensitive to densities of snow and sea ice, we have, however, included these quantities as parameters of the observation operator in the control vector (see Section 2.4). For ρ_I=910.0 kg/m^3, ρ_S=330.0 kg/m^3, ρ_w=1025.0 kg/m^3, the sensitivity of f_i, f_r, and f_l to a change in h_i/c is $a = 0.112$, and the respective sensitivities to a change in h_s/c are $b = -0.322$, $b = -0.542$, and $b = 0.678$.

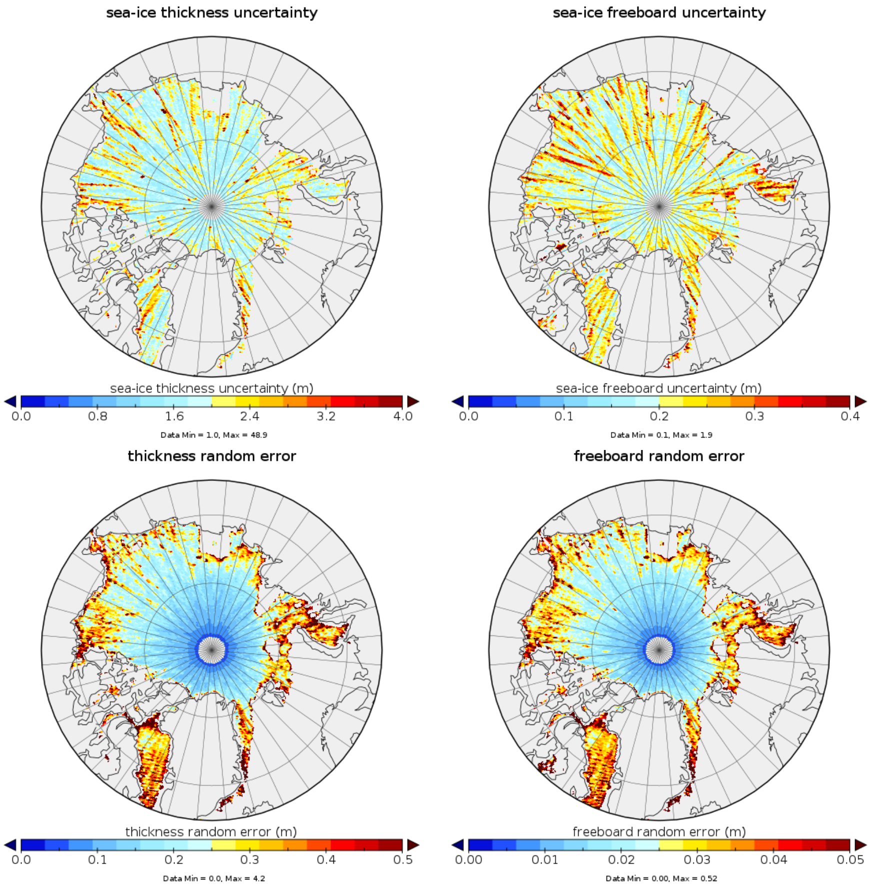

Figure 13. Uncertainty ranges of CryoSat-2 products: SIT (left), SIFB (right), total uncertainty (top), random component (bottom).

The CryoSat-2 product files provided by AWI directly contain monthly SIT and SIFB on the EASE 2.0 grid, respectively with random and total (random plus systematic) per-pixel uncertainty ranges. Figure 13 shows product uncertainties for April 2015. In our assessments we use the total uncertainties for the SIT and SIFB products, and for the RFB product the random uncertainty component of the SIFB product. We assume uncertainties are uncorrelated in space.

For our hypothetical monthly LFB products, we assume a coverage of the northern hemisphere with a retrieved value over each cell of the EASE 2.0 grid with SIC above 0.7. We explore two assumptions on the uncertainty of the products, a mission with a high accuracy (uniform uncertainty of 0.02 m) and a mission with low accuracy (uniform uncertainty of 0.20 m). In both cases uncertainties are uncorrelated in space.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

For our hypothetical monthly mean snow depth (SND) products, we also assume a coverage of the northern hemisphere with a retrieved value over each cell of the EASE 2.0 grid with SIC above 0.7. As for LFB we explore two assumptions on the uncertainty of the products, a mission with a high accuracy (uniform uncertainty of 0.02 m) and a mission with low accuracy (uniform uncertainty of 0.15 m). In both cases uncertainties are uncorrelated in space.

Table 2 provides an overview on the products we assess. For later use is also lists, for each product and three selected control regions the number of sampled EASE 2.0 grid cells and the corresponding regional average uncertainties. Finally it also shows the uncertainties on the spatial average of the sampled variable over all sampled EASE 2.0 grid cells based on the assumption of uncorrelated observational uncertainty.

Table 2. Overview on data sets, the # of sampled EASE 2.0 grids in control regions 5-7 (columns 2-4), the respective average uncertainties (columns 5-7), the uncertainty of the product aggregated over all sampled EASE 2.0 grid cells.

	n			average uncertainty			aggregated uncertainty
Product	5	6	7	5	6	7	[m]
SIT	937	1425	1377	1.86	1.95	1.94	0.0181
SIFB	937	1425	1377	0.21	0.20	0.21	0.00188
RFB	937	1425	1377	0.029	0.024	0.027	0.000364
LFB low accuracy	1104	1500	1429	0.20	0.20	0.20	0.00145
LFB high accuracy	1104	1500	1429	0.02	0.02	0.02	0.000145
SND low accuracy	1104	1500	1429	0.15	0.15	0.15	0.00108
SND high accuracy	1104	1500	1429	0.02	0.02	0.02	0.000145

3 Target and Observation Jacobians

We compute an observational Jacobian $\mathbf{M}'$ for each of the observational products we assess. For a given product, the observational Jacobian is computed in two steps. The first step performs the following actions: a reference run is performed using the prior control vector x, the input variables to the observation operator are stored over the observational period, aggregated to the model grid and the observation operator is applied to derive the observation equivalent $M(x)$ on the space-time grid of the observational product. In the second step, for each component of the control vector the following procedure is applied: A sensitivity run is performed with a control vector $x+p_i$ that is identical to the prior control vector but with the i component changed by a perturbation ϵ, and an observation equivalent $M(x+p_i)$ is computed in the same way as for the reference run. The Jacobian column is then computed as $\sigma_i(M(x+p_i)-M(x))/\epsilon$ where σ_i is the prior uncertainty of x_i. As a consequence of the normalisation by the prior uncertainty, each row in the Jacobian has the same unit as the respective observation. For a given product, column i of the corresponding observational Jacobian quantifies the sensitivity of the model-simulated equivalent to that product with respect to a 1-sigma change of the i component of the control vector x_i.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

For any given product the dimension of the observational Jacobian is the product of the dimension of the control space and the grid size of the observational product. As example, Figure 14 displays the row of the Jacobians for a April means of SIT, SIFB, RFB, LFB, and snow depth (SND) over a single point indicated by the black dot (and by the yellow cross on Figure 3)).

SIT sensitivity is dominated by the model's initial SIT in control region 6 but shows also considerable sensitivities to the initial SIC, the initial SND, the initial ocean temperature (TEMP) and the zonal wind stress (WIX). The negative sensitivity to SIC is caused by two mechanisms. The first mechanism is expressed by Equation (6): The observation $h_{i,eff}$ is the effective SIT (thickness averaged over the ice-covered grid cell) and is reduced if the initial SIC is increased (and vice versa) because the model conserves the total sea ice volume. The second mechanism is related to the formulation of sea ice growth in the model, which can grow more (less) sea ice if the SIC is reduced (increased). The small negative sensitivity of SIT to SND is caused by the strong insolation effect of snow, which hampers the growth of sea ice (or fosters the growth if SND is reduced).

The physical process behind the small negative sensitivities on the initial ocean temperature needs no further explanation; we recall, however, that, in the presence of sea ice, the control variable relates to a temperature change below the second model layer (in 17m depth). The negative sensitivity on the zonal wind stress (WIX) mirrors less advection of thick sea ice stemming from the Beaufort Gyre. SIT sensitivities on model parameters are very small compared to the sensitivities on the initial state or the atmospheric boundary conditions, as the short integration time (we sample the April mean of a model simulation starting on April 1) restricts the impact of the parameters.

The various freeboard products exhibit high sensitivity to initial SIT and SND. As SIT enters all freeboard observation operators in the same way (Section 2.5), the freeboard sensitivity on April mean SIT is equal for all products, which also renders their sensitivity to initial SIT almost equal. The LFB sensitivity on the initial SND is positive (LFB is the freeboard at the top of the snow layer) while the sensitivity of the RFB and SIFB is negative because an increased SND will reduce the radar and SIFB through the increased weight on the ice floe. (see Figure 11). Due to the definition of the observation operator for RFB (Equation (9)) its sensitivity to initial SND is larger than that of the SIFB (Section 2.5). The sensitivity of the freeboard products with respect to the parameters of the sea ice and ocean model is low. The impact of the sea ice density on the respective observation operators (Equation (8) to Equation (10)) is high, though, while the sensitivity on the snow density is much lower (because the sea ice thickness is much larger than the SND at the observational point). The SND shows only considerable sensitivity to the initial SND in control region 6 and some minor positive sensitivity on the precipitation in the same region.

Likewise we computed target Jacobians $\mathbf{N}'$ for each of the six target quantities (SIV and SNV each over 3 regions) described in Section 2.2. Each target quantity is a scalar and thus the Jacobian has one entry for each component of the control vector. As an example Figure 15 displays the Jacobians for SIV and SNV over the Outer New Siberian Islands (ONSI) region. The first point to note is that sensitivities of regional SIV and SNV to the control vector differ, so an observation must constrain different components of the control vector to perform well for one or the other.

SIV over the ONSI region is highly sensitive to initial SIT over control regions 5 and 6 which at least partly overlap with the target area. As the SIT observation and due to the same mechanisms discussed above, the SIV target quantity also exhibits a negative sensitivity to the initial SIC, SND, and zonal wind stress. It is interesting to note that SIV is also sensitive to initial and boundary conditions over more remote control regions. For example, it exhibits a positive sensitivity to the initial SIT in the control regions 1 and 7 from which thick sea ice is advected into the target region during the period from April 1 to May 28. This also explains the negative sensitivity to the zonal wind stress in region 7 and the meridional wind stress in region 1: For high enough concentration the sea ice almost behaves as an incompressible fluid allowing even for a sensitivity to wind stress changes in very remote control regions, e.g. the negative sensitivity to the zonal wind stress in region 8. The positive sensitivity to the zonal wind stress in region 1 (with thick ice) may be less obvious, as it follows the deflection of ice drift by about 20° to the right. The largest SIV sensitivity to model parameters is found for the snow albedo of freezing conditions (albsn), but still that sensitivity is low compared compared to the sensitivity with respect to the initial state and atmospheric boundary conditions.

SNV shows particularly high sensitivity to the initial SND but also considerable sensitivity with respect to the precipitation and air temperature (region 6). The largest model parameter sensitivity is found for the snow albedo for melting conditions: Increasing the snow albedo will reduce the melting.

4 Sea ice and snow volume uncertainty

Based on the products shown in Table 2, we conducted assessments for the 15 cases listed in rows 4-18 of Table 3. Row three ("prior") shows a reference case without observations, i.e. it shows the uncertainty in the target quantities that result from the prior uncertainty in the control vector. Then for each of the SIT/freeboard products described in Section 2.5:

- SIT
- SIFB
- RFB
- hypothetical low accuracy LFB
- hypothetical high accuracy LFB

there are three assessments, which cover the following cases:

- product evaluated individually
- product evaluated together with a hypothetical low accuracy SND product
- product evaluated together with a hypothetical high accuracy SND product

As explained in Section 2.1 the uncertainty component from the model error $\sigma(y_{\mathrm{mod}})$ in Equation (3) covers the residual uncertainty that remains with an optimal control vector, i.e. it reflects uncertainty from uncertain aspects not included in the model error and structural uncertainty reflecting wrong or missing process formulations. $\sigma(y_{\mathrm{mod}})$ is model-dependent and is probably the most subjective component in the prior and posterior uncertainties. $\sigma(y_{\mathrm{mod}})$ acts as an offset (for the respective variance) for all cases, and reduces the contrast between the cases. As the focus in our assessments lies on the differences between the cases, we exclude it from the target uncertainties in rows 3-18 and provide estimates in separate rows. To illustrate the subjective nature of this estimate and possible ranges, we derive two crude estimates (last two rows). The first estimate (denoted by $\sigma_{\text{mod, absolute}}$ and listed in the last but one row) assumes a model that perfectly simulates the same ice-covered area of all three regions as our model and that, over this area, achieves an uncertainty of 0.2 m for SIT and of 0.1 m for SND. The second estimate (denoted by $\sigma_{\text{mod, relative}}$ and listed in the last row) assumes a model that simulates the same SIV and SNV as our model with an uncertainty of 10% for SIV and 30% for SNV. We use a higher uncertainty for SNV because it has a stronger dependence on the surface forcing (mainly precipitation), for which the temporal and small-scale spatial structures are not resolved in the control vector.

Table 3. Prior and posterior uncertainties of sea ice volume (SIV, columns 4-6) and snow volume (SNV, columns 7-9) respectively for three regions in km^3. Column 1 indicates observation, column 2 indicates uncertainty range ("product" refers to uncertainty specification provided with product), column 3 indicates uncertainty range of additional hypothetical snow product ("–" means no snow product is used). In each of columns 4-9 the lowest uncertainty range is highlighted in bold face font. The two bottom rows give estimates for the uncertainty due to model error, i.e. the residual uncertainty with optimal control vector.

				SIV			SNV	
Observation	σ[m]	$\sigma(h_s)$[m]	WLS	ONSI	ESS	WLS	ONSI	ESS
Prior	-	-	136.5	131.6	289.6	62.3	63.3	110.1
SIT	product	-	28.7	34.3	94.4	59.5	61.3	107.9
SIT	product	0.15	19.8	22.4	62.6	11.0	11.8	21.4
SIT	product	0.02	12.4	10.4	24.1	2.4	2.5	4.5
Sea Ice Freeboard	product	-	86.4	84.1	203.4	40.4	39.8	75.2
Sea Ice Freeboard	product	0.15	21.5	25.0	67.7	11.0	11.8	21.4
Sea Ice Freeboard	product	0.02	12.6	11.0	25.3	2.4	2.5	4.5
Radar Freeboard	product	-	51.3	39.2	93.8	16.4	14.2	26.0
Radar Freeboard	product	0.15	8.8	10.9	34.7	8.0	8.3	16.6
Radar Freeboard	product	0.02	3.0	3.8	12.4	2.2	2.3	4.4
Laser Freeboard	0.20	-	81.0	67.0	143.9	17.7	17.1	30.8
Laser Freeboard	0.20	0.15	20.4	22.1	57.8	9.0	9.6	17.7
Laser Freeboard	0.20	0.02	12.2	10.7	24.8	2.3	2.4	4.5
Laser Freeboard	0.02	-	11.5	9.0	20.0	2.5	2.3	4.2
Laser Freeboard	0.02	0.15	6.6	6.0	14.6	1.9	2.0	3.7
Laser Freeboard	0.02	0.02	**2.4**	**2.7**	**8.3**	**1.3**	**1.4**	**2.6**
$\sigma_{\text{mod, absolute}}$	-	-	30.3	36.2	73.5	15.1	18.1	36.8
$\sigma_{\text{mod, relative}}$	-	-	48.7	70.8	165.9	10.2	11.4	5.3

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

Figure 16 shows the uncertainty reduction with respect to the prior case as defined in Equation (5) for both SIV and SNV and all three target regions. A value of 100% means that the product has resolved all uncertainty in the respective target quantity, while a value of 0% means that the product was not useful to improve the forecast of the target quantity. We first discuss the single product assessments, i.e. without additional use of a hypothetical snow product. For all three regions, the SIT has considerably better performance for SIV than for SNV. Between SIV and SNV the only difference consists in the target Jacobians, $\mathbf{N}'$. For example for target region ONSI, Figure 15 shows particularly high sensitivity of SIV to initial SIT and of SNV to initial SND in control regions 6 and 5. Hence, to constrain SIV (SNV) over that target region a product has to constrain primarily initial SIT (SND) over these two control regions. Figure 14 shows that, indeed, SIT provides a much stronger constraint on initial SIT than on initial SND. By contrast to SIT, SIFB has similar performance for SIV and SNV, over all target regions (Figure 16). Compared to SIT, SIFB shows a much lower sensitivity to initial SIT but a higher sensitivity to initial SND (Figure 14 - the sign of the sensitivity is irrelevant in this consideration), and thus a more balanced performance for SIV and SNV than the SIT product. RFB and the two hypothetical LFB products achieve a better performance for SNV than for SIV. The only difference between the RFB and SIFB Jacobians is the larger impact of h_s/c for RFB, as a consequence of the correction for the signal propagation through snow (see Section 2.5). This is the reason, why RFB shows a better performance for SNV than for SIV, while SIFB had about equal performance for SIV and SNV. LFB has the same sensitivity to initial SIT as RFB but an even larger sensitivity to initial SND. Consequently, for the low accuracy LFB product, the imbalance between the performance for SIV and SNV is even higher than for the RFB product. The high accuracy LFB product has so good performance on SIV already, that there is not much scope for yet better performance on SNV.

So far we have discussed for a given product the differences in performance for SIV and SNV. Next we address performance differences between products. The first thing to note is that the step from SIT to SIFB drastically reduces the performance for SIV. As explained in Section 2.5, on the retrieval side of Figure 12 the step from SIFB to SIT applies Archimedes' principle, with uncertain assumptions primarily on the input variables snow and ice density and snow depth, which yield an increase in product uncertainty by about an order of magnitude (Figure 13 and Table 2). On the modelling side of Figure 12 the step from SIT to SIFB is dealing with uncertainty on the same input variables (snow and ice densities and snow depth), which renders the simulation of SIFB more uncertain than that of SIT. In the model, the uncertainty in these variables is determined by the prior uncertainty of the control vector, either directly (snow and ice densities) or indirectly (snow depth) through their model-simulated dependency on the control vector. It appears that the increase in uncertainty, when going from SIT to SIFB on the modelling side, overcompensates for the reduction in uncertainty on the retrieval side, when going back from SIT to SIFB. In other words, on the modelling side, the assumptions on uncertain input appear more conservative than those on the retrieval side. On the retrieval side going (backwards) from SIFB to RFB consists in a reduction of product uncertainty by about another order of magnitude, as the retrieval of RFB does not require information on snow depth. Even with this further reduction of product uncertainty, the performance of RFB is inferior to that of SIT for SIV over WLS and ONSI, and only just superior for SIV over ESS.

Differences between target regions in the performance of the same product are the result of a complex interplay of the Jacobians $\mathbf{N}'$ for the target regions and product's constraint on the control vector quantified by $\mathbf{C}(x)$ (see Equation (3)). For each of the target regions a different (combination) of control regions is most relevant: For WLS this is control region 5 (not shown), for ONSI control region 5 and 6 (Figure 15) and for ESS on control region 6 and 7 (not shown). The ability of a product to constrain a particular control region is determined by the combination of the observational Jacobian of the product and the product uncertainty (see Equation (1)).

One is always tempted to explain regional performance differences in a simple way, just from differences in observational coverage and uncertainty. Technically this means to replace our observational Jacobian $\mathbf{M}'$ that is based on model dynamics by a drastically simplified representation based on the assumptions that only observations over a given control region do constrain that region (and no other region), and that the observational Jacobian for each product and control variable is spatially uniform. The constraint of a product on a control region would then be proportional to the square root of the number of samples n of that region and to the reciprocal of the average observational uncertainty $\overline{\sigma}$ over the region. Table 2 shows both impact factors for the most relevant control regions, i.e. 5-7. For RFB and compared to region 6, the relevant quantity $\sqrt{n}/\overline{\sigma}$ is about 41% lower in region 5 and 12% lower in region 7. This is at least quantitatively in line with the performance decrease for RFB and SIV from ONSI (most relevant regions: 6 and to smaller extent 5) to ESS (most relevant regions: 6 and to smaller extent 7) to WLS (most relevant region: 5). But the performance ranking for RFB and SNV is different, i.e. the simple explanation already fails. Also for SIT, the differences in $\sqrt{n}/\overline{\sigma}$ between the three control regions are smaller and fail to explain the performance decrease from WLS to ONSI to ESS. Such calculations demonstrate the limits of a performance assessment that is only based on observational coverage and uncertainty, while neglecting the model dynamics.

The two hypothetical LFB products have a slightly better spatial coverage of the most relevant control regions than the products derived from CryoSat-2 and use uniform data uncertainties that span the range from 2 cm (high accuracy LFB) to 20 cm (low accuracy LFB). We need to recall here that the specified data uncertainty combines (Equation (2)) the observational uncertainty (i.e. product uncertainty) with the residual model uncertainty due to structural errors and uncertain contributions not accounted for in the control vector. Only the high accuracy LFB can clearly outperform all CryoSat-2 products for both SIV and SNV and over all three regions, while the low accuracy LFB is between that of SIFB and RFB.

Next we discuss the effect of combining either of these five products with the two hypothetical SND products. The difference in the (sample row of the) respective product Jacobians shown in Figure 14 suggests complementarity of SND to the SIT and FB products. Indeed, the combination with SND considerably increases the performance of all SIT/freeboard products for SIV and SNV and over all regions. Most striking is the lift of the SIT performance for SNV. The combination with SND results in similar performance for SIT and SIFB, slightly better performance of low accuracy LFB, yet slightly better performance for RFB and the best performance for the high accuracy LFB. The assessment for SIV and in combination with low accuracy SND yields the same performance ranking of products, with slightly larger differences between products. Combining with the high accuracy SND product instead of the low accuracy SND product yields a performance gain for all products and for SIV and SNV over all regions.

Between the two LFB products, the increase in accuracy yields a considerable performance gain for SIV and SND over all regions, when assessed individually and in combination with SND. The combination of the high accuracy LFB with the low accuracy SND performs better for SIV and SNV over all regions.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

The Cryosphere Discussions
Open Access
EGU

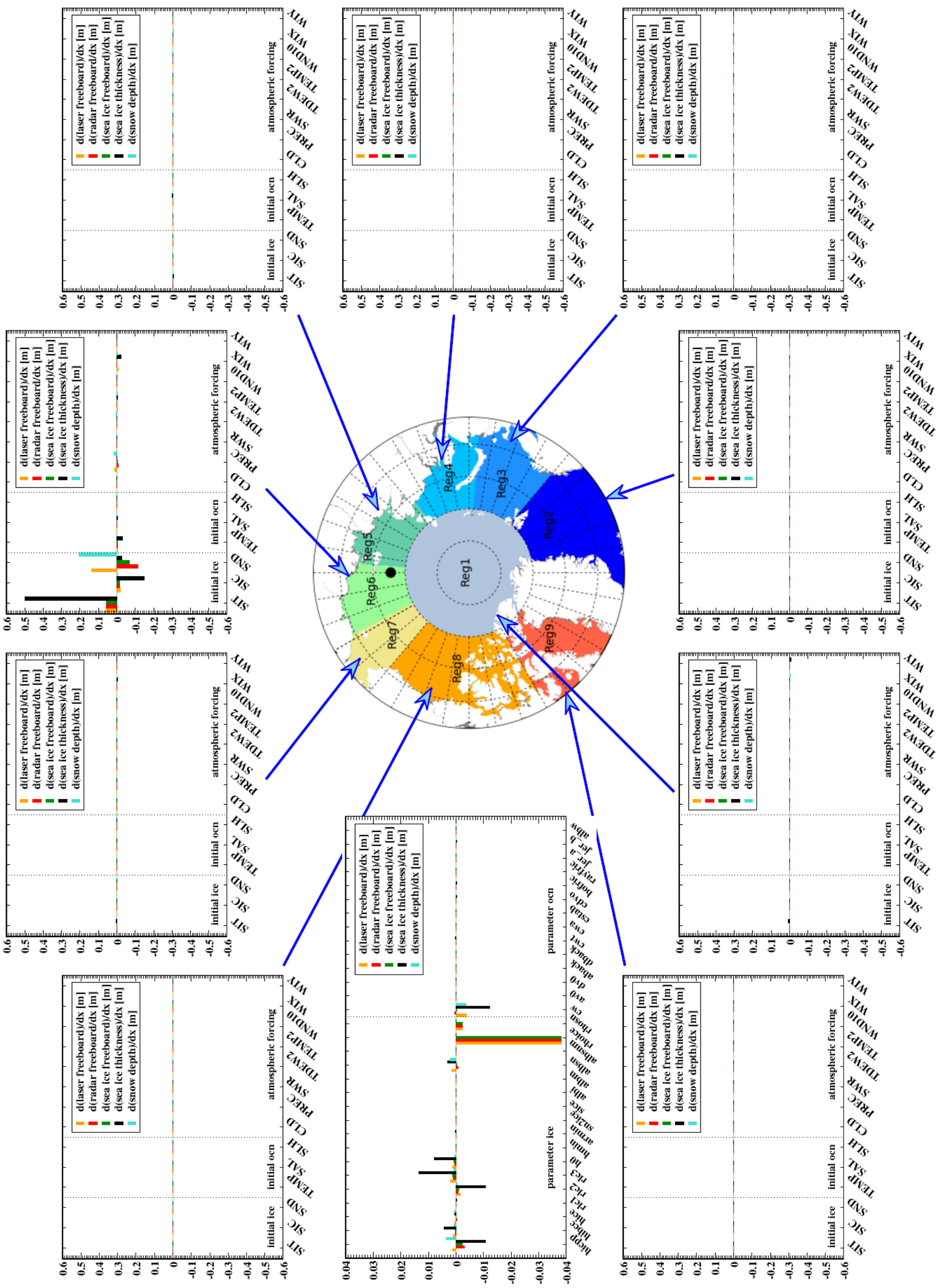

Figure 14. Jacobian rows for a April means of SIT, SIFB, RFB, and LFB over a single point indicated by the black dot (and by yellow cross on Figure 3).

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

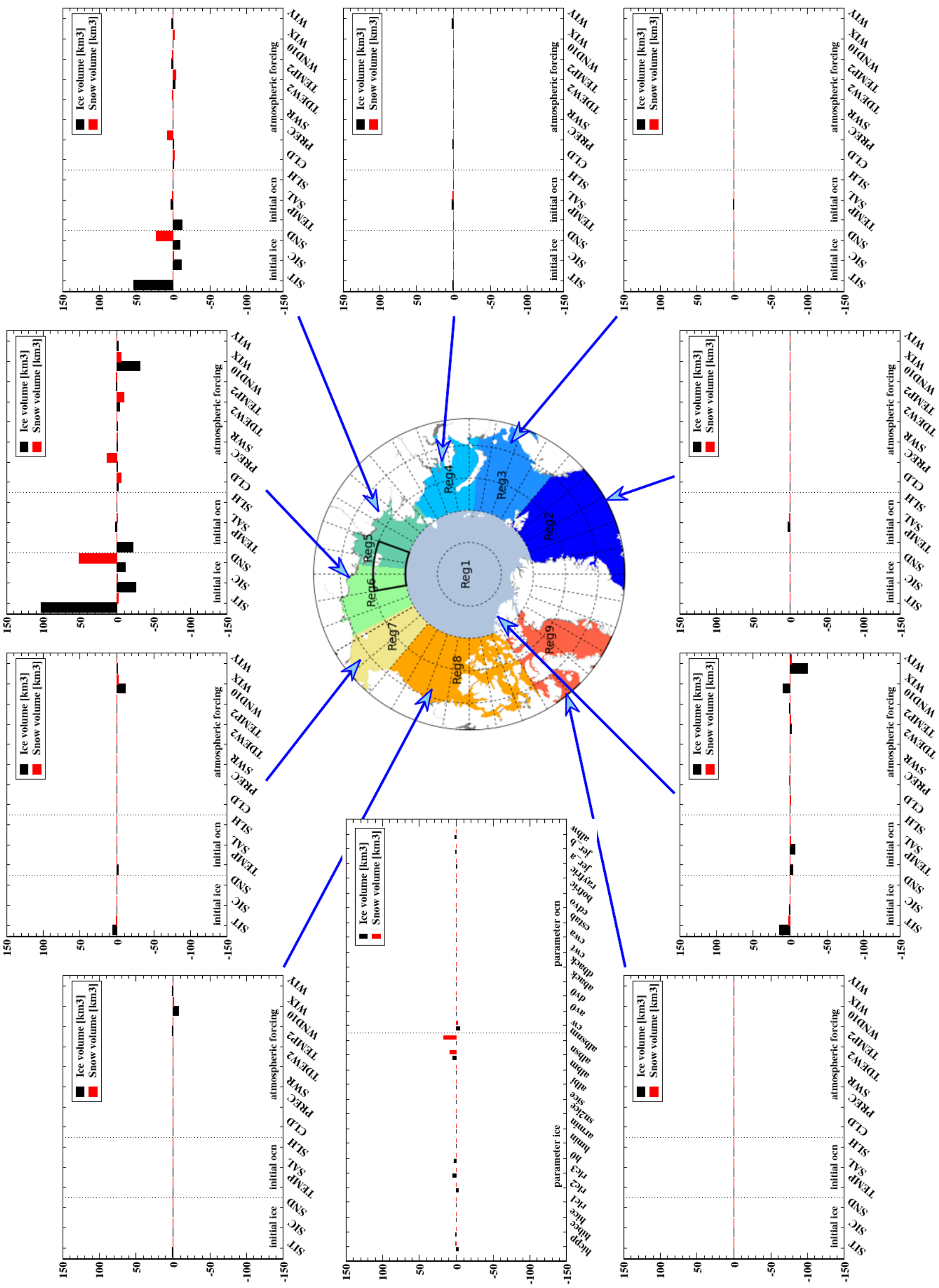

Figure 15. Jacobians for sea ice (SIV) and snow (SNV) volume over target region Outer New Siberian Islands (ONSI).

The Cryosphere
Discussions
Open Access

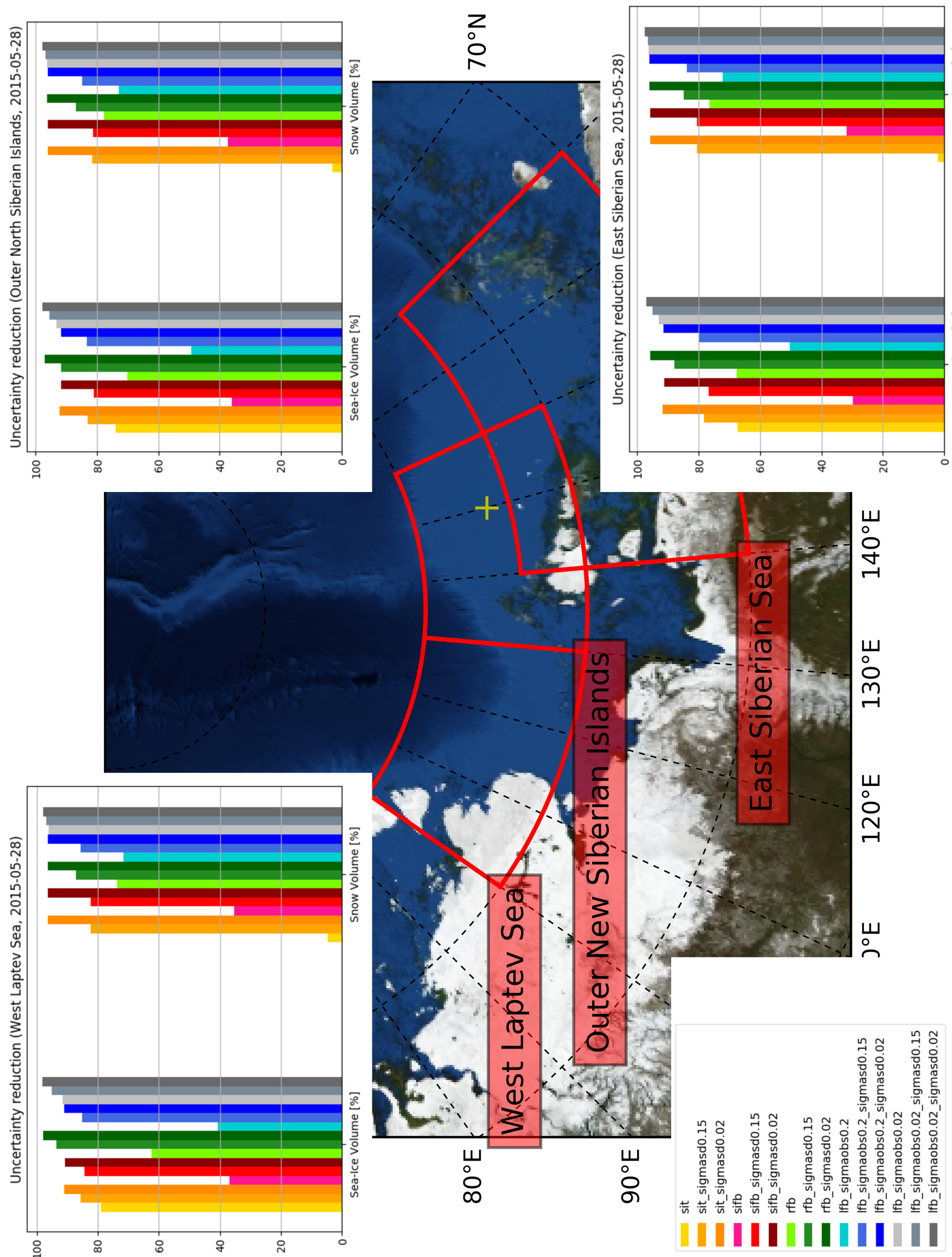

Figure 16. Uncertainty reduction for sea ice (SIV) and snow (SNV) volume over target regions.

5 Discussion

There are a number of factors in the setup of our ArcMBA system that impact our assessments. One of them is the model that is required to realistically compute the sensitivities (Jacobians) of the target quantities and of the observation equivalents to changes in the control vector. As detailed in Section 2.3, the MPIOM has a state of the art representation of processes, compares reasonably well with a range of observations, and the setup we are using has a spatial resolution below the grid size of the observations and well below the size of the target regions. The model thus appears appropriate for our study and the ArcMBA system in general. Nevertheless, through the Jacobians the results depend on the model, and it would be useful to confirm the robustness of the assessments through the use of a second model.

The study has investigated the performance of four-week forecasts in May 2015. It would be interesting to analyse if the relative performance of the products varies from year to year and with the length of the forecasting period, and how the products perform for different target regions and variables, e.g. SIC.

In our setup, the control vector has 157 components. In particular within any of our 9 control regions we do not resolve changes in the spatial patterns of the initial conditions nor in the spatio-temporal patterns of the forcing data. This means that we are ignoring uncertain aspects in the inputs that determine our simulation, which results in so-called aggregation errors (Trampert and Snieder, 1996; Kaminski et al., 2001) and renders the ArcMBA assessments of the product impacts too optimistic. As the target quantities are integrals over large regions, we expect, however, that our control regions can capture most of the uncertainty. Also the set of reasonable surface forcings is in practise limited by physical relations between variables, in space, and in time. Similar restrictions apply to the initial state. Further, we use the same control vector for all cases, so that the relative performance with respect to the prior (uncertainty reduction) and among products is more reliable. Nevertheless it appears useful to explore extended control vectors, for example with decreased sizes of the control regions, in particular in areas with high impact on observations or target quantities.

Another factor that impacts our assessments are correlations in the data uncertainty. These uncertainty correlations are difficult to estimate. We used zero correlation for each of the products, which is certainly the most optimistic assumption and yields the best performance. As we made this assumption consistently for all products, the relative performance between the products is less affected. To illustrate the implications of uncorrelated uncertainty in the products, we have computed the resulting uncertainty in the respective average of each observable over all sampled grid cells (last column of Table 2). This yields for the April 2015 mean of SIT about $2cm$, of SIFB about $2mm$, of RFB about $0.4mm$ and (using the respective high-accuracy versions) of LFB and SND about $0.1mm$.

The effect of uncertainty correlation on the assessments can be demonstrated also by the following simplified calculation: If we partition our product grid into groups of n by n pixels, and assume perfect uncertainty correlation and the same Jacobian for each observation within a given group, then we decrease the first term in Equation (1) by a factor of n^2. This case then yields the same results as a case with an uncertainty that is uncorrelated and increased by a factor of n. This means we can interpret the impact of the low resolution LFB product (uncorrelated uncertainty of 20 cm) as the impact of a high resolution LFB product (with 10 times lower uncertainty, i.e. 2 cm) in which the uncertainty within each 10 by 10 group of pixels is completely correlated. Likewise for the SND product and (roughly) 6 by 9 groups of pixels, because $(15\ \mathrm{cm}/2\ \mathrm{cm})^2$ is about 6×9. One reason for spatial uncertainty correlation would be a sensor footprint that exceeds the size of a 25km EASE grid cell. Likewise, for sensors with footprints considerably smaller than a 25km EASE grid cell, the procedure for upscaling from the sampled fraction of a grid cell to a grid-cell average could suffer from systematic errors that affect large scales in the same way, which would result in large-scale uncertainty correlations.

Our hypothetical products (LFB and SND) observe every pixel with SIC above 0.7. This is optimistic but, at least for snow, not totally unrealistic, depending on the mission concept. Recalling that the data uncertainty has to include also an uncertainty from model error, the value of 0.02 m for the high accuracy products (without spatial correlation) is extreme and unrealistic (as it is already challenging requirement on the observational uncertainty) but still useful as a sanity check.

The uncertainty specified with the SIT product is higher than the uncertainty specified with the SIFB products derived from CryoSat-2. This increase reflects the inclusion of uncertainty input quantities for the application of Archimedes' principle, in particular of climatological snow depth. In the assessment of SIFB Archimides' principle is applied in the observation operator, where the input quantities including snow depth are taken from the model. The fact that the impact of SIFB on SIV is lower than than that of SIT indicates that the assumptions on the uncertainty of input quantities for the application of Archimedes' principle is more conservative on the modelling side than those that were made on the retrieval side (yielding the respective product uncertainties). More conservative assumptions on the retrieval side would yield higher uncertainty in the SIT product. We also note that the use of the RFB product that does not rely on an external snow depth climatology but uses a consistently simulated snow depth may reduce biases.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

6 Summary and conclusions

The ArcMBA tool was used to assess the impact of a series of EO products on the quality of four-week forecasts of SIV and SNV over three regions along the Northern Sea Route in May 2015. The tool is built around the MPIOM, a coupled model of the sea ice-ocean system extended by observation operators that link the simulated variables to equivalents of SIT, SIFB, RFB, LFB, and SND products.

On the basis of the per-pixel uncertainty ranges that are provided with the CryoSat-2 SIT, SIFB, and RFB products, the SIT achieves a much better performance for SIV than the SIFB product, while the performance of RFB is more similar to that of SIT. For SNV, the performance of SIT is only low, the performance of SIFB higher and the performance of RFB yet higher. A hypothetical LFB product with low accuracy (20 cm uncertainty) lies in performance between SIFB and RFB for both SIV and SNV. A reduction in the uncertainty of the LFB product to 2 cm yields a significant increase in performance.

Combining either of the SIT/freeboard products with a hypothetical SND product achieves a significant performance increase. The uncertainty in the SND product matters: A higher accuracy product achieves an extra performance gain.

The provision of spatial and temporal uncertainty correlations with the EO products would be beneficial not only for assessments within systems like the ArcMBA, but also for assimilation of the products. For example, complete uncertainty correlation within each group of 10 by 10 pixels (with uniform Jacobians within each group) is equivalent to an uncertainty increase by a factor of 10.

The ArcMBA can be extended to cover further EO products and further target variables. In the setup used here the model can simulate a range of sea ice-ocean variables in addition to those considered in the present study (e.g. ice drift, mixed layer depth, freshwater, circulation). Switching to a more comprehensive model configuration would enable the investigation of yet further variables. For example the model can be operated with its biogeochemistry module HAMOCC (Ilyina et al., 2013) or in a mode coupled to an atmospheric general circulation model.

The study has investigated the performance of four-week forecasts in May 2015. It would be interesting to analyse how the relative performance of the products varies from year to year, with the length of the forecasting period, for other target regions, and with a different sea ice-ocean model.

The ArcMBA system is an ideal framework to assist the formulation of mission requirements or the development of EO products. Through an end-to-end simulation it can translate product specifications in terms of spatio-temporal resolution and coverage, accuracy, and precision into a range of performance metrics.

Acknowledgements. This work was funded by the European Space Agency.

References

Bovensmann, H., Bösch, H., Brunner, D., Ciais, P., Crisp, D., Dolman, H., Hayman, G., Houweling, S., and Lichtenberg, L.: Report for mission selection: CarbonSat - An earth explorer to observe greenhouse gases, Tech. rep., Noordwijk, The Netherlands, http://nora.nerc.ac.uk/514012/, freely available via Official URL link., 2015.

Chevallier, M., Smith, G. C., Dupont, F., Lemieux, J.-F., Forget, G., Fujii, Y., Hernandez, F., Msadek, R., Peterson, K. A., Storto, A., Toyoda, T., Valdivieso, M., Vernieres, G., Zuo, H., Balmaseda, M., Chang, Y.-S., Ferry, N., Garric, G., Haines, K., Keeley, S., Kovach, R. M., Kuragano, T., Masina, S., Tang, Y., Tsujino, H., and Wang, X.: Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project, Climate Dynamics, 49, 1107–1136, https://doi.org/10.1007/s00382-016-2985-y, https://doi.org/10.1007/s00382-016-2985-y, 2017.

Crisp, D., Atlas, R., Breon, F.-M., Brown, L., Burrows, J., Ciais, P., Connor, B., Doney, S., Fung, I., Jacob, D., Miller, C., O'Brien, D., Pawson, S., Randerson, J., Rayner, P., Salawitch, R., Sander, S., Sen, B., Stephens, G., Tans, P., Toon, G., Wennberg, P., Wofsy, S., Yung, Y., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission. Trace Constituents in the Troposphere and Lower Stratosphere, Advances in Space Research, 34, 700 – 709, https://doi.org/http://dx.doi.org/10.1016/j.asr.2003.08.062, http://www.sciencedirect.com/science/article/pii/S0273117704003539, 2004.

Day, J. J., Hawkins, E., and Tietsche, S.: Will Arctic sea ice thickness initialization improve seasonal forecast skill?, Geophysical Research Letters, 41, 7566–7575, https://doi.org/10.1002/2014GL061694, http://dx.doi.org/10.1002/2014GL061694, 2014.

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.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, https://doi.org/10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Eastwood, S., Jenssen, M., Lavergne, T., Sorensen, A., and Tonboe: EUMETSAT Ocean and Sea Ice Satellite Application Facility. Global sea ice concentration reprocessing product (v1.2), Product user manual, Technical Report, Norwegian and Danish Meteorological Institutes, Oslo, Norway and Copenhagen, Denmark, 14, 2079–2087, 2015.

Feng, L., Palmer, P. I., Bösch, H., and Dance, S.: Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman Filter, Atmospheric Chemistry and Physics, 9, 2619–2633, https://doi.org/10.5194/acp-9-2619-2009, http://www.atmos-chem-phys.net/9/2619/2009/, 2009.

Gent, P., Willebrand, J., McDougall, T., and McWilliams, J.: Parameterizing eddy-induced tracer transport in ocean circulation models, J. Phys. Oceanogr., 25, 463–474, 1995.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, Journal of Advances in Modeling Earth Systems, 5, 572–597, https://doi.org/10.1002/jame.20038, http://dx.doi.org/10.1002/jame.20038, 2013.

Griffies, S. M.: The Gent-McWilliams skew flux, J. Phys. Oceanogr., 28, 831–841, 1998.

Haak, H., Jungclaus, J., Mikolajewicz, U., and Latif, M.: Formation and propagation of great salinity anomalies, Geophysical Research Letters, 30, n/a–n/a, https://doi.org/10.1029/2003GL017065, http://dx.doi.org/10.1029/2003GL017065, 1473, 2003.

Hardt, M. and Scherbaum, F.: The Design of Optimum Networks for Aftershock Recordings, Geophys. J. Int., 117, 716–726, 1994.

Hendricks, S., Ricker, R., and Helm, V.: User Guide - AWI CryoSat-2 Sea Ice Thickness Data Product (v1.2), 2016.

Hibler, W.: A dynamic thermodynamic sea ice model, Journal Geophysical Research, 9, 815–846, 1979.

Houweling, S., Breon, F.-M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., and Ciais, P.: Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, Atmos. Chem. Phys., 4, 523–538, 2004.

Hungershoefer, K., Breon, F.-M., Peylin, P., Chevallier, F., Rayner, P., Klonecki, A., Houweling, S., and Marshall, J.: Evaluation of various observing systems for the global monitoring of CO_2 surface fluxes, Atmospheric Chemistry and Physics, 10, 10 503–10 520, https://doi.org/10.5194/acp-10-10503-2010, http://www.atmos-chem-phys.net/10/10503/2010/, 2010.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Núñez-Riboni, I.: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations, Journal of Advances in Modeling Earth Systems, 5, 287–315, https://doi.org/10.1029/2012MS000178, http://dx.doi.org/10.1029/2012MS000178, 2013.

Jungclaus, J., Giorgetta, M., Reick, C., Legutke, S., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Fischer, N., Glushak, K., Gayler, V., Haak, H., Hollweg, H., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Müller, W., Notz, D., Pohlmann, T., Raddatz, T., Rast, S., Roeckner, E., Salzmann, M., Schmidt, H., Schnur, R., Segschneider, J., Six, K., Stockhause, M., Wegner, J., Widmann, H., Wieners, K., Claussen, M., Marotzke, J., and Stevens, B.: CMIP5 simulations of the Max Planck Institute for Meteorology (MPI-M) based on the MPI-ESM-P model: The 1pctCO2 experiment, served by ESGF, WDCC at DKRZ, https://doi.org/10.1594/WDCC/CMIP5.MXEPc1, 2012.

Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J.-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM, Journal of Climate, 19, 3952–3972, https://doi.org/10.1175/JCLI3827.1, https://doi.org/10.1175/JCLI3827.1, 2006.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J.: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth system model, J. Adv. Model. Earth Syst., 5, 422–446, https://doi.org/10.1002/jame.20023, 2013.

Kadygrov, N., Maksyutov, S., Eguchi, N., Aoki, T., Nakazawa, T., Yokota, T., and Inoue, G.: Role of simulated GOSAT total column CO2 observations in surface CO2 flux uncertainty reduction, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, https://doi.org/10.1029/2008JD011597, http://dx.doi.org/10.1029/2008JD011597, 2009.

Kaminski, T. and Mathieu, P.-P.: Reviews and syntheses: Flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle, Biogeosciences, 14, 2343–2357, https://doi.org/10.5194/bg-14-2343-2017, http://www.biogeosciences.net/14/2343/2017/, 2017.

Kaminski, T. and Rayner, P. J.: Assisting the Evolution of the Observing System for the Carbon Cycle through Quantitative Network Design, Biogeosciences Discussions, 2017, 1–22, https://doi.org/10.5194/bg-2017-168, http://www.biogeosciences-discuss.net/bg-2017-168/, 2017.

Kaminski, T., Rayner, P., Heimann, M., and Enting, I.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4703, 2001.

Kaminski, T., Scholze, M., and Houweling, S.: Quantifying the Benefit of A-SCOPE Data for Reducing Uncertainties in Terrestrial Carbon Fluxes in CCDAS, Tellus B, https://doi.org/http://dx.doi.org/10.1111/j.1600-0889.2010.00483.x, 2010.

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, Atmospheric Chemistry and Physics, 12, 7867–7879, https://doi.org/10.5194/acp-12-7867-2012, https://www.atmos-chem-phys.net/12/7867/2012/, 2012.

Kaminski, T., Kauker, F., Eicken, H., and Karcher, M.: Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies, The Cryosphere, 9, 1721–1733, https://doi.org/10.5194/tc-9-1721-2015, 2015.

Kauker, F., Kaminski, T., Ricker, R., Toudal-Pedersen, L., Dybkjaer, G., Melsheimer, C., Eastwood, S., Sumata, H., Karcher, M., and Gerdes, R.: Seasonal sea ice predictions for the Arctic based on assimilation of remotely sensed observations, The Cryosphere Discussions, 9, 5521–5554, https://doi.org/10.5194/tcd-9-5521-2015, 2015.

Kwok, R. and Cunningham, G. F.: ICESat over Arctic sea ice: Estimation of snow depth and ice thickness, Journal of Geophysical Research: Oceans, 113, n/a–n/a, https://doi.org/10.1029/2008JC004753, http://dx.doi.org/10.1029/2008JC004753, c08010, 2008.

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

Lisaeter, K. A., Evensen, G., and Laxon, S.: Assimilating synthetic CryoSat sea ice thickness in a coupled ice-ocean model, Journal of Geophysical Research: Oceans, 112, n/a–n/a, https://doi.org/10.1029/2006JC003786, http://dx.doi.org/10.1029/2006JC003786, c07023, 2007.

Mikolajewicz, U., Sein, D., Jacob, D., Königk, T., Podzun, R., and Semmler, T.: Simulating Arctic sea ice variability with a coupled regional atmosphere-ocean-sea ice model, Meteorol. Zeitschrift, 14, 793–800, 2015.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, Geophysical Research Letters, 39, n/a–n/a, https://doi.org/10.1029/2012GL053326, http://dx.doi.org/10.1029/2012GL053326, l22707, 2012.

Niederdrenk, A.: The Arctic hydrologic cycle and its variability in a regional coupled climate model, PhD Thesis, Unversity Hamburg, pp. 1–186, 2013.

Niederdrenk, A. L., Sein, D. V., and Mikolajewicz, U.: Interannual variability of the Arctic freshwater cycle in the second half of the twentieth century in a regionally coupled climate model, Climate Dynamics, 47, 3883–3900, https://doi.org/10.1007/s00382-016-3047-1, https://doi.org/10.1007/s00382-016-3047-1, 2016.

NOAA: Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth, NOAA, National Geophysical Data Center, Boulder, Colorado, 1988.

Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., and Marotzke, J.: Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth system model, Journal of Advances in Modeling Earth Systems, 5, 173–194, https://doi.org/10.1002/jame.20016, http://dx.doi.org/10.1002/jame.20016, 2013.

Pacanowski, R. and Philander, S.: Parameterization of vertical mixing in numerical-models of tropical oceans, J. Phys. Oceanogr., 11, 1443–1451, 1981.

Patra, P. K., Maksyutov, S., Sasano, Y., Nakajima, H., Inoue, G., and Nakazawa, T.: An evaluation of CO2 observations with Solar Occultation FTS for Inclined-Orbit Satellite sensor for surface source inversion, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, https://doi.org/10.1029/2003JD003661, http://dx.doi.org/10.1029/2003JD003661, 2003.

Rayner, P., Michalak, A. M., and Chevallier, F.: Fundamentals of Data Assimilation, Geoscientific Model Development Discussions, 2016, 1–21, https://doi.org/10.5194/gmd-2016-148, http://www.geosci-model-dev-discuss.net/gmd-2016-148/, 2016.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Let., 28, 175–178, 2001.

Rayner, P. J., Enting, I. G., and Trudinger, C. M.: Optimizing the CO_2 Observing Network for Constraining Sources and Sinks, Tellus, 48B, 433–444, 1996.

Rayner, P. J., Utembe, S. R., and Crowell, S.: Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study, Atmospheric Measurement Techniques Discussions, 7, 1367–1392, https://doi.org/10.5194/amtd-7-1367-2014, http://www.atmos-meas-tech-discuss.net/7/1367/2014/, 2014.

Redi, M. H.: Oceanic isopycnal mixing by coordinate rotation, J. Phys. Oceanogr., 12, 1154–1158, 1982.

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

Röske, F.: An atlas of surface fluxes based on the ECMWF reanalysis—A climatological data set to force global ocean general circulation models, Tech. Rep. 323, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 2001.

Semtner, A.: A Model for the Thermodynamic Growth of Sea Ice in Numerical Investigastions of Climate, Journal pf Physical Oceanography, 6, 379–389, 1976.

Steele, M., Morley, R., and W. Ermold, W.: PHC: A global ocean hydrography with a high quality Arctic Ocean, J. Climate, 14, 2079–2087, 2001.

Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.M. (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, https://doi.org/10.1017/CBO9781107415324, 2013.

Tarantola, A.: Inverse Problem Theory and methods for model parameter estimation, SIAM, Philadelphia, 2005.

Trampert, J. and Snieder, R.: Model Estimations Biased by Truncated Expansions: Possible Artifacts in Seismic Tomography, Science, 271, 1257–1260, 1996.

UNESCO: Algorithms for computation of fundamental properties of seawater, UNESCO Technical Papers in Marine Science, 44, 1983.

Warren, S., Rigor, I., Untersteiner, N., Radionov, V., Bryaz-gin, N., Aleksandrov, Y., and Colony, R.: Snow depth on Arctic sea ice, J. Clim., 12, 1814–1829, 1999.

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

Yang, Q., Losa, S. N., Losch, M., Tian-Kunze, X., Nerger, L., Liu, J., Kaleschke, L., and Zhang, Z.: Assimilating SMOS sea ice thickness into a coupled ice-ocean model using a local SEIK filter, Journal of Geophysical Research: Oceans, 119, 6680–6692, https://doi.org/10.1002/2014JC009963, http://dx.doi.org/10.1002/2014JC009963, 2014.

The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-249
Manuscript under review for journal The Cryosphere
Discussion started: 30 November 2017

Data and Code Availability

Data Sets are available upon request to the corresponding author.