Project no. 265863

ACCESS

Arctic Climate Change, Economy and Society

Instrument: Collaborative Project
Thematic Priority: Ocean.2010-1 "Quantification of climate change impacts on economic sectors in the Arctic"

D1.83 – Report on the assessment of forecast skill

Due date of deliverable: **31/12/2014**

Actual submission date: 26/03/2015

Used Person/months: 16

Start date of project: **March 1st, 2011**   Duration: **48 months**

Organisation name of lead contractor for this deliverable: **OASys**

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

1 Introduction

The Arctic climate system is undergoing a rapid transition. Such changes, in particular reductions in sea-ice extent, are impacting coastal communities and ecosystems and are changing the conditions for resource extraction, shipping, and fisheries. In this context, high-quality predictions of the ice conditions are of paramount interest. Such predictions are typically performed by numerical models of the sea ice-ocean system. The skill of such predictions can be substantially improved through assimilation of observations. In this report, we describe the construction of an assimilation and prediction system of the Arctic sea-ice conditions.

Observational data streams for such a prediction system have to be available near real time. We used four data streams which fulfil this requirement, namely, the OSISAF sea-ice concentration and sea-surface temperature products, a snow thickness product provided by the University of Bremen, and the CryoSat-2 data product derived at AWI. The availability of the above data streams is limited to the period from 2012 to 2014. Also there is only one single two month period per year (March and April) for which the CryoSat-2 product is currently available. By contrast to our original plan we thus restricted our study to assimilation of the above four data streams in the spring of each of the three years and to prediction of the ice conditions in the following summer.

Initial tests indicated that our model was not sufficiently calibrated to achieve the required simulation quality. Hence, as an extra (unforeseen task) we calibrated the model using observations over an 19 year period. Furthermore, it turned out that the model was not capable to absorb the information in the above-mentioned ice thickness product to a sufficient degree. Through a set of additional assimilation experiments, we were capable of developing a so-called bias correction scheme that allowed to take full

advantage of this data stream. This was demonstrated through three seasonal forecasts of summer ice conditions with high skill.

2 The model and data assimilation system

2.1 NAOSIM

The model used for the present analysis is the coupled ice-ocean model NAOSIM (North Atlantic/Arctic Ocean Sea Ice Model, *Kauker et al.* (2003)). NAOSIM is based on version 2 of the Modular Ocean Model (MOM-2) of the Geophysical Fluid Dynamics Laboratory (GFDL). The version of NAOSIM used here has a horizontal grid spacing of 0.5° on a rotated spherical grid. The rotation maps the $30°W$ meridian onto the equator and the North Pole onto $0°E$. Hence, the model's x- and y-directions are different from the zonal and meridional directions. In the vertical it is resolved by 20 levels, the spacing of which increases with depth. The ocean model is coupled to a sea-ice model with viscous-plastic rheology. At the open boundary near 50° N the barotropic transport is prescribed from a coarser resolution version of the model that covers the whole Atlantic northward of 20° S (*Köberle and Gerdes*, 2003). Atmospheric forcing (10m-wind velocity, 2m-air temperature, 2m-specific humidity, total precipitation, and downward solar and thermal radiation) is taken from the National Center's for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (NCEP-CFSR) (*Saha and Co-authors*, 2010) for 1979 to 2010 and the NCEP Climate Forecast System version 2 (CFSv2) (*Saha and Co-authors*, 2014) for 2011 to the end of 2014. The initial state of January 1st 1979 is taken from a hindcast from January 1948 to end of December 1979, forced by the NCEP/NCAR reanalyses (*Kalnay et al.*, 1996), which was, in turn, initialised from PHC (*Steele et al.*, 2001) (ocean temperature and salinity) and a constant ice thickness of 2m with 100% ice cover where the air temperature is below the freezing temperature of the ocean's top layer.

2.2 Assimilation system

The variational assimilation system NAOSIMDAS (*Kauker et al.*, 2009, 2010) operates through minimisation of a cost function that quantifies the fit to all observations plus the deviation from prior knowledge on a vector of control variables x:

$$J(\tilde{\mathbf{x}}) = \frac{1}{2}\left[(M(\tilde{\mathbf{x}}) - \mathbf{d})^T \mathbf{C}(d)^{-1}(M(\tilde{\mathbf{x}}) - \mathbf{d}) + (\tilde{\mathbf{x}} - \mathbf{x_0})^T \mathbf{C}(x_0)^{-1}(\tilde{\mathbf{x}} - \mathbf{x_0})\right] \tag{1}$$

where M denotes the model, considered as a mapping from the control vector to observations, d the observations with data uncertainty covariance matrix $\mathbf{C}(d)$, x_0 the vector of prior values of the control variables with uncertainty covariance matrix $\mathbf{C}(x_0)$, and the superscript T is the transposed. The control variables are typically a combination of the initial state, the atmospheric forcing and the process parameters. In this study the control vector will only consist of initial state variables. The data uncertainty $\mathbf{C}(d)$ reflects the combined effect of observational $\mathbf{C}(d_{\text{obs}})$ and model error $\mathbf{C}(d_{\text{mod}})$:

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

$\mathbf{C}(d_{\text{mod}})$ captures all uncertainty in the simulation of the observations except for the uncertainty in the control vector, because this fraction of the uncertainty is explicitly addressed by the assimilation procedure through correction of the control vector.

3 Model calibration

For this study some relevant parameters of the sea ice-ocean model have been newly calibrated. Remotely sensed sea-ice concentration (OSISAF), sea-ice thickness (ICESat from JPL), and ice drift (low resolution from OSISAF) data have been utilised. For the calibration of the parameters one has to take the memory of the Arctic sea-ice system into account which is in the range of 7 to 10 years, i.e. the simulations used

to calibrate the model should be much longer then this time range. This does not allow to use the 4dVar data assimilation system for the calibration because the non-linearities in the system do not allow to use that large assimilation windows. Other methods like genetic algorithms (*Sumata et al.*, 2013) might be applicable but have not been used here.

Here we utilised conventional sensitivity experiments and adjusted the model by manual calibration of the parameters. We started the experiments on January 1979 and ran the model until end of December 2008. We skipped the first 10 years of the experiments and evaluated the equilibrium response for the residual roughly 20 years.

The most obvious discrepancy of the simulated ice thickness distribution in the standard simulation is too thick ice all over the year in the Beaufort Sea as is illustrated in Figure 1. Exemplarily the sensitivity of the model to changes in $c^\star$ (dependence of the ice strength on the ice concentration) is shown in Figure 2. In the standard simulation a value of 20 is used for $c^\star$ meaning that for 90% ice concentration only about 14% of the ice pressure is acting compared to a closed ice cover. If $c^\star$ is set to 30 about 5% of the ice pressure is acting and if $c^\star$ is set to 5 about 60% of the ice pressure is acting. The effect of increasing $c^\star$ is increasing the ice thickness slightly, decreasing $c^\star$ is decreasing the thickness especially in September and in the Beaufort Sea strongly - a decreased $c^\star$ makes the ice in the not totally ice covered regions more ridged resulting in reduced ice transport into the Beaufort Sea especially in Summer.

The calibration is performed mainly by calibrating parameters which influence the ice dynamics. The parameters are the atmospheric drag coefficient (*cdair*), the oceanic drag coefficient (*cdwat*), the ice strength parameter ($p^\star$), the already mentioned parameter ($c^\star$) and the parameter determining the ellipsoid form of the rheology (*eccen*) which represents the ratio of the normal stress and the shear stress. Additionally the vertical tracer mixing parameter $kappa_h$ is modified (Table 1).
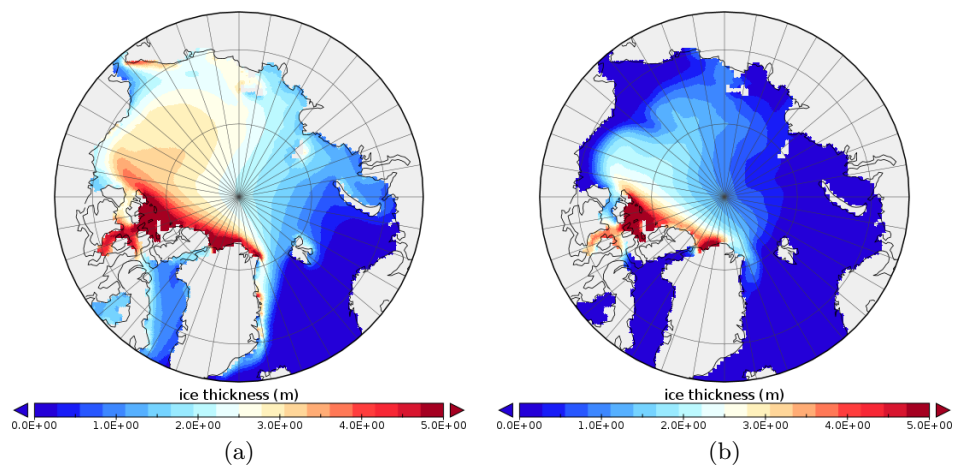

Figure 1: The ice thickness climatology [m] in a) March and in b) September for the standard simulation.
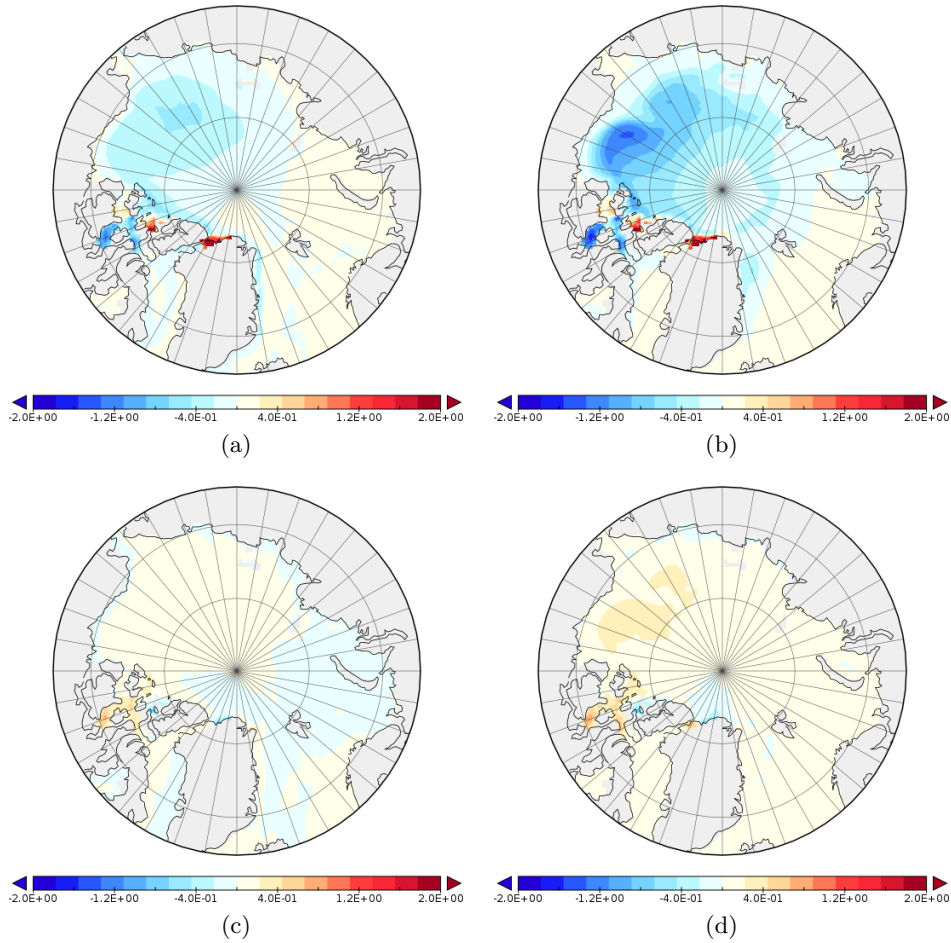
Figure 2: The difference of the ice thickness climatology [m] for March (left column; a) and c)) and September (right column; b) and d)) for changing $c^\star$ from 20 to 5 (top row; a) and b)) and from 20 to 30 (bottom row; c) and d)).

The performance of the model is evaluated in terms of its fit to observed monthly-mean sea-ice concentration, ICESat-JPL ice thickness, and ice drift from OSISAF. The only observation available for the entire validation period over the whole year is the the OSISAF ice concentration product which also comprises spatially and temporally varying uncertainty estimates (EUMETSAT Ocean and Sea Ice Satellite Application Facility. Global sea-ice concentration reprocessing data set 1978-2009 (v1.1, 2011). Norwegian and Danish Meteorological Institutes. http://osisaf.met.no.). For the ICESat-JPL ice thickness (available in Feb/March and Oct/Nov 2003 to 2008) no spatially and temporally variable uncertainties exists. A mean error of about $50cm$, corresponding to a relative error of about 40% is given by *Kwok and Cunningham* (2008). We apply a 40% relative uncertainty for the calculation of the cost function. Because it is well known that the error is larger for ice thinner than 1m, data thinner then $1m$ are excluded from the calculation. For the calculation of the OSISAF ice-drift cost function we employ uncertainties estimated by *Sumata et al.* (2014). Data with uncertainty are available from October 2003 to December 2006 for the winter half year (October to April).

According to the sensitivities the parameters are altered to decrease the cost function of the different observational streams. Because of the varying numbers of months and years with observations of ice concentration, ice thickness, and ice drift the three cost function terms are not simply summed but normalised before the summation, i.e. each data stream is given the same weight. Accordingly, the value of the total cost function is 3 for the standard configuration. Table 1 gives an excerpt of the used combinations of the parameters. Finally a setup is used in this study which reduced the cost function for ice thickness and ice drift strongly but increased the cost function for ice concentration, especially in

Aomip	Cdwat	kappa_h	Pstar	Cstar	Eccen	costfct concn.	costfct concn.	costfct ICESAT	costfct ICESAT	costfct drift	total	weighted
[2.475e-3]	[5.5e-3]	[1cm2/s]	[15kN/m2]	[20]	[2]	Nov-Apr	May-Oct	Feb/Mar	Oct/Nov	Nov-Mar	costfct	total
-	-	-	-	-	-	227645	437405	7166	10018	123605	682234	3.00
yes	-	-	-	-	-	260951	407270	15330	21497	46645	751694	3.53
yes	10	-	25k	5	1.5	258870	447429	11888	16645	27703	762535	2.95
yes	10	0.5	25k	10	1.5	346630	445950	3746	6367	25740	828433	1.99
yes	10	0.5	25k	5	1.5	318122	519937	5314	5221	29725	878319	2.11
yes	10	0.5	25k	10	1.5	331155	684573	4003	6752	30824	1057307	2.40
yes	3.8	0.5	20k	10	1.5	346630	445950	3746	6367	25740	828433	2.21
yes	3.8	0.5	25k	10	1.5	344803	465418	3549	6101	23438	843309	1.97
1.6	3.8	0.5	15k	10	1.5	357204	448409	4876	7460	46686	864636	2.31
1.6	3.8	0.5	20k	10	1.5	348046	444789	3851	6918	34455	838059	2.10
1.6	3.8	0.5	25k	10	1.5	363301	461067	3426	6705	27625	834499	2.05
Piomas2.1						-	-	2692	5115	46664	-	-

Table 1: The cost function for the three data streams separated by the season (columns 7 to 11) the sum of the terms (column 12) and the normalised sum (column 13) for different combination of parameters (column 1 to 6). For parameters the second row displays the values in the standard configuration. If in the first column 'yes' is listed, the atmospheric drag coefficient is replaced by a formulation used in the AOMIP project (http://www.whoi.edu/page.do?pid=30576), where *cdair* depends on the wind speed. The row with the green background gives the configuration used in this study. The orange coloured column lists the cost function of ice thickness and ice drift for PIOMAS 2.1.

winter where the ice margin is located too far south in the new setup. Because we are here interested in the seasonal predictions of summer ice conditions, this deficit is tolerated. Note, that the selected parameter setup is not optimal but much better than the standard setup.

PIOMAS 2.1 (*Zhang and Rothrock*, 2003), an Arctic sea ice-ocean model which uses Optimal Interpolation to assimilate ice concentration and sea-surface temperature, is often used as a reference, because it is well validated (see e.g. *Schweiger et al.* (2011)). The cost function for ice thickness and ice drift is given as well for PIOMAS2.1 in Table 1. The newly calibrated NAOSIM (newNAOSIM) performs similar as PIOMAS2.1 with respect to ice thickness and better for ice drift which can be seen as well in Figure 3 for monthly temporal resolution. The deviations of the climatologies of newNAOSIM and of PIOMAS2.1 from the IceSat-JPL climatology (Figure 5) reveals remarkable similarities between both models. Both show too thick ice in the Beaufort Sea in February/March and too thin ice north of the Canadian Archipelago and north of Greenland and north of Fram Strait in February/March and October/November when compared to ICESat-JPL. However, because both models rely on different model formulations and parameterisations and because the atmospheric forcing is different (NCEP-CFSR in case of NAOSIM and NCEP in case of PIOMAS2.2 which differ considerably for some variables, see e.g. *Lindsay et al.* (2014)) one might speculate that for instance north of Greenland or north of Fram Strait ICESat-JPL shows unreliable thick ice. We calculated the ice volume on the grid points were ICESat-JPL data exist (Figure 4). The total ice volume of stdNAOSIM and newNAOSIM are pretty close although the horizontal distribution differs considerably. A comparison of the models and OSISAFs mean winter ice drift (Figure 6) explains the drastic decrease of the ice drift cost function (Figure 3 b) - stdNAOSIMs ice drift was much to high. For newNAOSIM ice thickness climatology is depicted in Figure 7. Even more pronounced are the differences between stdNAOSIM and newNAOSIM when individual months are considered as for instance September 2007 (Figure 8). Additionally Figure 8 shows the ice thickness of PIOMAS2.1 in September 2007 and the ICESat-JPL thickness for October/November 2007. In summary, the local ice thickness and the local ice drift of newNAOSIM fits much better to the remotely-sensed observations.

The sea-ice extent and sea-ice area in September for stdNAOSIM and newNAOSIM compared with three different observed estimates (Figure 9) reveals considerable differences as well. In general newNAOSIM performs much better, but for ice extent in the most recent years stdNAOSIM was closer to the observation (note that this does not hold for ice area (Figure 9b)). However, in the early years newNAOSIM performs much better especially in the year 1990 where stdNAOSIM drastically underestimated the ice extent pointing to the importance of a reliable horizontal ice thickness distribution (and connected to that a reliable ice drift).
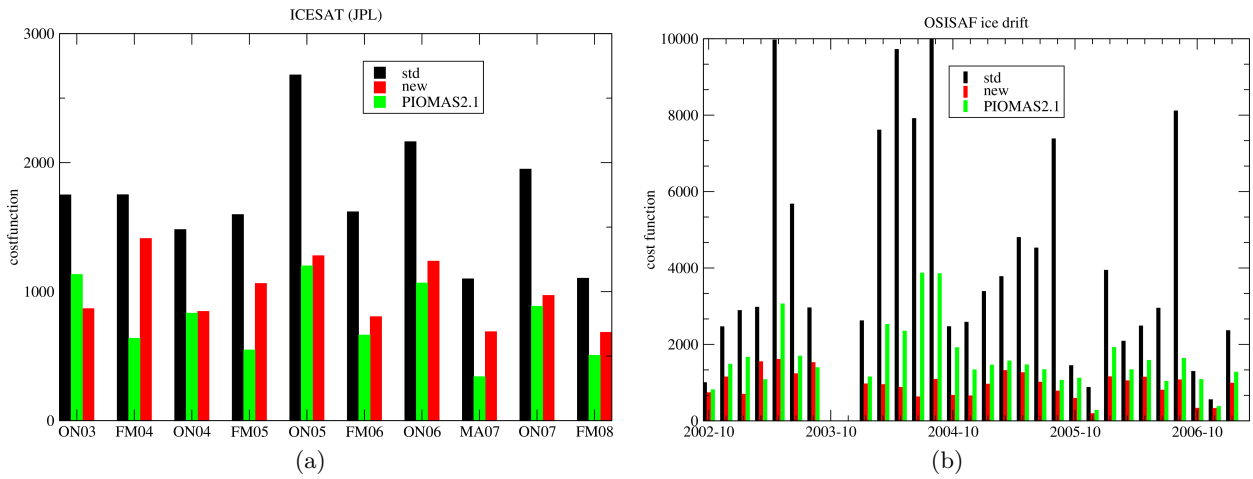
Figure 3: The monthly cost function a) for ICESat-JPL and b) OSISAF ice drift for NAOSIMS standard and newly calibrated setup and for PIOMAS2.1.
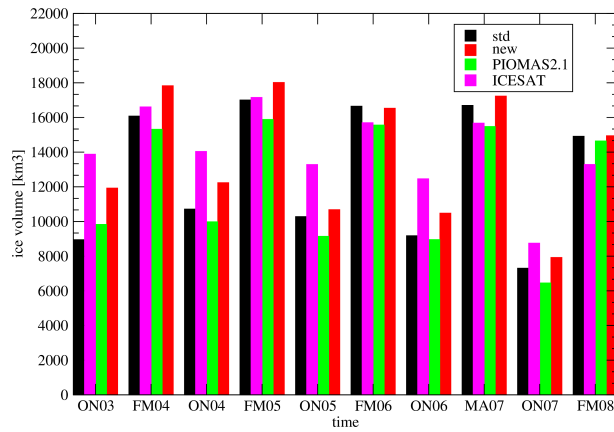

Figure 4: The ice volume on model $[km^3]$ grid points were ICESat-JPL data exists.
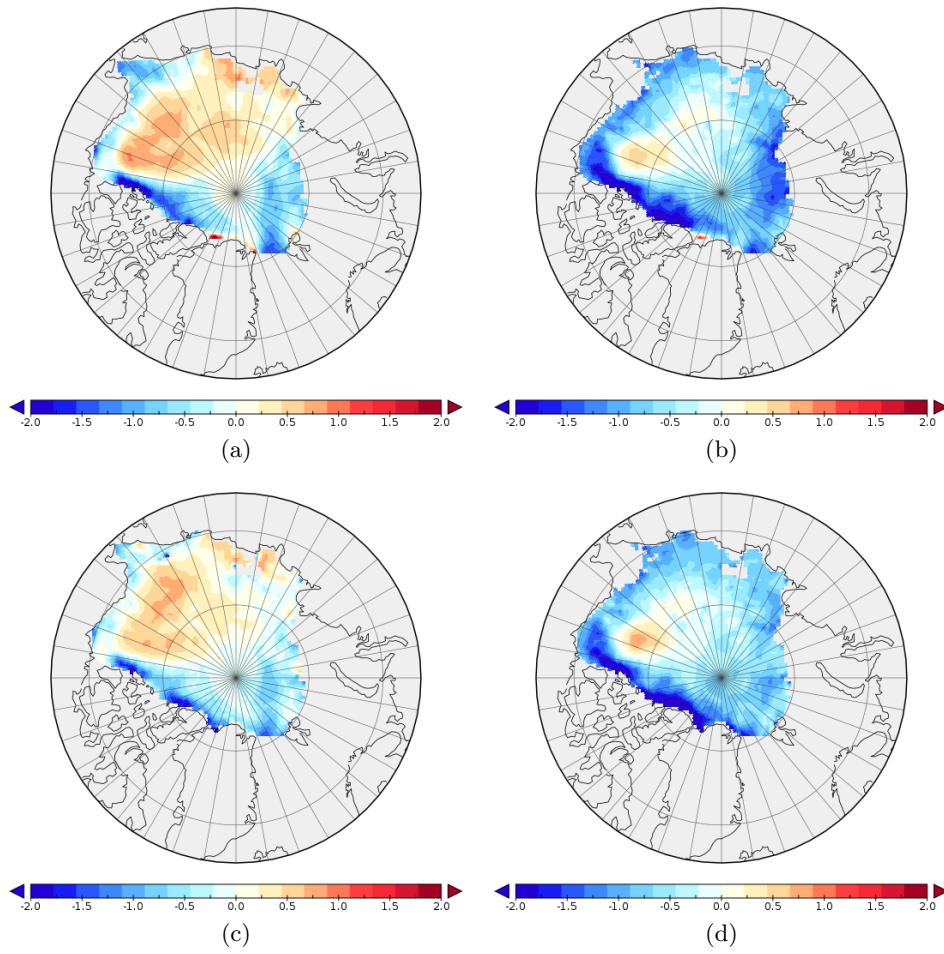
Figure 5: The difference of the modelled and observed (ICESat-JPL) ice thickness climatology [m] for February/March (left column; a) and c)) and October/November (right column; b) and d)) for the newly calibrated NAOSIM (top row; a) and b)) and PIOMAS2.1 (bottom row; c) and d)).
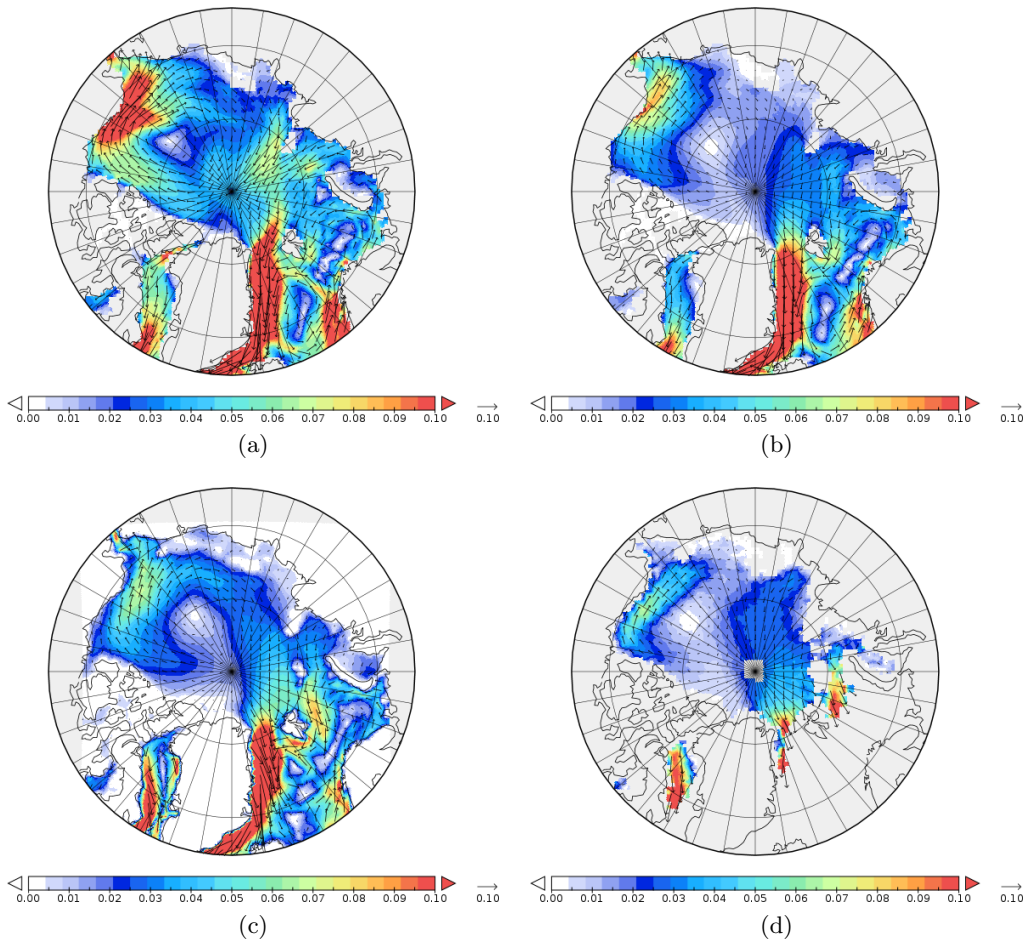
Figure 6: The mean ice drift [m/s] in the months where OSISAF ice drift exists for a) stdNAOSIM, b) newNAOSIM, c) PIOMAS2.1 and d) the OSISAF ice drift itself. The models calculate ice drift even if no ice is present. The so-called free drift solution is not masked out.
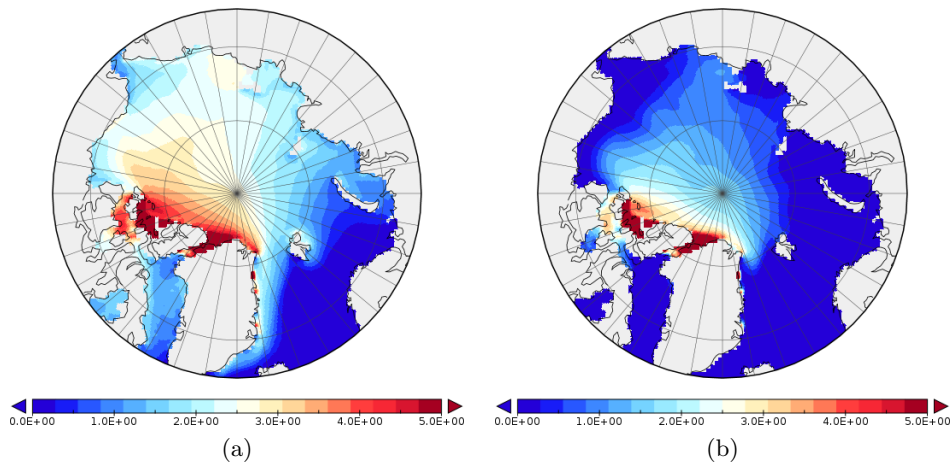

Figure 7: The ice thickness climatology [m] in a) March and b) September in the newly calibrated simulation.
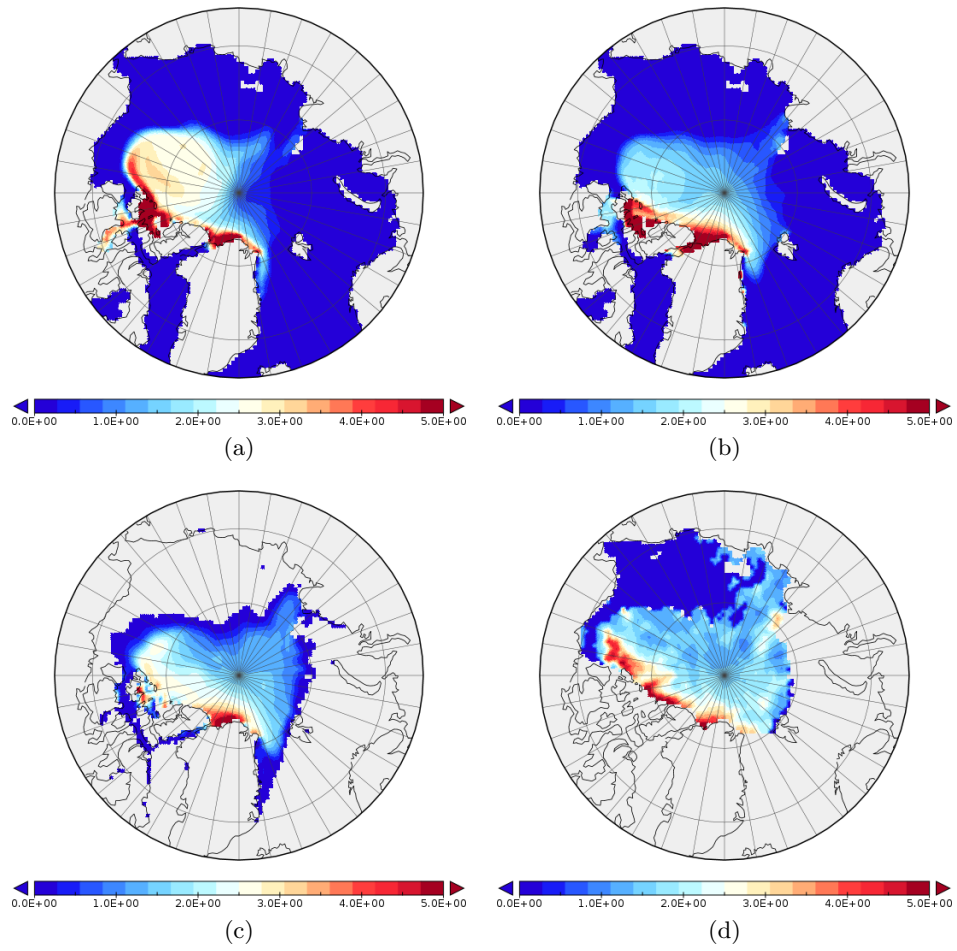
Figure 8: The ice thickness in September 2007 for a) standard NAOSIM b) newNAOSIM c) PIOMAS2.1 and as a reference d) the ICESat-JPL ice thickness in October/November 2007. For PIOMAS2.1 no thicknesses lower then 1cm are plotted.
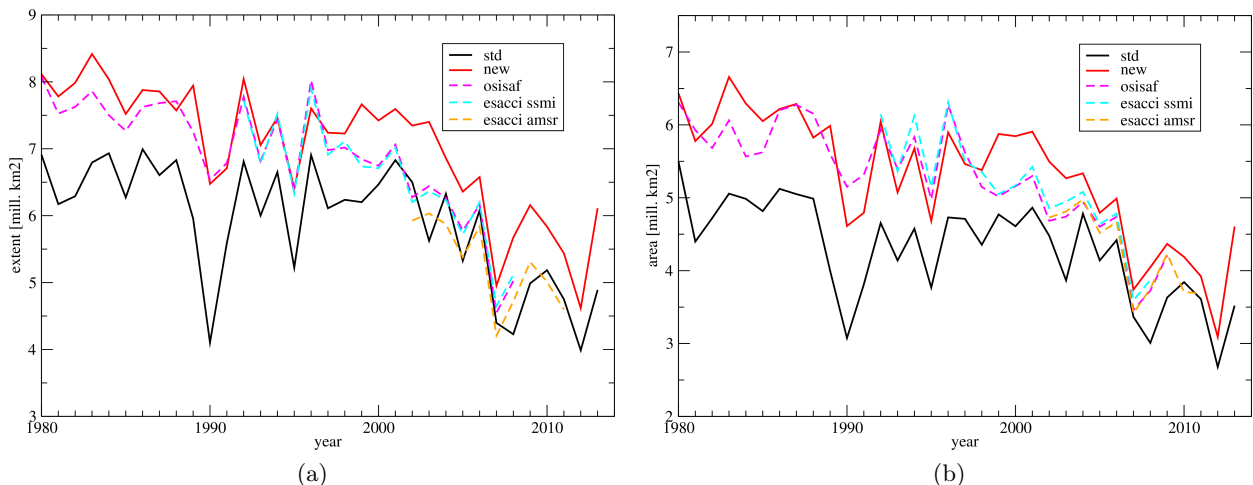
Figure 9: The September a) sea-ice extent and b) sea-ice area of stdNAOSIM, newNAOSIM, and three estimates based on remotely sensed ice concentration. OSISAF (EUMETSAT Ocean and Sea Ice Satellite Application Facility. Global sea-ice concentration operational data set. Norwegian and Danish Meteorological Institutes. Available from http://osisaf.met.no) and ESACCI based on two different sensors (SSMI/R and AMSR, available from http://icdc.zmaw.de/esa-cci_sea-ice-ecv0.html).

4   Observational data streams for the assimilation

This study assesses the potential of recent sea-ice and ocean remotely sensed observations to increase the skill of seasonal predictions of the Arctic sea ice-ocean system. Therefore we utilise data sets which are currently operationally available (ice concentration and sea-surface temperature) or will become operational in the near future (snow and ice thickness). Ice concentration is and has been available for about three decades. One of the most reliable daily ice concentration data sets is the operational OS-ISAF ice concentration (EUMETSAT Ocean and Sea Ice Satellite Application Facility. Global sea-ice concentration operational data set. Norwegian and Danish Meteorological Institutes. Available from http://osisaf.met.no). However, spatially and temporally variable uncertainty estimates are only provided for the post processed OSISAF ice concentration, not for the near real time product.OSISAF also provides daily Atlantic high resolution Sea-Surface Temperature. The product covers the time period from 2012 to today. Currently available remotely-sensed ice thickness products are derived from the SMOS (*Tian-Kunze et al.*, 2014) and CryoSat-2 (*Ricker et al.*, 2014) missions. The SMOS product delivers daily ice thickness but is only reliable for thicknesses lower than about $0.5m$. Therefore, the product is only useful near to the ice edge and is not used here (see *Yang et al.* (2014) for a data assimilation study where the product is applied). The CryoSat-2 ice thickness product is assumed to be reliable for ice thickness exceeding about $1m$. Only winter monthly mean CryoSat-2 ice thickness are available from 2011 to 2014. As has been shown by e.g. *Castro-Morales et al.* (2014) snow thickness is important for the freezing and melting of sea-ice. Although remotely-sensed snow thicknesses are still experimentell a product by the University of Bremen (pers. comm. Christian Melsheimer) has been utilised to constrain the model. Daily data exist around the year with spatially and temporally variable uncertainties. The above described data sets allow to perform data assimilation experiments starting in March for the years 2012 to 2014 and will be described more detailed below.

4.1   Sea-ice concentration

The operational OSISAF ice concentration contains no spatially and temporally variable uncertainties. The uncertainties are calculated following a procedure that has been applied in the ESA Climate Change Initiative Sea Ice (http://esa-cci.nersc.no/) project (L. Toudal Pedersen, pers. comm.). Thereby, the so-

called 'smearing uncertainty' and the algorithmic uncertainty are assumed to be independent and added. The algorithmic uncertainty σ_a exists only for the post processed OSISAF concentration. Evaluating the algorithmic uncertainty in that product reveals that it is normally lower then 6%. This value is applied for all days and locations in the operational product as a conservative estimate of the algorithmic uncertainty. The smearing uncertainty is calculated from the 3x3 standard deviation sd (grid point and eight surrounding grid points) and the empirical formula $\sigma_s = \sqrt{sd}/2.2$ (see Figure 10) is applied. Assuming the independence of both uncertainties the total uncertainty is given by $\sigma = \sqrt{\sigma_a^2 + \sigma_s^2}$. Exemplarily, the ice concentration and the total uncertainty on September, 15$^{\text{th}}$ 2014 is shown in Figure 11.
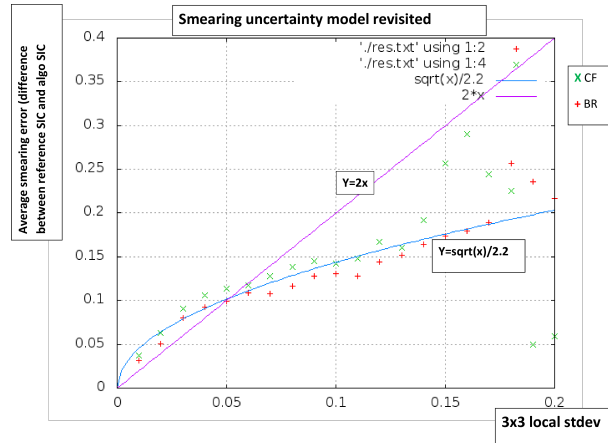

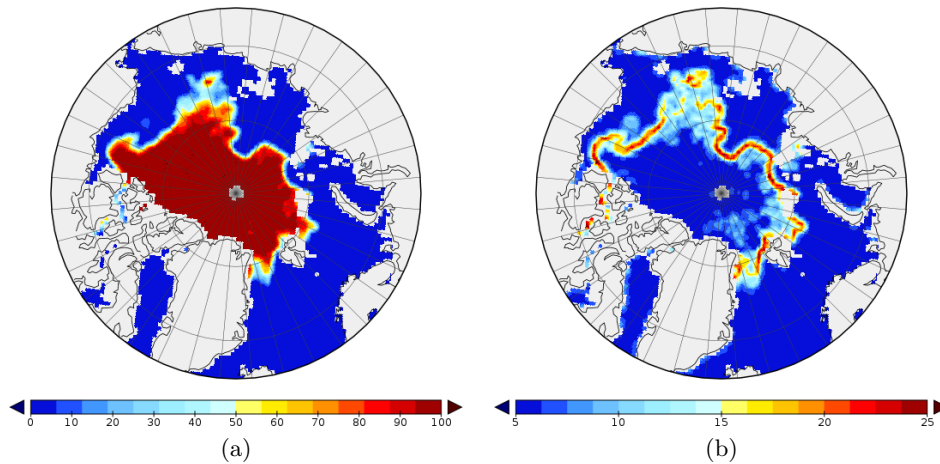Figure 10: The empirical fit of the smearing uncertainty (L. Toudal Pedersen, pers. comm.)

Figure 11: The a) ice concentration [%] and b) the corresponding total uncertainty [%] on September, 15$^{\text{th}}$ 2014.

4.2 Sea-surface temperature

The high latitude Sea-Surface Temperature (SST) product is developed at met.no in cooperation with DMI (http://osisaf.met.no/docs/osisaf_ss2_pum_ahl-sst_v2p1.pdf). The high latitude SST product has a resolution of 5km and is produced twice daily at 00 UTC and 12 UTC. It covers the Atlantic Ocean from 50N to 90N. Polar orbiting satellites with the AVHRR instrument are used for this product. The EUMETSAT METOP-A satellite and NOAA satellites NOAA-18 and NOAA-19 are currently used. The primary purpose of these instruments is to monitor clouds, and it operates in the visible and infrared spectral bands. Figure 12 depicts two examples of daily mean SST on March, 1$^{\text{st}}$ 2014 and on September, 15$^{\text{th}}$ 2014. Light grey areas represents land, sea-ice, or clouds. No spatially and

temporally variable uncertainties are available. Therefore a global uncertainty of $0.5C$ is applied (see http://osisaf.met.no/docs/osisaf_ss2_valrep_ahl-sst_v1p0.pdf).
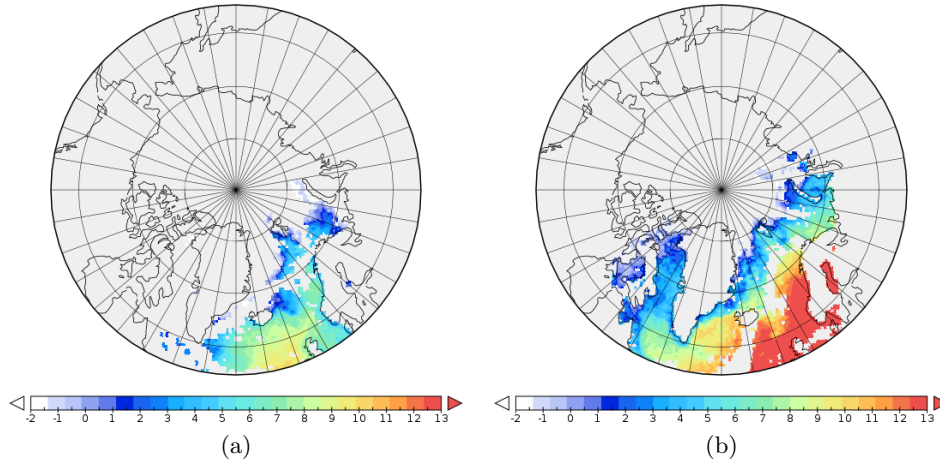

Figure 12: The sea-surface temperature $[C]$ on a) March, 1st 2014 and b) on September, 15th 2014.

4.3   Snow thickness

Snow on sea-ice significantly reduces the heat flow between the ocean and the atmosphere and thus influences the growth and melting of sea-ice. Therefore, snow on sea-ice is an important quantity for modeling and predicting sea-ice. The thickness of the snow layer on sea-ice is estimated using satellite instruments, namely microwave radiometers, microwave scatterometers and optical sensors (University of Bremen, 'Ice Routing Optimisation 2' (IRO2) Report 'Schneeauflage auf Meereis', IRO2 is a German national project funded by the German Ministry of Economy and Transportation, available via Technische Informationsbibliothek Hannover (TIB)).

The basis is an algorithm that uses the influence of the snow layer on the thermal microwave emission by the underlying sea-ice at 19 and 37 GHz (*Markus and Cavalieri*, 1998). This algorithm, however, is applicable to level first-year ice under non-melting conditions. To mask out areas where these conditions are not met, the radar back scatter at 5.3 GHz (C band) is used as it is increased in melting conditions and for rough ice. Here, a threshold of -13dB for the normalised radar back scattering cross section is used. On level ice the uncertainty is first set to a minimum of $6\,cm$ as estimated from a comparison with the NASA ice bridge mission (*Kurtz et al.*, 2013a,b). Then, the uncertainty is derived from two other sources: (1) using Gaussian error propagation to derive the error of the retrieved snow thickness form the radiometric error of the input data (satellite brightness temperatures) and the uncertainty of the retrieval parameters; (2) the standard deviation of the most recent five days of retrieved snow depths. The larger of those two values is taken, provided is is above the initial $6\,cm$. On multi-year ice and under melting-conditions a very large uncertainty of 5m is applied.

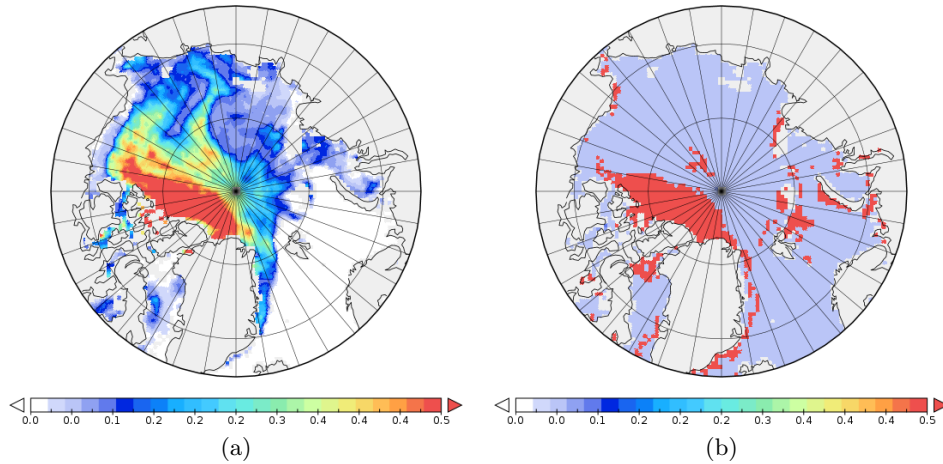Exemplarily the daily mean snow thickness and its uncertainty is shown in Figure 13 on March, 1st 2014.

Figure 13: The a) snow thickness [m] and its b) uncertainty [m] on March, 1$^{\text{st}}$ 2014.

4.4 Sea-ice thickness

For seasonal prediction of sea-ice thickness information are essential which was stressed by many authors (see e.g. *Day et al.* (2014); *Tietsche et al.* (2014)). However, all these studies are perfect model studies meaning that no observations are used but pseudo-observations generated by the models itself.

Satellite remote sensing of sea-ice thickness is based on altimeter measurements of elevation of the snow or ice surface and the elevation of the local sea level. The elevation difference, which is called free board, can be converted into sea-ice thickness using Archimedes principle for which in addition information about snow depth and densities of sea-ice, snow and water is included. Naturally, a number of sources of uncertainties exists. Among others uncertainties in the estimation of the sea level from the geoid, the estimation in the snow cover and its density for the transformation of free board into ice thickness and for some instruments like CryoSat-2 in the estimation of the radar back scatter which is connected to the selection of the retracker algorithm (*Ricker et al.*, 2014).

Before we will analyse the CryoSat-2 ice thicknesses we will show some comparisons of 'historic' (from the 2000s) space-borne remotely-sensed ice thicknesses.

4.4.1 'Historical' space-borne remotely-sensed ice thickness

We used ice thickness estimates based on ICESat (Ice, Cloud, and land Elevation Satellite) before for the calibration of NAOSIM because the products of the Jet Propulsion Laboratory (JPL) is assumed to be the best validated product (*Kwok and Cunningham*, 2008). The JPL product used a combination of SAR images and values of reflectivity to detect leads. The 17 ICESat data sets contain nearly month-long observations of sea-ice surface elevation and other radar characteristics for fall, winter, and spring periods covering the time period of winter 2003 to autumn 2008, both in the Arctic and Antarctic regions. Near-surface measurements are used to determine the sea-ice free board (height of the surface above the ocean level) to infer sea-ice thickness using model estimates of snow depth on top of the floating ice.

Another product based on ICESat is produced by the Goddard Space Flight Center (GSFC). Here, in contrast a lowest-elevation approach to retrieve tie points is used (*Zwally et al.*, 2008). Another difference between JPL and GSFC processing is the usage of different snow depth fields for the free board-to-thickness conversion. While *Kwok and Cunningham* (2008) use ECMWF fields to calculate snow accumulation, the GSFC product applies the Warren snow climatology (*Warren et al.*, 1999).

The difference between ICESat-GSFC and ICESat-JPL ice thickness climatologies for February/March and October/November (Figure 14) is remarkably large. In February/March the differences are negative almost everywhere with a mean difference of about $50cm$ in the central Arctic. Especially large are the differences north of the Canadian Archipelago and north of Greenland and around Franz-Joseph-Land where differences of more then $2m$ can be reached. Interestingly, these are the areas were we found the

largest negative differences when we compared newNAOSIM and PIOMAS2.1 with ICESat-JPL (cmp. figure 5) leading to the conclusion that most likely ICESat-JPL is overestimating the ice thickness in these areas strongly. In Autumn (October/November) similar differences between ICESat-GSFC and ICESat-JPL can be found but more regions have slightly positive values, and the large differences north of Greenland are strongly reduced.

In the framework of ESAs Climate Change Initiative (ESACCI - http://ionia1.esrin.esa.int) sea-ice variables are investigated (SICCI - http://esa-cci.nersc.no/). The objectives of the Sea Ice CCI (SICCI) are to provide quality-controlled ice concentration data sets for the Arctic and Antarctic from 1979 to present based on passive microwave data and to provide Arctic sea-ice thickness data sets based on radar altimeter data from 1993 to present with the best possible validation and error characterisation. Advances in the processing of ERS and Envisat radar altimeter data have though shown that these instruments can be used calculate low resolution sea-ice thickness from measurements of ice free board. We calculated difference of this newly generated product (still prototype version 0.9) with ICESat-JPL (Figure 15) for February/March and October/November. In October/November the differences of the climatology can reach $2m$ in the central Arctic. In the SICCI Product Validation and Intercomparison Report (pers. comm. Stefan Kern - public but not available online yet) comparisons with Russian drift stations NP-37 and NP-38 and EM-Bird measurements reveal a positive bias of about $1m$ and comparisons with ULS data from the Beaufort Gyre Exploration Project (BGEP) reveal a positive bias of $0.7m$ to $1m$ of the SICCI thickness. Nevertheless, north of the Canadian Archipelago a negative bias can be found in the difference to ICESAT-JPL pointing again towards a positive bias of ICESat-JPL in that region (cmp. figure 14 and figure 5). In October/November the SICCI and ICESat-JPL differences are much lower but can reach locally $2m$ as well.

For completeness the differences between SICCI and ICESat-GSFC climatologies are shown in Figure 16.

In summary we conclude that very large biases between the 'historical' sea-ice thickness products exists which exceeds the differences between ICESat-JPL and either model (PIOMAS2.1 or newNAOSIMj) and also between both models.
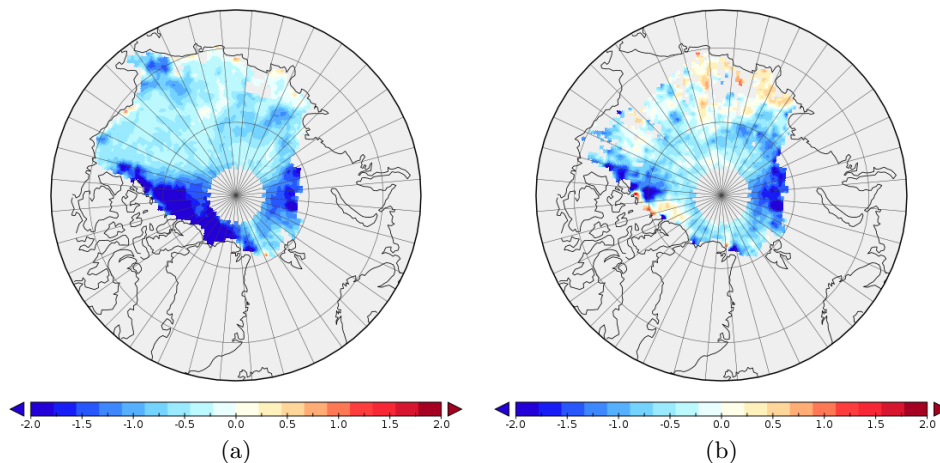

Figure 14: The difference between ICESat-GSFC and ICESat-JPL ice thickness climatologies [m] for a) February/March and b) October/November.
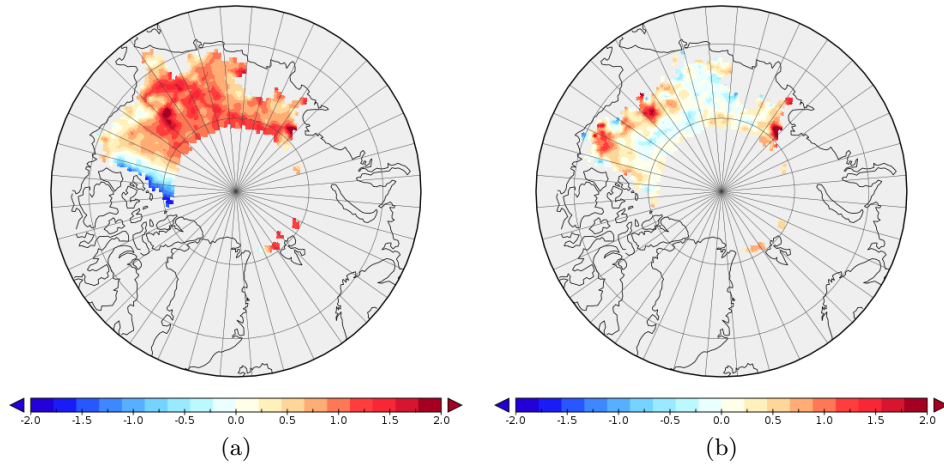
Figure 15: The difference between SICCI and ICESat-JPL ice thickness climatologies [m] for a) February/March and b) October/November.
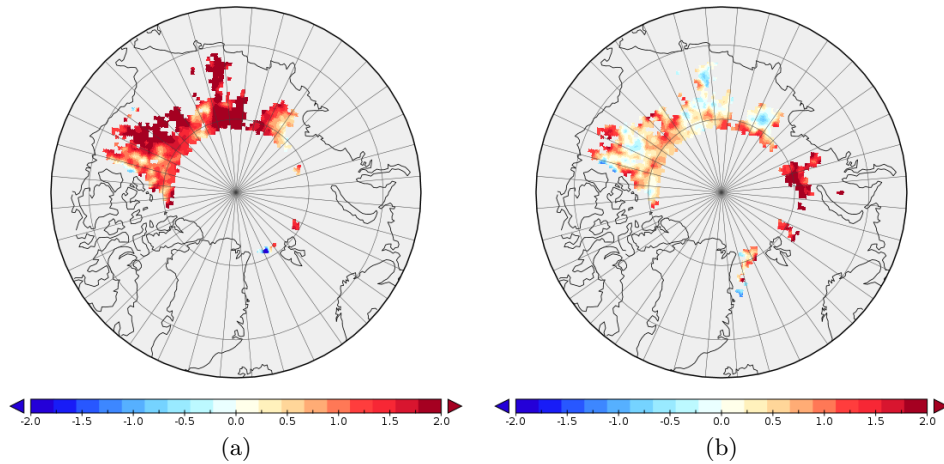

Figure 16: The difference between SICCI and ICESat-GSFC ice thickness climatologies [m] for a) February/March and b) October/November.

4.4.2 CryoSat-2 retracker uncertainty

One major source of uncertainty with respect to the CryoSat-2 ice thickness is the selection of the retracker. *Ricker et al.* (2014) used three different thresholds with 40, 50, and 80% of the first maximum of radar power echo which spans the range of values used in the current literature. The effect on the estimated ice thickness is shown exemplarily in Figure 17 for March 2013. While the 40% and 50% threshold gives relatively similar thicknesses the 80% threshold gives drastically smaller thicknesses. Figure 18 shows the mean of the ensemble of the three different retracker thresholds and the ensemble standard deviation for March 2013. On level ice the uncertainty (the standard deviation of the ensemble is used here) is about 20% of the mean value but on multi-year ice 50% and more can be reached.
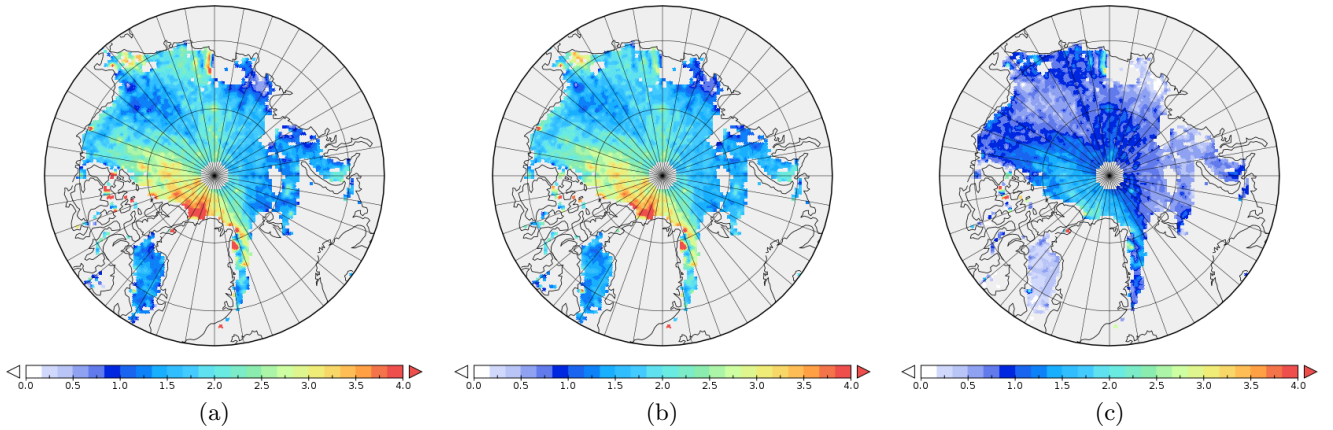
Figure 17: The CryoSat-2 ice thickness [m] for March 2013 estimated by a retracker threshold of a) 40% b) 50%, and c) 80%.
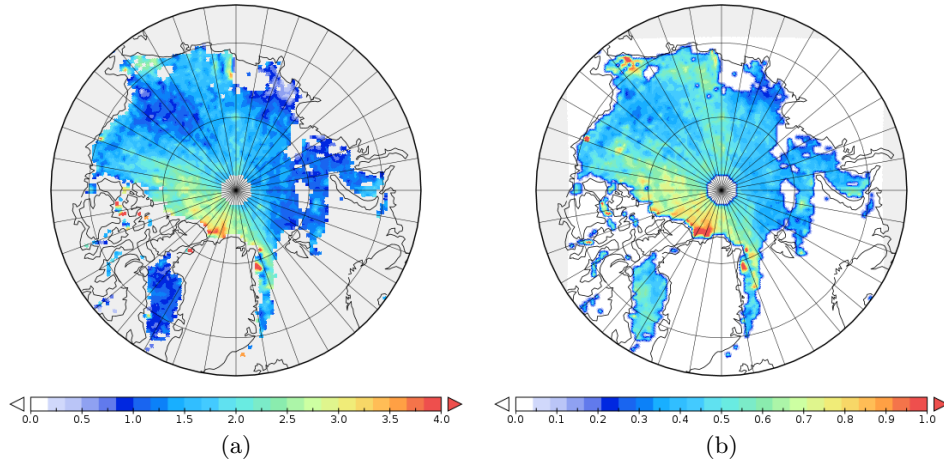

Figure 18: The ice thickness [m] of the a) ensemble mean and b) the standard deviation [m] of the ensemble on March 2013.

4.4.3 On CryoSat-2 snow thickness uncertainty

Another source of uncertainty in the estimation of the ice thickness is the snow cover. *Ricker et al.* (2014) used a modified Warren climatology (*Warren et al.*, 1999) to transform the free board into ice thickness. Thereby the Warren climatology is modified on level ice because it has been shown by many authors that the Warren climatologies snow depth is too large there (see e.g. *Kern et al.* (2015)). Figure 19 exhibits for March 2012, March 2013, and March 2014 the CryoSat-2 ice thickness of the 50% threshold (top row; Figure 19 a), b), and c)), the modified Warren climatology (second row; Figure 19 d), e), and f) - level ice can be clearly distinguished from multi-year ice by the sharp horizontal gradient), the snow depth estimates of University Bremen (UB) (see section 4.3) (third row; Figure 19 g), h), and i)), and the difference between the snow depth from the University of Bremen and the modified Warren climatology (bottom row; Figure 19 j), k), and l)). As has been discussed in section 4.3 snow depth (UB) is only validated on level ice with a mean error of about $6cm$. However, the difference between the snow depth (UB) and the modified Warren climatology exceeds this error in many areas even for level ice. A striking feature of the ice thickness in March 2014 (Figure 19 c)) is the strong increase north of the Canadian Archipelago of up to $1m$ compared to March 2012 and March 2013. This change goes along with a large increase in the difference of the snow depth (Figure 19 l)) of about $30cm$ in almost the same area north

of the Archipelago where the ice thickness is increased in March 2014. Although, the snow depth (UB) is not validated on multi-year ice one might speculate that the increase in ice thickness in March 2014 north of the Archipelago is caused by a misinterpretation of snow as ice. This demonstrates the current limits of space-borne remotely-sensed ice thickness products.
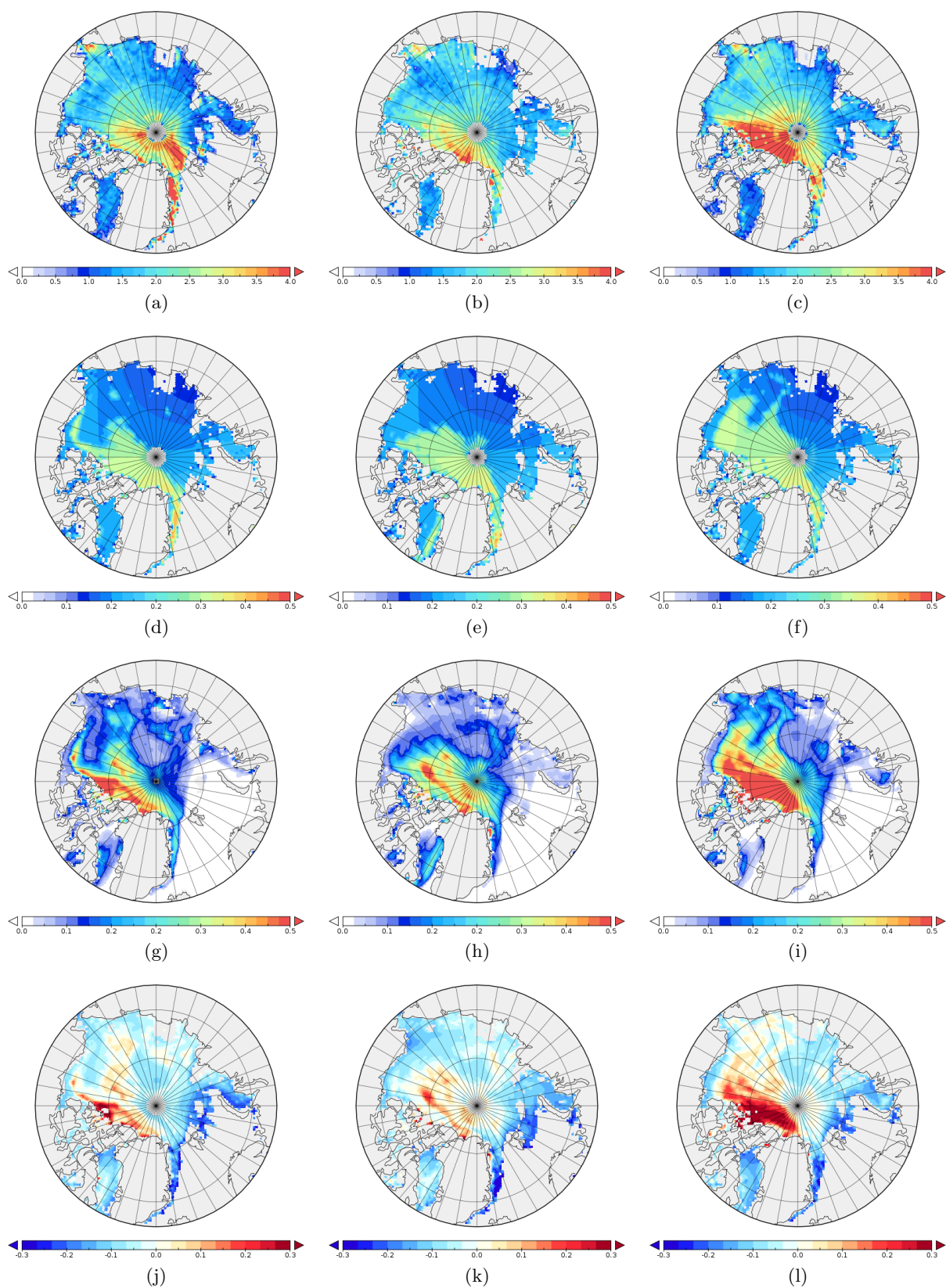
Figure 19: The CryoSat-2 ice thickness [m] of the 50% threshold for a) March 2012, b) March 2013, and c) March 2014, the modified Warren snow depth climatology [m] for d) March 2012, e) March 2013, f) March 2014, the snow depth [m] of the University of Bremen for g) March 2012, h) March 2013, and i) March 2014, and the difference between the snow depth of the University of Bremen and the modified Warren climatology [m] for j) March 2012, k) March 2013, and l) March 2014.

4.4.4 Bridging the gap between CryoSat-2 and ICESat-JPL

A direct comparison of the CryoSat-2 ice thickness with e.g. the ICESat-JPL product is not possible because the data do not overlap in time. However, numerical models can be used to bridge the temporal gap between both data sets. This has been done for newNAOSIM and PIOMAS2.1. The mean difference of ice thickness between newNAOSIM and PIOMAS2.1 and the ensemble mean of the three retracker thresholds of CryoSat-2 for March 2011 to 2013 shows a large positive bias of both models in almost the whole Arctic except north of Fram Strait (Figure 20 a) and c)). Most pronounced is the positive bias north of the Canadian Archipelago and north of Greenland. The corresponding November difference (Figure 20 b) and d)) shows for both models a much smaller bias. Especially PIOMAS2.1s deviations are in almost all regions lower then $0.5m$. The analogous differences for the 50% threshold in March (Figure 21 a) and c)) are much lower then the ensemble mean differences and are similar to the corresponding differences with ICESat-JPL (Figure 5 a) and c)) except north of the Canadian Archipelago and north of Greenland where ICESat-JPL exhibits much thicker ice. The November differences for the 50% thresholds ((Figure 21 c) and d)) are only slightly lower than corresponding differences of the ensemble mean ((Figure 20 b) and d)).
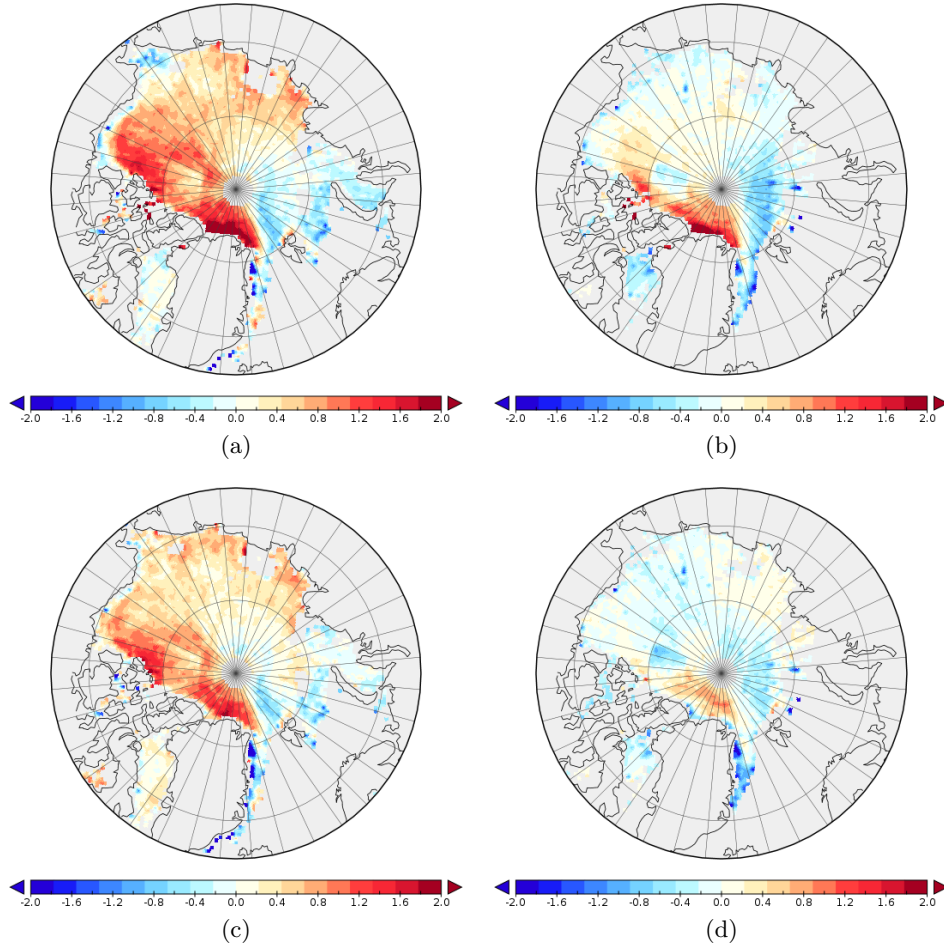

Figure 20: The mean difference of the ice thickness [m] between the models and the CryoSat-2 ensemble mean for March 2011 to 2013 (a) newNaosim, b) PIOMAS2.1) and for November 2011 to 2013 (c) newNaosim, d) PIOMAS2.1).
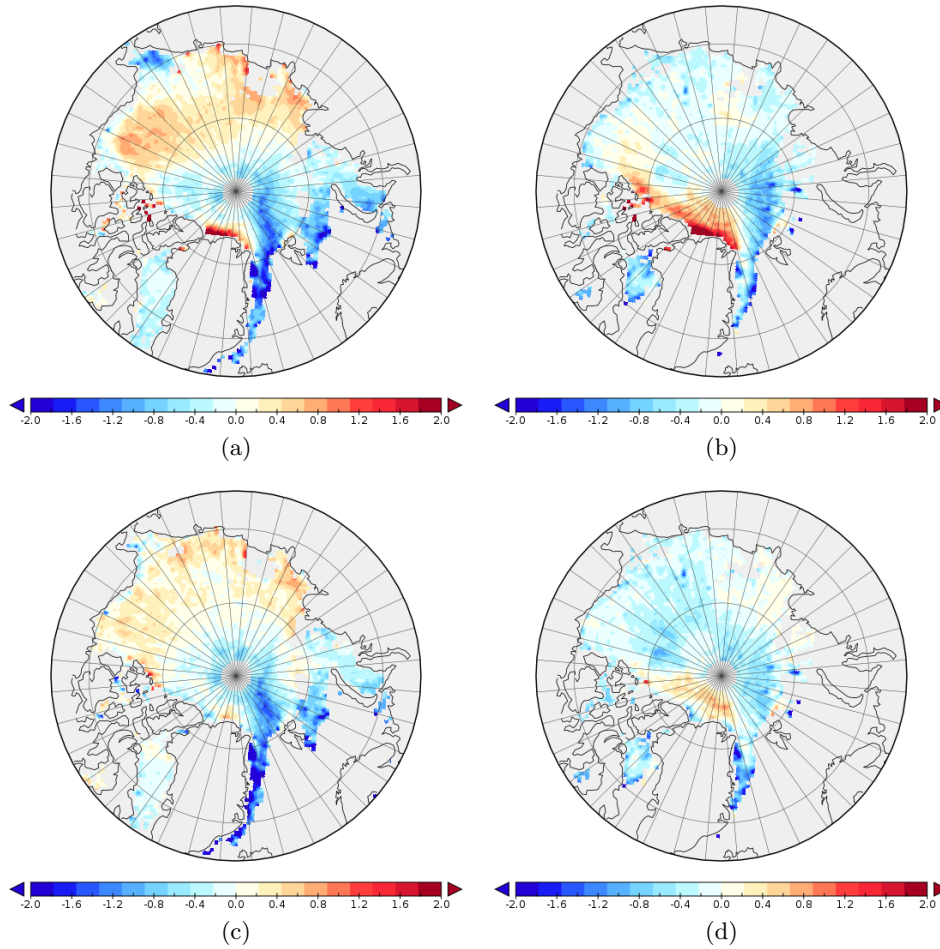
Figure 21: The mean difference of the ice thickness [m] between the models and the CryoSat-2 50% threshold for March 2011 to 2013 (a) newNaosim, b) PIOMAS2.1) and for November 2011 to 2013 (b) newNaosim, d) PIOMAS2.1).

4.4.5 CryoSat-2 and data assimilation

It is outside the scope of this study to compare CryoSat-2 ice thickness with in-situ observations. This has been done e.g. within the ACCESS report D1.29 'Report on altimeter sea-ice thickness errors due to ice type, geometry and snow pack effects'. One conclusion of this report was: 'CryoSat appears to over-estimate the thickness of level ice and under-estimates the thickness of deformed ice' (However, the 40 % threshold was used in this study. With the 50 % threshold that we use in our study we get a slightly lower sea-ice thickness, but the spatial thickness distribution remains.).

We discussed in the previous sections some caveats of the product but we showed also that the 50% retracker thresholds ice thickness is consistent with the ICESat-JPL ice thickness when models are used to bridge the temporal gap between the two data sets in large areas except north of Canadian Archipelago and north of Greenland. Therefore we will use the 50% threshold data for the data assimilation purpose. It remains to estimate the uncertainties applied for the data assimilation. Because the uncertainty in the retracker threshold is a major source of uncertainty we will be guided by the quotient of the ensemble standard deviation and the mean state (exemplarily shown for March and November 2012 in Figure 22). In March a sharp gradient in the quotient allows to distinguish level and multi-year ice. On level ice a uncertainty of about 20% is appropriate whereas on multi-year ice 50% uncertainty is reached in many locations and on some locations even 100%. In November a clear separation of level and multi-year ice is not visible anymore. Similar patterns are yielded for 2011 and 2013. Based on this analysis we applied an uncertainty of 20% on level ice and 50% on multi-year ice all the year round in the data assimilation
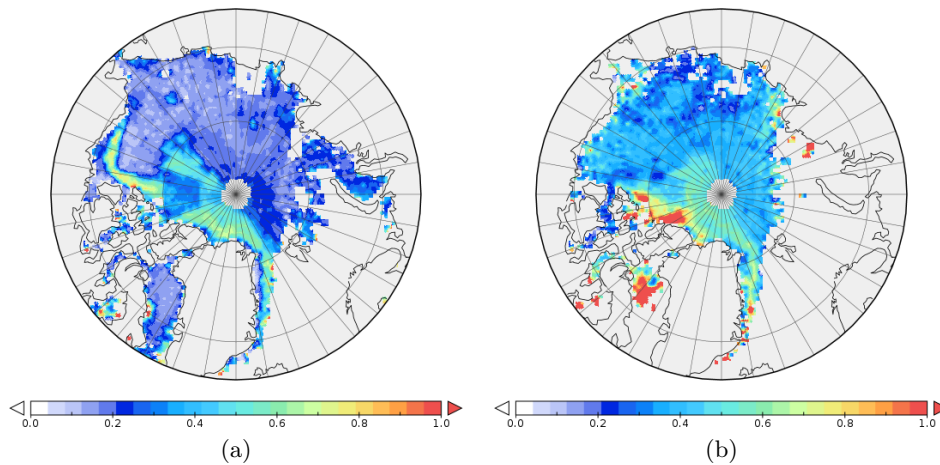
experiments.

Figure 22: The quotient of the ensemble standard deviation of ice thickness and the ensemble mean ice thickness for a) March 2012 and b) November 2012.

5   Data assimilation

In this section three different sets of experiments are described: First, experiments assimilating all four data streams, then experiments to derive a bias corrections scheme for the ice thickness product, and finally experiments applying the bias correction scheme. The prior fields x_0 in equation (1) are taken from the newly calibrated model run decribed in section 3 at appropriate times.

5.1   Assimilation of CryoSat-2 ice thickness

In the assimilation experiments CryoSat-2 ice thickness (50% retracker threshold), ice concentration, snow depth and SST with uncertainties as described above are utilised. We start the assimilations at March 1st of the years 2012 to 2014 and use an assimilation window of two months i.e. from March 1st until April 30th. For calculation of the cost function regarding the CryoSat-2 ice thickness a monthly-mean model ice thickness is computed. Clearly the weight of this data stream would be lower than for the data streams with daily availability. As one of the main objectives of this study is the performance of this data stream, we decided to substantially increase its weight in the cost function through multiplication with a factor of 180. Note that the control vector consists only of the initial state at March 1st of the ice concentration, ice thickness, and snow depth but also of the ocean temperature and salinity at all model levels.

Figure 23 depicts the evolution of the cost function and the norm of the gradient over the iteration number for all three years. The number of iterations is different but all assimilations reach a considerably small gradient after about 50 to 70 iterations. Figure 24 shows the total cost function and the contributions of all four data streams and the prior separately before (a priori) and after the last iteration (a posteriori) of the optimisations for each year. Especially the cost function for ice thickness and snow depth are decreased strongly. The reduction with respect to ice concentration and SST is only moderate.
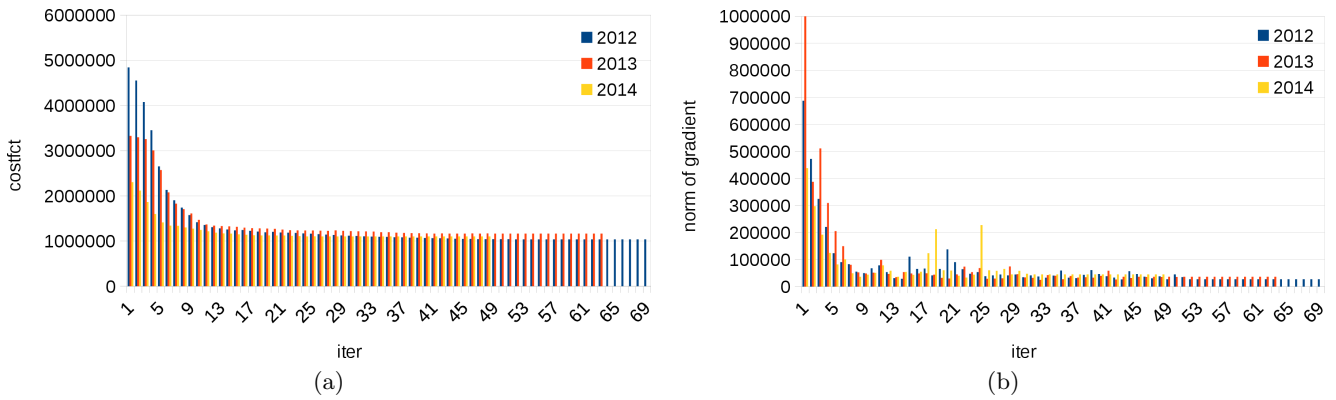
Figure 23: The evolution of a) the cost function and b) the norm of the gradient over the iteration for the assimilation experiments of the years 2012 to 2014.
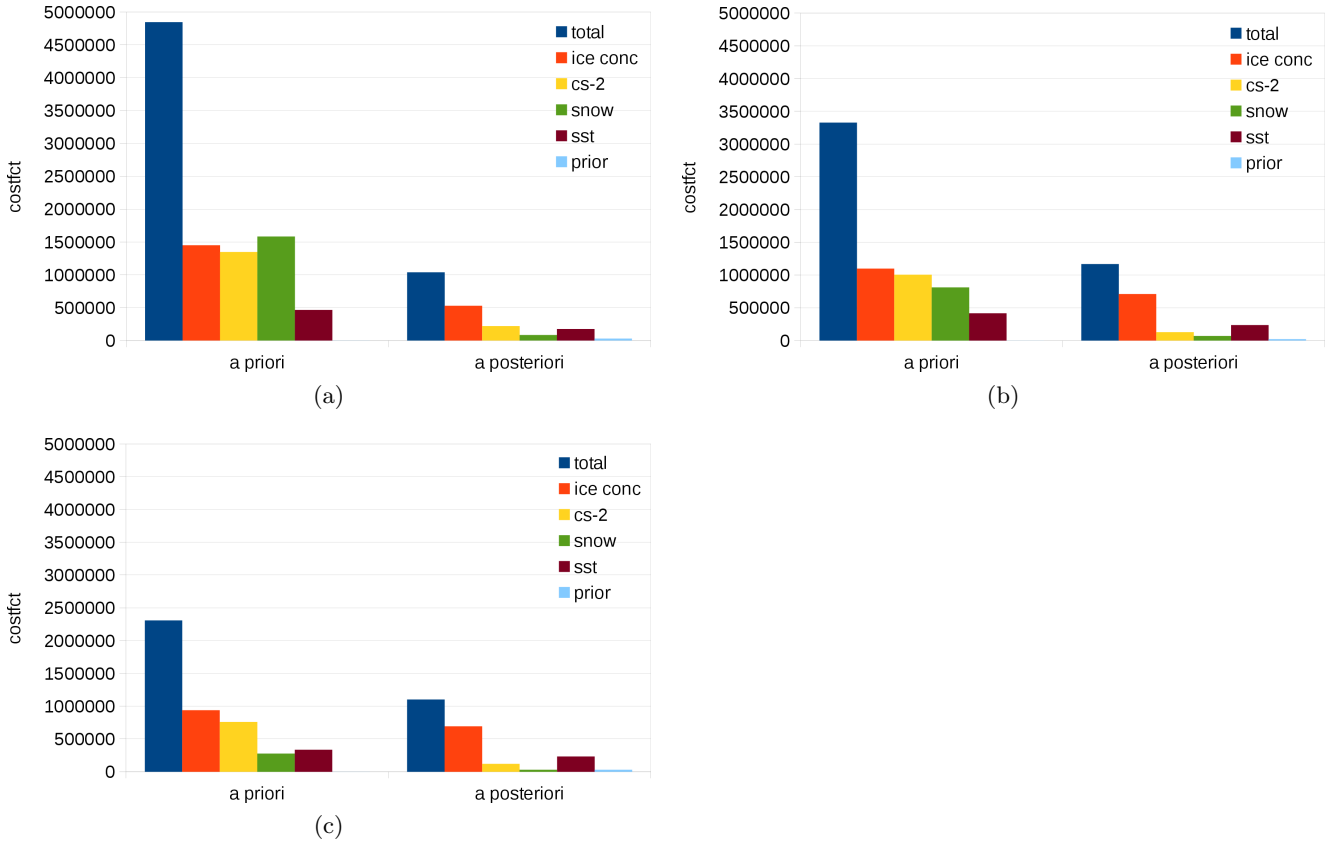

Figure 24: The terms of the cost function before and after the optimisation for a) 2012 b) 2013 c) 2014.

As we assimilate only data in March and April, the observations for the subsequent months (as long as data exist) provide independent information that we can use to assess the forecast skill prior and posterior to the assimilation. Our skill score is the squared misfit weighted by the squared uncertainty (as in the definition of the cost function, but without the extra scaling factor for the ice thickness contribution). We regard this as a better measure of the skill compared to the root mean square error (which handles every misfit equally regardless of the observational uncertainties) because our measure takes the observational uncertainties into account. We run the model from March until end of April of the following year (except for 2014 where we could run the model only until end of December for obvious reasons). The cost function

22

for all data streams is depicted in Figure 25 calculated for every month separately for the control run (no data assimilation) and the optimised run. For March and April the cost function of the CryoSat-2 data is strongly reduced in the optimised run for all years. But in November the skill of the optimised run is lower for 2012 and 2013 and only slightly better for 2014. In the following March and April the optimised runs show some improvements (for 2012 and 2013). The ice concentration of the optimised run has a higher skill in March and April but is loosing the advantage pretty fast. In summer (July to September) only in 2012 some improvement is visible while in 2013 the skill is worse than in the control run especially in September. In 2012 and 2013 the skill is slightly but consistently improved from December to April of the following year. The cost function for snow depth is strongly reduced from March to June for all years. From July to September almost no ice is left. From January to April a slight improvement is recognisable. The cost function of the SST shows only in March and April some improvement. Overall, the monthly-mean cost functions do not show any improvement of the assimilation in summer, i.e. do not show any improvement of the skill with respect to seasonal prediction.
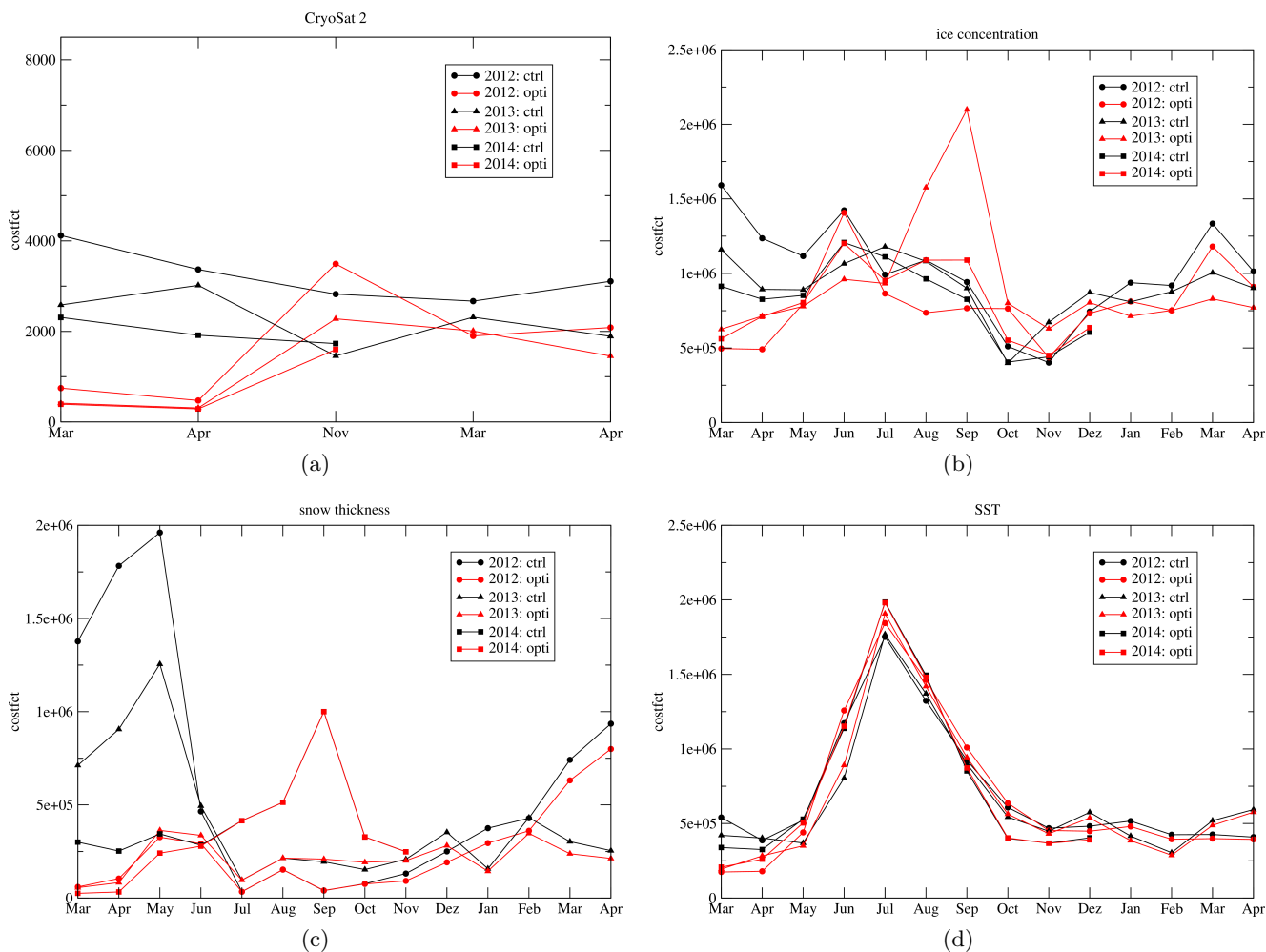

Figure 25: The temporal evolution of the monthly summed up cost function from March to April of the following year (for 2014 to December of the same year) for a) the CryoSat-2 ice thickness (data are currently only available for March, April and November; not scaled, see text), b) the OSISAF ice concentration, c) the snow depth (UB), and d) the OSISAF SST.

The March ice thickness of the control run and the optimised run are shown for every year in Figure 26 together with the misfit between the optimised March ice thickness and the CryoSat-2 ice thickness. The largest differences between the optimised and the control run ice thickness in March can be seen in the Beaufort Sea, the East Siberian Sea and in the Kara and Laptev Sea. North of the Canadian Archipelago

and north of Greenland both, the optimised and the control run, depict much thicker ice in March 2014 compared to 2012 and 2013 (consistent with the CryoSat-2 estimates). The misfit between the optimised and the CryoSat-2 March thickness shows only in March 2012 large differences from Fram Strait towards the pole. All other locations show only minor differences (except east of Greenland which is not important here). The reason for the strong reduction of the ice thickness in and north of Fram Strait in March (and April) 2012 can be explained by the ice concentration assimilation. As already discussed in section 3 newNAOSIM is overestimating the ice concentration in winter in the Barents and Greenland Sea resulting in a too far south ice margin (Figure 27 a) which is reflected as well in the cost function listed in Table 1. In the optimised run the misfit is strongly reduced (Figure 27 b)) which is achieved by reducing the ice thickness north of Fram Strait towards the pole, reflecting the pathway of the Transpolar Drift in the model (cmp. figure 6 b)). In other words: the prize we have to pay for a more reliable ice margin is a misfit in the CryoSat-2 ice thickness. Especially in March 2012 CryoSat-2 is showing very thick ice next to and along the described pathway (Figure 19 a)) which is obviously not consistent with the model's ice margin (and the model's ice dynamic).
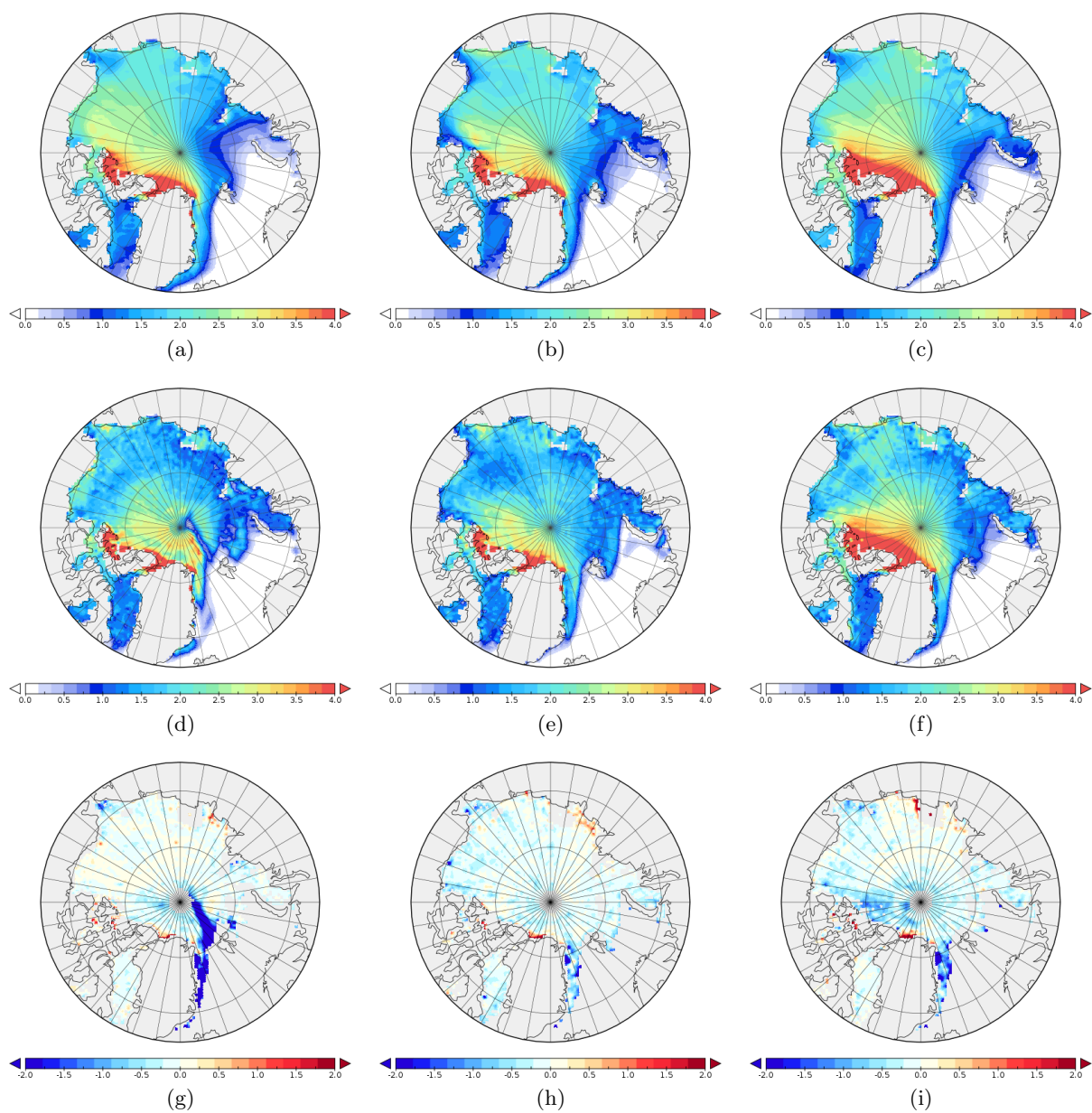
Figure 26: The ice thickness [m] prior to (top row) and after the optimisation (second row) and the difference between the optimised ice thickness and CryoSat-2 [m] for March 2012 (a), d), and g)), March 2013 (b), e), and h)), and March 2014 (c), f), and i)).
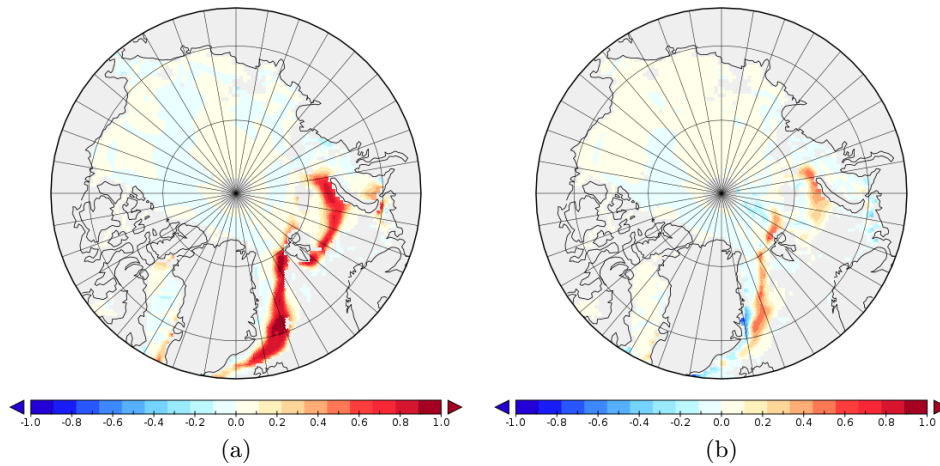
Figure 27: The misfit [0-1] of the a) control run and and b) the optimised run to the OSISAF ice concentration for April 2012.

Because we do not have any ice thickness data in September we can only compare the ice concentration during this time period of lowest ice extent (Figure 28). In general, the model is underestimating the ice concentration even in the central Arctic by about 0.1 which we attribute to the model physics. The corresponding differences of the control run and the optimised run to the OSISAF ice concentration (Figure 29) shows in the control run a strong overestimation in the Beaufort Sea and over the Chukchi Plateau (less pronounced in September 2014) and an underestimation over the Eurasian Basin slope for all years. In the optimised run the overestimation is gone which we attribute to the reduction in the initial ice thickness on March 1st in the Beaufort Sea and over the Chukchi Plateau (cmp. Figure 26 d), e) and f)). However, the reduction is especially in 2013 too large, resulting in an underestimation of the ice concentration in September there. Over the European Basin slope no improvement can be seen, on the contrary, especially in September 2013 the underestimation is exaggerated which is responsible for the very large cost function in September 2013 (cmp. Figure 25 b)). In summary, the assimilation of the CryoSat-2 ice thickness improves the predictability in September in the Beaufort Sea (and over the Chukchi Plateau in September 2012) but reduces the predictability everywhere else.
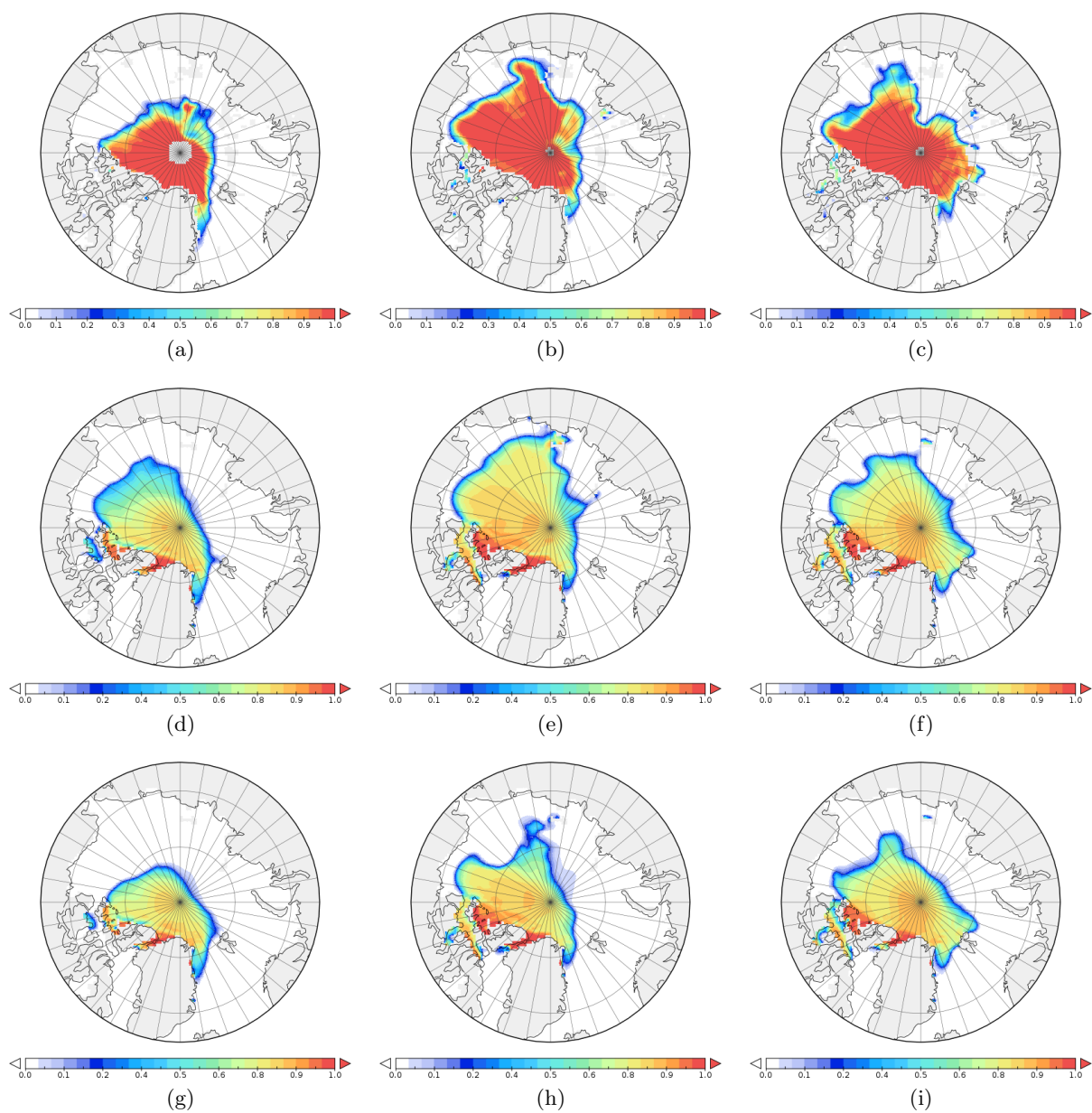
Figure 28: The ice concentration [0-1] as observed by OSISAF (top row), prior to (second row) and after the optimisation (bottom row) for September 2012 (a), d), and g)), September 2013 (b), e), and h)), and September 2014 (c), f), and i)).
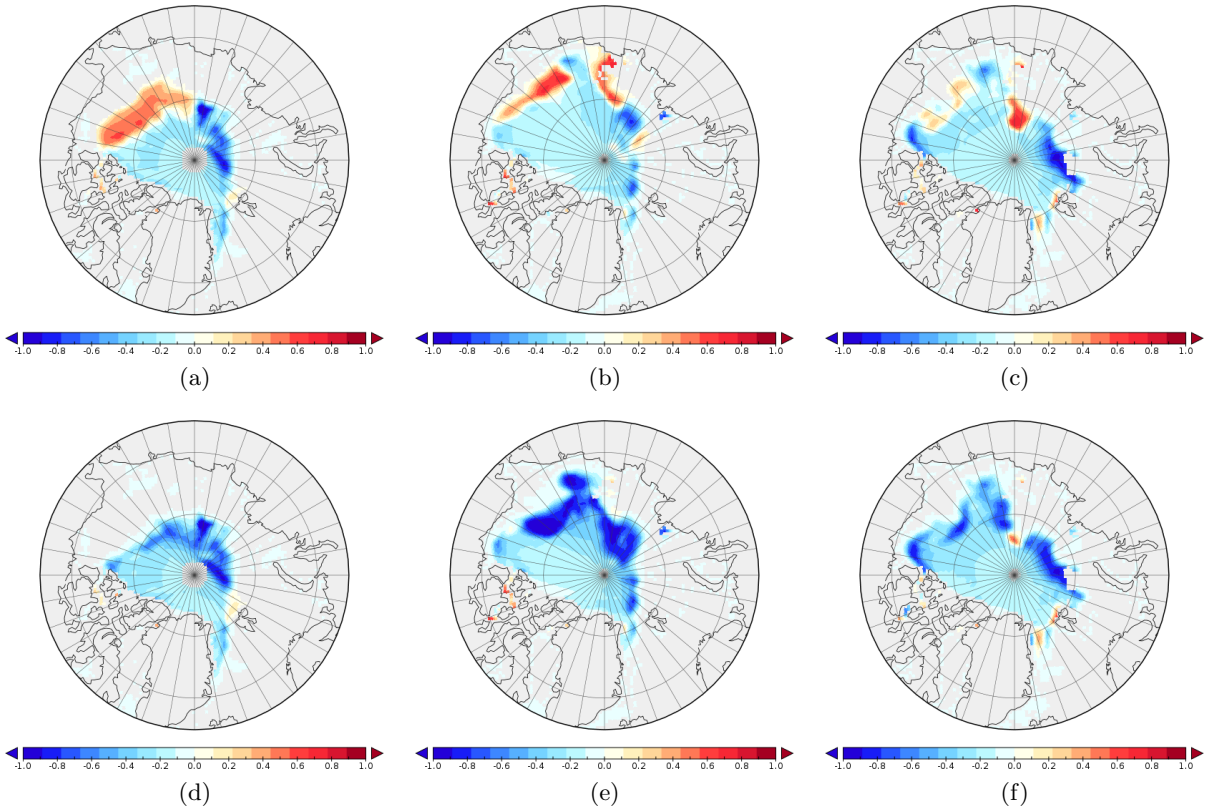
Figure 29: The misfit of the ice concentration prior to (top row) and after the optimisation (bottom row) and the OSISAF ice concentration for September 2012 (a) and d)), September 2013 (b) and e)), and September 2014 (c) and f)).

5.2 Reconstructed initial ice thickness

Our next set of experiments explores the feasibility of inferring an initial ice thickness distribution on March 1st that is consistent with the summer ice concentration. We, hence, use an extended assimilation window from March 1st to end of September and assimilate ice concentration from July to the end of September together with snow depth and SST which we assimilate from March to September. We do not use any ice thickness observations.

The performance of the assimilation system can be found in Figure 30 and Figure 31. In contrast to the assimilation with CryoSat-2 data the norm of the gradient can reach very large values for larger number of iterations. Nevertheless, all terms of the cost function are reduced considerably (Figure 31) for all years.
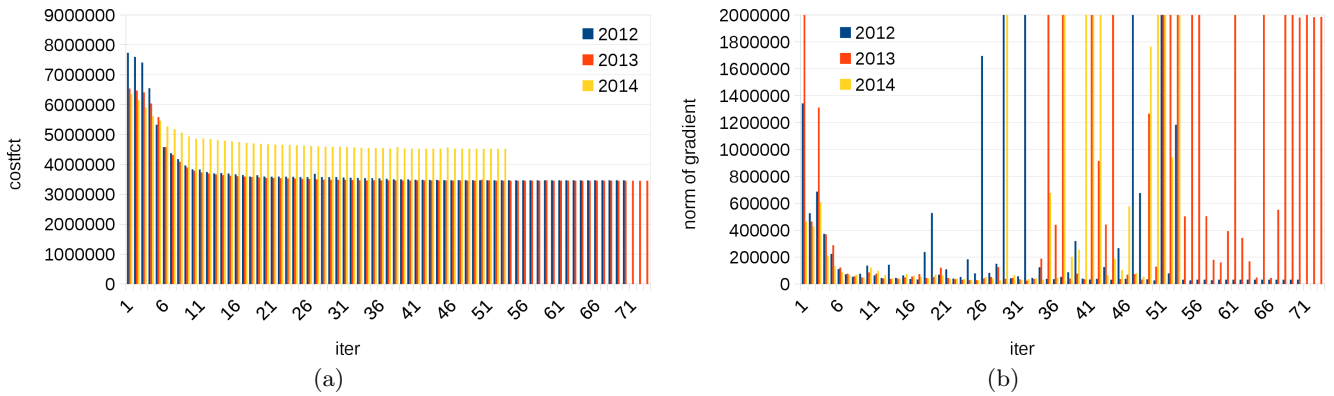
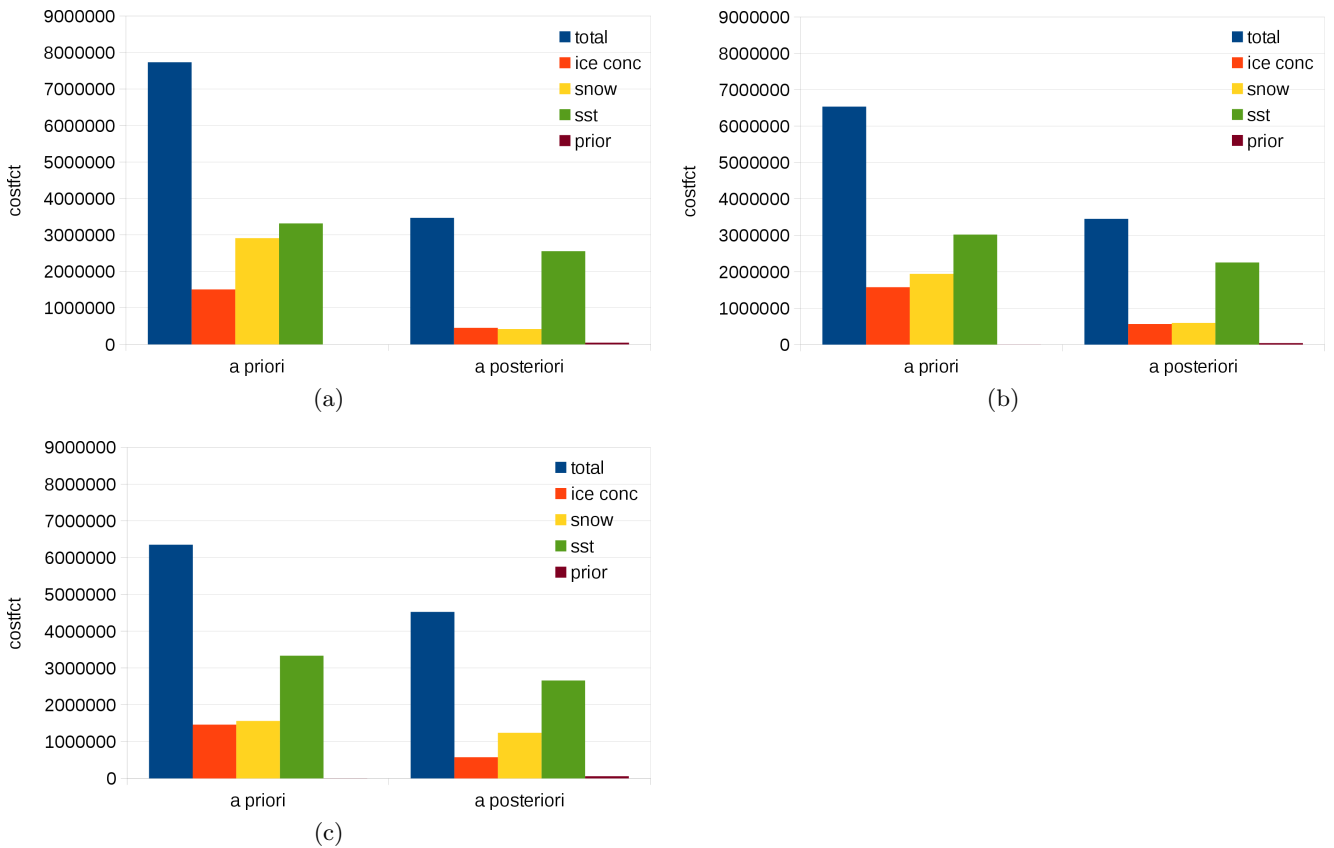Figure 30: As Figure 23 but for the reconstruction experiment.

Figure 31: As Figure 24 but for the reconstruction experiment.

The analysed monthly cost function (cmp. Figure 25) for ice concentration (Figure 32 b)) shows a large reduction from July to September but also the cost function from March to June is reduced although we have not assimilated ice concentration in that time period. However, already in October the improvement is gone which we attribute to the freezing of sea-ice which is not constrained by any observations. The monthly cost function for the CryoSat-2 data (Figure 32 a)) is strongly increased for March and April but gets closer to the cost function of control run for November and March and April of the following year. The monthly cost function of the snow depth is very similar to the CryoSat-2 assimilation experiments but the cost function of the SST shows some improvement as well from May to August (This is not trivial although we assimilated the data because we modified only the initial state.). The setup of the experiments does not allow to deduce if it is a direct effect of the SST assimilation or

29

an indirect effect by the ice concentration effect but we suspect that it is the latter.
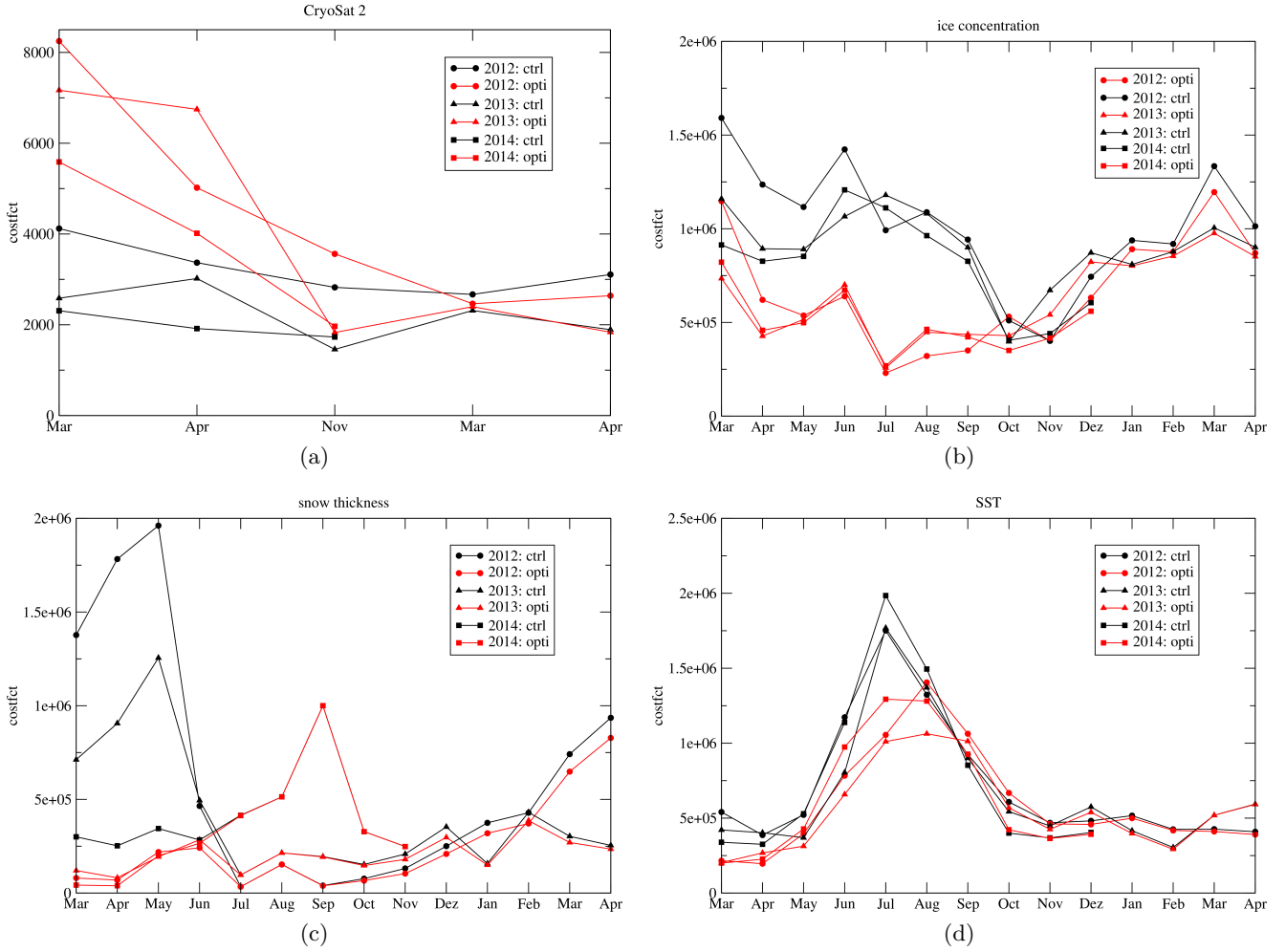

Figure 32: As Figure 25 but for the reconstruction experiment.

The optimised March ice thickness shows in very large areas of the Arctic larger values than the CryoSat-2 data (Figure 33 c)) in all three years. Areas with similar values as CryoSat-2 are the southern parts of the Beaufort Sea, the Chukchi Plateau, and the Kara and Laptev Sea. Lower values can be found over the Eurasian Basin slope except for areas north of the Laptev Sea. Obviously the model needs the high ice thickness values to optimise the ice concentration in summer which is shown in Figure 34 for September. Note that almost all large misfits at the ice margins could be removed by the optimised initial state. Only the bias of about 0.1 in the central Arctic remains. By looking into animations of the ice thickness monthly fields one can identify that the thick ice north of the Laptev Sea is essential for the bias reduction over the Eurasian Basin slope in all years.

Given the degrees of freedom that our assimilation setup leaves to the model (surface boundary conditions and process parameters fixed, only initial state subject to optimisation), the model appears to require an initial ice thickness well above that observed by CryoSat in large parts of the central Arctic. However, the reconstruction of the initial fields allows to reduce the misfit of the ice concentration in September to a large extent. We will use this in the next experiments.
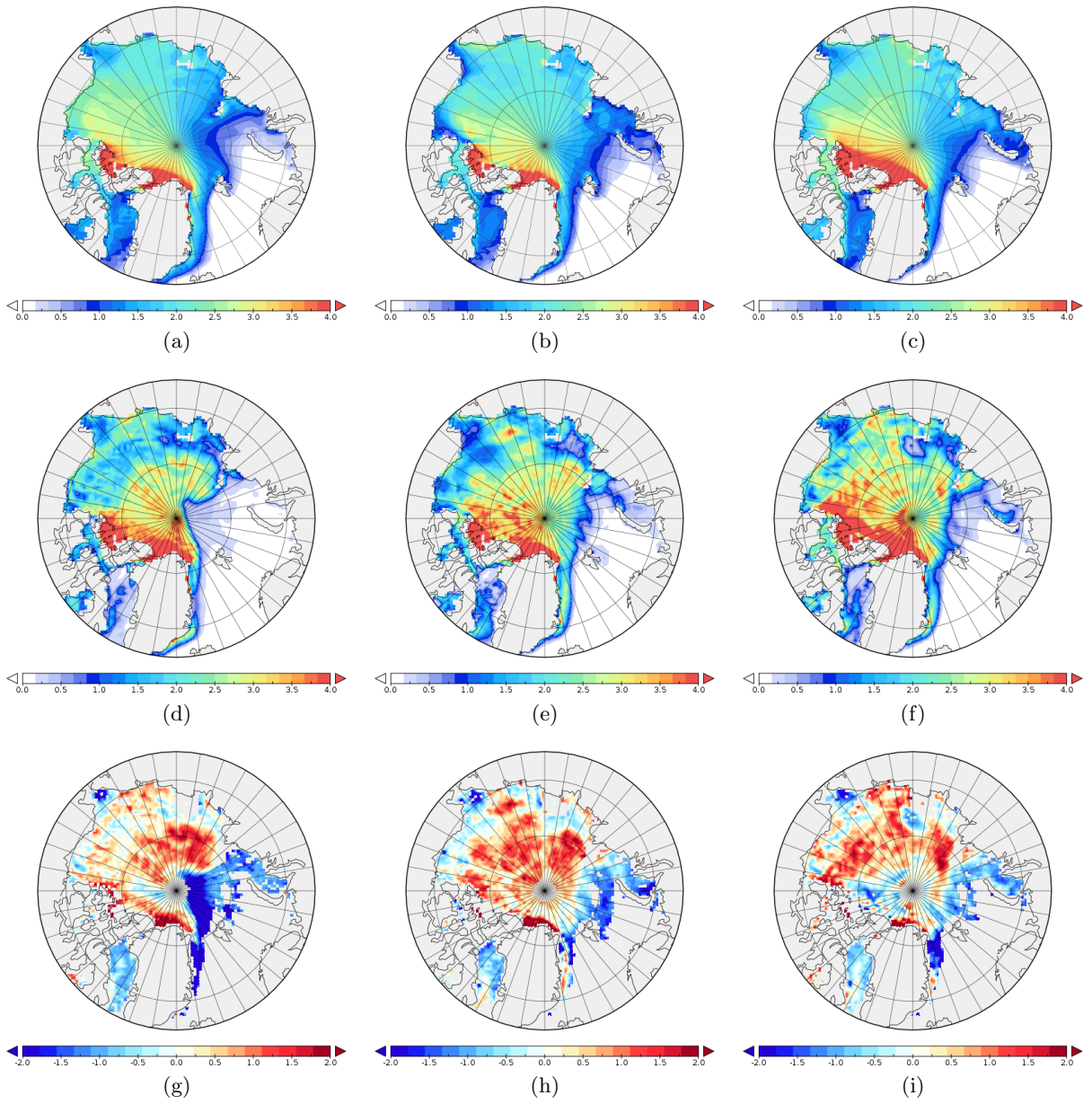
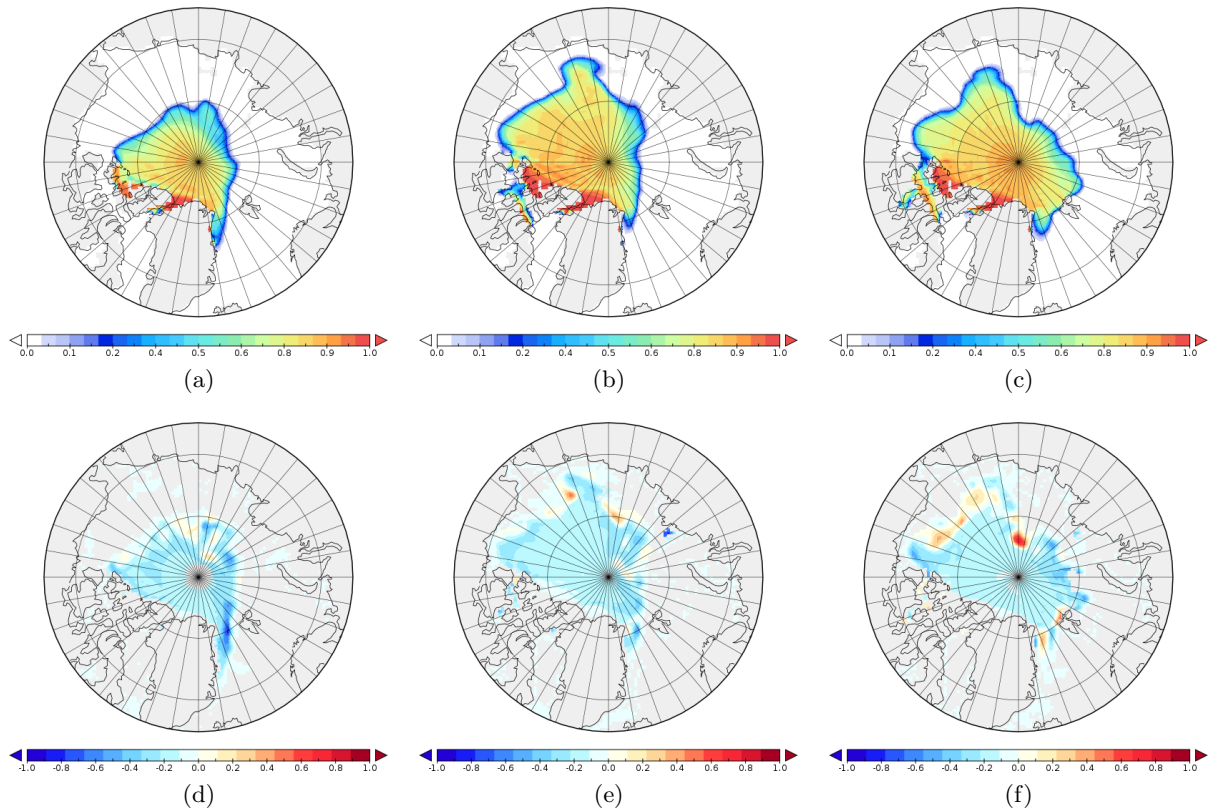Figure 33: As Figure 25 but for the reconstruction experiment.

Figure 34: The ice concentration [0-1] after optimisation (top row) and the misfit to the OSISAF ice concentration for September 2012 (a) and d)), September 2013 (b) and e)), and September 2014 (c) and f)).

5.3 Assimilation with bias-corrected CryoSat-2 ice thickness

The ratio of the reconstructed and the CryoSat-2 ice thickness shows remarkably large similarities between the three years (shown in Figure 35 for March). We used this to apply a bias correction. First we averaged the ratio fields over the three years 2012 to 2014 (shown in Figure 36 for March and April). Second we multiplied the CryoSat-2 ice thickness for March and April by the corresponding ratio fields yielding a bias-corrected CryoSat-2 ice thickness. Then we performed an assimilation similar to that described in section 5.1 except that we apply the bias correction procedure to the CryoSat-2 ice thickness data. The uncertainties are retained.
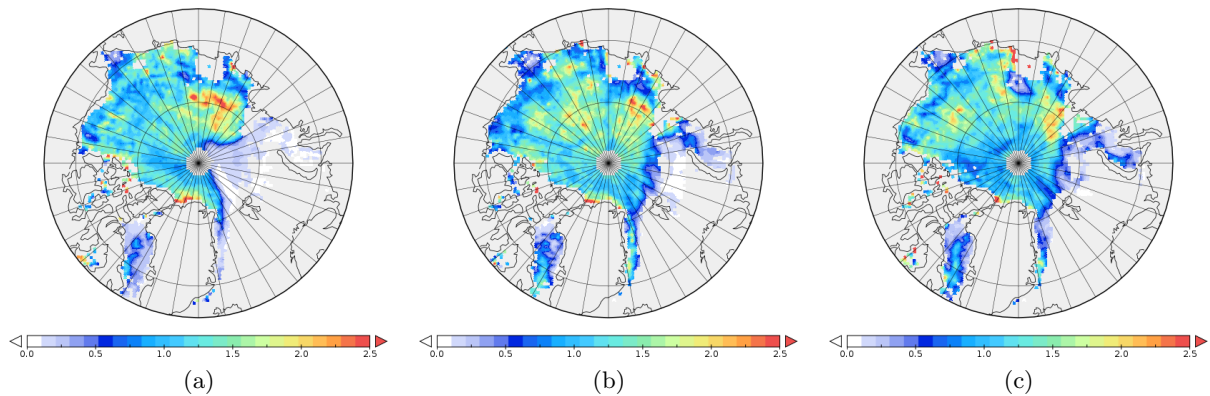

Figure 35: The ratio of the reconstructed and the CryoSat-2 ice thickness a) March 2012, b) March 2013, and c) March September 2014.
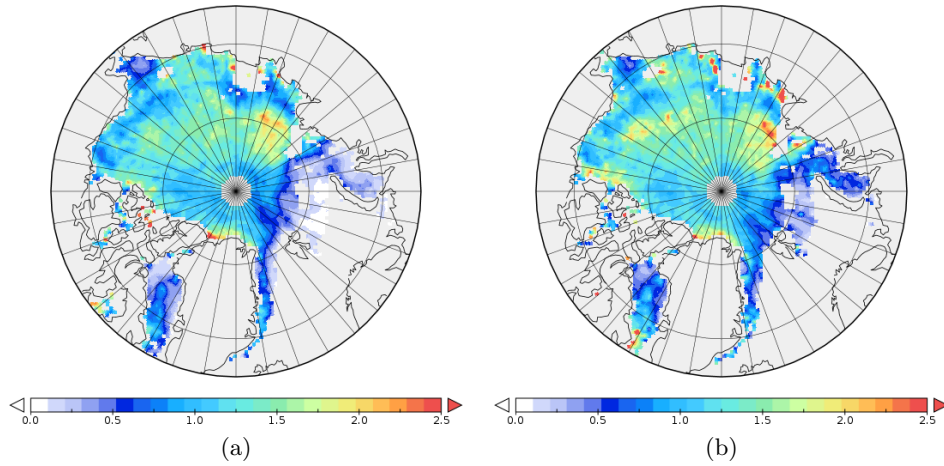
Figure 36: The mean ratio for a) March and b) April.

The evolution of the cost function and the norm of the gradient of the bias-corrected experiment is shown in Figure 37. Figure 38 depicts the total cost function and all five terms of it separately before (a priori) and after the last iteration (a posteriori) of the assimilations for each year. In general, the bias-corrected assimilation performs very similar to the original assimilation. But the analysed monthly cost functions from Mar to the April of the following years show remarkable improvements (Figure 39) for some variables. While the cost function with respect to the ice thickness shows of course a large increase, the cost function of the ice concentration shows a large reduction over almost the whole time. Only in October no improvement is found. The cost function of the SST is now reduced from March until July.
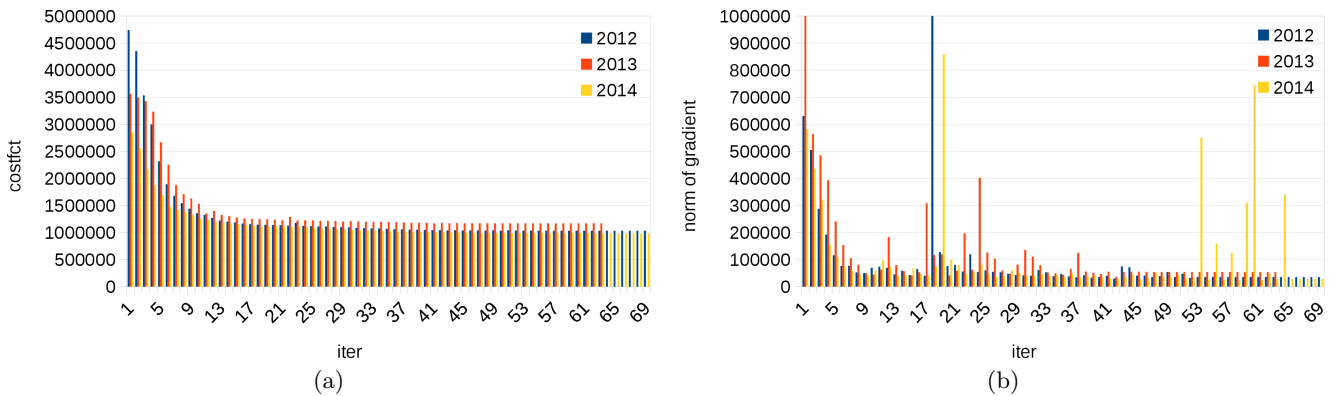

Figure 37: As Figure 23 but for the bias-corrected assimilation experiment.
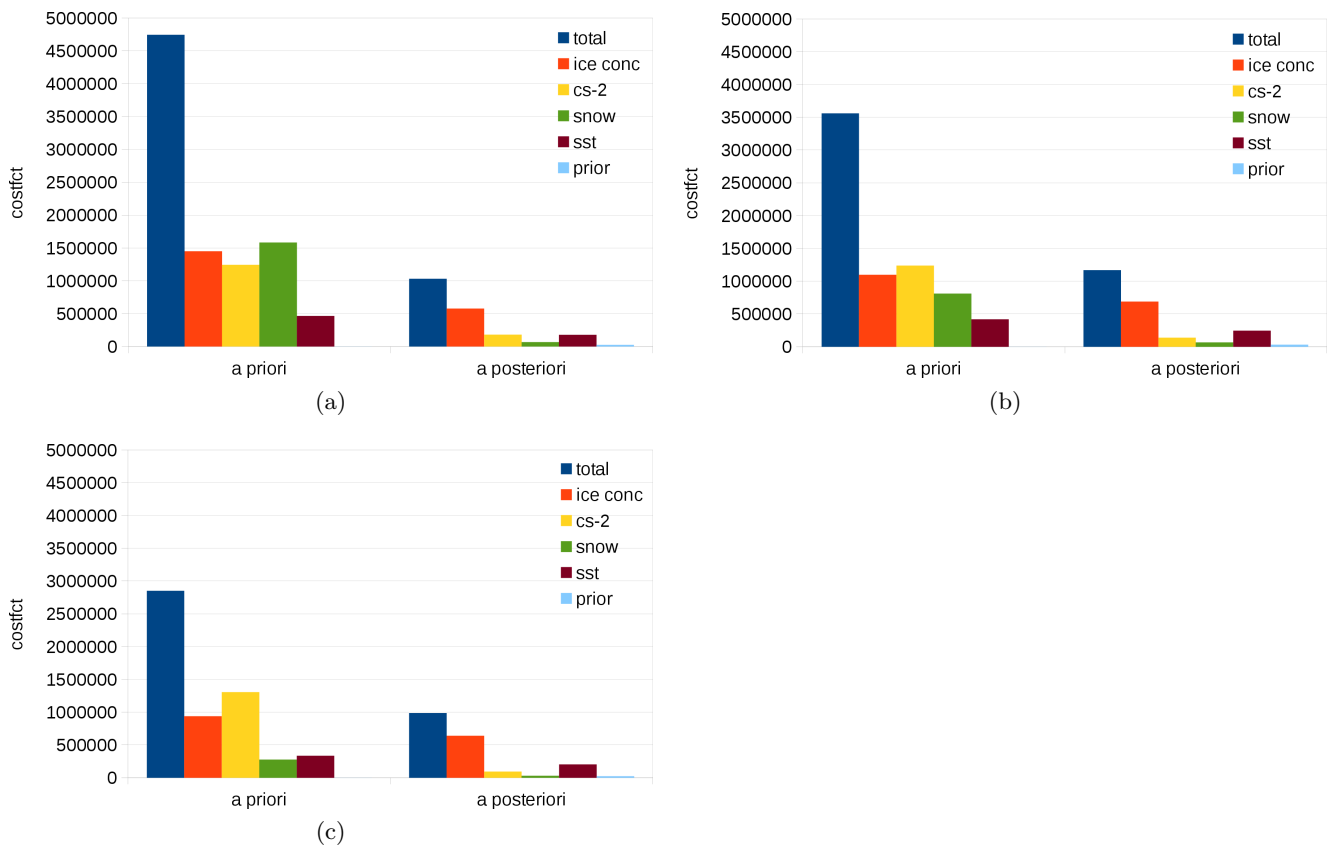
(a)

(b)

(c)

Figure 38: As Figure 24 but for the bias-corrected assimilation experiment.
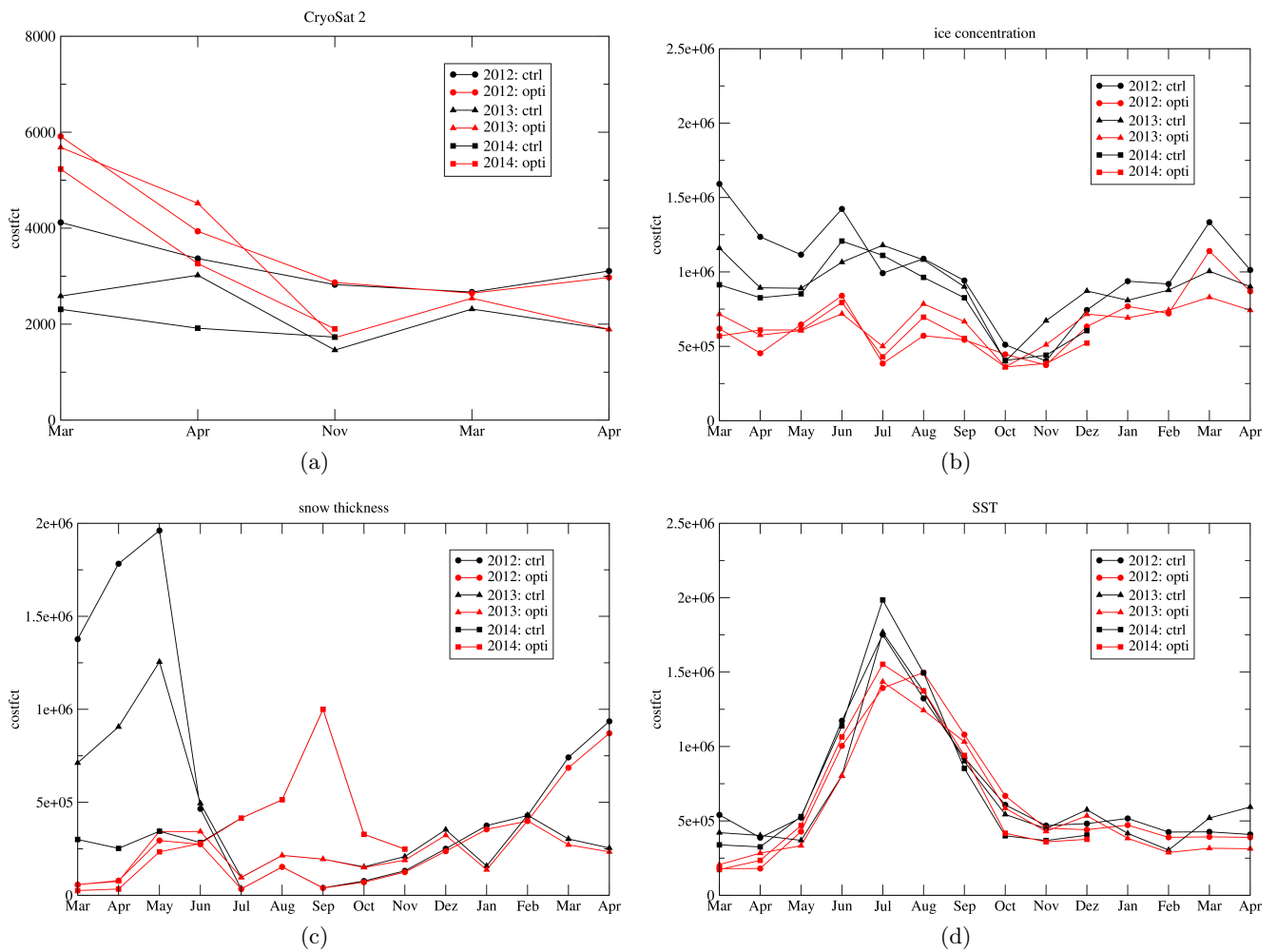
Figure 39: As Figure 25 but for the bias-corrected assimilation experiment.

The misfit of the ice concentration in September is now strongly reduced for all years (Figure 40, cmp. Figure 29). By applying the bias-correction to the CryoSat-2 ice thickness we were able to increase the predictive skill of the seasonal forecast considerably. Because all data streams used here will be available also in May 2015 we will apply the assimilation system to optimise the initial conditions for the Sea Ice Outlook 2015. Additionally we will correct the September ice concentration by scaling it by a factor of 1.2 to compensate for the mean bias. For the years 2012 to 2014 this results in the ice concentration fields depicted in Figure 41.

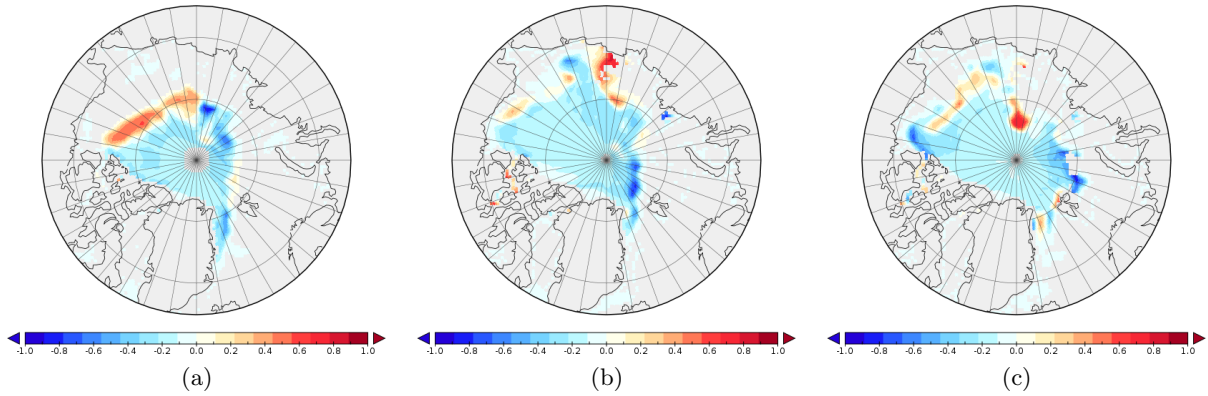Figure 40: The misfit of the ice concentration prior to (top row) and after the optimisation (bottom row) and the OSISAF ice concentration for September 2012 (a) and d)), September 2013 (b) and e)), and September 2014 (c) and f)).
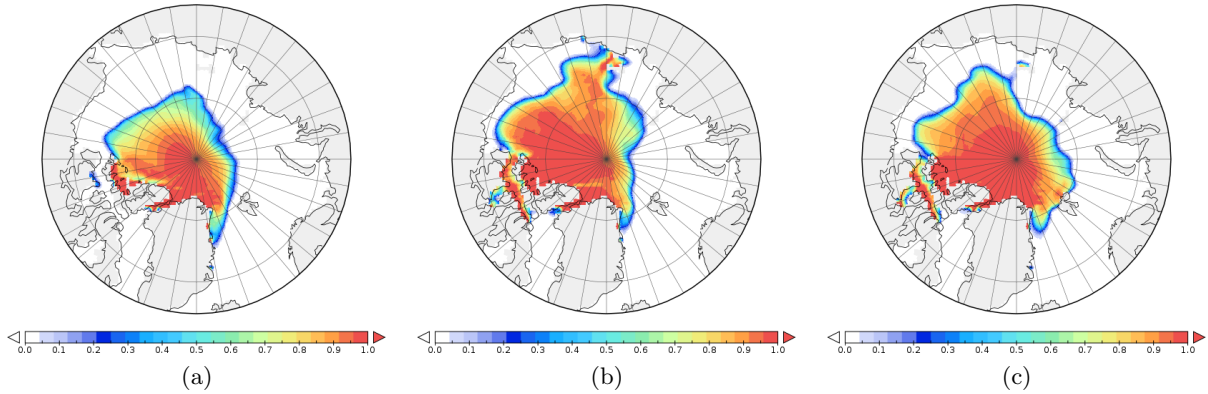

Figure 41: The ice concentration of the proposed forecasting system for September 2012 (a) and d)), September 2013 (b) and e)), and September 2014 (c) and f)).

6 Discussion and summary

NAOSIM has been newly calibrated using observations from 1990 to 2008. By contrast to the standard NAOSIM version (stdNAOSIM) we call the calibrated version newNAOSIM. We restricted the calibration to parameters which control the sea-ice dynamics (but also the ocean dynamics) resulting in a horizontal ice thickness distribution much closer to the ICESat-JPL observations. A positive bias in the Beaufort Sea was strongly reduced as well a negative bias over the Eurasian Basin slope. This is connected to a reduction of the ice drift speeds which are now much closer to the ice drift provided by OSISAF. It has been shown that the horizontal ice thickness distribution for single events like the September 2007 sea-ice minimum is also improved strongly. Also the time series of September sea-ice extent and area follow now much closer the observations. While stdNAOSIM produces a minimum in extent and area in 1990 which almost reaches the 2007 event, newNAOSIM's 1990 simulation is much more realistic. This underlines the importance of a realistic horizontal ice thickness distribution to simulate were added as part of this study. For the OSISAF sea-ice concentration we calculated uncertainties as proposed by the ESACCI project which is at the moment certainly the best approach. The snow depth of the University of Bremen is only validated on level ice. We used a very large uncertainty on multi-year ice during the assimilation. For the OSISAF SST we applied a temporally and spatially fixed uncertainty.

It is well-known that sea-ice thickness information is essential for seasonal predictions because, on this time-scale, the initial conditions of the upper ocean layer and of the sea-ice thickness have the strongest

impact on the sea-ice prediction. We paid special attention to the CryoSat-2 data which are currently the only source of Arctic-wide sea-ice thickness estimates for ice thicker than about $1m$. Among many other sources of uncertainty in the CryoSat-2 retrieval algorithm the selection of the retracker threshold is essential. We demonstrated that the 50% retracker threshold thicknesses are most consistent with two sea ice-ocean models, newNAOSIM and PIOMAS2.1. For the data assimilation, the estimation of the uncertainties was guided by the uncertainty of the determination of the retracker threshold. This is certainly not the only source of uncertainty but the largest.

Three assimilation experiments with all four data sets (including CryoSat-2 ice thickness data) over an assimilation window from March and April, for each of the years 2012 to 2014, were carried out and a forecast of the summer ice conditions was performed. To focus on the effect of constraining the initial state of the sea ice-ocean system, we assumed to have perfect seasonal atmospheric forecast providing perfect surface boundary conditions. It turned out that the assimilation could only improve the summer conditions for some regions. Arctic-wide the forecast in summer could not be improved through the use of the sea-ice and ocean observations in March and April.

To investigate whether the model is at all capable of simulating a substantially improved summer ice concentration, we conducted a second set of experiments with the following modifications with respect to the first set of experiments: We extended the assimilation window to the 7 month period from March to September, we excluded the CryoSat-2 ice thickness from the data steams to be assimilated and restricted the use of ice concentration to the period from July to September. Indeed, each of these three experiments was able to determine an initial state that yields a strong improvement in the simulated summer sea-ice conditions. However, the model still underestimates the ice concentration by about 0.1 overall. We attribute this to model deficits in the governing partial differential equation for the ice concentration under melting conditions. Also, the initial ice thickness is in large areas much thicker than the CryoSat-2 ice thickness. This does not necessarily mean that the CryoSat-2 ice thickness is biased but could point also to an overestimated seasonality in the model. This in turn could be connected to the systematic underestimation of the summer ice concentration in the model - if the ice concentration decreases too strongly in the melting season too much energy is captured by the ocean surface which increases the basal melting of the sea-ice (Ice-albedo-feedback). The improvement of the thermodynamics of the model will be a future research topic and will hopefully achieve a simulated ice thickness closer to the one observed by CryoSat-2.

Since the second set of experiments made use of the summer ice conditions, we called the inferred posterior ice thickness fields 'reconstructed'. The ratio of this 'reconstructed' and the CryoSat-2 ice thickness fields for March and April is very similar for all three years. This allows us to develop a bias correction scheme, which scales the CryoSat-2 ice thickness fields by the monthly three year average of the above ratio. Then we performed a set of assimilation experiments for March and April similar to the first set of assimilations but used the above bias correction scheme for the CryoSat-2 ice thickness. This yields a considerable improvement in forecast skill for sea-ice from July to September for all three years. We note that our prediction target, namely the summer ice conditions of 2012 to 2014, have entered the assimilation procedure that was used to derive the ice thickness ratio in our bias correction scheme. However, the bias correction scheme can now also be applied to years outside the period from 2012 to 2014. One of these applications will be the Sea Ice Outlook 2015.

7   Acknowledgements

References

Castro-Morales, K., F. Kauker, M. Losch, S. Hendricks, K. Riemann-Campe, and R. Gerdes, Sensitivity of simulated arctic sea ice to realistic ice thickness distributions and snow parameterizations, *Journal*

of Geophysical Research: Oceans, *119*, 559–571, 2014.

Day, J., E. Hawkins, and S. Rietsche, Will arctic sea ice thickness initialization improve seasonal forecast skill?, *Geophys. Res. Let.*, 2014.

Kalnay, E., et al., The ncep/ncar 40-year reanalysis project, *Bulletin of the American meteorological Society*, *77*, 437–471, 1996.

Kauker, F., R. Gerdes, M. Karcher, C. Köberle, and J. Lieser, Variability of Arctic and North Atlantic sea ice: A combined analysis of model results and observations from 1978 to 2001, *Journal of Geophysical Research Oceans*, *108*, 13–1, 2003.

Kauker, F., T. Kaminski, M. Karcher, R. Giering, R. Gerdes, and M. Voßbeck, Adjoint analysis of the 2007 all time arctic sea-ice minimum, *J. Geophys. R.*, 2009.

Kauker, F., R. Gerdes, M. Karcher, T. Kaminski, R. Giering, and M. Voßbeck, June 2010 Sea Ice Outlook - AWI/FastOpt/OASys, Sea Ice Outlook web page, 2010.

Kern, S., K. Khvorostovsky, H. Skourup, E. Rinne, Z. S. Parsakhoo, V. Djepa, P. Wadhams, and S. Sandven, The impact of snow depth, snow density and ice density on sea ice thickness retrieval from satellite radar altimetry: results from the esa-cci sea ice ecv project round robin exercisesnow depth on arctic sea ice, *The Cryosphere*, *9*, 37–52, 2015.

Köberle, C., and R. Gerdes, Mechanisms determining the variability 2003: Mechanisms determining the variability of Arctic sea ice conditions and export, *J. Clim*, *16*, 2843–2858, 2003.

Kurtz, N. T., J. Richter-Menge, S. Farrell, M. Studinger, J. Paden, J. Sonntag, and J. Yungel, Icebridge airborne survey data support arctic sea ice predictions, *EOS, Transactions American Geophysical Union*, *94*, 2013a.

Kurtz, N. T., S. L. Farrell, M. Studinger, N. Galin, J. P. Harbeck, R. Lindsay, V. D. Onana, B. Panzer, and J. G. Sonntag, Sea ice thickness, freeboard, and snow depth products from operation icebridge airborne data, *The Cryosphere*, *7*, 1035–1056, 2013b.

Kwok, R., and G. F. Cunningham, Icesat over arctic sea ice: Estimation of snow depth and ice thickness, *Geophys. Res. Let.*, 2008.

Lindsay, R., M. Wensnahan, A. Schweiger, and J. Zhang, Evaluation of seven different atmospheric reanalysis products in the arctic, *J. Climate*, 2014.

Markus, T., and D. Cavalieri, Snow depth distribution over sea ice in the southern ocean from satellite passive microwave data, in *Antarctic Sea Ice: Physical Processes, Interactions and Variability*, edited by M. Jeffries, no. 74 in Antarctic Research Series, pp. 19–39, American Geophysical Union, AGU, Washington, D.C., 1998.

Ricker, R., S. Hendricks, V. Helm, H. Skourup, and M. Davidson, Sensitivity of cryosat-2 arctic sea-ice freeboard and thickness on radar-waveform interpretation, *The Cryosphere*, *8*, 1607–1622, 2014.

Saha, S., and Co-authors, The ncep climate forecast system reanalysis, *Bull. Amer. Meteor. Soc.*, *91*, 10151057, 2010.

Saha, S., and Co-authors, The ncep climate forecast system version 2, *J. Climate*, *27*, 21852208, 2014.

Schweiger, A., D. Lindsay, J. Zhang, M. Steele, and H. Stern, Uncertainty in modeled arctic sea ice volume, *J. Geophys. R.*, 2011.

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

Sumata, H., F. Kauker, R. Gerdes, C. Köberle, and M. Karcher, A comparison between gradient descent and stochastic approaches for parameter optimization of a sea ice model, *Ocean Sci.*, *9*, 609–630, 2013.

Sumata, H., T. L. F. Girard-Ardhuin, N. Kimura, M. Tschudi, F. Kauker, M. Karcher, and R. Gerdes, An intercomparison of arctic ice drift products to deduce uncertainty estimates, *Journal of Geophysical Research: Oceans*, *119*, 4887–4921, 2014.

Tian-Kunze, X., L. K. N. Maaß, M. Mäkynen, N. Serra, M. Drusch, and T. Krumpen, Smos-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, *The Cryosphere*, *8*, 997–1018, 2014.

Tietsche, S., J. Day, V. Guemas, W. Hurlin, S. Keeley, D. Matel, R. Msadek, M. Collins, and E. Hawkins, Seasonal to interannual arctic sea ice predictability in current global climate modules, *Geophys. Res. Let.*, *41*, 2014.

Warren, S., I. Rigor, N. Untersteiner, V. Radionov, N. Bryaz-gin, Y. Aleksandrov, and R. Colony, Snow depth on arctic sea ice, *J. Climate*, *12*, 18141829, 1999.

Yang, Q., S. Losa, M. Losch, X. Tian-Kunze, L. Nerger, J. Liu, L. Kaleschke, and Z. Zhang, Assimilating smos sea ice thickness into a coupled ice-ocean model using a local seik filter, *Journal of Geophysical Research: Oceans*, *119*, 6680–6692, 2014.

Zhang, J., and D. A. Rothrock, Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates, *Mon. Wea. Rev.*, 2003.

Zwally, H., D. Yi, R. Kwok, and Y. Zhao, Icesat measurements of sea ice freeboard and estimates of sea ice thickness in the weddell sea, *J. Geophys. R.*, *113*, 2008.