RESEARCH ARTICLE

Correspondence to:
T. A. S. Rasmussen,
tar@dmi.dk

Impact of Assimilation of Sea-Ice Surface Temperatures on a Coupled Ocean and Sea-Ice Model

Till A. S. Rasmussen[1] , **Jacob L. Høyer[1]**, **Darren Ghent[2]** , **Claire E. Bulgin[3]** , **Gorm Dybkjær[1]**, **Mads H. Ribergaard[1], Pia Nielsen-Englyst[1], and Kristine S. Madsen[1]**

[1]Danish Meteorological Institute, Copenhagen East, Denmark, [2]University of Leicester, Leicester, UK, [3]Department of Meteorology, University of Reading, Reading, UK

Abstract We establish a methodology for assimilating satellite observations of ice surface temperature (IST) into a coupled ocean and sea-ice model. The method corrects the 2 m air temperature based on the difference between the modeled and the observed IST. Thus the correction includes biases in the surface forcing and the ability of the model to convert incoming parameters at the surface to a net heat flux. A multisensor, daily, gap-free surface temperature analysis has been constructed over the Arctic region. This study revealed challenges estimating the ground truth based on buoys measuring IST, as the quality of the measurement varied from buoy to buoy. With these precautions we find a cold temperature bias in the remotely sensed data, and a warm bias in the modeled data relative to ice mounted buoy temperatures, prior to assimilation. As a consequence, this study weighted the modeled IST and the observed IST equally in the correction. The impact of IST was determined for experiments with and without the assimilation of IST and sea-ice concentration. We find that assimilation of remotely sensed data results in a cooling of IST, which improves the timing of the snow melt onset. The improved snow cover in spring is only based on observations from one buoy, thus additional good quality observations could strengthen the conclusions. The ice cover and the sea-ice thickness are increased, primarily in the experiment without sea-ice concentration assimilation.

1. Introduction

Arctic sea-ice and snow surface temperatures (IST) are of interest for both short term forecasts and in evaluation of the impact of long term climate trends on the ice-atmosphere boundary (Hansen et al., 2010; Wang & Key, 2003). This study focuses on the potential for assimilating Arctic IST from a remotely sensed product into a coupled ocean and sea-ice model. A short term forecast ($\sim days$) is an initial value problem. Improvements to the initial state of the model will therefore impact the skill of the model over this timescale. Assimilation of surface temperatures has been widely used to improve the boundary temperature over ocean (sea surface temperature [SST]) in atmospheric and coupled models (While & Martin, 2013) and over land (land surface temperature) in atmospheric models (McNider et al., 1994). Satellite IST is a relatively novel product and few experiments have been carried out that include satellite observations of IST in modeling systems (e.g., Merchant et al., 2013).

IST governs the upper boundary of the sea-ice thermodynamics and is important for the heat flux into the sea ice and snow (Sturm et al., 2001, 2002), and as a consequence of this the growth and decay of these layers. The lower sea-ice boundary is governed by the ocean, which in most cases is at or near the freezing point. Arctic in situ observations are sparse and they are associated with relatively large uncertainties (Dybkjær et al., 2012; Rigor et al., 2000), thus it is hard to establish a good data set for verification and validation. Automatic weather stations are the most accurate type of in situ observation; however, the spatial coverage is limited and if placed on sea ice they are vulnerable to changing seasons and ice drift. The Danish Meteorological Institute (DMI) has deployed a weather station on fast ice in a fjord near Qaanaaq in the North Western part of Greenland for 3 winter periods (2014–2016) extending from approximately January to May. These observations show a large temporal variability in the temperature 1 m above the snow or ice surface (see, e.g., Høyer et al., 2017). The large variability is also found in the snow surface skin temperature but it is quickly reduced when entering the snow and ice. This reduction in temperature variability occurs mainly

due to the insulating effect caused by a snow layer on sea ice, which is a well-known behavior for snow covered surfaces (Schneebeli & Sokratov, 2004; Sturm et al., 1997).

One type of in situ observation of IST is made by ice mass balance buoys (IMBs), that use thermistor strings to measure temperature profiles through the snow and sea ice (Jackson et al., 2013). The IMBs are used as a validation tool as they follow the ice pack and provide observations scattered in time and space. One of the main challenges with the IMBs is that it can be hard to determine the interfaces between air, snow, ice, and water accurately. Large biases can occur if the observation is taken at the wrong vertical position, as the temperature can vary significantly with depth, particularly in the snow, but also in the sea ice.

Observations of IST can also be made by remote sensing. Satellites provide an opportunity for unique spatial coverage of sea-ice regions with several overpasses every day (see, e.g., Dybkjær et al., 2012; Hall et al., 2004; Key & Haefliger, 1992). However, the infrared satellite observations are limited by cloud cover, which can be persistent in the Arctic. The combination of large diurnal temperature variations and irregular sampling intervals can cause averaged satellite observations to differ considerably from in situ point measurements and modeled output of instantaneous temperatures. In addition, missing clouds in the cloud mask may cause a cold bias in the IST product (Dybkjær et al., 2012). Remotely sensed observations of ocean and land surface temperatures have been made for years, but sea-ice and snow surface temperature products from satellites are less mature.

This study assesses the potential for integration of satellite IST observations into forced coupled ocean and sea-ice models. We present two new advancements: (1) the generation of a multisatellite gap-free SST and IST analysis product and (2) the inclusion of the satellite IST observations in a coupled ocean and ice modeling system with an assessment of the impact. The satellite observations are used for bias correction of the surface forcing applied from an atmospheric forcing field. This approach requires an a priori examination of the modeled IST error statistics obtained from a model simulation without IST assimilation.

The paper consists of six sections. Section 2 describes the satellite observations, section 3 the model setup, section 4 the method applied to implement the remotely sensed temperatures in the model, and in section 5 the results and discussion. Finally, the conclusions can be found in section 6.

2. Satellite Observation

To generate a consistent satellite data set that covers the ice covered areas including the Marginal Ice Zone (MIZ) and the open ocean, several satellite SST, and IST data sets are included in the analysis. The satellite data are listed in Table 1 with an indication of spatial resolution and the surface type of observations included in the data product.

The MODIS satellite products are obtained from the ESA GlobTemperature project, which utilizes the MOD29/MYD29 1 km IST only products available from NASA (Riggs et al., 2006) and reformats them into a common format for all GlobTemperature products. The Metop IASI satellite products have been obtained from EUMETSAT version 6 of the IASI processor (T. August, personal communication, 2016) using Piece-Wise Linear Regression retrievals. The product is the first guess estimate from the algorithm and provides all sky IST observations. A reprocessed and improved product is now available from EUMETSAT upon request,

Table 1
Overview of the Level 2 (L2) Satellite Surface Temperature Products Considered

Sensor	Obs type	Resolution (km)	Provider
Metop AVHRR	SST, MIZT, IST	1	EUMETSAT OSI-SAF
MODIS Aqua	IST	1	NASA + GlobTemperature
MODIS Terra	IST	1	NASA + GlobTemperature
Metop IASI	IST	12–40	EUMETSAT
NOAA AVHRR	SST	4	NAVOCEANO
NOAA AVHRR	SST	2	EUMETSAT OSI-SAF

Note. Obs Type is the surface measured. SST is the sea surface temperature, MIZT is the marginal ice zone temperature, and IST is the ice surface temperature. Metop IASI IST is the only product that observes surface temperatures under cloudy conditions. This product is not included in the final product.

using an additional Optimal Estimation algorithm that retrieves clear-sky observations. The Metop AVHRR product contains SST, IST, and Marginal Ice Zone Temperature (MIZT) observations and is a preoperational version of EUMETSAT OSI-205 product, described in Dybkjær et al. (2017). The NOAA AVHRR SST product from the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI-SAF) is the North Atlantic regional product, OSI-202-b. The NAVOCEANO SST product is produced by the U.S. Naval Oceanographic Office (NAVOCEANO) from NOAA AVHRR observations (May et al., 1998). This SST product has been shown to be one of the most accurate SST products in the Arctic and has been used to reference other SST products (Høyer et al., 2012, 2014). All satellite observations are level 2 (L2) observations and are limited by cloudy conditions, except for the IASI product, which uses Microwave observations to infer the surface temperature under cloudy conditions. The use of several satellite products increases the data coverage and the sampling throughout the day which is particularly beneficial over sea ice to reduce the sampling effects. Note that the Metop AVHRR product is the only one that covers both open ocean and sea ice.

Separate level 3 (L3) satellite fields were generated by aggregating the L2 orbit observations onto a fixed spatial grid with a resolution of 0.05°. One L3 field was generated for each satellite product in a preprocessing step to the level 4 (L4) Optimal Interpolation (OI). The daily L3 fields were obtained by aggregating the L2 satellite products within 36 h from the analysis time. The 72 h aggregation window was chosen from analysis of hourly in situ observations and taking into account the sampling characteristics associated with clear-sky polar orbiting satellite observations. The L3 aggregation also included quality control on the different products. Quality levels 4 or 5 were accepted for the SST products (OSI-SAF NAR and NAVOCEANO) in the Group for High Resolution SST (GHRSST) L2P format. This corresponds to observations being classified as good or excellent by the satellite producer (see, e.g., Donlon et al. (2007) for definitions of GHRSST quality levels). Only MODIS observations with a good sea-ice quality flag were allowed. After communication with EUMETSAT, it was decided to discard IASI IST observations with a quality indicator above 3 (not related to the GHRSST definition). The Metop AVHRR SST and IST observations classified as cloud free by the Polar Platform Systems (PPS) cloud mask were included in the processing (Dybbroe et al., 2006).

A daily-varying surface mask was generated from a combination of the static land/sea mask and the daily Sea Ice Concentration (SIC) from OSI-SAF (Eastwood et al., 2011) to further assist in the processing, merging, and interpolation. The surface mask identifies areas with land, sea ice (SIC > 70%), MIZ (30% < SIC ≤ 70%) and open water (SIC ≤ 30%). These ice concentration limits are used within the OSI-SAF to distinguish between open water, open ice, and closed ice. In the L3 Metop AVHRR aggregation, all satellite observations using the SST algorithm over sea-ice and all ice algorithm observations over open ocean were discarded. See Dybkjær et al. (2012) for the different algorithm types. For all the other satellite products, observations warmer than 3°C over sea ice and colder than −4°C in open ocean were discarded.

Several studies have reported the uncertainties of the satellite SST products to be around 0.5°C when compared against in situ observations from drifting buoys (Castro et al., 2016; Høyer et al., 2012). Similarly, the performance of the satellite IST products has been assessed in several studies (Dybkjær et al., 2012; Hall et al., 2004; Key et al., 1994), but the lack of good reference in situ data, as discussed above, makes these IST validation results less reliable. For a first check of the consistency, we therefore compared the different L3 satellite products over the sea ice. Averaged time series were extracted for several areas throughout the Arctic, and an example is shown in Figure 1 from the Lincoln Sea. The figure shows the variations in IST with time scales of several days, as expected from the aggregation window of 72 h. Maximum temperatures are reached in July and August and minimum temperatures in February. In general, all of the satellite products agree well on the temporal variability induced by weather events, but the IASI product displays a significantly higher absolute temperature and lower variability compared to the purely infrared 1 km products from Metop AVHRR and MODIS. Note that some of the cold events seen in the infrared products could be a result of cloud contamination. An offset is evident during all seasons, and it occasionally exceeds 5°C. This offset may originate from the microwave radiation algorithm that is enabled when one of the three cloud tests are positive, because the signal originates from below the snow surface (Tonboe et al., 2011). The difference might also be due to the clear-sky effect present in the purely infrared products. However, analysis of in situ observations of IST in the Arctic revealed that the average temperature difference between cloudy and clear-sky cases for 72 h averages is about −0.5°C (clear-sky being colder than all sky). The clear-sky effect can thus not completely explain the observed difference between the IASI and the purely infrared satellite products.
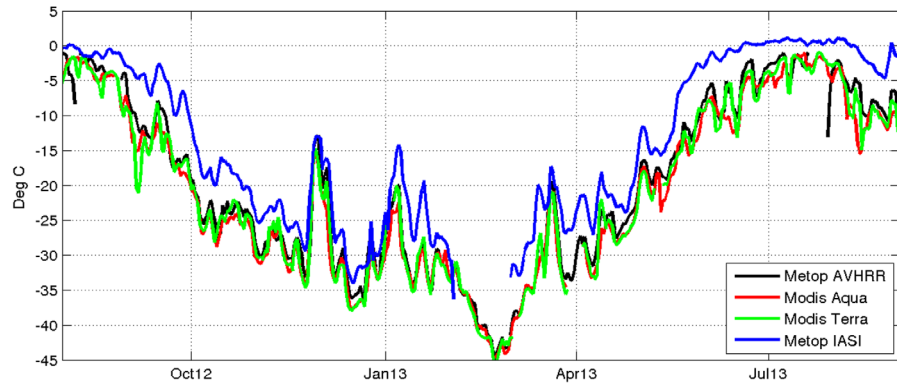
Figure 1. The different L3 satellite IST products for a region in the Lincoln Sea, North of Greenland (84–87°N, 58–62°W). See Table 1 for the different satellite products and their characteristics.

The OI method used to generate the L4 product assumes that the data are anomalies from a first guess field and do not contain systematic biases. This is most likely not fulfilled for the IASI observations and it was therefore decided not to include these data in the L4 generation, as it would violate the interpolation assumptions and probably introduce inconsistencies in the L4 data set.

2.1. Construction of the Level 4 Product

The final L4 analysis product is a merged and interpolated daily field with a 0.05° resolution in latitude and longitude that covers surface temperatures of the ocean, sea ice and MIZ. It was generated from 1 August 2012 to 30 September 2013. The OI method used to construct the merged and gap-free SST/IST analysis is taken from the high latitude SST DMI processing scheme described in Høyer and She (2007) and Høyer et al. (2014). As mentioned above, the OI estimation assumes we are working with anomalies on the ith observation, f_i', with associated errors, ϵ_i. The estimation at the interpolation point, f_0', can be described as

$$f_0' = \sum_{i=1}^n (f_i' + \epsilon_i)p_i + I_0 \tag{1}$$

where p_i are the weights that have to be determined, I_0 is the interpolation error, and n is the number of observations. The primes indicate anomalies from a first guess value. The mean square interpolation error is defined as:

$$\overline{I_0^2} = \overline{\left(\sum_{i=1}^n (f_i' + \epsilon_i)p_i - f_0' \right)^2} \tag{2}$$

The optimal weights that result in the minimum interpolation error is found by finding the minimum of equation (2). This is done by setting the derivatives of the error with respect to the weights equal to zero and solving the equation. If it is assumed that the noise in one observation is not correlated with the noise in the other observations, with the true function, f, we can derive the central equations for the estimation problem

$$\sum_{j=1}^n \overline{f_i' f_j'} p_j + \sigma_{\epsilon i}^2 p_i = \overline{f_i' f_0'} \qquad (i = 1, 2, \ldots n) \tag{3}$$

where p_i denotes the optimal weights that are solved for and $\sigma_{\epsilon i}^2$ is the error variance of the ith observation. Equation (3) shows that the optimal estimation problem involves covariance between the individual observations $\overline{f_i' f_j'}$ and between the observations and the estimation point $\overline{f_i' f_0'}$. These covariances, C_{ij}, have been calculated prior to the prediction and assumed to be constant in time. It is assumed they take the form of

$$C_{ij} = \exp(-\lambda * dist_{ij}^\gamma) \tag{4}$$

where $dist_{ij}$ is the distance in km between point i and j. The satellite observations were used to derive spatially varying parameters, λ and γ, for the open ocean regions and spatially uniform values for the sea-ice
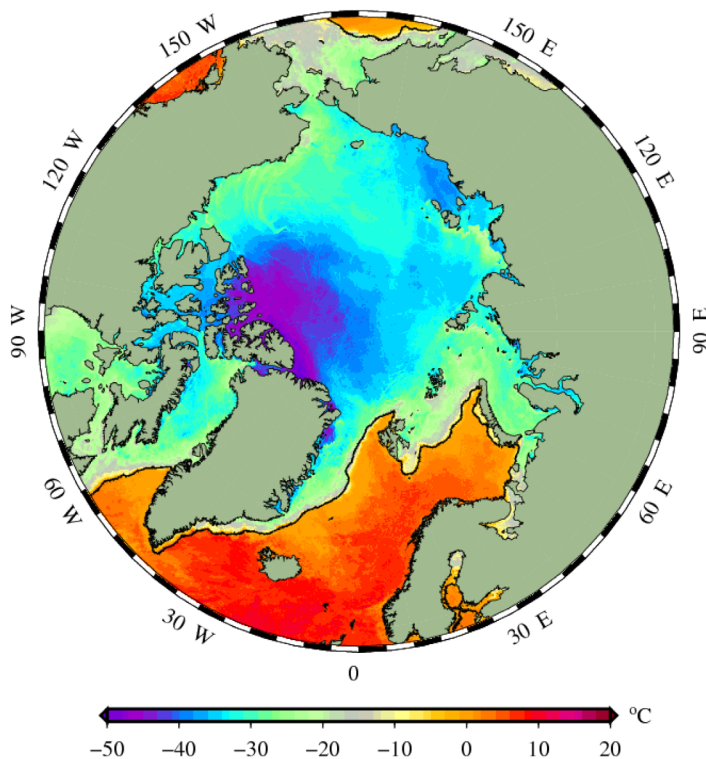
covered regions. On average, γ was set to 0.62 for the open ocean and 1 for the ice covered regions. Correspondingly, λ is 0.19 $(km^{\gamma})^{-1}$ for open ocean and 0.01 $(km^{\gamma})^{-1}$ for ice covered regions. The difference between the ocean and ice covariance parameters reflects the different physical conditions, where the ocean processes occur on significantly smaller spatial scales than the atmospheric processes over sea ice.

In the current implementation, we use the previous day's analysis as first guess field. The SST and IST observations from the current day are therefore interpreted as anomalies with respect to the field from the previous day. The search radius for the OI method is set to 100 km and the maximum number of satellite observations included in the optimal estimation is 16, where up to four observations having the highest correlation with the estimation point are selected for each quadrant. The average number of satellite observations included in the analysis is about 15.9, indicating very high satellite data coverage within the search limits. Due to the different physical conditions for ocean and sea-ice surface temperature variability, separate guess variance statistics have been derived for the open ocean and the sea-ice covered regions. The SST first guess variance is similar to what was presented in Høyer and She (2007) and Høyer et al. (2014). Over sea ice, the first guess IST variability was derived using 1 year of Metop AVHRR L3 aggregated observations. The previous day L4 field was subtracted from the L3 IST observations and a spatial two-dimensional field of standard deviations were calculated for 1 year of anomalies.

Figure 2. Example of a L4 SST/IST analysis field for 1 March 2013. Colorbar indicates the temperature. The black contour line shows the sea-ice concentration of 15%, which separates the IST and the SST.

The first guess field and the error covariance used in the MIZ are obtained from a weighted linear combination of the open water and ice values, where the ice concentration was used as the weighting factor. An example of a daily SST/IST L4 field is shown in Figure 2 for 1 March 2013, which is the time of the year, where the maximum sea-ice extent occurs. To spin up the L4 product for the 1 year generation, the initial field was started from a climatology from 1 August 2012 and run for a month. Subsequently, the 1 September 2012 field was then used as the first guess field for the run from 1 August 2012. This procedure ensured that the first guess field was consistent with the conditions in year 2012 and the satellite products included in the analysis. The temperature differences from the beginning to the end of August were relatively small and only affected the first few days of the runs.

3. Ocean and Sea-Ice Model

Operational state of the art Arctic Ocean and sea-ice models describe the three-dimensional flow of the ocean, the horizontal flow of the sea ice and the thermodynamics of the ocean and the sea ice. In the case of multicategory sea-ice models, they also describe the redistribution of sea ice between the different ice thickness categories. The ocean and sea-ice model system at DMI (Madsen et al., 2016) consists of the HYbrid Coordinate Ocean Model (HYCOM), e.g., (Chassignet et al., 2007) and the Community Ice CodE (CICE (v4.0) (Hunke, 2001; Hunke & Dukowicz, 1997)) coupled with the Earth System modeling Framework (ESMF) coupler (Collins et al., 2004). The horizontal resolution is approximately 10 km in order to ensure an eddy-permitting ocean model and resolve the coastal shelf waters, while balancing computational resources. The domain covers the Arctic Ocean and the Atlantic Ocean to approximately 20°S. The model domain is shown in Figure 3.

The dynamics of the sea-ice model is driven by drag from wind and ocean, surface tilt of the ocean, Coriolis force, and the internal strength of sea ice that will resist movement of the ice pack. The internal strength is based on the Elastic-Viscous-Plastic (EVP) sea-ice rheology (Hunke, 2001), that originates from the Viscous-Plastic (VP) described by Hibler (1979). CICE includes multiple categories of sea ice within each gridded cell
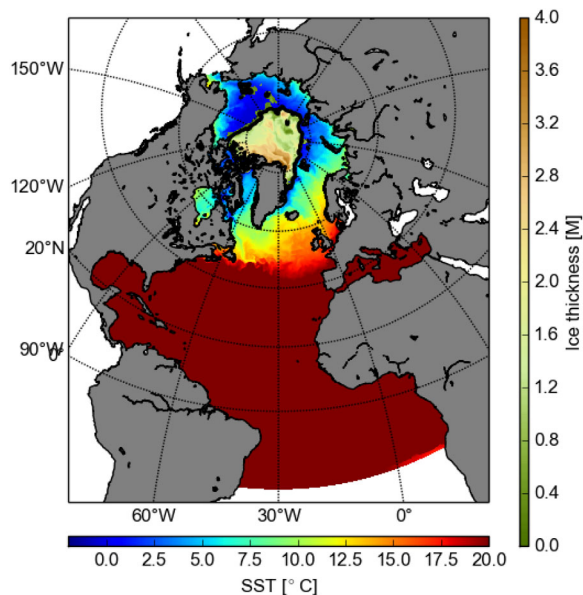
Figure 3. SST and sea-ice thickness for the model domain. Gray is land and white is ocean outside of the domain. The horizontal colorbar indicates the SST in all areas with ice concentration less than 30%. The vertical colorbar describes the ice thickness in areas where the ice concentration is larger than 30%. The black line indicates the 30% ice concentration based on OSI-SAF.

in order to be able to describe the inhomogeneity. The setup for these experiments includes five ice thickness categories defined according to Lipscomb (2001) (see Table 2).

The thermodynamics of CICE prescribes a vertical temperature profile with a resolution of four layers of sea ice and one layer of snow for each sea-ice category (Bitz & Lipscomb, 1999). Snow is very important for the thermodynamics of sea ice as it insulates sea ice from the atmosphere. The lower boundary is governed by the upper ocean temperature which normally is at the freezing point (t_{frz}) as defined in equation (5), where s is the salinity measured in practical salinity unit (psu):

$$t_{frz} = -0.054 \cdot s \qquad (5)$$

The upper boundary is governed by the heat transfer from the atmosphere. The net heat flux is calculated based on the 2 m atmospheric temperature, humidity, incoming long and short wave radiation, and 10 m wind all provided by ERA-Interim (Dee et al., 2011) and the state of the surface of the sea-ice model. These inputs are converted to four components of the net heat flux, namely latent heat, sensible heat, net shortwave radiation, and net longwave radiation. Due to the harsh environment in situ observations needed for validation have limited temporal and spatial resolution. In consequence, the uncertainty of the atmospheric forcing is generally high and improved control of the upper boundary is thus desirable.

The ocean model, HYCOM v2.2.98, explores a hybrid vertical coordinate system, combining isopycnals with z-level coordinates and sigma coordinates. The vertical mixing is defined by the K-Profile Parameterization (KPP) scheme (Large et al., 1994). The DMI setup of HYCOM has 40 vertical levels and the atmospheric forcing originates from ERA-Interim as it is the case for the sea-ice model.

The model is initialized in summer 1997 with a combined climatology, using the Polar Science Center Hydrographic Climatology (PHC; Steele et al., 2001) in the Arctic and World Ocean Atlas 2001 0.25° (Conkright et al., 2002) in the Atlantic, with a 100 km linear transition. The model domain has two open boundaries in the Bering Strait and South Atlantic Ocean, where temperature and salinity are prescribed from the combined climatology, while a constant barotropic component was prescribed at the boundaries. Similarly, tidal forcing is prescribed from Oregon State University TOPEX/Poseidon Global Inverse Solution (TPXO 8.2) (Egbert & Erofeeva, 2002). Tides are induced at the boundary and internally within the ocean model using eight constituents with periods of approximately 12 and 24 h.

The assimilation of sea ice and SST within the model system is based on the nudging scheme described in equation (6):

$$A_{newmod} = (1-\alpha) \cdot A_{oldmod} + \alpha A_{obs} \qquad (6)$$

"A" is the parameter being assimilated and α is a weight function defined by equation (7). This parameter must be in the range between 0 and 1:

Table 2
List of the Five Sea-Ice Categories Used in the Model

Category	Open water (cm)	1 (cm)	2 (cm)	3 (cm)	4 (cm)	5 (cm)
Min thickness	0	0	60	140	240	360
Mean thickness	0	30	100	190	300	
Max thickness	0	60	140	240	360	

Note. Category 5 is not bounded.

$$\alpha = \frac{dt}{E \cdot sday} \tag{7}$$

where dt is the time step of the model, E is the e-folding time in days, which is set to 10 for both SST and sea ice. sday is seconds per day (86,400 s/d).

Observations of sea-ice concentration assimilation are based on OSI-SAF (Eastwood et al., 2011). Only nonflagged OSI-SAF reprocessed sea-ice observations are used, therefore sea ice is not assimilated in a 20–50 km zone near the coast where the satellite observations are affected by land. We also do not assimilate where the modeled and observed sea-ice concentration agree within 10%, to limit observationally induced noise effects. In these areas, the sea-ice concentrations are determined purely by the sea-ice model physics. In order to maintain the same distribution of ice categories the constant, c_{assim} is defined in equation (8)

$$c_{assim} = \frac{A_{newmod}}{A_{oldmod}} \tag{8}$$

For each sea-ice state variable that is corrected (ice concentration, ice thickness, snow thickness, and the snow/sea-ice energy) the new value for each category is described in equation (9). The aim is to keep the ice thickness, snow thickness, and the energy per volume of snow and ice constant:

$$A_{newmod,i} = c_{assim} \cdot A_{oldmod,i} \tag{9}$$

Index i is the category from 1 to 5. This method of estimating the ice thickness is similar to the one labeled as "red" in Smith et al. (2016). Here it was recognized that this method could add too much sea-ice volume to the system in cases where the thinnest ice category was removed for the wrong reasons.

Observed SST used for assimilation originates from the Operational SST and Sea Ice Analysis (OSTIA) system (Donlon et al., 2011) for 1997–2010. From 2011 and onward SST originates from the global DMI_OI Level 4 analysis (Høyer et al., 2014). Surface salinity is relaxed toward climatology with a 30 day relaxation time. More than 100 rivers are included as monthly climatological discharges obtained from the Global Runoff Data Centre (GRDC, http://grdc.bafg.de) and scaled as prescribed by Dai and Trenberth (2002). In addition we have merged the data set using globally gridded Core v2 runoff data (Large & Yeager, 2008) for Greenland, the Canadian Archipelago, Svalbard, and islands within the Arctic Ocean. Finally for the Faroe Islands, data from the hydrological catchment model HYPE (Lindström et al., 2010) are used.

4. Method and Experiments

Three different estimates of IST are used within the experiments and the validation of these. Each source has different characteristics, which are listed in Table 3.

The remotely sensed IST is the temperature of the top few microns of the surface. Due to the large temporal variations of this temperature, direct assimilation requires firm knowledge of the timing of the observations. The modeled IST is primarily determined by input from an atmospheric model and the ability of the model to convert this to a net heat flux. The observational constraints on the atmospheric forcing are limited, thus it is expected that there are biases within the forcing. The approach chosen for this study is therefore to do a bias correction, based on the difference between the observed and the modeled IST, and to apply this to the atmospheric temperatures that are used to force the coupled ocean and sea-ice model.

Table 3
Description of Data Sources Used in These Experiments

Type of data	Model	Buoys	Remote sensed
Frequency (hours)	1	6	24
Valid	Snapshot	Snapshot	72 h average
Comments	Gridded 2-D	lat, lon, and time, point data	Gridded 2-D

Note. The columns describe the update frequency, the duration each update covers and whether it is a point observation or a grid average.

Table 4
Simulations Carried Out and the Data That is Included

Exp	Ice concentration	IST correction
AC-NI	Yes	No
NC-NI	No	No
AC-AI	Yes	Yes
NC-AI	No	Yes

Note. All experiments include SST assimilation over open ocean. Acronyms are: assimilate ice concentration (AC), assimilate IST (AI), no assimilation of ice concentration (NC), and no assimilation of IST (NI).

Four different model experiments have been carried out in order to assess the impact of the correction. The experiments start from a model hindcast that was initiated in 2004. The hindcast assimilates ice concentration and SST. On 1 August 2012, all four experiments were started with characteristics as described in Table 4.

Experiments AC-NI and AC-AI are designed to assess the difference between the current operational setup and the operational setup, plus the atmospheric correction using the L4 IST product. Experiments AC-NI and AC-AI both assimilate ice concentration, and as a result of this the ice concentration is partly predefined. Therefore, experiments NC-NI and NC-AI are designed to assess the effect of the correction of the atmospheric temperatures without the assimilation of the ice concentration. The steps are outlined in the list below and carried out for both of the baseline experiments (AC-NI and NC-NI).

1. Modeled IST data available every hour from experiment AC-NI or NC-NI are averaged for 72 h centered around the times where remote sensing data are available. This assumes that the observations are spread equally across the entire period, which is not necessarily the case. This corresponds to the black line of Figure 4.

2. The difference between the remotely sensed IST (red line, Figure 4) and the modeled IST (black line, Figure 4) is found. Based upon the limited conclusions from the ground truth (IST from buoys, described in sections 5.3 and 5.4), the satellite and the model results are prescribed a similar error. Therefore, the correction is set to half the difference of the modeled and the observed IST, as defined in

$$\Delta = \frac{mod_{IST} - obs_{IST}}{2} \tag{10}$$

3. The 2 m temperature of the atmosphere (magenta line, Figure 4) generally correlates well with the modeled IST (black line, Figure 4). Therefore, the obtained Δ is interpolated to the time step of the atmospheric forcing (3 hourly) and used to correct the atmospheric temperatures where the modeled sea-ice concentration exceeds 85% as described in

$$t2m_{new} = t2m_{old} - \Delta \tag{11}$$

   where t2m is the atmospheric 2 m temperature.

4. Experiments AC-NI and NC-NI have been repeated with the updated atmospheric temperature as experiments AC-AI (blue line, Figure 4) and NC-AI. The period for the experiments is the 1 August 2012 to the 1 August 2013.

5. The result of experiment AC-AI is shown as the green line in Figure 4.

Figure 4. Time series of atmospheric forcing temperatures before the corrections (experiment AC-NI, magenta) and after the correction (experiment AC-AI, blue). Also shown are IST from the reference experiment AC-NI (black line), the experiment with corrected forcing (green) and the observed IST based on remote sensing observations (red line). All time series are from a position at 87.1°N and 184.5°W. The atmospheric forcing is plotted every third hour; both the modeled and remotely sensed IST are plotted as 72 h averages.

Figure 4 shows the effect of the atmospheric correction in experiment AC-NI on a selected model grid point with a varying ice cover. The observed IST is generally colder than the modeled IST. The atmospheric temperatures maintain the variability of the high frequency atmospheric forcing (hours to 1 day) whereas the low frequency IST observations correct the atmospheric temperatures on longer time scales. In addition from mid-June and throughout summer, the difference between the atmospheric temperatures before and after the correction is zero (magenta and blue lines), which is when the ice concentration drops below 85%.

5. Results and Discussion

This section will present and discuss the effect of assimilating remotely sensed ice surface temperatures into a coupled ocean and ice model based on a correction of the atmospheric surface temperatures. At first, focus will be on the magnitude of the change in the atmospheric near surface temperatures and the resulting ice surface temperatures. Second, the model results will be compared with independent observations from buoys that drift around the Arctic Ocean sampling temperature profiles of the snow, the ice, and the upper ocean. Finally, two independent parameters, that forecast skill typically is based on, will be evaluated; these are the ice concentration and the ice thickness.

5.1. Baseline for IST Corrections

The differences between the observed and the modeled snow or ice surface temperatures are the baseline for creating the correction to the 2 m air temperature. Figure 5 shows the monthly averaged ice or snow surface temperatures. The first row is based on the model results, the second row is the observations, and the last row is the differences (model-obs). The black contour line is the 85% ice concentration, which also outlines the area where the correction of the atmospheric temperatures is applied.

From Figure 5, it is clear that the modeled surface temperature is significantly warmer than the observations both in September 2012 and March 2013 with differences exceeding 8°C in both months. There are spatial differences and the MIZ seems to be less biased relative to the remotely sensed IST. In March, parts of the observed MIZ are warmer than the model. A reason for this could be that the satellite observations in this area are a mixture of observations from the warm ocean and the cold sea ice, whereas the modeled IST only represents the ice. One exception from the general picture is the Beaufort Gyre in March. This area is completely ice covered and the modeled temperatures are a few degrees Celsius colder than the observed temperatures. This is most likely related to the atmospheric forcing as discussed in section 5.2.

5.2. Atmospheric Temperatures

As explained in section 4, the atmospheric temperatures are corrected using the observed IST fields. Figure 6 shows the original atmospheric 2 m temperatures and the corrections applied based on the differences in ice surface temperatures shown in Figure 5.

The left column of Figure 6 shows September 2012. It is clear that the assimilation requirement of more than 85% ice coverage limits the area where the atmospheric correction is applied (see Figure 5a). This requirement is even more evident in Figures 6c and 6e, where the correction has limited effect. One could argue that it would be more realistic to use the ice concentration to weight the IST as is done with the remotely sensed product instead of a hard cut off at 85%.

Figures 6d and 6f show the corrections in March, which are similar for both runs. Both simulations generally have a negative correction meaning that the corrected 2 m temperature is colder than the original 2 m temperature. It is not surprising that there are biases in the atmospheric reanalysis and that ERA-Interim has a warm bias as this has been reported in several publications (e.g., Jakobson et al., 2012). This publication compares observations from a floating ice station from April 2007 to August 2007 and concludes that all Arctic reanalysis have biases in 2 m temperatures, winds, humidity, etc. The fact that there is a warm bias makes the approach of a temperature correction with time scales of several days suitable. There is one exception from this pattern in the area around the Beaufort Gyre where the model is colder than the observation, thus a warm correction is applied. Figure 6b shows that the area around the Beaufort Gyre is the coldest area in the atmospheric forcing in March. This spatial variation can also be seen in the modeled IST of experiment AC-NI (see Figure 5b). The area of the Beaufort Gyre is not significantly colder compared to for instance the area north of Greenland in the observed IST (see Figure 5d). The reason for this is unknown
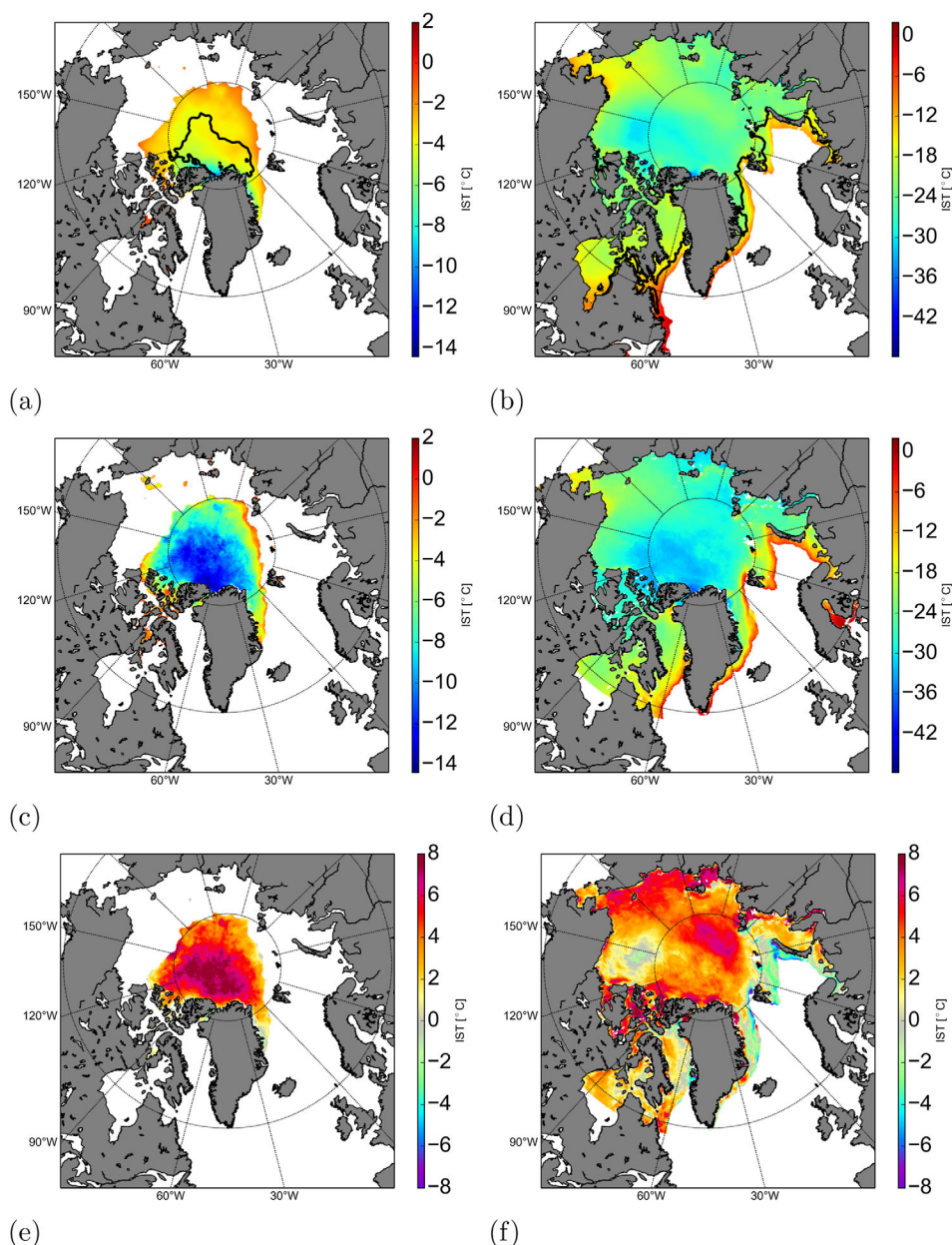
Figure 5. Monthly averages of IST shown for ice concentration larger than 15% based on experiment AC-NI. *Left column is September 2012 and right column* is March 2013. First row is modeled IST with black contour indicating the 85% ice concentration, second row is the observed IST, and the last row is the differences (Model-Obs).

but it could be linked to differences in the modeled ice and snow thicknesses, compared to the true conditions, which may introduce systematic differences in the surface temperatures.

5.3. Validation of the Remotely Sensed IST Product With Respect to Buoys

Validation of satellite-based IST observations against in situ observations is not straightforward. There are several complicating factors, as described in Dybkjær et al. (2012), such as the data sparseness, the quality of the observations, and the representativeness of the observations in describing the skin surface temperature. Recently, efforts have been made to gather more accurate radiometric observations of the temperature of the sea-ice surface to be used for satellite validations. However, these observations are not available for the time period when this study has been conducted. Instead, in situ observations used for validation here include IMB buoys from the North Atlantic Arctic coupling in a changing climate: impacts on ocean
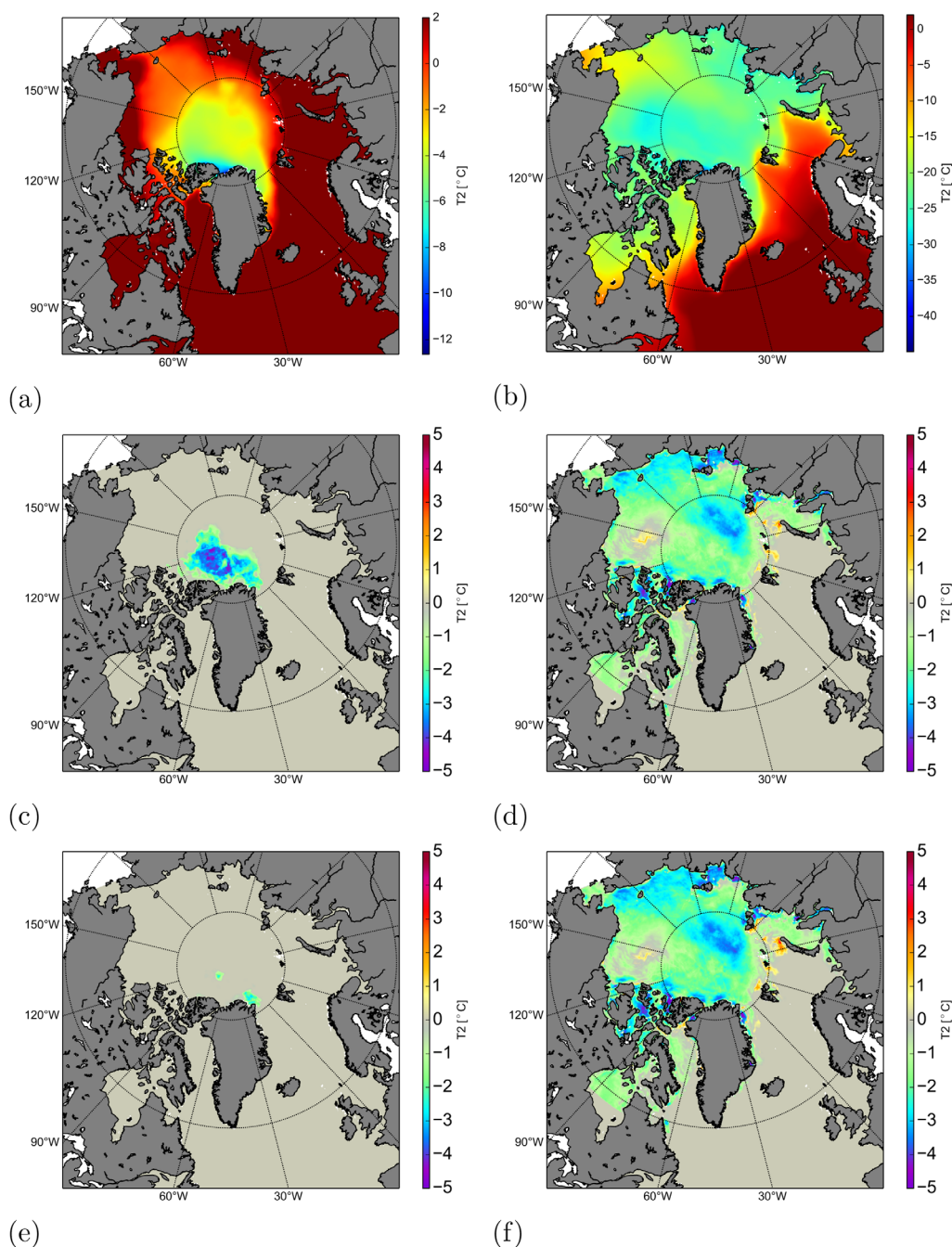
Figure 6. Monthly averaged atmospheric 2 m temperature as obtained from Era-Interim interpolated to the model grid is shown for (a) September and (b) March. Note the different color scales. Row two shows the mean monthly correction applied in AC-AI. The third row shows the mean monthly correction applied in NC-AI. Left column is September and the right column represents March. The correction is calculated as new 2 m temperature − old 2 m temperature.

circulation carbon cycling and sea-ice (NAACOS) project, the Alfred Wegener Institute (AWI), the International Arctic Buoy Programme (IABP), and Cold Regions Research and Engineering Laboratory (CRREL) (see trajectories in Figure 7). The validation results are shown in Table 5.

Before validating the L4 product, the in situ observations are averaged over a temporal window of 72 h, similar to the satellite aggregation window. The L4 values are extracted for the averaged in situ position, and the matching satellite versus in situ pairs are used to derive the validation statistics. Table 5 shows
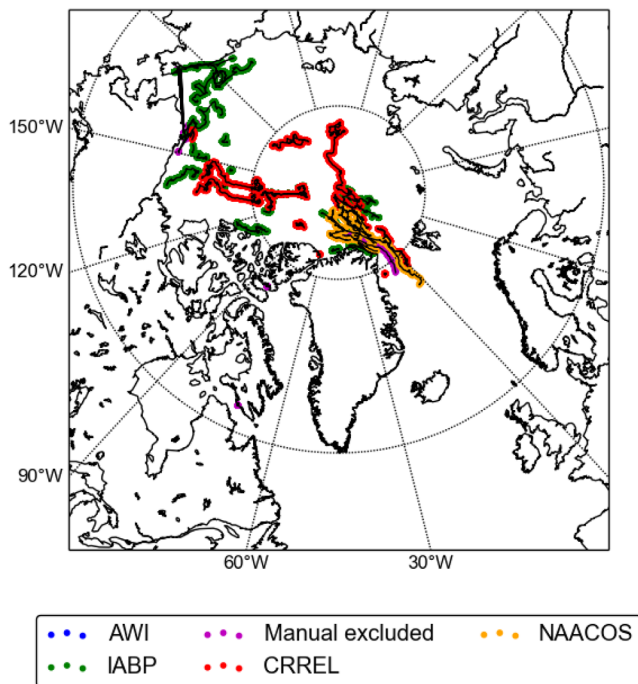
Figure 7. Trajectories for buoys that measure surface temperatures between August 2012 and July 2013. NAACOS, AWI, IABP, and CRREL buoys are shown. AWI trajectories are hidden underneath the other buoy trajectories.

typically a cold bias of the satellite product when compared against in situ observations. The mean difference depends to a large degree on the type of in situ observations, where a closest match is found for the IABP buoys and the worst match is for the AWI IMBs. Note that a significant part of the bias may be due to the in situ observations not reporting the surface temperature. Snow is a very effective insulator and if the temperature sensors are covered by just a few centimeters of snow it will usually lead to a significant warming of the sensor compared to the surface conditions due to an almost permanent heat flux from the ocean toward the air. In the case of warm air convection, the insulation results in a cold bias. Some of the IMBs have about 10 cm spacing between the temperature sensors, and it is therefore hard to identify the surface with great accuracy. The standard deviation of variability also includes the sampling effects from differences in satellite and in situ sampling. These effects include the temperature difference between the in situ point measurement and the larger satellite footprints, and the temperature differences arising from the in situ observations not measuring the actual surface skin temperature. In addition, the irregular satellite sampling compared to the regular in situ observational cycles also introduces discrepancies that are included in the standard deviations of differences.

In general, the validation results are in agreement with the results presented in Dybkjær et al. (2012) for the Metop AVHRR ice surface temperature observations. It is likely that the majority of the differences arise from differences in the satellite and in situ sampling, and representativeness effects. The size of these effects may overshadow real algorithm issues, and underlines the need for good reference data from, e.g., ice mounted infrared radiometers.

5.4. Validation of the Modeled IST Product With Respect to Buoys

The model has been validated against buoys using the same metrics as was described in section 5.3 for the remotely sensed product. Coincident and colocated time series have been extracted from both model and buoys in order to compare the results. A more complete picture is calculated using bias, standard deviation, root mean square (RMS) difference, and the Pearson correlation. These values have been calculated for the AWI, CRREL, NAACOS, and IABP buoys as described in section 5.3. It should be noted that the approach of comparing models with buoys and remotely sensed surface temperatures is slightly different, partly due to different sampling frequencies of the model and the remotely sensed surface temperatures. This means that the remotely sensed temperatures do not represent the diurnal variations. The results can be seen in Table 6.

The baseline for the two twin experiments is the experiments without assimilation of IST, namely NC-NI and AC-NI. Based on Table 6, the bias, standard deviation, and correlations are similar whether or not they include assimilation of sea ice. The bias in NC-NI varies from −4.4°C to 1.7°C; however, the standard

Table 5
Validation Statistics for the L4 IST Product (L4-In Situ), for the Different Sources of In Situ Observations

Type	Number	Bias	Std.	RMS	Corr.
IABP	1,792	−0.69	5.4	5.4	0.82
CRREL	5,171	−2.3	4.8	5.3	0.90
NAACOS	1,367	−3.3	4.6	5.7	0.86
AWI	2,261	−5.3	2.6	5.9	0.91

Note. Std is the standard deviation of the differences, RMS is the root mean square difference, and Corr is the Pearson correlation. Validation statistics are grouped according to their origin but are otherwise aggregated across buoys and locations.

Table 6
Bias (Model-Buoy), Standard Deviation of Differences, RMS and Correlation for the Different Model Experiments Against the Buoy Observations

Buoy group	N obs	Exp. AC-NI				Exp. NC-NI			
		Bias	Std.	RMS	Corr.	Bias	Std.	RMS	Corr.
AWI	2,909	−4.5	4.5	6.3	0.80	−4.4	4.4	6.2	0.79
CRREL	11,145	1.3	3.6	3.8	0.96	1.7	3.7	4.1	0.95
IABP	63,809	−3.3	8.0	8.6	0.73	−3.0	7.9	8.4	0.73
NAACOS	2,591	0.23	3.4	3.4	0.96	0.47	3.2	3.3	0.96
Buoy group	N buoys	Exp. AC-AI				Exp. NC-AI			
AWI	2	−5.5	4.6	7.2	0.79	−5.4	4.6	7.1	0.79
CRREL	13	0.16	3.4	3.4	0.96	0.49	3.5	3.5	0.96
IABP	22	−4.3	7.9	9.0	0.74	−4.0	7.9	8.9	0.74
NAACOS	4	−1.2	3.3	3.5	0.96	−0.56	3.1	3.1	0.96

Note. The numbers have been calculated for the time of the experiment (1 August 2012–2013). Trajectories can be seen in Figure 7. N obs is the number of individual measurements obtained by N buoys.

deviation and the RMS error are smaller and the correlation is higher for the buoys with a warm bias. This indicates that the quality of the observations with the CRELL and the NAACOS buoys are better. Another reason for the difference especially in the bias can be the area where the buoys are deployed and drift. The spatial patterns for IST can be seen in Figure 5f. Table 6 shows that the model has a colder bias compared to the buoys for the experiments with IST assimilation. This is expected as the surface forcing of these experiments is colder. This improves the skill in some cases and in other cases it decreases the skill of the model. In general, the statistics of Tables 5 and 6 cannot be used to point to either the model or the remotely sensed product to be superior. Thus it is difficult to make a firm validation of the experiments based on the buoy data. The standard deviation for all four groups of buoys is more or less unchanged in the two twin experiments (AC-NI, AC-AI) and (NC-NI, NC-AI). This is expected as the assimilation of the IST is based on averages of 72 h and fast variations are not captured by this. In addition, the buoys are independent of both the remotely sensed observations and the modeled surface temperatures. Compared to Table 5, the standard deviations are smaller with the exception of the AWI buoys. A reason for this can be that the time stamp of the model output and the buoys are well determined, whereas the time stamp of the remote sensed temperatures is given as averages of 72 h irregularly sampled satellite observations. An additional point is that the bias of the experiments with and without assimilation of ice concentration is different. This is a result of the method used when assimilating sea-ice concentration. When sea ice is added or removed energy is also removed/added. This correction of energy may result in a slight change of the temperature profile within the ice and snow.

In general, the results can be split in two groups of buoys. The first group consists of CRREL and NAACOS. These have similar measures and a good correlation. The standard deviation is between 3°C and 4°C and slightly reduced in the experiments that include assimilation of IST (AC-AI and NC-AI). The second group is AWI and IABP buoys which have significantly worse statistics. The results of these cover individual buoys that have the same statistics as the first group and individual buoys that have more or less no correlation with the model. A thorough quality filter of these buoys would probably improve the result; however, it is not always simple to judge whether the buoy suffers from failures (instrumentation, vertical location of the temperature measurements, etc.) or the statistics reveal real problems with the modeled temperature.

5.5. Temperature Profiles
DMI has deployed eight buoys (labeled as NAACOS) in the Arctic Ocean that measures the temperature profile of the snow and ice they are mounted in. Four of these have provided useful results and these are, to the author's knowledge, the only buoys within the study period that have been manually evaluated. This means that the interfaces between air, snow, ice, and ocean have been determined with better accuracy based upon the heating cycle response in the IMBs. In addition, these have a high vertical resolution (2 cm) from the air throughout the ice and into the waters. Based on this, extra attention has been given to these buoys.

The sea-ice model splits the thermodynamic calculations into five horizontal categories for each grid cell based on the ice thickness. For each category, a vertical profile is produced that consists of one snow temperature and four vertical ice layers. Boundary conditions are prescribed from the atmospheric forcing at the surface and the ocean temperature at the bottom; thus all variations are local within the ice.

Figure 8 shows the observed and the modeled temperature profiles as functions of time for buoy C from the NAACOS deployments. The individual trajectories are not shown in this article; however, all tracks can be found in Figure 7.

Figure 8 shows the annual cycle where the sea-ice temperatures are close to zero in summer and decreasing in winter. The effect of snow as an insulation for the ice pack with much colder winter temperatures
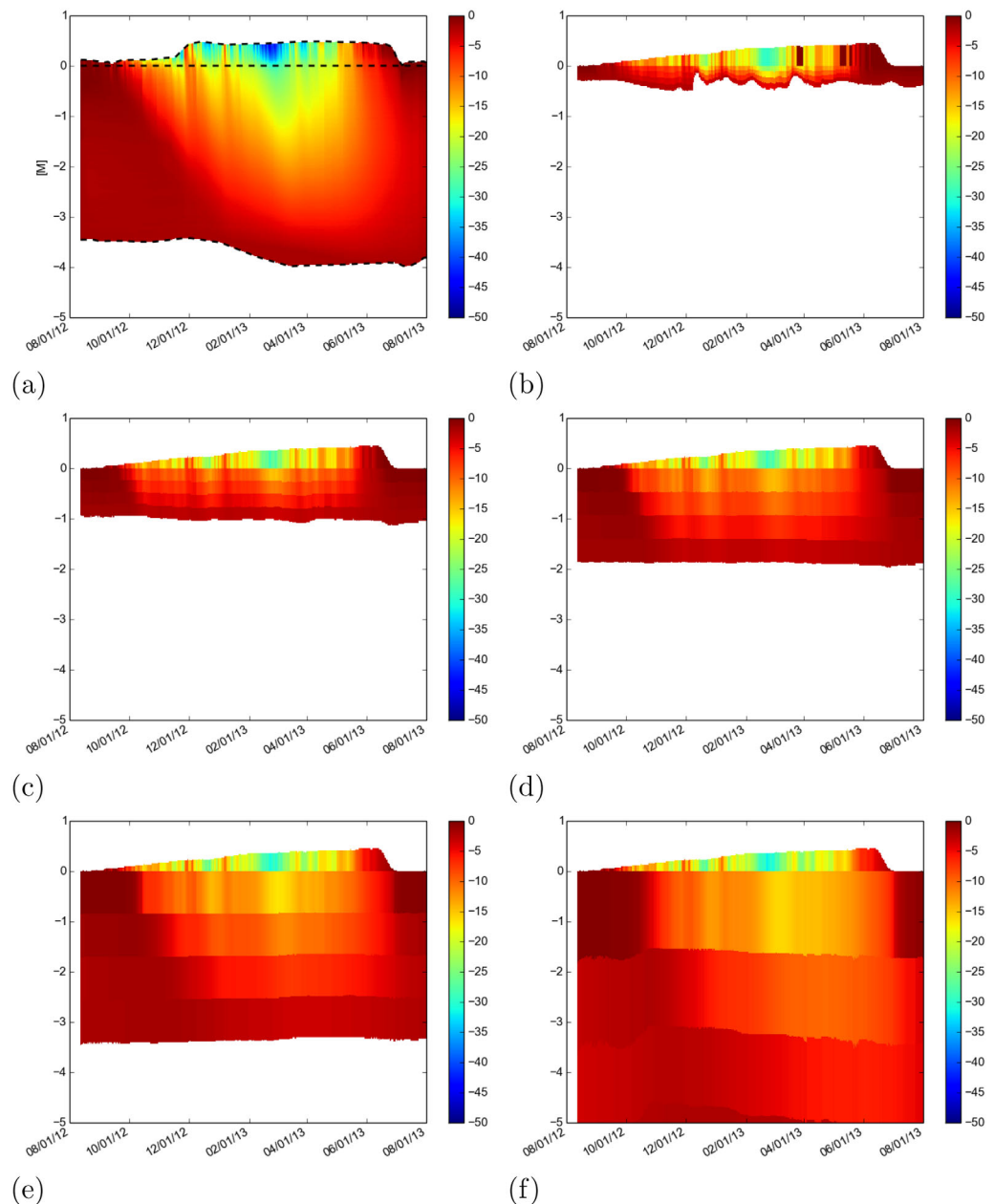
Figure 8. (a) The observed temperature profile from buoy C of the NAACOS deployments. The *y* axis shows the snow thickness above zero and the ice thickness below zero in meters. (b–f) The temperature profiles for all categories defined in Table 2 along the track of NAACOS buoy C. Note that the profiles have been cut off at 5 m depth despite the ice being thicker in category 5.
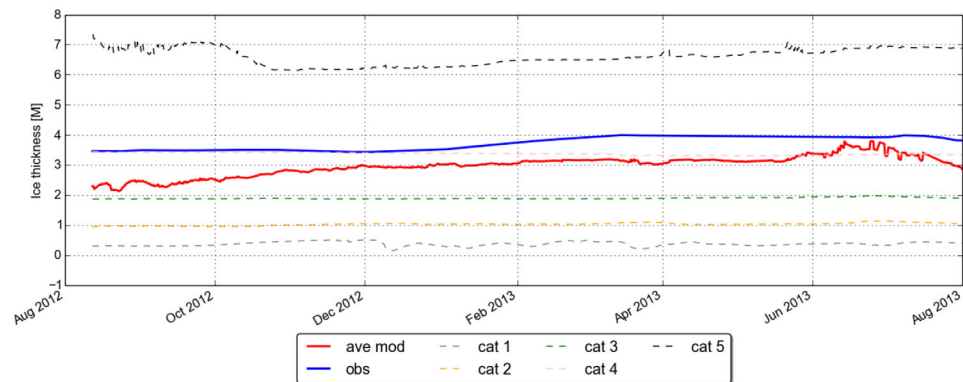
Figure 9. Ice thickness for NAACOS buoy C (see Figure 7). The red solid line shows the average modeled ice thickness. The solid blue line shows the average observed ice thickness. Dashed lines show the ice thickness of each sea-ice category described in Table 2.

compared to the sea ice is clearly seen in both observation and modeled profiles. The snow layer is removed during summer when the Arctic surface is characterized by melt ponds and limited snow cover. The influence of the snow on the temperature profiles of the snow and sea-ice underlines the importance of understanding the growth and decay of the snow cover; however, this is not a trivial task.

In order to compare with buoys, the model ice category with an ice thickness similar to the buoy is selected. The temporal evolution of the modeled ice thickness for each category, along with averaged ice thickness from the buoy and the observations, are shown in Figure 9. It can be seen that the average sea-ice thickness of each category is near the mean of the intervals in Table 2; exceptions are the thinnest and the thickest ice categories.

The average ice thickness of the model and the observation is reasonably close to each other; however, the settings for these two are different as the ice thickness of the model is based on an average of 100 km^2, whereas the buoy is a point measurement. Both model and observations are relatively close to the ice thickness of category 4 (green dashed line in Figure 9). Based on this, the temperature profile of category 4 from the model is compared with the temperature profile of the buoy. The observed temperature profile is much finer grained compared to the modeled temperature. Therefore, the observations are interpolated vertically to the grid of the model with one snow layer and four temperature layers. The difference between the modeled profile of category 4 and the observed profile are shown for all four experiments in Figure 10.

The temperature bias is largest in winter and spring, which confirms the comparisons at the surface of both remotely sensed observations and buoy surface profiles. The model is generally warmer than the observations, and the temperature differences are reduced when the position, within the sea ice approaches the ocean surface. This is not surprising, as the ocean surface temperature is at the freezing point just below the sea ice. This means that the temperature difference at the bottom of the sea ice is constrained to the approximation of the freezing point of the surface of the ocean.

Figure 10 shows that the ice thickness increases underneath the buoy but not in category 4 of the model. This is due to the multicategory approach of the ice model, where the ice thickness of the individual categories remains constant (see Figure 9) but the relative distribution of the different categories may change. An important thing to note is that the growth of the snow increases slowly in all four experiments compared to observations. Snow is added to the grid cells of the model when the atmospheric model prescribes precipitation and when the temperature is lower than 0°C. Therefore, it is assumed that the slow buildup of snow is caused by rates of precipitation in the atmospheric forcing that are too low. One alternative reason is that the buoy observes points whereas the model is an average of a large area. Another reason could be that the buoy may have experienced snow that was redistributed due to winds. The different scales may cause problems when comparing the two measures.

An important difference is seen when the snow disappears from the ice surface in the beginning of the melt season. The two experiments that include surface temperature correction (exp. AC-AI, Figure 10c and
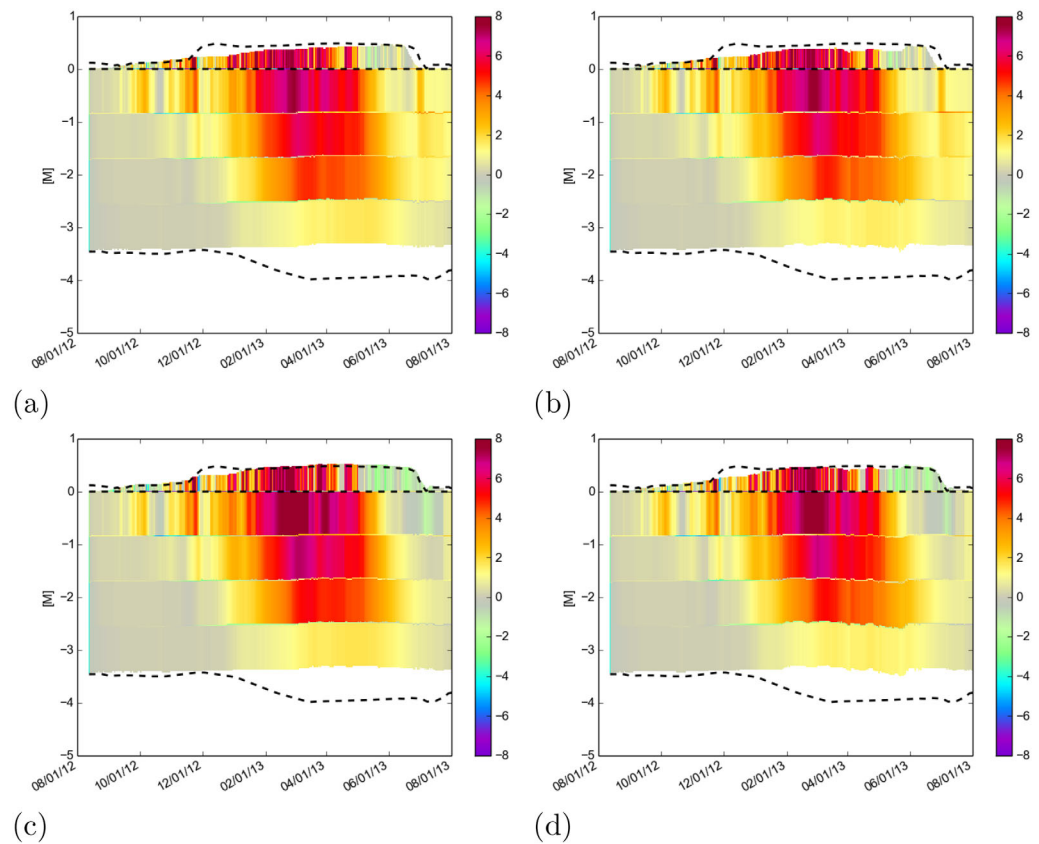
Figure 10. The difference (MOD-OBS) between category 4 of all four experiments and the observed temperature profile is shown for NAACOS buoy C, see Figure 7. (a) AC-NI, (b) AC-AI, (c) NC-NI, and (d) NC-AI. The dashed lines indicate the air-snow, snow-ice, and ice-water interfaces, as observed by buoy C. The observations have been interpolated to the layers of the model assuming a constant temperature throughout each layer.

NC-AI, Figure 10d) show a better timing of the reduction of the snow cover. The snow cover is reduced when the temperature reaches 0°C. As a result, the temperature correction improves the timing of the melt season. This also removes a warm bias within the upper part of the sea ice. This conclusion would be stronger if there were more available buoy observations with temperature profiles covering the full season.

5.6. Ice Concentration

The sea-ice concentration and the ice thickness are important physical variables as these have huge operational (e.g., shipping) and climate impact. In relation to short term forecasting these determine whether a ship is able to pass through an area or not. Sea-ice concentration is the parameter which can best be determined both from ice charts, and from automated retrievals using primarily microwave sensors which results in good estimates of the ice cover especially in winter time. There are two main drivers for the sea-ice concentration and the sea-ice thickness. These are the thermodynamics and the ice dynamics. This study only deals with the thermodynamics and corrections of the bias of the upper boundary condition (air versus ice and snow surface temperature). Figure 11 shows the ice concentration of the four experiments and the differences with and without assimilation. The black contour is OSI-SAF sea-ice concentration of 15%.

There is generally a good match with OSI-SAF data for all four experiments. The match is slightly better for experiments AC-NI and AC-AI, which is expected as these experiments use ice concentration for assimilation. Applying IST assimilation (AC-NI to AC-AI) results in a slight increase in the ice concentration near the coast of Alaska and the margin between the Laptev Sea and the Kara Sea. These are areas with medium range ice concentrations in experiment AC-NI.

There is a clear effect on the sea-ice concentration between the two experiments without assimilation of ice concentration (NC-NI and NC-AI) especially near the Bering Strait. The effect is most prominent within the
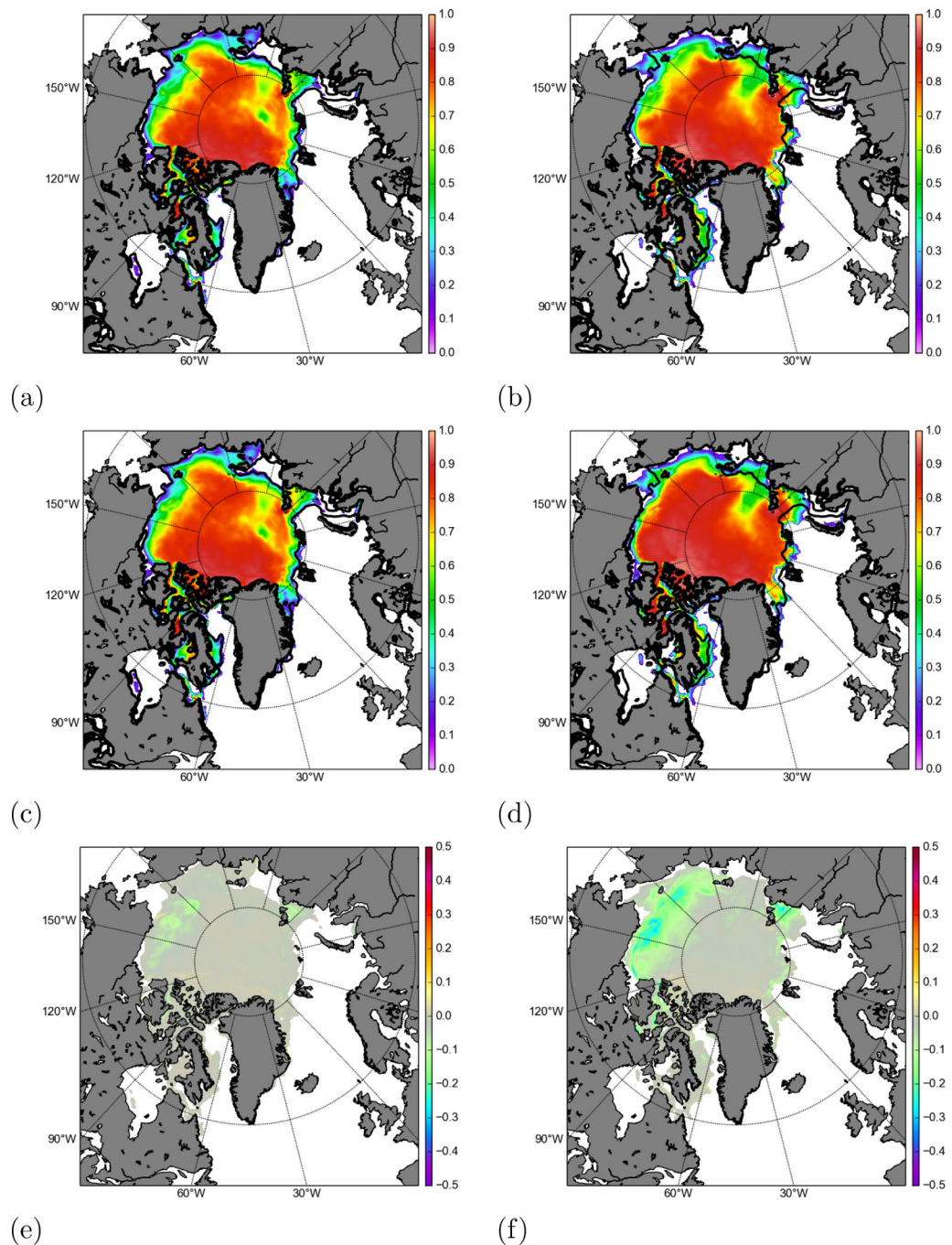
Figure 11. Monthly mean fractional sea-ice concentration in July 2013 for the four experiments from Table 4. (a) AC-NI, (b) AC-AI, (c) NC-NI, and (d) NC-AI. (e) ((AC-NI) − (AC-AI)) and (f) ((NC-NI) − (NC-AI)) show the effect of IST assimilation on the sea-ice concentration. The black contour is OSI-SAF ice concentration of 15%.

ice pack, where the criterion of an ice concentration of 85% is fulfilled. This shows that there is an effect of assimilating IST in the model and that it results in an increase in the ice concentration, which seems realistic. The following sections will discuss the influence further.

5.7. Ice Thickness

The ice thickness represents the memory of the system; however, accurate ice thickness observations are mainly available outside the melt season. Remotely sensed ice thickness based on SMOS (Kaleschke et al., 2012) and Cryosat-2 (Laxon et al., 2013) are available from October to April with full Arctic coverage. In

addition, Operation Ice Bridge (OIB) observes along selected flight tracks in March/April. The difference in ice thickness between the experiments with and without corrections of the atmospheric forcing is limited to winter of this 1 year experiment (see Figure 12b). Therefore, this analysis will not try to judge which experiment has the best distribution, but simply show and describe the changes. The four experiments and the differences between experiments AC-NI/AC-AI and NC-NI/NC-AI are shown in Figure 12.

The ice thickness of the experiments including assimilation of sea-ice concentration (Figures 12a and 12c) is generally thicker and the brown colored region extends further into the central Arctic than the two
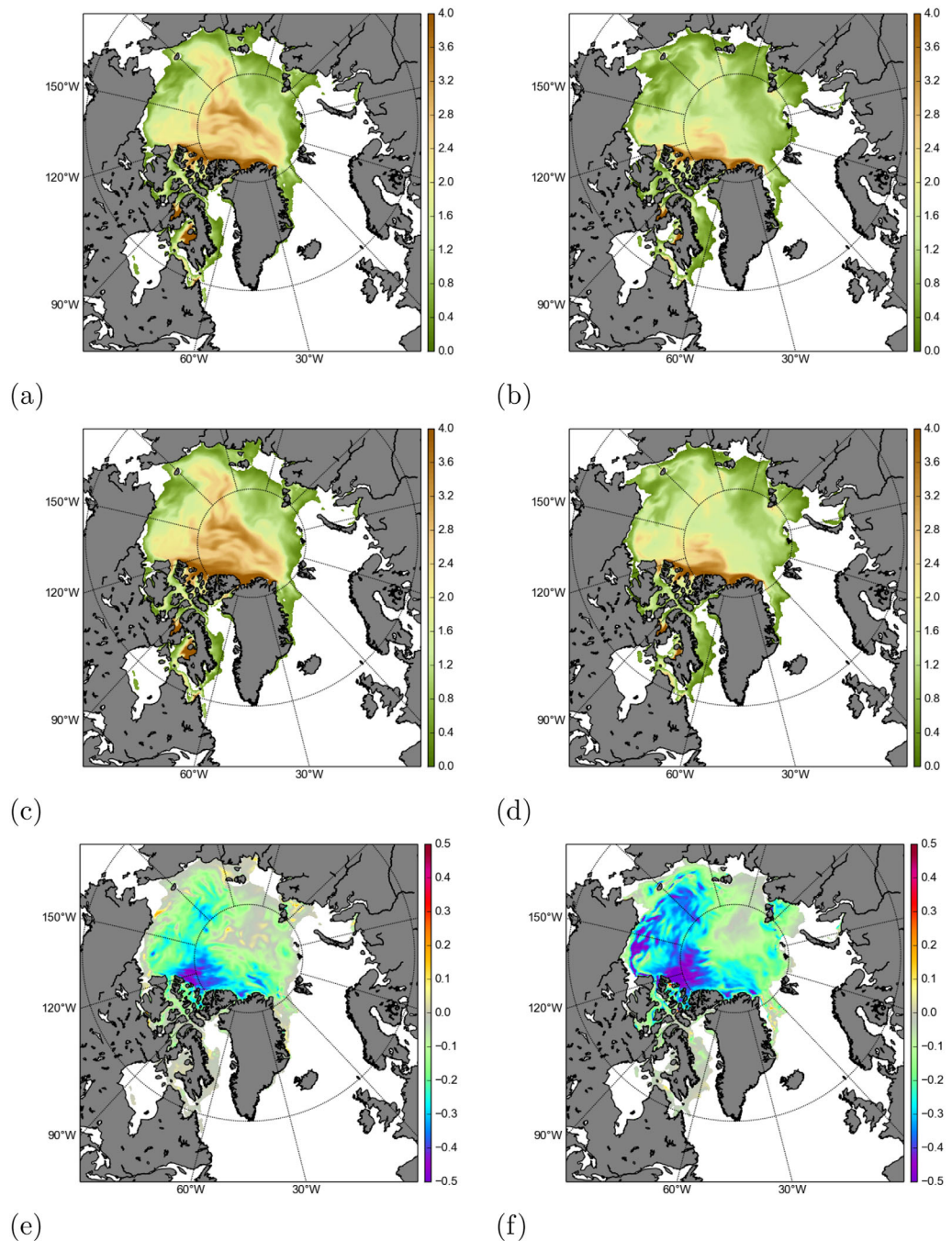
Figure 12. Monthly mean sea-ice thickness in July 2013. (a) Experiment AC-NI, (b) exp NC-NI, (c) exp AC-AI, and (d) exp NC-AI. Figure 12e (exp AC-NI − exp AC-AI) and 12f (exp NC-NI − exp NC-AI) show the difference in ice thickness between the experiments without IST assimilation and the experiments with IST assimilation.
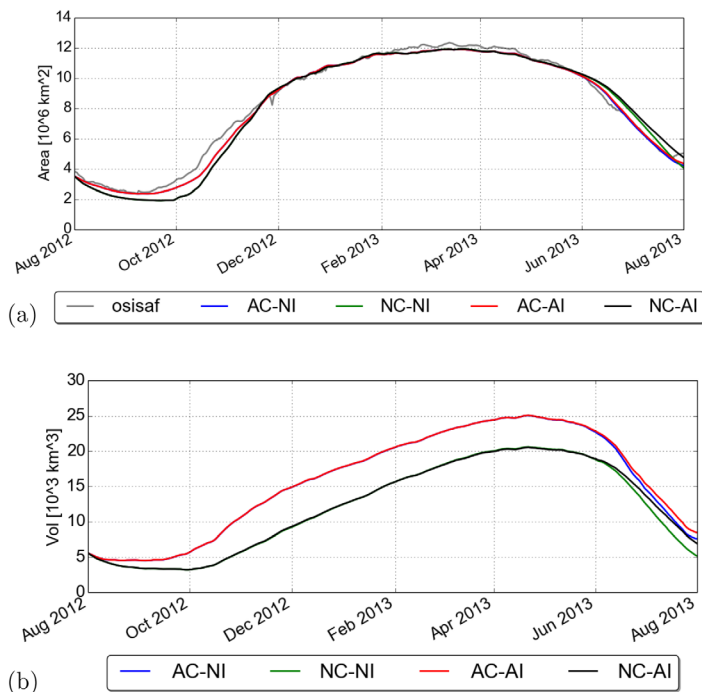
Figure 13. (a) The integrated sea-ice cover throughout the period of the experiments. Grid cells with ice cover less than 15% have been excluded. (b) The integrated ice volume.

experiments without the ice concentration assimilation (Figures 12b and 12d). This confirms the general impression that the model is too warm. The assimilation of ice concentration also minimizes the reduction of the ice thickness. Another reason for the increased ice thickness with assimilation of ice concentration is described in Smith et al. (2016), where the method labeled as "red" is similar to the method used in this study. The article describes how adding ice concentration can add ice volume. Both with and without assimilation of ice concentration the ice thickness just north of the Canadian Archipelago is increased by up to 0.4 m. The effect of the surface temperature correction is generally largest for the experiments without ice assimilation. The increase of ice thickness is seen in large parts of the Arctic Ocean.

5.8. Ice Cover and Ice Volume

The integrated model ice cover and ice volume for the four experiments and OSI-SAF have been derived. As the model does not cover the ice covered areas on the Pacific side of Bering Strait this area has been filtered out from the OSI-SAF ice cover to allow for intercomparisons of the integrated values. The results can be seen in Figure 13.

From approximately August to October, both the ice cover and the ice volume based on the experiments without assimilation are reduced compared to the ice cover and volume with sea-ice assimilation. From October to November, the ice concentration for the sea-ice assimilated experiments is slightly smaller compared to OSI-SAF. This is due to the relatively weak assimilation. During winter, the atmosphere is sufficiently cold to bring the ice cover to the same level as it is for the observations based on OSI-SAF. The difference between the ice volume of the two different twin experiments shows that the volume has a longer memory and that the difference in volume is more or less constant. From June, or approximately the beginning of the melt season, the ice concentration is reduced slightly faster in the experiments with assimilation of ice concentration. From July NC-NI reduces the ice cover faster than the other experiments, and by the end of July the ice cover is similar to the ice cover of AC-NI and AC-AI. From June onward, the ice volume of NC-AI and AC-AI is reduced slightly slower compared to the corresponding results of NC-NI and AC-NI. The result is that experiments AC-NI and NI-AC have approximately the same ice volume. This indicates that the effect of the ice temperature assimilation on the ice concentration and the volume is largest in summer, where a small change of the heat flux balance may change conditions from melting to freezing.

6. Conclusions

This study demonstrates an innovative approach for combining several different satellite IST products into a multisensor, gap-free, interpolated L4 SST, and IST product. Given the satellite sampling characteristics, the warm biased atmospheric forcing, and the operational ocean and sea-ice model physics, these data were considered suitable for assimilation into an operational model as a bias correction. We found that the bias correction of satellite ISTs improved the timing of snow melt over sea ice and as a consequence of this it improved the temperature profile of the snow and sea ice. This conclusion is only based on comparison with one buoy, thus it may be stronger if more buoys were included in the comparison. The remotely sensed IST is generally colder than the modeled IST based on experiments AC-NI and NC-NI; however, there are spatial and temporal differences. This mainly results is a negative correction to the atmospheric 2 m temperature. As a consequence of this, the ice thickness increased from June and throughout the summer when the snow melts and the temperature approaches freezing point. The ice cover is only increased from experiments NC-NI to NC-AI as experiments AC-NI and AC-AI assimilate the ice cover directly. Quantifying the impact of IST assimilation remains challenging due to the quality and limited number of in situ observations routinely made in polar environments. Future studies could look at multiyear model runs assimilating IST, with focused campaigns for gathering in situ data at higher spatial and temporal resolution. In addition,

the temporal variation of IST is relatively high compared to the differences in IST. Thus it would be beneficial to use level-2 data for the assimilation in order to be able to collocate the modeled and observed IST in time.

References

Bitz, C. M., & Lipscomb, W. H. (1999). An energy-conserving thermodynamic model of sea ice. *Journal of Geophysical Research*, *104*(C7), 15669–15677.

Castro, S. L., Wick, G. A., & Steele, M. (2016). Validation of satellite sea surface temperature analyses in the Beaufort Sea using UpTempO buoys. *Remote Sensing of Environment*, *187*, 458–475.

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwel, G. R., Hogan, P. J., Wallcraft, A. J., et al. (2007). The HYCOM (Hybrid Coordinate Ocean Model) data assimilative system. *Journal of Marine Systems*, *65*, 60–83.

Collins, N., Theurich, G., DeLuca, C., Suarez, M., Trayanov, A., Balaji, V., et al. (2004). Design and implementation of components in the earth system modeling framework. *International Journal of High Performance Computing Applications*, *19*(3), 341–350. https://doi.org/10.1177/1094342005056120

Conkright, M. E., O'Brien, T. D., Stephens, C., Locarnini, R. A., Garcia, H. E., Boyer, T. P., & Antonov, J. I. (2002). *World Ocean Atlas 2001, Volume 6: Chlorophyll*. In Levitus, S., (Ed.). (NOAA Atlas NESDIS 54, CD-ROMs, Technical Report, 46 pp.). Washington, DC: U.S. Gov. Printing Office.

Dai, A., & Trenberth, K. (2002). Estimates of freshwater discharge from continents: Latitudinal and seasonal variations. *Journal of Hydrometeorology*, *3*, 660–687.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*, 553–597.

Donlon, C., Rayner, N., Robinson, I., Poulter, D., Casey, K., Vazquez-Cuervo, J., et al. (2007). The global ocean data assimilation experiment high-resolution sea surface temperature pilot project. *Bulletin of the American Meteorological Society*, *88*(8), 1197–1213.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2011). The operational sea surface temperature and sea ice analysis (OSTIA). *Remote Sensing of the Environment*, *116*, 140–158. https://doi.org/10.1016/j.rse.2010.10.017

Dybbroe, A., Hörnquist, S., Lavanant, L., & Marguinaud, P. (2006). Cloud masking for the O&SI SAF global Metop/AVHRR SST product. In *Proceedings of the 2006 EUMETSAT meteorological satellite conference* (48 pp.). EUMETSAT, Helsinki.

Dybkjær, G., Tonboe, R., & Høyer, J. (2012). Arctic surface temperatures from Metop AVHRR compared to in situ ocean and land data. *Ocean Science*, *8*(6), 959–970.

Dybkjær, S. E., Pfeiffer, G. R., & Howe, E. (2017). OSI SAF high latitudes L2 sea and sea ice surface temperature product user manual OSI 205 (Eumetsat technical REPORT). Retrieved from http://osisaf.met.no/docs/osisaf_cdop3_ss2_pum_hl-l2-sst-ist_v1p1.pdf

Eastwood, S., Larsen, K., Lavergne, T., Nielsen, E., & Tonboe, R. (2011). Global sea ice concentration preprocessing—Product user manual product OSI-409 (Eumetsat technical REPORT). Retrieved from http://osisaf.met.no/docs/pum_seaicereproc_ss2_v1p3.pdf

Egbert, G. D., & Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic Technology*, *19*(2), 183–204. https://doi.org/10.1175/1520-0426(2002) 019⟨0183:EIMOBO⟩2.0.CO;2

Hall, D. K., Key, J. R., Casey, K. A., Riggs, G. A., & Cavalieri, D. J. (2004). Sea ice surface temperature product from MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(5), 1076–1087.

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, *48*, RG4004. https://doi.org/10.1029/2010RG000345

Hibler, W. B. (1979). A dynamic thermodynamic sea ice model. *Journal of Physical Oceanography*, *9*, 817–846.

Høyer, J. L., Borgne, P. L., & Eastwood, S. (2014). A bias correction method for arctic satellite sea surface temperature observations. *Remote Sensing of the Environment*, *146*, 201–213.

Høyer, J. L., Karagali, I., Dybkjær, G., & Tonboe, R. (2012). Multi sensor validation and error characteristics of arctic satellite sea surface temperature observations. *Remote Sensing of Environment*, *121*, 335–346.

Høyer, J. L., Lang, A. M., Tonboe, R., Eastwood, S., Wimmer, W., & Dybkjær, G. (2017). Towards field inter-comparison experiment (FICE) for ice surface temperature (ESA Tech. Rep. FRM4STS OP-40). Retrieved from http://www.frm4sts.org/wp-content/uploads/sites/3/2017/12/OFE-OP-40-TR-5-V1-Iss-1-Ver-1-Signed.pdf

Høyer, J. L., & She, J. (2007). Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea. *Journal of Marine System*, *65*(1–4), 176–189.

Hunke, E. C. (2001). Viscous-plastic sea ice dynamics with the EVP model: Linearization issues. *Computational Physics*, *170*, 18–38.

Hunke, E. C., & Dukowicz, J. (1997). An elastic-viscous plastic model for sea ice dynamics. *Journal of Physical Oceanography*, *27*, 1849–1867.

Jackson, K., Wilkinson, J., Maksym, T., Meldrum, D., Beckers, J., Haas, C., et al. (2013). A novel and low-cost sea ice mass balance buoy. *Journal of Atmospheric and Oceanic Technology*, *30*(11), 2676–2688.

Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., & Jaagus, J. (2012). Validation of atmospheric reanalyses over the central Arctic Ocean. *Geophysical Research Letters*, *39*, L10802. https://doi.org/10.1029/2012GL051591

Kaleschke, L., Tian-Kunze, X., Maß, N., Maekynen, M., & Drusch, M. (2012). Sea ice thickness retrieval from SMOS brightness temperatures during the arctic freeze-up period. *Geophysical Research Letters*, *39*, L05501. https://doi.org/10.1029/2012GL050916

Key, J., & Haefliger, M. (1992). Arctic ice surface temperature retrieval from AVHRR thermal channels. *Journal of Geophysical Research*, *97*(D5), 5885–5893.

Key, J., Maslanik, J., Papakyriakou, T., Serreze, M. C., & Schweiger, A. (1994). On the validation of satellite-derived sea ice surface temperature. *Arctic*, *47*, 280–287.

Large, W. G., & Yeager, S. (2008). The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*, *33*, 341–364. https://doi.org/10.1007/s00382-008-0441-3

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, *32*, 363–403.

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., et al. (2013). Cryosat-2 estimates of Arctic sea ice thickness and volume. *Geophysical Research Letters*, *40*, 732–737. https://doi.org/10.1002/grl.50193

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. (2010). Development and test of the HYPE (hydrological predictions for the environment) model a water quality model for different spatial scales. *Hydrology Research*, *41*(3–4), 295–319.

Lipscomb, W. H. (2001). Remapping the thickness distribution in sea ice models. *Journal of Geophysical Research*, *106*(C7), 13989–14000.

Madsen, K. S., Rasmussen, T. A. S., Ribergaard, M. H., & Ringgaard, I. M. (2016). High resolution sea-ice modelling and validation of the arctic. *Polarforschung*, *85*(2), 101–105.

May, D. A., Parmeter, M. M., Olszewski, D. S., & McKenzie, B. D. (1998). Operational processing of satellite sea surface temperature retrievals at the naval oceanographic office. *Bulletin of the American Meteorological Society*, *79*(3), 397–407.

McNider, R. T., Song, A. J., Casey, D. M., Wetzel, P. J., Crosson, W. L., & Rabin, R. M. (1994). Toward a dynamic-thermodynamic assimilation of satellite surface temperature in numerical atmospheric models. *Monthly Weather Review*, *122*(12), 2784–2803.

Merchant, C. J., Matthiesen, S., Rayner, N. A., Remedios, J. J., Jones, P. D., Olesen, F., et al. (2013). The surface temperatures of earth: Steps towards integrated understanding of variability and change. *Geoscientific Instrumentation, Methods and Data Systems*, *2*(2), 305–321.

Nicolaus, M., Hoppmann, M., Arndt, S., Hendricks, S., Katlein, C., König-Langlo, G., et al. (2017). Snow height and air temperature on sea ice from Snow Buoy measurements. https://doi.org/10.1594/PANGAEA.875638

Riggs, G. A., Hall, D. K., & Salomonson, V. V. (2006). *MODIS sea ice products user guide to collection 5*. Boulder, CO: NSIDC.

Rigor, I. G., Colony, R. L., & Martin, S. (2000). Variations in surface air temperature observations in the arctic, 1979–97. *Journal of Climate*, *13*(5), 896–914.

Schneebeli, M., & Sokratov, S. A. (2004). Tomography of temperature gradient metamorphism of snow and associated changes in heat conductivity. *Hydrological Processes*, *18*(18), 3655–3665.

Smith, G. C., Roy, F., Reszka, M., Surcel Colan, D., He, Z., Deacu, D., et al. (2016). Sea ice forecast verification in the Canadian global ice ocean prediction system. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 659–671. https://doi.org/10.1002/qj.2555

Steele, M., Morley, R., & Ermold, W. (2001). PHC: A global ocean hydrography with a high-quality arctic ocean. *Journal of Climate*, *14*(9), 2079–2087. https://doi.org/10.1175/1520-0442(2001)014<2079:PAGOHW>2.0.CO;2

Sturm, M., Perovich, D. K., & Holmgren, J. (2002). Thermal conductivity and heat transfer through the snow on the ice of the Beaufort Sea. *Journal of Geophysical Research*, *107*(C10), 8043. https://doi.org/10.1029/2000JC000409

Sturm, M. J., Holmgren, J., König, M., & Morris, K. (1997). The thermal conductivity of seasonal snow. *Journal of Glacial*, *43*, 26–41.

Sturm, M. J., Holmgren, J., & Perovich, D. K. (2001). Spatial variations in the winter heat flux at Sheba: Estimates from snow-ice interface temperatures. *Annals of Glaciology*, *33*, 213–220.

Tonboe, R. T., Dybkjær, G., & Høyer, J. L. (2011). Simulations of the snow covered sea ice surface temperature and microwave effective temperature. *Tellus, Series A*, *63*(5), 1028–1037.

Wang, X., & Key, J. R. (2003). Recent trends in arctic surface, cloud, and radiation properties from space. *Science*, *299*(5613), 1725–1728.

While, J., & Martin, M. (2013). Development of a variational data assimilation system for the diurnal cycle of sea surface temperature. *Journal of Geophysical Research: Oceans*, *118*, 2845–2862. https://doi.org/10.1002/jgrc.20215