Ocean Science

Arctic surface temperatures from Metop AVHRR compared to in situ ocean and land data

G. Dybkjær, R. Tonboe, and J. L. Høyer

Center for Ocean and Ice, Danish Meteorological Institute, Lyngbyvej 100, 2100 København, Denmark

Correspondence to: G. Dybkjær (gd@dmi.dk)

Abstract. The ice surface temperature (IST) is an important boundary condition for both atmospheric and ocean and sea ice models and for coupled systems. An operational ice surface temperature product using satellite Metop AVHRR infra-red data was developed for MyOcean. The IST can be mapped in clear sky regions using a split window algorithm specially tuned for sea ice. Clear sky conditions prevail during spring in the Arctic, while persistent cloud cover limits data coverage during summer. The cloud covered regions are detected using the EUMETSAT cloud mask. The Metop IST compares to 2 m temperature at the Greenland ice cap Summit within STD error of 3.14 °C and to Arctic drifting buoy temperature data within STD error of 3.69 °C. A case study reveals that the in situ radiometer data versus satellite IST STD error can be much lower (0.73 °C) and that the different in situ measurements complicate the validation. Differences and variability between Metop IST and in situ data are analysed and discussed. An inter-comparison of Metop IST, numerical weather prediction temperatures and in situ observation indicates large biases between the different quantities. Because of the scarcity of conventional surface temperature or surface air temperature data in the Arctic, the satellite IST data with its relatively good coverage can potentially add valuable information to model analysis for the Arctic atmosphere.

1   Introduction

The Ice Surface Temperature (IST) is one of the most important components in the Arctic surface-atmosphere energy balance. The surface temperature strongly affects the atmospheric boundary layer structure, the turbulent heat exchange

and the ice growth rate (Maykut, 1986). Advanced thermodynamic ice models treat the temperature of the snow surface as a vital parameter for the development of sea ice in models (e.g., Fichefet and Maqueda, 1997; Bitz and Lipscomb, 1999). Steffen et al. (1993) estimated that a systematic surface temperature change of 1 °C corresponds to an outgoing long wave radiation change of approximately 5 Wm-2. From a modelling point-of-view a systematic year round 5 W energy flux anomaly can be sufficient to change the sea ice regime from seasonal to perennial sea ice, or vice versa (Björk and Söderkvist, 2002). The surface temperature is a boundary condition in numerical weather and climate prediction models and distributed observations of the surface temperature are, therefore, of great value for building the initial temperature boundary conditions. This paper presents an operational IST product available in near real time, based on METOP AVHRR satellite infra-red radiometers.

The drifting of Arctic sea ice constantly causes opening and closing of the sea ice cover and changes in ice cover of only a few percent can influence the heat flux between ocean and atmosphere drastically (Maykut, 1978; Marcq and Weiss, 2012). For a model to produce realistic initial surface temperature boundary fields, detailed information of ice concentration and ice drift is needed. The ice concentration fields that are assimilated in, for example, the global deterministic NWP model at ECMWF, have uncertainties of up to 10 % (Andersen et al., 2006) and contribute to surface and air temperature uncertainties of several degrees (Lüpkes et al., 2008). A sparsely distributed Arctic buoy observation network can not resolve these variations on the spatial scales on which these changes are occurring, thus, emphasising the potential of using satellite observations to estimate Arctic ice surface temperatures. The extreme conditions in the Arctic

complicate the deployment of instruments and limit their life-time. Because of these difficulties the temperature observation coverage in the Arctic Ocean is very sparse. We identified 30 valid drifters with real-time data transmission during 2011, resulting in a density of in situ buoys in the Arctic ocean of approximately 1 per $500\,000\,\text{km}^2$. Furthermore, temperature measurements from drifters may be dubious, because the drifters nesting on the ice may be buried in snow or even be solar heated. Satellite observations of the snow and ice surface temperatures can complement the in situ observations in order to increase the coverage of surface temperature observations. The IST data analysed here are estimated using Thermal Infra-Red sensors (TIR) from the polar orbiting METOP satellite under clear sky conditions. The 6 GHz microwave radiometer data have elsewhere been used for IST estimation during all sky conditions, but these data provide an integrated snow pack temperature rather than the surface temperature, because of the microwave's penetration in to the snow and ice (Tonboe et al., 2011; Hwang and Barber, 2008). There are other satellite infra-red IST products like those based on MODIS data (Hall, 2004b) and the AVHRR Polar Pathfinder data by Fowler et al. (2012). These products have been validated and described by, for example, Scambos et al. (2006) and Hall et al. (2004b) and used for climate and case studies. The Pathfinder dataset is well suited for climatologically studies, but can not be used for recent or real-time ice surface temperature analysis, due to irregular dataset updates. Furthermore, the Polar Pathfinder spatial resolution is 5 km, which makes it less suitable for fine scale mapping and analysis. The MODIS IST product has very similar characteristics to the METOP IST product (see Sect. 6), with product timeliness and sensor continuity as the main differences. Timeliness and data continuity are essential issues for the model communities to setup data validation and assimilation schemes (Stammer et al., 2007). The MODIS sea ice products have a time lag of days, from observation to product availability, and timeliness of the present IST product is a couple of hours. The METOP AVHRR data stream that is used for this IST production is guaranteed continuity and is scheduled until at least 2020, in contrast to the MODIS data stream that will end with the current Aqua and Terra missions. No satellite IST products are, to our knowledge, used in Numerical Weather Prediction (NWP) models or in sea ice models despite a potential for improving the model predictions. We think that this may be due to the lack of highest quality, future proofed, fully validated operational IST products, in near real time. The objectives of this study is to present and validate a new high resolution (1 km) IST product for the Arctic, that meets these requirements.

The paper describes the composite Arctic Surface Temperature algorithm, that operates in the MyOcean Sea Ice and Wind TAC in near real time. The product is fully operational and intended for use by operational meteorological and oceanographic agencies. Twelve months of validation and comparison is performed versus 3 types of in situ observations: (1) a TIR radiometer mounted on the sea ice, (2) a 2 m temperature record from the Greenland ice cap Summit, and (3) drifting buoy and ship temperature records from the Arctic Ocean and surrounding seas. In addition, the IST data are also compared to model surface temperature fields from the global ECMWF weather prediction model. The paper is organised with Sect. 2 describing the study areas, followed by a presentation of the IST algorithm in Sect. 3. In Sect. 4 the in situ and NWP model data are presented along with the data match-up procedures. Results are shown in Sect. 5. Finally, the results are discussed in Sect. 6. Description of acronyms that are specific for this paper is given in Table 1.

2 Study areas

The work presented here show IST validation and comparison results from 3 different Arctic environments and in situ instruments: (1) drifter temperatures from sea ice in the Arctic Ocean and the seas around Greenland and the Canadian Archipelago, (2) radiometer temperatures from a study site in the Inglefield Bredning next to Qaanaaq in North East Greenland, and (3) air temperatures from the synoptic station on the Greenland ice cap Summit. The positions of sites and areas are illustrated in Fig. 1. The field site near Qaanaaq is indicated by a red rectangle and the location of Summit on the Greenland ice cap is indicated by the blue star. Photos from Summit and Inglefield Bredning are shown in Fig. 2.

The Arctic Ocean and the adjacent seas are partly covered with perennial ice and partly with seasonal ice. The sea ice surface is often associated with packed ice and open leads, occasionally allowing extensive flux of heat from the ocean to the atmosphere. Especially during cloud free periods in the Arctic winter the ice surface is colder than the air, due to long wave radiative cooling, but also the annual mean surface temperature is lower than the air temperature (Tonboe et al., 2011; Radionov et al., 1997). The surface temperature of the ice covered Arctic Ocean ranges between approximately 0 and $-50\,°\text{C}$.

The field work site in the fjord of Inglefield Bredning, where an in situ TIR radiometer was deployed during a 4 days field work, was located 4 km off the coast near the town of Qaanaaq. The sea ice on the fjord was level seasonal ice with complete ice cover (see Fig. 2). The temperature range during field work was -17 to $-24\,°\text{C}$. The study site is indicated with the blue star in Fig. 3, where the surface temperature is shown at two occasions in the Inglefield Bredning fjord system.

The third in situ temperature record included in this study is collected from the Summit synoptic meteorological station on the Greenland ice cap. This site is located at 3200 m altitude with small and relative homogeneous surface roughness (see Fig. 2) and with temperatures ranging from approximately $0\,°\text{C}$ to $-70\,°\text{C}$. Summit is indicated with a blue star in Fig. 1.

Fig. 1. Overview of the 3 data sites, Arctic Ocean and adjacent seas, Qaanaaq field site and Summit. The green and blue dots and tracks indicate positions of applied drifter and ship observations throughout the study period. The red rectangle indicates the position the Inglefield Bredning by Qaanaaq where the in situ radiometer was deployed. The blue star is the position of the Summit synoptic weather station.

3 Metop Arctic Surface Temperature product

The Metop AVHRR Arctic Surface Temperature product (MAST) is an integrated IST, marginal ice zone temperature and high-latitude/Sea Surface Temperature product (SST), developed in the MyOcean project in collaboration with other projects and made available through the Sea Ice and Wind Thematic Assembly Center (SIWTAC) since January 2011 (MyOcean_data, 2011). Two algorithms are deployed in MAST, the Metop IST algorithm (MIST) for sea ice temperature estimation and a high latitude SST algorithm.

The MAST product is intended for data assimilation schemes in ocean, ice and atmosphere models as a supplement to traditional drifter and air temperature measurements. Timeliness, resolution and accuracy is, therefore, considered important for the product development, which has led to a product timeliness of 2–3 h, spatial resolution of 1.1 km at nadir and temporal sampling frequency of up to 10 passes per day at $\pm 80°$ latitude and 14 daily passes at the poles. MAST data from the past month can be downloaded through ftp and thredds as level 2 data in 3-minute segments from the MyOcean data repository (MyOcean_data, 2011). Further documentation can be found at the MyOcean web page (MyOcean_doc, 2011).
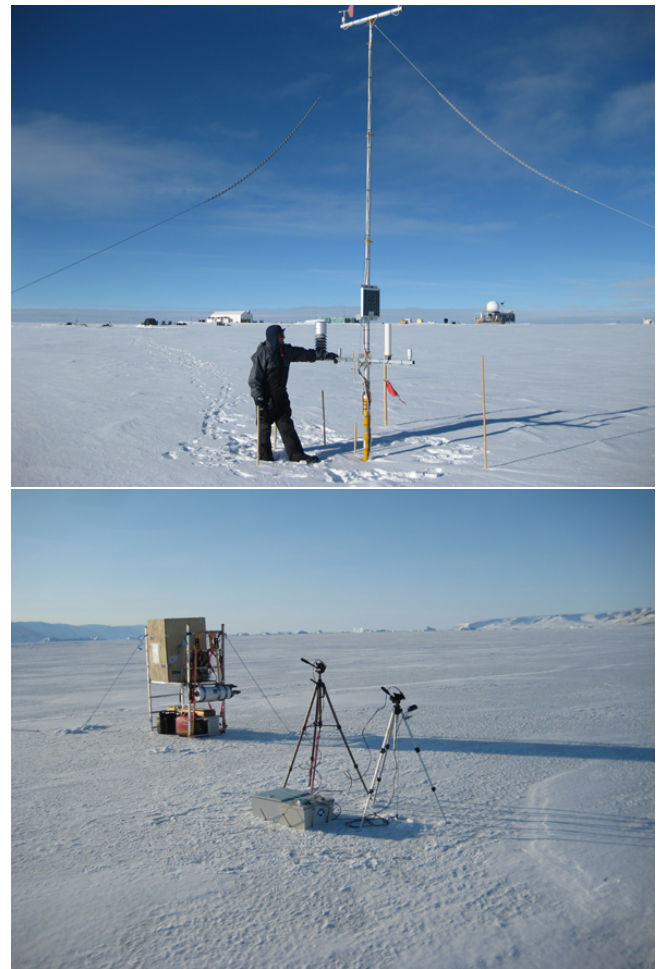

Fig. 2. The synoptic weather station at Summit (top). The photo is taken before the maintenance team has lifted the instruments to proper height after one year of snow fall. The ISAR radiometer set-up during measurements on the Inglefield Bredning fjord, in North East Greenland (white cylinder on scaffold, bottom photo).

The concept of MAST is taken from Vincent et al. (2008), where SST and IST algorithms are alternately deployed, depending on the AVHRR channel 4 brightness temperatures (T_{11}). For T_{11} temperatures warmer than $-2.2\,°\mathrm{C}$ the SST algorithm is deployed and for temperatures colder than $-4.2\,°\mathrm{C}$ the MIST algorithm is deployed. Surfaces with intermediate T_{11} temperatures are considered marginal ice zone and the marginal ice zone temperature is calculated from a linear combination of the MIST and the SST algorithms, scaled by the T_{11} temperature. The MIST algorithm and calibration is adopted from Key et al. (1997):

$$\mathrm{MIST} = a + b \cdot T_{11} + c \cdot (T_{11} - T_{12}) + d \cdot [(T_{11} - T_{12}) \cdot S(\mathrm{scan})] \quad (1)$$

where a-d are calibration constants based on a NOAA12 AVHRR calibration, T_{11} and T_{12} are AVHRR IR channels 4 and 5 (with centre wavelength at approximately 11 and 12 microns, respectively), $S(\mathrm{scan}) = \sec(\mathrm{scan}) - 1$ and scan

is the satellite scan angle. The calibration coefficients a-
d for Metop-AVHRR data are not available at the present
and the coefficients from the AVHRR instrument on board
NOAA12 were applied for this version of the algorithm.
The NOAA12 calibration coefficients are retrieved from
RTM modelled brightness temperatures for the AVHRR in-
frared channels and related to model skin temperatures (Key
et al., 1997). The channel centres, width and spectral re-
sponse functions of the NOAA12 and METOP AVHRR in-
struments are nearly identical. We, therefore, considered the
applied calibration equally valid for METOP AVHRR data
than for NOAA AVHRR data. Different sets of coefficients
are used for 3 brightness temperature intervals; see Key
et al. (1997) and Vincent et al. (2008). The T_{11} tempera-
ture intervals are: $-2.2\,°C > T_{11} > = -13.15\,°C$; $-13.15\,°C$
$> T_{11} > = -33.15\,°C$ and $-33.15\,°C > T_{11}$.

A description of the SST algorithm applied in the MAST
production is outside the scope of this paper. A detailed de-
scription of the SST algorithm and calibration is given in EU-
METSAT (2012) and a comprehensive validation and inter-
comparison analysis with other SST products is done by
Høyer et al. (2012).

A sample of the MAST product covering the Arctic, is
shown in Fig. 4 (top), as a six days mean temperature field
from March 3rd to March 9th, 2010. The corresponding sur-
face temperature field from the operational NWP model at
ECMWF is plotted in the bottom panel for comparison. The
NWP field is the mean of all bi-daily analysis fields from
all 6 days and the MAST field is the mean temperature from
the cloud free regions and periods only. The two temperature
fields are, therefore, not necessarily quantitatively compara-
ble. The temperature plots reveal that the general patterns
from the MAST data are also present in the NWP data and
that the MAST data are negatively biased in the central Arc-
tic. The latter can either illustrate an actual bias between the
2 temperature fields or an artefact caused by the difference in
data sampling methods.

3.1 Input TIR and cloud-flag data

The input TIR satellite data used by MAST and the cloud de-
tection algorithms are AVHRR swath data, received through
the EUMETCast global Metop data stream as 3 min seg-
ments. The 3 min segments are processed using the NWC
SAF PPS software (Thoss, 2009). Cloud-flags, sun-satellite
geometry information and AVHRR TIR data are subse-
quently used as input to the MAST processing chain and
as supplementary information in the match-up datasets, de-
scribed below. The spatial resolution of the cloud-flag data,
the AVHRR TIR data and the MAST product is 1.1 km at
nadir and approximately 2.5 km near the edges of the swath.

All MIST data used in the present analysis are associated
with the most likely clear sky cloud-flag or the second most
likely clear sky cloud-flag, "clear sky" or "clear sky, pos-
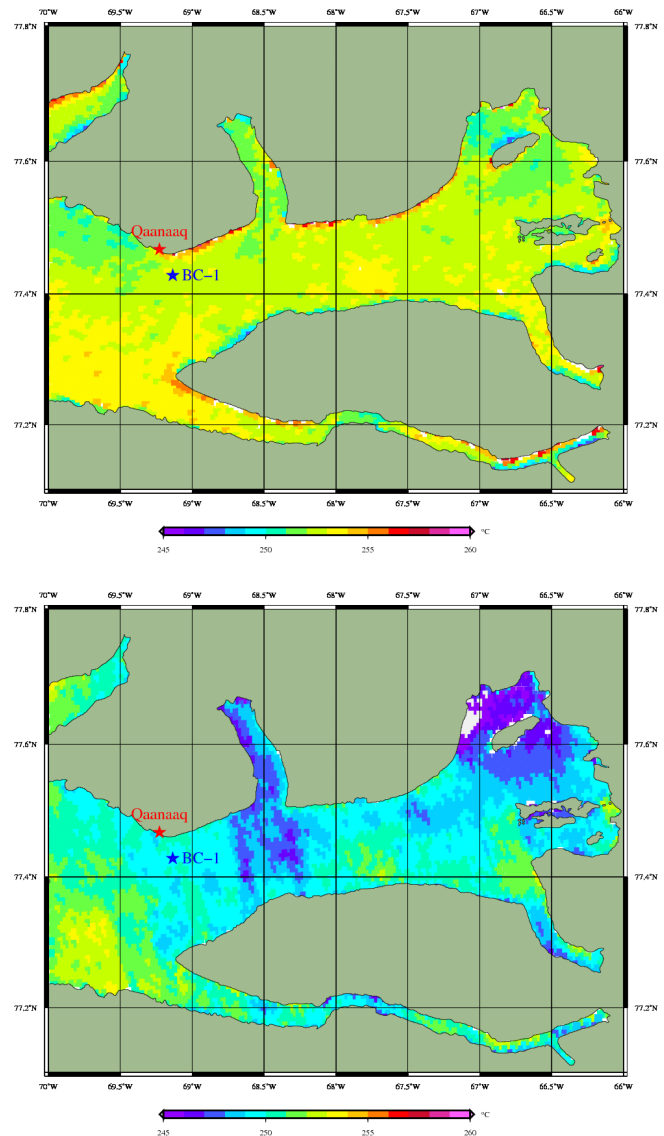sibly contaminated by surface ice or snow", the cloud-flags

Fig. 3. MIST temperature plot for Inglefield Bredning on 3 April:
top 13:45, bottom 22:22. The town of Qaanaaq and the position
of the ISAR in situ radiometer are marked with red and blue star,
respectively. This subset corresponds to the red rectangle in Fig. 1.
The colour bars range between $-28\,°C$ and $-13\,°C$.

11 and 14, respectively. The reason for including the second
best "clear sky" cloud category (cloud-flag 14) is to increase
data volume for the monthly error statistics shown in Fig. 5.
Annual statistics will be calculated for both cloud flag val-
ues and for cloud-flag 11 data only, to demonstrate the best
performance of MIST.

4 Data description, post processing and match up

Three types of in situ observations have been used in this
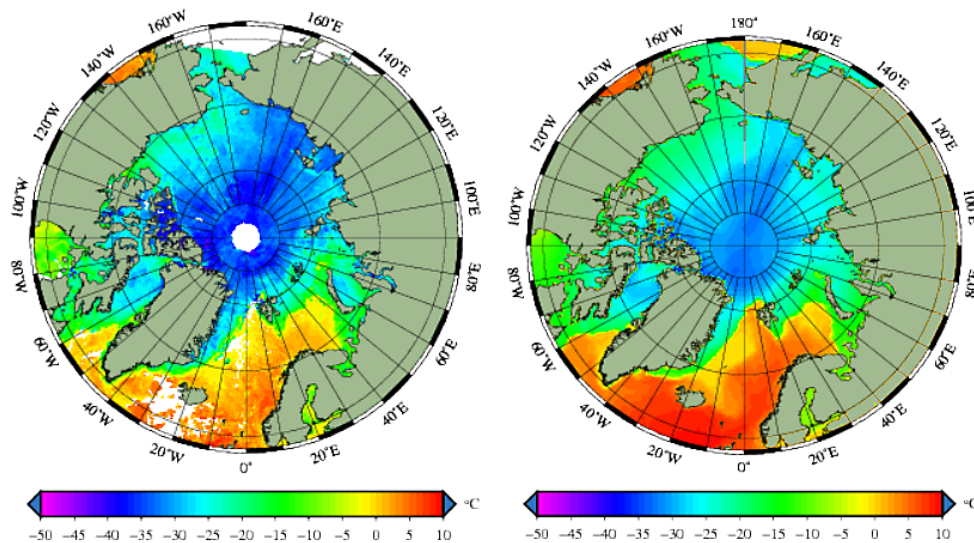study to provide more detailed error characteristics of MIST.

Fig. 4. Six day mean MAST surface temperature from March 3rd to 9th, 2010 (left) and corresponding ECMWF skin temperatures (right).

Table 1. Description of acronyms and abbreviations for data and datasets.

Data acronym's and abbreviations	
MIST	The METOP AVHRR Ice-Surface Temperature data
$MIST_{RE\_CAL}$	Re-calibration of MIST the OBS_{ISAR} in situ data
MU_{ARCTIC}	Arctic ocean match up dataset for MIST and OBS_{ARCTIC} data
MU_{ISAR}	Field work match up dataset for MIST and OBS_{ISAR} data
MU_{SUMMIT}	Summit match up dataset for MIST and OBS_{SUMMIT} data
OBS	Either or all of the 3 applied observation datasets
OBS_{ARCTIC}	In situ temperature data from ships and buoys collected via GTS data transmission system
OBS_{SUMMIT}	In situ temperature data from synop station 04416 (Greenland Summit) collected via GTS data transmission system
OBS_{ISAR}	In situ temperature data from portable thermal-infrared radiometer
NWP	Numerical Weather Prediction – general term
$NWP_{SURFACE}$	NWP Ice surface temperature from current global deterministic model at ECMWF
NWP_{2MT}	NWP 2 meter air temperature from current global deterministic model at ECMWF

In addition to in situ observations, also model analysis temperature fields are included for comparison. Finally, algorithm re-calibration and automated quality filters have been tested to check MIST sensitivity to calibration and to data outliers. A Match-Up (MU) dataset for each of the three in situ datasets is generated and is the basis for the validation and comparison exercise.

4.1 In situ observations

In situ temperature observations from drifters and ships (hereafter denoted OBS_{ARCTIC}) from the GTS data stream (GTS, 2012), are used to validate MIST on sea ice in the Arctic Ocean and adjacent seas. The OBS_{ARCTIC} were initially collected without quality check and data filtering was performed subsequently. Data appearing on blacklists from the UK Met Office are rejected (Parrett, 2011) and observations colder than $-70\,°C$ and warmer than $-1\,°C$ are removed. In addition, all drifter and buoy data with non-physical variability are removed. This three-step procedure reduces the number of automatically registered drifter and ship platforms in the GTS data stream, north of $70°\,N$ throughout 2011, from 56 to 30.

In situ observations from the synoptic station on Greenland Summit (WMO-04416) have also been used. The Summit station is a standard WMO synoptic station with 2 m air temperature measurements, at a fixed position in 3200 m altitude. These observations have also been obtained from the GTS and were checked visually. No data were removed. These observations are denoted OBS_{SUMMIT}.

Finally, four days of in situ TIR radiometer measurements from cloud free and nearly cloud free conditions have been collected. The in situ radiometer data, subsequently denoted OBS_{ISAR}, are obtained from an ISAR radiometer mounted on levelled first-year ice in the Inglefield Bredning, next to the town of Qaanaaq in North-East Greenland (Dybkjær et al., 2011). The ISAR instrument is a narrowband self calibrating single channel sensor, developed at National Oceanography Center, Southampton (NOCS). It is comparable to the channel 4 of the AVHRR sensors and provides an accurate observation of the surface skin temperature, accounting for the contribution from the sky (Donlon et

al., 2008). The emissivity used to convert brightness temperature to ice surface temperatures including atmospheric reflection is 0.99. This corresponds to sea water emissivity for a target angle of 25 degrees, and it is in agreement with values of sea ice emissivity used by Dozier and Warren (1982) and Key and Haefliger (1992).

The different in situ data sources can result in rather dubious validation results as surface and air temperatures can differ by several degrees. This is discussed in Sect. 6.

4.2 NWP data

Model fields of sea ice surface temperature ($NWP_{SURFACE}$) and 2 m air temperature (NWP_{2MT}) have been retrieved from the European Center of Medium-Range Weather Forecasts (ECMWF) as auxiliary data in the error analysis. The $NWP_{SURFACE}$ and NWP_{2MT} data are model analysis fields from the current global model (ECMWFdoc, 2012). The data are re-sampled to a regular 0.5° grid. All 00:00z and 12:00z analysis fields are used.

4.3 Ice concentration

The ice concentration data used in the match-up procedure (see below) are the 10 km sea ice concentration fields from OSISAF sea ice project (OSISAF, 2011).

4.4 Quality filtering

Erroneous outliers, often caused by non-detected clouds, are inevitable in present and similar datasets, due to the opacity of the atmosphere to TIR data. We test a simple quality filter based on the residual between NWP and MIST data. The NWP based quality filter removes all records with MIST-NWP errors larger (smaller) than $\pm 3 \cdot$ standard deviation of the mean MIST-NWP error. $NWP_{SURFACE}$ data were used to filter the MIST-OBS_{ARCTIC} data, removing 121 records of the 7930 cloud flag 11 records (see Fig. 8). NWP_{2MT} data were used to filter the MIST-OBS_{SUMMIT} data, removing 17 records of the 607 cloud flag 11 records (see Fig. 9).

4.5 Re-calibration tests

To assess errors associated with the use of NOAA 12 calibration coefficients in the MIST algorithm, two re-calibration tests of the MIST algorithm were performed against OBS_{ARCTIC} and OBS_{ISAR} data. Hence, the re-calibration is not performed to establish new calibration coefficients, but to compare the best possible empirical calibration from the Arctic buoy and the ISAR measurements to the operating setup. If the re-calibration tests do not improve the performance significantly, the dominant errors are associated with other issues than algorithm calibration. The re-calibrations are determined from least square fit to in situ datasets and the biases of these data are consequently zero. Improvements of the MIST quality are, therefore, solely

assessed by the standard deviation of errors (STDE). Re-calibrated MIST data are indicated $MIST_{RE\_CAL}$.

4.6 Match-up criteria

The MIST satellite data are matched with OBS data and auxiliary information like time lag, distance to observation, ice concentration, NWP temperatures, AVHRR brightness temperatures, and scan and sun angles. The match-up procedure varied between the different types of observations. Therefore, the match-up datasets were treated separately for the different types of in situ observations. The match-up dataset with drifter and ship observations covering the Arctic Ocean and adjacent seas is abbreviated MU_{ARCTIC}. The dataset containing MIST match-up with in situ radiometer data from the field work in Inglefield Bredning is called the MU_{ISAR} dataset and the MIST match-up with Summit air temperatures is subsequently denoted MU_{SUMMIT}. The match-up criteria are:

MU_{ARCTIC}:

– Period: 11 months – February to December 2011.

– Maximum time lag: 1 h

– Maximum distance to OBS: 2 km

– Maximum scan-angle: 45 degrees

– Maximum MIST value: −4.2 °C (see MIST definition above)

– Minimum ice concentration: 90 %

– Cloud flag 11 and 14 allowed

MU_{ISAR}:

– Period: 4 days in April 2011.

– Maximum time lag: 2 min

– Maximum distance to OBS: 2 km

– Cloud conditions were clear or almost clear and ice cover 100 % during data recording.

MU_{SUMMIT}:

– Period: All 2011.

– Maximum time lag: 1 h

– Maximum distance to OBS: 2 km

– Maximum scan-angle: 45 degrees

– Maximum MIST value: −4.2 °C (see MIST definition above)

– Cloud flag 11 and 14 allowed

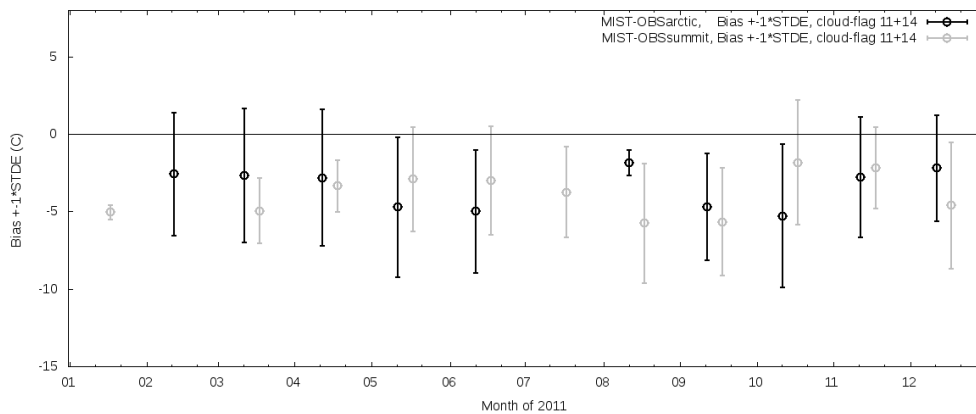Abbreviation and acronyms for data and datasets are listed and briefly explained in Table 1.

Fig. 5. Monthly mean bias values of MIST minus OBS_{ARCTIC} and OBS_{SUMMIT} are plotted with open black and grey squares, respectively, and error bars are ± 1 STDE. Insufficient match-up data for January and July was found to substantiate robust error analysis for MIST-OBS_{ARCTIC} comparison and likewise for February for the MIST-OBS_{SUMMIT} comparison.

5 Results

The MU_{ARCTIC} dataset contains more than 20 000 records complying with the match-up criteria described above. Of these data pairs there are up to 16 MIST records for each OBS record, because of the 2 km square search radius match-up criteria. This validation strategy was based on experience from the MU_{ISAR} data. The MU_{ISAR} error statistics was analysed using both mean of all individual data pairs, mean and median MIST values, without clear indications of a best measure. Thus, it was decided to treat all MIST-OBS data pairs individually. The validation and inter-comparison results are divided into annual and monthly data and differences are described by their standard deviation of error (STDE), bias and correlation coefficient (R).

The initial MIST performance is based on the entire match-up datasets. Mean monthly error statistics from the MU_{ARCTIC} and MU_{SUMMIT} datasets are plotted in Fig. 5 and the corresponding initial quality of the full match-up datasets is written in Table 2. The annual STDE of MIST-OBS_{ARCTIC} is 4.29 °C with a cold bias of −3.43°C. From the monthly error statistics in Fig. 5, we find the smallest errors during the Arctic summer where cloud detection algorithms in general seem to perform best. The quality of the re-calibrated MU_{ARCTIC} dataset showed practically no STDE reduction. A change from 4.29 °C to 4.27 °C (not shown), indicates that other sources of errors are much larger than the errors originating from the calibration of the MIST algorithm, i.e., indicating that the adopted NOAA 12 calibration coefficients work well for the METOP AVHRR instrument and that erroneous cloud screening is a dominating source of error. However, a reliable assessment of the algorithm calibration must be performed on a large dataset, with measured or manually screened cloud information. By applying the NWP quality filter to remove major outliers in cloud-flag 11 data of the

MU_{ARCTIC} dataset (see Fig. 8), the overall STDE improves to 3.69 °C (Table 3).

In contrast to the errors of the MU_{ARCTIC} dataset, we find highly accurate MIST match-up data with the in situ radiometer data in the MU_{ISAR} dataset. Here we also find further improvements in the re-calibrated data, thus, reducing STDE from 1.02 °C to 0.73 °C, for the MIST and $MIST_{RE\_CAL}$ data match-up with OBS_{ISAR} data, respectively (Tables 2 and 3).

The OBS_{ISAR}, MIST and $MIST_{RE\_CAL}$ data are plotted in Fig. 6, where data from all scan angles are shown to illustrate the diurnal temperature variation. The error bars on the MIST data represent the minimum and maximum values of all MIST values within the 2 km range of the ISAR observations. During the first two days of the match-up period the coherency between MIST and OBS_{ISAR} data is particularly high, whereas data from the last two days are slightly less correlated. This may coincide with optically thin atmospheric disturbances, as assessed from the less smooth ISAR data during that period. STDE for the $MIST_{RE\_CAL}$ data for the period 1 April to 2 April is 0.42 °C and bias is 0.32 °C.

A day and a night MIST snapshot from Inglefield Bredning are shown in Fig. 3, where also the location of the ISAR instrument during field work is marked *BC-1*. The two MIST plots are separated by approximately 10 h on 3 April, showing the mid-day situation in the top panel and the evening situation in the bottom panel. A close look at the day situation reveal heating of the South oriented steep and rocky coastline and relatively homogeneous surface temperatures elsewhere in the fjord. The evening plot shows general cooling of the sea ice surface with strong cooling in certain areas. These inhomogeneous cooling effects are most likely caused by advection of cold air from the glaciers.

The MU_{SUMMIT} match-up dataset is different from the 2 previously described datasets, in the sense that MIST data

Table 2. Preliminary MIST error statistics for the complete MU_{ARCTIC}, MU_{ISAR} and MU_{SUMMIT} datasets.

Initial quality	STDE (C)	Bias (C)	R	Counts
MIST-OBS$_{ARCTIC}$	4.29	-3.43	0.91	21251
MIST-OBS$_{ISAR}$	1.02	-1.81	0.77	28
MIST-OBS$_{SUMMIT}$	3.48	-3.35	0.96	1578

Table 3. Best quality MIST error statistics. The NWP based quality filter is applied to cloud flag 11 data of the MU_{ARCTIC} and MU_{SUMMIT} datasets (see Figs. 8 and 9), and the MU_{ISAR} dataset is re-calibrated.

Best quality	STDE (C)	Bias (C)	R	Counts
MIST-OBS$_{ARCTIC}$	3.69	-2.76	0.89	7809
MIST$_{RE\_CAL}$-OBS$_{ISAR}$	0.73	0	0.78	28
MIST-OBS$_{SUMMIT}$	3.14	-3.22	0.95	590

are matched with air temperatures from a high elevation site, at approximately 3200 m altitude.

In Fig. 5, STDE and bias values for the MU_{SUMMIT} dataset are plotted as monthly mean values, showing a year round negative bias between approximately $-2\,°C$ and $-5\,°C$, similar to bias values found for the MU_{ARCTIC} data, but with generally lower STDE values. The mean annual STDE and bias of the MIST-OBS$_{SUMMIT}$ analysis is $3.48\,°C$ and $-3.35\,°C$, respectively (Table 2). As was the case for the MU_{ARCTIC} error analysis, a markedly improvement of MIST performance was found for quality filtered cloud-flag 11 data from the MU_{SUMMIT} dataset (Fig. 9), namely to STDE of $3.14\,°C$ and a bias of $-3.22\,°C$ (Table 3).

A comparison of the Summit air temperatures (OBS$_{SUMMIT}$) with the corresponding NWP_{2MT} values, reveals very large annual error of $5.71\,°C$ (Table 4), emphasising a need for additional ground truth to generate realistic model analysis fields. The NWP_{2MT}-OBS$_{SUMMIT}$ bias is small as expected, because the OBS$_{SUMMIT}$ are one of the few 2 m temperature observations on the Greenland ice cap that is available for model data assimilation.

From the MU_{ARCTIC} dataset, we have also calculated the annual error statistics between MIST and NWP data. The results are shown in Table 4 showing MIST-$NWP_{SURFACE}$ / NWP_{2MT} bias values, around $-3.5\,°C$ and STDE values of $3.92\,°C$ and $3.49\,°C$, respectively. Also in Table 4, we see that the biases between $NWP_{SURFACE/2MT}$-OBS are small, as the OBS$_{ARCTIC}$ data are the in situ surface temperature data used to build the NWP analysis fields.

6 Discussion

When comparing remotely sensed data with ground measurements, it is assumed that the spatial and temporal characteristics of a given parameter are comparable, regardless

of the parameter measured from space or on the ground. In this case, where a temperature estimate representing more than $1\,km^2$ is compared to a point measurement, it is assumed that the autocorrelation length of the surface temperature is larger than the satellite footprint, and similarly that the temporal autocorrelation of the surface temperature is longer than the MIST sampling frequency. Veihelmann et al. (2001) estimated the standard deviation of the surface temperature inside a $4.5\,km^2$ area in the Weddell Sea to be approximately $0.5\,°C$. The corresponding value inside the double search radius (a 4 by $4\,km^2$ square) was calculated to be approximately $1\,°C$, based on data from the Qaanaaq field experiment. Also from the MU_{ISAR} dataset, we estimate the maximum temporal temperature gradient to be approximately $0.9\,°C\,h^{-1}$, during sunrise. The temporal and spatial sampling issues contribute to the overall MIST error, but they are assumed not to contribute to bias. Due to the relatively rigid match-up criteria used to generate the match-up datasets, the sampling errors are estimated to be around $1\,°C$ for the MU_{ARCTIC} and MU_{SUMMIT} datasets, and less for the MU_{ISAR} dataset, because of practically no time lag between MIST and OBS$_{ISAR}$ data.

Spatial variance of snow and ice surface emissivity is another issue that contribute to IST estimation errors. In earlier works by Warren (1982) and Dozier and Warren (1982) emissivity variations caused by snow grain size and liquid water content were considered negligible, and only a slightly decreasing impact from increasing snow pack density was identified. Emissivity may decrease approximately by 0.005 when the snow density increases from about $200\,kg\,m^{-3}$ to $300\,kg\,m^{-3}$ (Dozier and Warren, 1982). Dozier and Warren (1982) considered the view angle to be the most important variable for emissivity variations. In more recent works by Cheng et al. (2010) and Salisbury et al. (1994) it is acknowledge that also increasing snow grain size have markedly lowering effects on the snow emissivity for the TIR wavelength used here. At nadir a grain size increase from 300 to 550 microns can decrease the emissivity by approximately 0.005, thus, adding another $\sim 1\,°C$ to uncertainty on an IST estimate. Distributed information of snow density and grain size on Arctic scale does not exist and empirical IST algorithms adapt to average snow properties. Associated errors are, therefore, anticipated; no biases from these errors are expected.

The MIST algorithm seems to account successfully for scan angle dependent emissivity. Figure 7 shows the contribution of the scan-angle term of the MIST algorithm as a function of scan-angle. The slightly quantified angular correction is caused by the different calibration constants used for each of the T_{11} regimes in which MIST is working. The spread around the main lines is induced by the $T_{11} - T_{12}$ factor of the MIST algorithm. The average temperature correction at 45 degree scan-angle is approximately $1\,°C$, which is in good agreement with corresponding angular emissivity reduction found for the 11 micron channel on ATSR (Stroeve
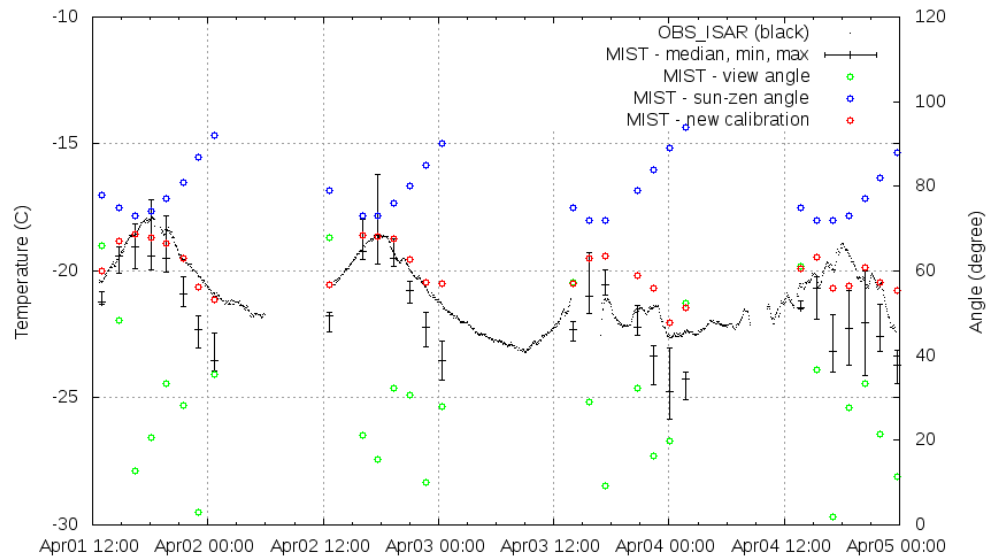
Fig. 6. OBS$_{ISAR}$ (thin black dots) and MIST (black crosses with error bars) temperatures from Qaanaaq field experiment, 2011. MIST data are plotted as the median value of all MIST measurements inside 2 km of the ISAR instrument and error bars are the corresponding minimum and maximum value. Sun-zenith and scan angles are indicated with blue and green dots, respectively. MIST$_{RE\_CAL}$ data are plotted as average of all used pixels with red circles. Maximum diurnal temperature amplitude is 4.6 K from midday 2nd April to morning 3rd April.

et al., 1996). It is essential to mention that MIST errors are uncorrelated to scan-angle (not shown).

The assumed three largest error contributors to satellite based IST estimates are erroneous cloud detection, algorithm errors from general simplification and the in situ data errors. With respect to the latter, the radiometric surface temperature can be significantly different from the thermodynamic temperature measurements from drifters and ships. This difference is largest in cloud free conditions caused by long-wave radiative cooling, and maximum differences are measured to range between 4 and 7 °C by Radinov et al. (1997) and Veihelmann et al. (2001) and confirmed by model estimates by Tonboe et al. (2011). On average, the surface is colder than the air (Maykut, 1986; Radinov, 1997). Hence, we expect a physically induced negative MIST-OBS$_{ARCTIC}$/OBS$_{SUMMIT}$ bias, but we do not have sufficient documentation to quantify this. A quantification of the surface-air temperature difference is further complicated because the buoys do not necessarily measure the air temperature. A buoy thermometer can be buried in snow and, thus, measure internal snow temperature or a thermometer inside a buoy can be warmed up by radiative heating from the sun on the buoy housing (Key and Haefliger, 1992).

The presence of non-detected clouds will contribute to increased STDE and will in general result in a cold MIST bias, because cloud tops and other atmospheric constituents in general are colder that surface and air temperatures. Figure 8 is a scatter plot of OBS$_{ARCTIC}$ data versus MIST for cloud flag 11 data from the MU$_{ARCTIC}$ dataset. The scatter of the data show a clear cold bias of MIST and a number of extreme outliers probably from erroneously cold cloud contaminated
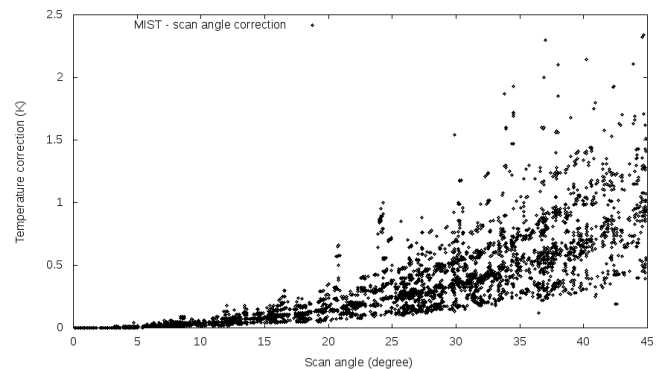

Fig. 7. Magnitude of the MIST scan-angle correction term from the MU$_{ARCTIC}$ dataset, plotted as a function of scan-angle.

MIST estimates. For very cold surface temperatures the scatter is small as these temperatures most likely are associated with clear skies where MIST is working well. The cold MIST bias for the extreme low temperatures is unambiguous and most likely reflecting real physical properties and, thus, confirming the discussion of the negative long-wave surface energy balance.

By imposing the NWP based quality filter to cloud-flag 11 data, from the MU$_{ARCTIC}$ dataset, the annual mean STDE improves from 4.29 °C to 3.69 °C and bias changes from −3.43 °C to −2.76 °C. The data removed by the quality filter are shown in Fig. 8 by red circles. The quality filter alone lowered the STDE value by 0.3 °C, and the rest of the error improvement is derived from the exclusion of cloud flag
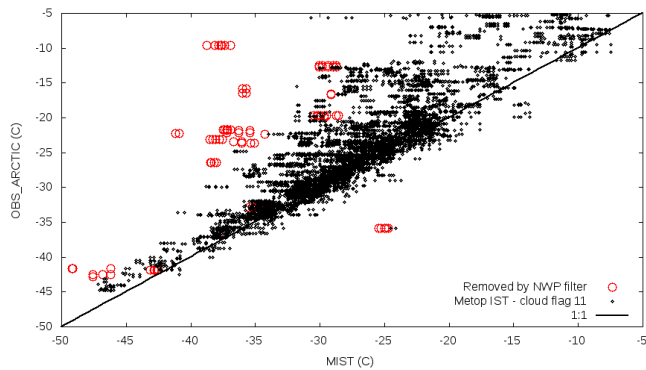
Fig. 8. Scatter plot of OBS$_{ARCTIC}$ as a function of MIST, based on MU$_{ARCTIC}$ data and for cloud flag 11 data. Red circles indicate data removed by the NWP based quality filter. The black line is the 1 : 1 line.

Fig. 9. Scatter plot of OBS$_{SUMMIT}$ as a function of MIST, based on MU$_{SUMMIT}$ dataset and for cloud flag 11 data. Red circles indicate data removed by the NWP based quality filter. The black line is the 1 : 1 line.

14 data. The re-calibration of the cloud flag 11 data against OBS$_{ARCTIC}$ data showed only negligible improvements of STDE, indicating that the errors caused by calibration inaccuracy are small compared to other errors. A product similar to MIST, the MODIS 1 km IST product, performs similarly with STDE of 3.7 °C and an associated bias of −2.1 °C (Hall et al., 2004b). This comparable performance indicates the level of quality that can be expected from fully automated satellite IST products.

Error statistics from the MU$_{ISAR}$ dataset is relevant in context of the error discussion above. Despite the limited amount of data available for this analysis, it reveals the potential and limitations of surface temperature estimation from space-borne TIR radiometers. The MU$_{ISAR}$ data are collected along with manual cloud screening, with no time lag between satellite and in situ observation, snow and ice surface conditions are relatively homogeneous (see photo in Fig. 2) and both MIST and in situ observations are skin temperatures. Main errors are, therefore, assumed to originate from the spatial sampling of MIST, to some extent also from varying snow properties and from a non-optimal calibration. The STDE and bias of the MU$_{ISAR}$ dataset is 1.02 °C and −1.81 °C, respectively. The re-calibrated data, MIST$_{RE\_CAL}$, clearly shows improved performance with STDE of 0.73 °C (Table 3) and even as low as 0.42 °C for the 2 first days of the MU$_{ISAR}$ dataset (see Fig. 6). This significant quality improvement obtained from the re-calibration, suggests that the current MIST calibration is not optimal. However, substantially more data points are needed in order to conclude on this point and, thus, to carry out a new calibration. Similar improvement from the re-calibration of the MU$_{ARCTIC}$ dataset was not found, implying that a proper re-calibration of the MIST algorithm must be based on visually cloud screened in situ surface temperature data.

In similar experiments with IST data from MODIS, AVHRR and ATSR compared to in situ TIR radiometer data, the corresponding errors were 1 °C, 1.4 °C and 0.2 °C (Scam-
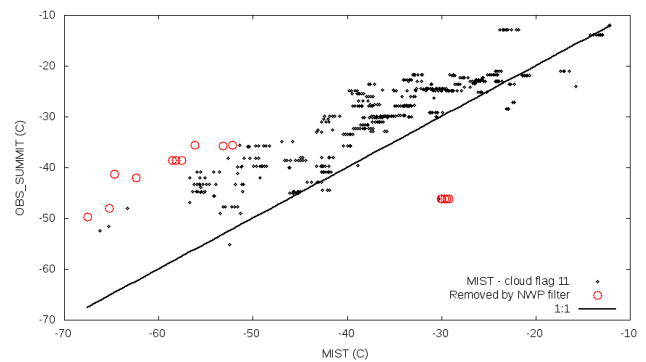
bos et al., 2006; Stroeve et al., 1996). These coherent results from comparable case studies substantiate the assumption that the quality of a well-calibrated IST algorithm basically comes down to proper match-up routines and to high quality cloud masking procedures.

NWP models have problems reproducing realistic temperature variability in the analysis fields for the Greenland ice cap as evident from the comparison to the Summit 2 m temperature measurements in Table 4. The reason is that vast areas of the ice cap are represented by a poorly distributed observation network. Most operational synoptic stations on Greenland are located along the coastline several hundred kilometres from Summit. The objective of comparing MIST data to air temperature on the Greenland ice cap is, therefore, to examine the feasibility of using MIST data in model assimilation schemes for snow and ice covered Arctic land areas. An inter-comparison of surface temperature observations, NWP skin temperatures and MIST would be ideal, but the only available long-term in situ observations on Summit are 2 m temperature and air pressure. However, 2 m and surface temperatures are comparable at Summit, due to level and homogeneous surface conditions, which results in a very high correlation between Summit 2 m and surface temperatures (Hall et al., 2004a).

Unlike the conditions at sea level, one can expect relative constant snow grain size distribution on the Greenland ice cap because of the persistent cold conditions and constant wind stress (Stroeve and Steffen, 1998). Furthermore, we also expect less interference with the atmosphere and clouds at Summit. These factors are assumed to contribute to smaller MIST errors on the ice cap than on the sea ice, and the less pronounced cloud cover will also reduce bias. The errors of the MIST-OBS$_{SUMMIT}$ comparison confirmed this, with an annual STDE of 3.14 °C, but the bias of −3.22 °C is slightly higher than the of the MU$_{ARCTIC}$ data (Table 3). A similar study of surface and air temperatures at Summit, using a visual cloud screening procedure, showed an annual

Table 4. Error statistics for MIST and OBS comparison with NWP data. The MIST statistics is based on cloud mask flag 11 data from the MU$_{ARCTIC}$ dataset.

	STDE (C)	Bias (C)	R
MIST-NWP$_{SURFACE}$	3.92	-3.50	0.81
MIST-NWP$_{2MT}$	3.49	-3.43	0.85
NWP$_{SURFACE}$-OBS$_{ARCTIC}$	4.58	0.60	0.86
NWP$_{2MT}$-OBS$_{ARCTIC}$	3.55	0.53	0.92
NWP$_{2MT}$-OBS$_{SUMMIT}$	5.71	0.70	0.93

surface-air temperature bias of $-1\,°C$ (Hall et al., 2004a). The discrepancy between the bias of $-1\,°C$, observed by Hall et al. (2004a) and the annual bias of this study, underline the significant impact of cloud screening.

Stammer et al. (2007) considered $4\,°C$ to be the error threshold for model assimilation schemes to benefit from satellite based IST, as complementary temperatures to the traditional observation network. MIST is well below that threshold and clearly below the STDE of $5.71\,°C$ from the NWP$_{2MT}$ comparison with OBS$_{SUMMIT}$.

7 Future works

Future development and operations of this product will be a split between MyOcean2 and the EUMETSAT's Ocean and Sea Ice–SAF (OSISAF). MyOcean2 will operate the integrated MAST algorithm to a level 3 product for the Baltic Sea and the MAST set up for the Arctic will migrated to the OSISAF as an operational level 2 production. Further development of MAST will be done in the OSISAF. This will include a future re-calibration of the IST algorithm, enhancement of cloud screening procedures and testing of other surface type classification procedures to optimise algorithm decision making.

Edited by: P.-Y. Le Traon

References

Andersen, S., Tonboe, R., Kern, S., and Schyberg, H.: Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using numerical weather prediction model fields: An intercomparison of nine algorithms, Remote Sens. Environ., 104, 4, 374–392, doi:10.1016/j.rse.2006.05.013, 2006.

Bitz, C. and Lipscomb, W.: An energy-conserving thermodynamic model of sea ice, J. Geophys. Res.-Oceans, 104, 15669–15677, 1999.

Björk, G. and Söderkvist, J.: Dependence of the Arctic Ocean ice thickness distribution on poleward energy flux in the atmosphere, J. Geophys. Res.-Oceans, 107, 3173, doi:10.1029/2000JC000723, 2002.

Cheng, J., Liang, S., Weng, F., Wang, J., and Li, X.: Comparison of Radiative Transfer Models for Simulating Snow Surface Thermal Infrared Emissivity, IEEE J. Sel. Top. Appl., 3, 323–335, 2010.

Donlon, C., Robinson, I. S., Wimmer, W., Fisher, G., Reynolds, M. R., Edwards, R. T. J., and Nightingale, T. S.: An infrared sea surface temperature autonomous radiometer (isar) for deployment aboard volunteer observing ships (vos), J. Atmos. Ocean Tech., 25, 93–113, 2008.

Dozier, J. and Warren, S. G.: Effect of Viewing Angle on the Infrared Brightness Temperature of Snow, Water Resour. Res., 18, 1424–1434, 1982.

Dybkjær, G., Høyer, J., Tonboe, R., Olsen, S., Rodwell, S., Wimmer, W., and Søbjærg, S.: QASITEEX 2011 – The Qaanaaq sea ice thermal emission experiment Field report, Danish Meteorological Institute, Technical Report, tr11-18, 27 pp., available at: http://www.dmi.dk/dmi/tr11-18.pdf, 2011.

ECMWFdoc: ECMWC atmospheric model cycle descriptions, http://www.ecmwf.int/products/data/technical/model_id/index.html, last access: 22 October 2012, 2012.

EUMETSAT: Low Earth Orbiter Sea Surface Temperature Product User Manual, SAF/OSI/CDOP/M-F/TEC/MA/127 available at: http://www.osi-saf.org, last access: 23 October 2012.

Fichefet, T. and Maqueda, M. A. M.: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics, J. Geophys. Res.-Oceans, 102, 12609–12646, 1997.

Fowler C., Maslanik, J., Emery, W., Haran, T., Scambos, T., and Key, J.: Polar Pathfinder EASE-grid Composites, available at: http://nsidc.org/data/docs/daac/nsidc0066_avhrr_5km.gd.html, last access: 23 October 2012.

GTS: MANUAL ON THE GLOBAL TELECOMMUNICATION SYSTEM, WMO Publication No. 386, http://www.wmo.int/pages/prog/www/TEM/GTS/ManOnGTS_en.html, last access: 22 October 2012.

Hall, D. K., Williams, R. S., Steffen, K., and Chien, J. Y. L.: Analysis of Summer 2002 Melt Extent on the Greenland Ice Sheet using MODIS and SSM/I Data, Geoscience and Remote Sensing Symposium, 2004 IGARSS '04 Proceedings, 2004 IEEE International, 5, 3029–3032, 2004a.

Hall, D. K., Key, J. R., Casey, K. A., Riggs, G. A., and Cavalieri, D. J.: Sea Ice Surface Temperature Product from MODIS, IEEE T. Geosci. Remote., 42, 1076–1087, 2004b.

Hwang, J. H. and Barber, D. G.: On the Impact of Ice Emissivity on Sea Ice Temperature Retrieval Using Passive Microwave Radiance Data, IEEE T. Geosci. Remote Lett., 5, 448–452, 2008.

Høyer, J. L., Karagali, I., Dybkjær, G., and Tonboe, R.: Multi sensor validation and error characteristics of Arctic satellite sea surface

temperature observations, Remote Sens. Environ., 121, 335–346, doi:10.1016/j.rse.2012.01.013, 2012.

Key, J. and Haefliger, M.: Arctic Ice Surface Temperature Retrieval from AVHRR Thermal Channels, J. Geophys. Res.-Oc. Atm., 97, 5885–5893, 1992.

Key, J. R., Collins, J. B., Fowler, C., and Stone, R. S.: High-Latitude Surface Temperature Estimates from Thermal Satellite Data, Remote Sens. Environ., 61, 302–309, 1997.

Lüpkes, C., Vihma, T., Birnbaum, G., and Wacker, U.: Influence of leads in the sea ice on the temperature of the atmosphere boundary layer during polar night, Geophys. Res. Lett., 35, L03805, doi:10.1029/2007GL032461, 2008.

Marcq, S. and Weiss, J.: Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean and the atmosphere, The Cryosphere, 6, 143–156, doi:10.5194/tc-6-143-2012, 2012.

Maykut, G. A.: Energy Exchange Over Young Sea Ice in the Central Arctic, J. Geophys. Res.-Ocean, 83, 3646–3658, doi:10.1029/JC083iC07p03646, 1978.

Maykut, G. A.: The surface heat and mass balance. In: N. Untersteiner (Ed.): The geophysics of sea ice. NATO ASI Series, plenum Press, New York and London, 395–464, 1986.

MyOcean_data: MyOcean SIWTAC data repository, http://thredds.met.no/thredds/catalog/myocean/siw-tac/siw-dmi-temp-arc/catalog.html, last access: 23 October 2012, 2011.

MyOcean_doc: MyOcean interactive data catalog, http://www.myocean.eu/web/24-catalogue.php last access: 23 October 2012, 2011.

OSISAF: Ocean and Sea Ice, Satellite Application Facility, http://www.osi-saf.org and http://osisaf.met.no, 2011.

Parrett, C.: Drifter and Ship blacklist, UK meteorological Office, personal communication, Oktober 2011.

Radionov, V. F., Bryazgin, N. N., and Alexandrov, E. I.: The Arctic Snow Cover of the Arctic Basin, Arctic and Antarctic Research Institure, Russia, Translation from Department of Science, Technical Report, APL-UW TR 9701, University of Washington, 1997.

Salisbury, J. W., D'Aria, D. M., and Wald, A.: Measurements of thermal infrared spectral reflectance of frost, snow, and ice, J. Geophys. Res.-Sol. Ea., 99, 24235–24240, 1994.

Scambos, T. A., Haran, T. M., and Massom, R.: Validation of AVHRR and MODIS ice surface temperature products using in situ radiometers, Ann. Glaciol., 44, 345–351, 2006.

Stammer, D., Johanessen, J., LeTraon, P.-Y., Minnett, P., Roquet, H., and Srokosz, M.: Requirements for Ocean Observations Relevant to post-EPS, EUMETSAT Position Paper: AEG Ocean Topography and Ocean Imaging, 10 January 2007, version 3, 2007.

Steffen, K., Abdalati, W., and Stroeve, J.: Climate sensitivity studies of the Greenland ice sheet using satellite AVHRR, SSMR, SSM/I, and in situ data. Meteorol. Atmos. Phys., 51, 239–258, 1993.

Stroeve, J. and Steffen, K.: Variability of AVHRR-Derived Clear-Sky Surface Temperature over the Greenland Ice Sheet, J. Appl. Meteorol., 37, 23–31, 1998.

Stroeve, J., Haefliger, M., and Steffen, K.: Surface Temperature from ERS-1 ATSR Infrared Thermal Satellite Data in Polar Regions, J. Appl. Meteorol., 35, 1231–1239, 1996.

Thoss, A.: Algorithm Theoretical Basis Document for SAFNWC/PPS Cloud Mask. Nowcasting and very short range forecasting, Satellite Application Facility, EUMETCAST, SAF-NWC-CDOP-SMHI-PPS-SCI-ATBD-1_v2_2_1.doc, available at: http://research.metoffice.gov.uk/research/interproj/nwpsaf/index.html, last access: 23 October 2012, 2009.

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

Veihelmann, B., Olesen, F. S., and Kottmeier, C.: Sea Ice surface temperature in the Weddell Sea (Antarctica), from drifting buoy and AVHRR data, Cold Reg. Sci. Technol., 33, 19–27, 2001.

Vincent, R. F., Marsden, R. F., Minnett, P. J., Creber, K. A. M., and Buckley, J. R.: Arctic waters and marginal ice zones: A composite Arctic sea surface temperature algorithm using satellite thermal data, J. Geophys. Res.-Oceans, 113, C04021, doi:10.1029/2007JC004353, 2008.

Warren, S. G.: Optical properties of Snow, Rev. Geophys. Space Ge., 20, 67–89, 1982.