General author response

We would like to thank both reviewers for their time and efforts put into reviewing this manuscript. There are many good comments and questions, which can lead to a much stronger and clearer manuscript.

However, there are some misunderstandings that we would like to clarify. The paper presents a clear sky T2m product based on clear sky satellite observations, similar to the clear sky T2m products which have previously been derived over ocean, lakes and land from satellite observations (e.g. Good et al., 2016). Obviously, we have failed to address this clearly, since it is pointed out by both reviewers (even though in the abstract (P1L14), we do state: "The satellite derived T2m product including estimated uncertainties covers clear sky snow and ice surfaces in the Arctic region during the period 2000-2009").

Below is a summary list of why we believe the paper and the derived T2m dataset are highly relevant for publication:

- Infrared surface temperature retrievals are not possible in cloudy conditions. This is similar to all other infrared retrievals of e.g. Sea Surface Tempeature and Land Surface Temperature (for the other surface types). These clear sky observations are used extensively in models and aggregated products and are among the data sources that give the largest impact and improvement on model forecasts.
- Similar satellite based T2m relationship models have been derived for land surfaces with great success and large uptake (Good, et al., 2016) and the investigation of the ISTskin versus T2m relationships over ice and the derivation of T2m from ISTskin were identified in Merchant et al. (2013) as very important areas for improving the understanding of the surface temperature of the Earth.
- The number of satellite observations in the Arctic region is much higher than what is obtained from traditional in situ observations.
- The product described here is not meant to replace in situ observations or existing T2m products, but should be considered as complementary to existing observations.
- We have added results demonstrating that that we have a 94% daily average coverage throughout the years 2003-2009 of T2m, representing all-sky temperatures for the GrIS, considering a 1x1 degree grid. For the sea ice region, the same number is 81%. The coverage is stable from 2003-2009 and somewhat lower before 2003.The 2003-2009 period is chosen here to represent the coverage in the recent satellite observing constellation.
- The days when the satellite derived T2m product is available, it represents the all-sky T2m, since it has been regressed towards in situ measurements obtained both in cloudy and clear sky conditions (see discussion in manuscript).
- We recognize that filling the gaps due to clouds is an important topic, which is outside the scope of this paper. The starting point, however, before producing a gap filled product is the product we describe here.
- The work is part of a Horizon 2020 project EUSTACE (lead by UK Metoffice) and will feed into a new global T2m analysis similar to e.g. the CRUTEMP and GISS, but also including satellite information (Rayner et al., 2019). The data set we present here will thus improve the global T2m estimates, which are among the most used and cited climate data sets worldwide.
- A revised discussion has been added to the methodology section, clarifying that the seasonal signals in the T2m regression models are included in all the models (through the IST relation) and that the extra terms in equations 9-12 represent the anomalies, as suggested by reviewer #2. This has been clarified in the text

Based upon the arguments listed above and the clarifications in the revised version (see attached manuscript with track-changes below), it is our sincere hope that the reviewers will realize the large potential and value of the T2m data set we describe here. Below, we have responded to the reviewers' comments point by point.

RC1 – Authors response

*Author response is blue.

The authors present a remote sensed 2m temperature product for the Arctic ocean and the Greenland Ice sheet. For this aim, they use Arctic and Antarctic Ice Surface Temperatures from thermal Infrared satellite Sensors (AASTI) data set and apply a correction to convert surface temperatures into 2 m temperatures. The derived temperatures are compared with in situ observations and the data set performance is compared to the performance of T2m from ERA-Interim.

First of all, I apologize for the long time I needed to complete this review, which added to the long time that the editor needed to find reviewers. Sadly, I'm not convinced that the authors properly resolve all the challenges that occur when remote sensed surface temperatures are converted into t2m temperatures.

My largest objection is that the authors fail to properly resolve the "cloud problem". Infrared satellites can only measure surface temperatures in cloud free conditions. However, cloudy conditions lead totally different weather than cloud-free conditions, especially in winter. Too little attention is paid to this problem in the methodology and results section. A publishable data set must resolve the cloud-data-gap problem and it should be explicitly shown that reasonable estimates can be provided for cloudy conditions, even if these conditions last for days or more. As the authors do not aim to present a clear-sky T2m product, which would be of limited value, this shortcoming in the methodology and discussion must be resolved prior acceptance can be considered.
We acknowledge and regret that it has not been stated clearly enough that we derive a clear sky T2m product (see general author response). Days with clouds and few clear sky observations (as explained in detail at P7L4:14) are not considered in this analysis, as we do not have any or sufficient observations to provide an estimate of the daily IST (i.e. to resolve the diurnal cycle). Please see revised abstract, introduction, results, discussion and conclusion sections, which hopefully makes it clear that this product covers clear sky conditions only.

We do not agree that clear sky products are of limited value (see general author response and the introduction of the revised manuscript). We believe that it is of great value to establish a relationship between surface air temperature measurements and satellite-based estimates of the ice surface temperature, with the aim of estimating surface air temperature in regions where no in situ observations are available (and thus drastically increasing the density of surface air temperature information globally). The derived T2m_sat product including well-characterized uncertainties can improve existing reanalysis products in regions with limited in situ observations (such as the Arctic), and thus provide more complete temperature fields.

Specific major and minor comments:
P3L13: which "data"? I presume satellite data, as AWSs do not move. This must be clearly worded.
Unfortunately, we cannot identify what the reviewer is referring to here. The in situ data is described in Section 2.1 and the satellite data in Section 2.2.

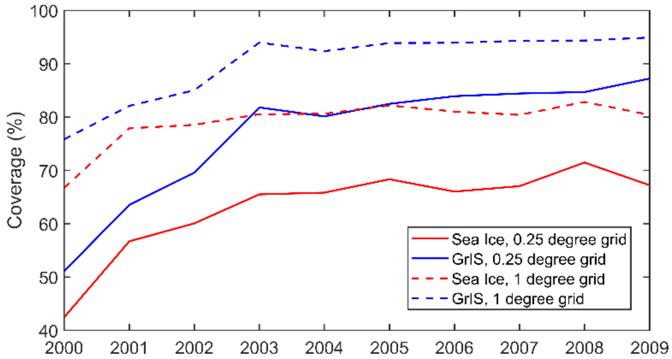
P4L10: Snow on sea ice have a major effect on the measured temperature as snow is a very good insulator (e.g. Graham et al, 2019, https://doi.org/10.1038/s41598-019-45574-5). Hence, if the buoy thermistor has a smaller diurnal cycle as the T2m sensor, snow cover is affecting the observation and buoy thermistor should be discarded as valid surface temperature observation. Unless "unrealistic data artifacts" includes damped daily cycles - which then should be stated explicitly -, I believe data scarcity should not be an excuse for retaining incorrect data.

We agree on this and we have now removed these data from the analysis. The variability method can only work for periods where significant diurnal variability is present and is thus not a robust method. No data of from these types of observations are included in the results. Thank you for pointing this out.

P5L4: All three citations listed to introduce the AASTI refer to technical documents; they do not refer to peer reviewed papers. This is not extremely relevant in itself, but it raises, in my humble opinion, the necessity that the authors restate briefly the methods to compile this dataset. Furthermore, from the title of Høyer 2019, this dataset is only providing clear-sky ice surface temperatures. This must be restated when the AASTI dataset is introduced.
Thanks, we agree on this. "Clear sky" has been added in the text, where the AASTI dataset is introduced. Furthermore, a brief description of the AASTI IST algorithm has been added following your advice.

P6L10: It does not become clear to me how these 3-hourly bin averages are aggregated into one daily value. The procedure should be added and described plainly.
The 3-hourly bin averages are not aggregated into one daily value. All the available observations within a day and 0.25 degrees are aggregated to produce the daily estimates. And similarly for the 3-hourly averages. The 3-hourly bin averages are only used to estimate the satellite sampling during the day and to gain confidence in the daily cycle estimates (as stated in P6L14-15).
We have reformulated this part to make it more clear. Also, we have expanded on the description of the AASTI data (and algorithm) to make it more clear that the original AASTI data is swath based (L2).

P7F3: Although the figure is somewhat instructive, I would be more interested to see (also) **a)** the ratio between cloudy and cloud free observations, as the current figure is clouded by the variations in observation density and **b)** the percentage of days with one or more valid observations within every time interval. Increasing from 1 to 25 observations in a 3-hour interval improves the measurement accuracy, decreasing from 1 to 0 leads to a data gap. By the way, I am puzzled by the fact that even far North (>75 N), where the polar night and midnight sun periods are long and the daily cycle weak, such a strong daily signal in the mean number of observations is found.



We agree with the reviewer that it is a good idea to provide some information on the coverage of the $T2m_{sat}$ product (and thus the days where the T2msat is not available due to clouds). This information has now been added to the manuscript in the discussion and conclusion. The coverage is stable from 2003-2009 and somewhat lower before 2003 (see figure above). The average daily coverage is 84% and 67% for land ice and sea ice, respectively, considering the stable 2003-2009 period and the 0.25 degree grid. When considering a 1 degree grid resolution these numbers increase to 94% and 81%, respectively. The 2003-2009 period is chosen here to represent the coverage in the recent satellite observing constellation. The high percentages in coverage demonstrate that the gaps due to cloudy days are limited and that the data set contains a significant amount of information on the all-sky daily T2m even though it is based upon clear sky satellite observations.
Note that the daily cycle in the satellite sampling depends both on the daily variations in cloud cover and the quality of the cloud screening, which is significantly improved for day light conditions. Also northwards of 80 degrees there is little daily variations, as expected.

P8L2: Here we are at the end of the description of the skin-temperature data treatment and there is nothing about treatment of data gaps introduced by cloud cover. Please correct me if I am wrong, but if I am right, that is a major omission. Given this absence, my presumption is - I cannot find any clarification in this manuscript - data gaps are left open; if one has no method to fill data gaps these gaps remain gaps. In favor of my presumption, Figure 6 has also regions with no data. Introducing gaps when your method fails positively bias your method performance and introduces an unknown bias in the final result. Again, please correct me if I am wrong; but if I am not, again, this data set cannot be used as all-weather T2m dataset and the paper cannot be accepted for publication until the cloud problem is resolved.
You are right that days with clouds (as defined in the paragraph P6L10-P7L14) do not contain any T2m estimates (see general author response). That is indeed the reason you see gaps in the example figure (Figure 6). However, this is similar to all other infrared retrievals of e.g. Sea Surface Temperature or Land Surface Temperature (for the other domains), and these clear sky observations are used extensively in models and aggregated products with large positive impacts on model forecasts.

Furthermore, no comments are made in how sub-tile temperature variations due to topography are dealt. I thus presume it is ignored, fine, but state explicitly. It does affect your correction procedure of 3.1.
Thanks for pointing this out. This has now been stated explicitly and discussed in section 5.

P8L6: The discussion of the comparison with in situ observations is a missed chance. It allows you to understand why remote sensed skin temperatures are deviating. Does the correlation improve if the exercise is repeated using valid 3-hourly estimates? If so, then it is a data gap (= cloud) problem.
We know from other validation studies, that the largest component of uncertainty in the satellite retrievals is typically the influence of undetected clouds. Clouds are very hard to detect over sea ice and will typically introduce a cold bias.
Repeating the exercise with the aggregated 3-hour IST_L3_skins will give us all-sky 3-hourly estimates of T2m. The aim of this paper is to derive a daily all-sky values and not 3-hourly values. However, a validation of 3-hourly T2m will likely be improved (compared to the daily values), since each 3-hour value will be validated against in situ observations taken within 3 hours from the satellite measurement. For daily estimates, we allow (only) two 3-hour bins (night/day) with available satellite observations in each day and this may result in larger differences, if these are not representative of the entire day. However, this effect has been mitigated by the checks listed in Section 2.2.

One way to investigate the effect from data gaps due to clouds (within the derived daily T2m product) is to look at the performance against the number of filled 3-hour bins (empty 3-hour bins means that clouds are present). The number of filled 3-hour bins increase over the period (see Figure 8) due to the increasing amount of satellites (see Figure 2). However, we do not see an increase in the performance of the estimated T2m over the period (Figure 7), which means that more filled 3 hour bins not necessarily improve the performance of our product (but possible improve the coverage). This has been clarified in manuscript.

Furthermore, as surface temperatures are very elevation dependent, this must be discussed as many of the PROMICE AWSs are close to the ice sheet margin, thus in terrain in which the elevation potentially varies more than 1000 m in a 0.25_ degree tile.
We agree that the elevation change within the matchup distance will introduce differences. The average slope of the 8 PROMICE stations is 1.49°, and this (together with the matching distance) increase the uncertainty in a pixel-to-point-measurement comparison, but as discussed in RC-2 P9 the matching distance has been set to obtain a robust number of in situ observations. Thanks for pointing this out. It has been stated now explicitly in the discussion.

P9L11: It is very common in comparable studies to cut your dataset into 3. In that case, you can perform the training-validation cycle three times; all three data subsets are used once in a training-validation cycle for validation and the remaining two are used for training in that cycle. Why is this approach not applied here?
We believe that the method used here, where we divide the data into two independent subsets (one for training and one for validation) is a suitable approach for this problem, which gives us realistic and independent information about the performance of the regression method. The limited amount of in situ observations makes it critical to divide the data into even more subsets.

P10: Equations 4 to 7 provide an elegant approach to evaluate rather simple correction functions, Eqs. 8-12. Still, this is not the best you can do. Why is not a state-ofthe-art method like a neural-network approach used? Furthermore, as there are data gaps due to clouds, how are they dealt here? Are these days neglected?
Days with clouds are neglected (see general comment and comment to P8L2). We agree that the use of advanced statistical methods to fill the gaps is very interesting and in the discussion we state that in order to fill gaps (due to cloudy conditions) different methods can be used such as a statistical technique or the use of atmospheric models and assimilation (P22L8). However, we believe that deriving a gap-free T2m product is outside the scope of this paper.

P17F8: The paper gives me no real clue how these data gaps are filled. As you can see, I am not convinced of the scientific soundness of the approach to convert clear-sky remote sensed skin temperatures into a continuous daily T2m data set. If I would have to do this, I would have taken the following approach: Take the discontinuous 3-hourly dataset of clear-sky skin temperatures, the continuous dataset of the cloudy/clear sky observation ratio, and other sensible remote sensed data products (like cloud properties, shortwave fluxes, atmospheric temperatures, sea ice state, local topography) and put it all into a neural network method (or any other AI) to order to produce a continuous 3-hourly T2m time series, trained with and evaluated against the in situ dataset. In a final step the 3-hourly data are averaged to daily means.

See general comment and comment to P8L2. As mentioned in P8L6, repeating the exercise with the aggregated 3-hourly satellite ISTskins, would result in an all-sky 3-hourly T2m estimate. Averaging the all-sky 3-hourly T2m estimates will lead to a (unknown) clear-sky bias in the daily averaged value (in cases where only a few 3-hour bins are filled within a day), and it cannot be used as an all-weather daily T2m estimate (in contrast to what we derive here, by tuning to daily all-sky in situ T2m).

I have read the results section as if the authors had produced a sensible daily T2m dataset, as it is possible that they indeed did this, but failed to convey that to me. At the other hand, since my objections to the method (description) are so major, it makes no sense to do detailed suggestions how the results, discussion and conclusions sections may be improved.
We hope that the reviewer will read and provide suggestions to the discussion and conclusions in the revised version. We have changed the discussion and conclusions to take into account the general comments from both reviewers and believe this has led to improved sections.

The results section analyses if there are systematic biases as function of the estimated T2m. In a renewed submission, the authors should analyze separately the performance for clear-sky, mixed sky and fully cloudy conditions. It should be proven that a reasonable method is found to estimate T2m for all conditions.
See also general comment and comment to P8L2. The days when the satellite derived T2m product is available, it represents the all-sky T2m average for that day, since it has been regressed towards in situ measurements obtained both in cloudy and clear sky conditions. This has been clarified in manuscript.

Table 8 and Figure 11 show in my humble view that the data set presented here is not good enough to be used. As e.g. Batrak and Müller (https://doi.org/10.1038/s41467-019-11975-3P) demonstrate, ERA-Interim and ERA5 and other reanalyses do a very poor job over the Arctic ocean due to missing snow cover over sea ice and misrepresented sea ice thickness. For Greenland, I have no paper at hand that does a similar analysis and I am neither aware that ERA-Interim is doing an extremely poor job there too. As you did not mention anything about applying an elevation correction on the reanalysis data – which is essential for a fair comparison, I suspect that overlook might be part of the poor performance of ERA-Interim over the ice sheet.
Nonetheless, a useful t2m product derived from remote sensing should be able to beat easily a flawed model product – and yours does not.
Furthermore, as these reanalyzes fail to represent Arctic T2m, they should not be used as benchmark. The data set should be benchmarked against reanalysis results of RCMs optimized for either the Arctic or Greenland. I know there are several colleagues at your institute that can help you in selecting appropriate RCMs and retrieving the data.

It is indeed a fact that current reanalyses do a very poor job in the Arctic. This is one of the main reasons that using satellite observations to provide an alternative and independent estimate of the near surface air temperatures is of great value in particular for the Arctic region. This has now been clarified in the introduction.
We think the comparison against ERA-I is relevant and fair, since no elevation correction is applied to the satellite derived T2m product either. Part of the discrepancy between the in situ observations and the T2m estimates (both ERA-I and the satellite derived T2m) is the point to pixel comparison, where the point measurement may not be representative of the entire pixel. The differences here arise indeed both because of differences in elevation but also due to the matchup distance. Even though the satellite derived products beats ERA-I over the Greenland Ice Sheets, the most important take-away message is what is (as stated in the conclusion P24L4) that the derived T2m product should not replace the already existing air temperature measurements, but rather to supplement these e.g. in areas where no in situ observations are currently available. Since, the errors in ERA-I and the $T2m_{sat}$ are independent and uncorrelated a combination of the two datasets would lead to an even better T2m estimate. We have clarified this when the comparison against ERA-I is performed.

ERA-I is one of the most used reanalysis products and the T2m has been used extensively for many publications, also in the Arctic. It is therefore of general interest to compare the performance against this product. ERA-5 and C3S Arctic regional reanalysis were not available at time of the analysis but will be considered for future work.

Finally, the discussion leaves me puzzled by the fact that the authors are aware of the cloud-gap problem, but try to publish a data set in which this problem is not fully solved (as there are data gaps in the presented data set) and failed to present properly the measures they have undertaken to mitigate the "cloud problem".
See general author response.

RC2 – Authors response

General: This study aims to produce a 2-m air temperature product for the Arctic region by establishing multiple linear regression models with satellite-derived surface temperature measurements and other covariates. The goal of this study is well stated, and the product is potentially very useful.
We are glad to hear that the reviewer acknowledge that the product potentially is very useful. We hope that the comments below and the changes in the manuscript resolve the issues that the reviewer has pointed out.

However, the authors tend to provide over-whelming details on some data processing techniques that already exist in two previous papers while the most important issue on cloud influences are much less discussed. Some parts of methodology and validation may have some mis-interpretation and still needs careful discussions. I am suggesting the following major/minor comments for authors to revise their manuscript.
Point-by-point response is given below.

Major: It is a bit concerning to me that many methodology materials have already been presented in Nielsen-Englyst et al. (2019) and Hoyer et a. (2018), but the authors spent many pages describing the same amount of details in this paper. I would suggest the authors to only retain the most important information, and then delete all other repeating materials by citing those previous studies (authors should start from already pre-processed data, direct readers to read those two papers, and then leave more room for remaining story). I think it's not suitable to publish the same amount of details in this paper again, especially considering Nielsen-Englyst's other paper was also published in The Cryosphere.
The methodology section has been shortened as suggested, and in particular the repetitions of the Nielsen-Englyst et al. (2019) results have been reduced. We do keep most of the information from Hoyer et al. (2018), since this is not a peer reviewed paper, following the arguments of reviewer #1 (P5L4).

P9 Methods: Why choose 0.5 degree as the matching threshold? I think this may be too large. If the authors reduce this threshold and reduce number of matching pairs, will the results be better? Please show some analyses, or at the very least, this needs to be carefully discussed.
The matching threshold was chosen to ensure a sufficient number of matchups to derive and validate the different regression models with sufficient statistical confidence. This has been added now, when it is introduced. We do mention in the discussion that the limited number of matching pairs is one of the greatest challenges in this study (and similar studies). This part has now been modified and extended to explicitly mention the matching threshold and elevation effects as the main limitations in the derivation and validation of the regression model. Thanks for pointing this out.

P12L12: I think the highest correlation in ISTskinSeason may be wrong interpretation, because what the authors computed are raw data correlation. Without removing the seasonal cycle, this calculation of correlation will be dominated by seasonal cycle, and of course ISTskinSeason will have better correlation. Since the authors' goal is to produce a T2m product, this product should aim at achieving high accuracy in anomalous T2m days, and also being able to capture general characteristics such as seasonal cycle (for general analysis) and trend (for global warming analysis). In my opinion, I think the authors should aim at achieving the highest anomaly correlation when training their models, and revise relevant parts when interpreting model performance.
We agree with the reviewer that it is important to consider the strong seasonal cycle existing in the T2m. All the regression models we test here include a linear relation between the Tskin and T2m and thus all of them capture the large seasonal cycle in Tskin (through the α_1 term). The close coupling between the Tskin and T2m on (synoptic (2-5 days) and seasonal time scales) is the main reason why we generally get very high correlations for all the regression models tested. The high correlations are thus expected and indeed dominated by the seasonal cycle in both temperatures. The following sentence has been added in the $IST_{skin\text{-}L3}$ validation in Section 2.2.: "The generally high correlations are dominated by the synoptic (2-5 days) and seasonal variations, which are pronounced in both IST and T2m."
The regression models listed in Eq.9 through Eq. 12 are included to examine how to best represent the residual variability not related to the synoptic and seasonal variations, as none of these terms depend upon the Tskin. To our understanding, this is actually what the reviewer asks for. As evident from Table 4, the ability to best represent the anomalous T2m days are actually found by including the seasonal component that only works on correcting the T2m-Tskin differences (and is not related to the Tskin itself, as these are represented by the α_1 term). Table 4 shows that we get an improved correlation and RMS by using this model compared to the other. This discussion and clarification has been clarified in the text (see Section 3.1) and in conclusion.

There seems to be overwhelming technical details provided, however, the most important issue is on how the cloud days were considered, but this seems to receive the least attention? Can the authors reduce some details as I mentioned in Comment #1, and leave more room for how the cloud days were treated? Those are the most useful and challenging issues for this product.
The technical details have been reduced as suggested. Days with clouds (or days where we do not have observations in both the night/day bin as explained in detail in the paragraph P6L10-P7L14) are not considered in this analysis, as we do not have sufficient observations to provide an estimate of the daily IST. The final product only provides at T2m estimate for days where daily ISTs are estimated in respect to the criteria presented in P6L10-P7L14. The days when the satellite derived T2m product is available, it represents the all-sky T2m, since it has been regressed towards in situ measurements obtained both in cloudy and clear sky conditions. Please see changes in abstract, introduction, results, discussion and conclusion sections, which hopefully makes it clear that this product covers clear sky conditions only (e.g. when satellite observations are available).

P1L16: again, if using this raw correlation, my guesses are that if you simply add a systematic bias (e.g., +0.3 degree C) to the satellite surface temperature to derive T2m, you will probably also get similar high correlation. Can authors perform this simple calculation and demonstrate more clearly on the gains of their regression model?
(This question is related to P12L12)
Yes, this is exactly what we have shown in first row in Table 4 (from Eq. 8). The first regression model simply adds a systematic bias and a scaling factor of the IST and we do get an improved RMS and correlation (see Table 4) compared to using the IST alone (see Table 2). Table 4 also shows the results of using different predictors (including the seasonal cycle). Only, the inclusion of the SWd radiation and the seasonal cycle shows significant improvements from using a simple correction as suggested by the reviewer.

Minor: P2L26: change to “significantly different”
Implemented. Thanks!

P7L4-5: I am not familiar with this (18-6, 6-18) way of presenting time. Is this following any convention? If not, I suggest the authors to make revisions.
This has been rewritten in the manuscript to: “In order to best resolve the diurnal cycle with satellite information we require data during both night (between 18 and 6 local solar time) and day (between 6 and 18 local solar time) in order to calculate ISTskin_L3.”
This is following the mathematical notation and terminology in the manuscript preparation site on the TC.net, date and time (https://www.the-cryosphere.net/for_authors/manuscript_preparation.html).

P9L5: delete “the” before “different regions”
Implemented

P9: Unclear presentation as to what is the percentage split for training/testing. Please revise presentation in the format of e.g., 80%/20% split.
The information has been added. Thanks.

Table 3: It would help if the authors can use stars to mark significance level of the correlation (e.g., *: p<0.05). . . Plainly presenting the correlation is less informative
This table has been removed in response to the reviewer’s major/first comment.

P11L7: Warming effects are mainly resulted from high clouds, but low clouds can cool the surface. Was this differentiated? Can authors provide some discussions on this?
In Nielsen-Englyst et al. (2019), there was not differentiated between low/high clouds. But overall it was found that the Tskin increased in cloudy conditions compared to clear sky conditions. This is in agreement with Intrieri (2002), who found that clouds tend to have an overall warming effect in the Arctic (considering and discussing both the warming effect from high clouds and the cooling effect from low clouds). We refer to Intrieri (2002) for further discussion on this topic. We have left out this part (repeated from Nielsen-Englyst et al., 2019), since it is not crucial information in this paper (in response to reviewer’s general comment to the methodology section), where we do not consider longer periods (>1 day) with clouds (see general comment). In this paper, the different cloud types are not important, as there are no satellite observations of the surface in any case.

P23L21: again, this highest correlation may not be a good indicator for this product to be reliable. Please see my earlier comments and provide some revisions or discussions on this matter.
Please see our responses to P12L12 and P1L16. We have added “and the lowest RMS” here.

Abstract: authors should mention their data product's spatial resolution in the abstract.

We agree on this and the information has been added.

Deriving Arctic 2 m air temperatures over snow and ice from satellite surface temperature measurements

Pia Nielsen-Englyst[1,2], Jacob L. Høyer[2], Kristine S. Madsen[2], Rasmus T. Tonboe[2] and Gorm Dybkjær[2]

[1] Technical University of Denmark (DTU), DK-2800 Kongens Lyngby, Denmark
[2] Danish Meteorological Institute (DMI), DK-2100 Copenhagen Ø, Denmark

Correspondence to: Pia Nielsen-Englyst (pne@dmi.dk)

Abstract.

The Arctic region is responding heavily to climate change, and yet, the air temperature of Arctic, ice covered areas is heavily under-sampled when it comes to in situ measurements, and large uncertainties exist in weather- and reanalysis products. This paper presents a method for estimating daily mean clear sky 2 meter air temperatures (T2m) in the Arctic from satellite observations of skin temperature, using the Arctic and Antarctic ice Surface Temperatures from thermal Infrared (AASTI) satellite dataset, providing spatially detailed observations of the Arctic. The method is based on a linear regression model, which has been ~~developed using~~tuned against in situ observations to estimate daily T2m based on ~~and daily mean~~ clear sky satellite ice surface skin temperatures combined with a seasonal variation~~ to estimate daily T2m~~. The daily satellite derived T2m product including estimated uncertainties covers clear sky snow and ice surfaces in the Arctic region during the period 2000-2009, estimated on a 0.25 degree regular latitude-longitude grid. Comparisons with independent in situ measured T2m give~~s~~ average correlations of 95.5% and 96.5% and average root mean square errors of 3.47°C and 3.20°C for land ice and sea ice, respectively. The reconstruction provides a much better spatial coverage than the sparse in situ observations of T2m in the Arctic, is independent of numerical weather prediction model input and it therefore provides an important alternative to simulated air temperatures to be used for assimilation or global surface temperature reconstructions. A comparison between in situ T2m versus T2m from satellite and ERA-Interim shows that the T2m derived from satellite observations validate similar or better than ERA-Interim estimates in the Arctic.

1 Introduction

The Arctic climate is changing rapidly with surface temperatures rising faster than other regions of the world due to Arctic amplification (Graversen et al., 2008; Pithan and Mauritsen, 2014). Meteorological measurements show that the 2000s were the warmest decade in Greenland since meteorological measurements started in the 1780s (Box et al., 2019; Cappelen, 2016; Masson-Delmotte et al., 2012).

The Arctic surface air temperature is one of the key ~~an important~~ climate indicators used to assess ~~of~~ regional and global climate changes (Hansen et al., 2010; Pielke et al., 2007) and both ~~as~~ ~~most~~ model simulations and observations ~~by global~~

~~climate models~~ indicate that ~~any~~ warming in the global climate ~~will~~ is~~be~~ amplified at the northern high latitudes (e.g. Collins et al., 2013; Holland and Bitz, 2003; Overland et al., 2018)~~Holland and Bitz, 2003; Overland et al., 2018)~~. Traditionally, near surface air temperatures have been measured at the height of 1-2 m using automatic weather stations (AWSs) or buoys (Hansen et al., 2010; Jones et al., 2012; Rayner, 2003; World Meteorological Organization, 2014)~~(Hansen et al., 2010; Jones et al., 2012; Rayner, 2003)~~. Extreme temperatures, winds and the remoteness of the Arctic make in situ observations in the Arctic temporally and spatially sparse (Reeves Eyre and Zeng, 2017), and challenging. In particular, it is difficult to achieve climate-quality temperature records for this region.~~.~~

The global near surface air temperature datasets that are currently most widely used (HadCRUT4, NOAAGlobalTemp and GISTEMP) are derived only by using in situ observations (Hansen et al., 2010; Morice et al., 2012; Smith et al., 2008; Vose et al., 2012). To increase the coverage and quality of the surface temperature products, polar orbiting satellites can offer a very good supplement to the in situ observations through a high spatial and temporal coverage of all regions in the Arctic. ~~and d~~ Daily near surface air temperatures derived from satellites therefore have the potential to increase the amount of information in the data sets and improve the quality of the climate records, as recognized in Merchant et al. (2013) and Rayner et al. (2019).

~~The most practical way to get continuous and spatially broad measurements of the data-sparse Arctic is through satellite remote sensing.~~ However, satellites with infrared sensors in the atmospheric window region of 10-12 micron wavelength measure the ice surface skin temperature (IST_{skin}) during clear skies whereas the current global temperature products estimate the near surface air temperature as are measured AWSs and buoys. ~~The standard measurement height is 2 m, but it varies with snow depth at the sites (World Meteorological Organization, 2014). However, all measurements are relatively near the surface and are therefore also often called "surface air temperatures".~~ The surface skin temperature may differ considerably from the surface air temperature during melting conditions, but during other conditions the skin and surface air temperature may be more or less the same (Nielsen-Englyst et al., 2019).

To benefit from the good coverage of satellite surface temperature data, we have explored the relationships between the surface air temperature and the satellite measurements. Several studies have compared satellite retrieved IST_{skin} and T2m from AWSs located on the Greenland Ice Sheet (GrIS; Dybkjær et al., 2012a; Hall et al., 2008, 2012; Koenig and Hall, 2010; Shuman et al., 2014) and over the Arctic sea ice (Dybkjær et al., 2012) and found temperature differences of which a significant part could be attributed to the temperature difference between T2m and IST_{skin}. Previously, work has been done to investigate the relationship between the surface and near-surface air temperature over ice using in situ observations (Adolph et al., 2018; Hall et al., 2008, 2004; Hudson and Brandt, 2005; Nielsen-Englyst et al., 2019; Vihma et al., 2008). Nielsen-Englyst et al. (2019) found that on average T2m is 0.65-2.65°C warmer than IST_{skin} with variations depending on location of the measurement i.e. in the lower ablation zone, upper-middle ablation zone, accumulation zone, seasonal snow cover and sea ice. The T2m-IST_{skin} difference was found to vary with season with smallest differences around noon and early afternoon during spring, fall and summer during non-melting conditions. Furthermore, wind speed and cloud cover were identified as key parameters determining the T2m- IST_{skin} difference.

Given the difficulties of operating equipment in the harsh Arctic conditions, the potential for using satellite IST_{skin} to estimate T2m is large in this region. The greatest limitation of satellite-derived infrared surface temperatures is cloud cover. Hence, a satellite-derived, clear-sky, surface temperature record can be significantly different from an all-sky surface temperature record (Koenig and Hall, 2010; Nielsen-Englyst et al., 2019). ~~The investigation of the IST_{skin} versus T2m relationships over ice and the derivation of T2m from IST_{skin} were also identified in Merchant et al. (2013) as important areas for improving the understanding of the surface temperature of the Earth.~~

This work, starting with ~~(~~Nielsen-Englyst et al.~~,~~ (2019), has been initiated to estimate clear sky T2m from satellite observations (whenever these are available) covering the snow and ice covered parts of the Arctic, in order to provide spatially-detailed observations for the areas unobserved by in situ stations and to supplement the in situ observations already available. ~~The investigation of the IST_{skin} versus T2m relationships over ice and the derivation of T2m from IST_{skin} were also identified in Merchant et al. (2013) as important areas for improving the understanding of the surface temperature of the Earth.~~

A~~In this paper a~~ regression-based approach has been used to estimate daily T2m using satellite IST_{skin} and a seasonal cycle function as predictors, based upon the work presented in Høyer et al. (2018). The derived product covers only days with none or limited clouds, where satellite skin temperature observations are available. However, for those days when the satellite derived T2m product is available, it provides an estimate of the daily averaged all-sky T2m, ~~since~~ as it has been regressed towards in situ measurements from both clear and cloudy conditions. In order to further facilitate the usage of the derived product in modelling and for monitoring purposes, each satellite retrieved T2m estimate comes with uncertainties. Similar efforts have been done to estimate clear sky near surface air temperatures (and corresponding uncertainties) over land, ocean and lakes using satellite observations to cover all surfaces of the Earth (Good, 2015; Good et al., 2017; Høyer et al., 2018). The previous work has mostly been done as a part of the European Union's Horizon2020 project EUSTACE (EU Surface Temperatures for All Corners of Earth, 2015-2019, https://www.eustaceproject.org), with the aim to produce a globally complete gap-free daily near surface temperature analysis since 1850. ~~.~~It is outside the scope of this paper to produce a continuous gap-free daily near surface temperature analysis. However, within EUSTACE this has been done using a statistical model to combine ~~combination of~~ satellite derived near surface air temperatures (as derived here over ice) and in situ observations, and their respective uncertainty estimates (Morice et al., 2019; Rayner et al., 2019).

This paper is structured such that Sect. 2 describes the in situ data and the satellite data. Section 3 presents the method used to estimate daily clear sky T2m and uncertainties. The resulting T2m dataset and its validation are presented in Sect. 4 and discussed in Sect. 5. Conclusions are given in Sect. 6.

2 Data

2.1 In Situ data

In situ observations of near surface air temperatures have been collected from weather stations, expeditions and campaigns covering ice and snow surfaces to assemble the DMI-EUSTACE database. The database includes quality controlled and uniformly formatted temperature observations covering ice and snow surfaces, during 2000-2009 (Høyer et al., 2018). Over Arctic land ice/snow we use the Programme for Monitoring of the Greenland Ice Sheet (PROMICE) data provided by the Geological Survey of Denmark and Greenland (GEUS; Ahlstrøm et al., 2008), the Atmospheric Radiation Measurement (ARM) Program data from the North Slope of Alaska (Ackerman and Stokes, 2003; Stamnes et al., 1999), and the Greenland Climate Network data (GC-Net; Kindig, 2010; Shuman et al., 2001; Steffen and Box, 2001). Only PROMICE data from the middle-upper ablation zone and accumulation zone have been used to ensure that data are only acquired over permanently snow or ice covered surfaces. Data on Arctic sea ice are primarily retrieved from the meteorological observation archive at the European Centre for Medium-Range Weather Forecasts (ECMWF) MARS data storage facility, providing 196 unique data series from drifting buoys~~ of which 11 measure IST$_{skin}$ as well~~. These sea ice data are supplemented with data from 10 U.S. Army Cold Regions Research Engineering Laboratory (CRREL) mass balance buoys (Perovich et al., 2016; Richter-Menge et al., 2006) and observations from the research vessel, POLARSTERN, operated by the Alfred-Wegener-Institute, operating the sea ice covered parts of the Arctic Ocean (Knust, 2017). We also use air temperature measurements obtained from ice buoys deployed in the Fram Strait region within the framework of the Fram Strait Cyclones (FRAMZY) campaigns during the years 2002, 2007, and 2008 as well as air temperatures from the Arctic Climate System Study (ACSYS) campaign in 2003 (Brümmer et al., 2011b, 2011c, 2012b, 2012a). Finally, we use data from two ice buoy campaigns operated by the Meteorological Institute of the University of Hamburg within the framework of the integrated EU research project DAMOCLES (Developing Arctic Modelling and Observing Capabilities for Long-term Environmental Studies; Brümmer et al., 2011a).

The different in situ types measure the air temperature at different heights that furthermore differ over time depending on the amount of snow fall, snow drift and snow melt. Here, we will refer to T2m for all observation types regardless of these variations. Nielsen-Englyst et al. (2019) showed small changes (<0.22°C) in T2m-IST$_{skin}$ differences when using only observations within the measurement range of 1.90-2.10 m in height compared to using all measurements (ranging in measurement height from 0.3 m to 3 m). The accuracy of the air temperature sensors for all observation sites is approximated to 0.1°C (Hall et al., 2008; Høyer et al., 2017b). Few data sources provide both skin and air temperatures e.g. the PROMICE and ARM stations~~ over land ice and 11 of the buoys from ECMWF operational data stream over sea ice~~. The PROMICE skin temperatures have been calculated from up-welling longwave radiation, measured by Kipp & Zonen CNR1 or CNR4 radiometer, assuming a surface longwave emissivity of 0.97 (van As, 2011). ~~The 11 buoy time series from ECMWF that provide IST$_{skin}$ observations as well are likely not measuring the actual skin temperature. The buoy thermistor is placed on the underside of the buoy which is set out on top of the ice. The buoys typically get buried in snow and the measured~~

~~temperature becomes an internal snow temperature. However, in this analysis the buoy temperature measurements will be treated and counted as IST_{skin} measurements, as we have no information on the snow depth on top of the buoys.~~ All in situ data have been screened for spikes and other unrealistic data artefacts by visual inspection. Afterwards, the in situ observations have been averaged to daily temperatures using all available observations. Figure 1 shows the number of daily averaged in situ observations each year during 2000-2009 of IST_{skin} and T2m over Arctic land ice and sea ice~~, respectively~~. In total 65,810 observations with daily T2m and 7,~~681~~ 057 observations with daily IST_{skin} are available over land ice. ~~However, only 624 of these cover Arctic sea ice.~~ See Table 1 for more information on the in situ observations used in this study.

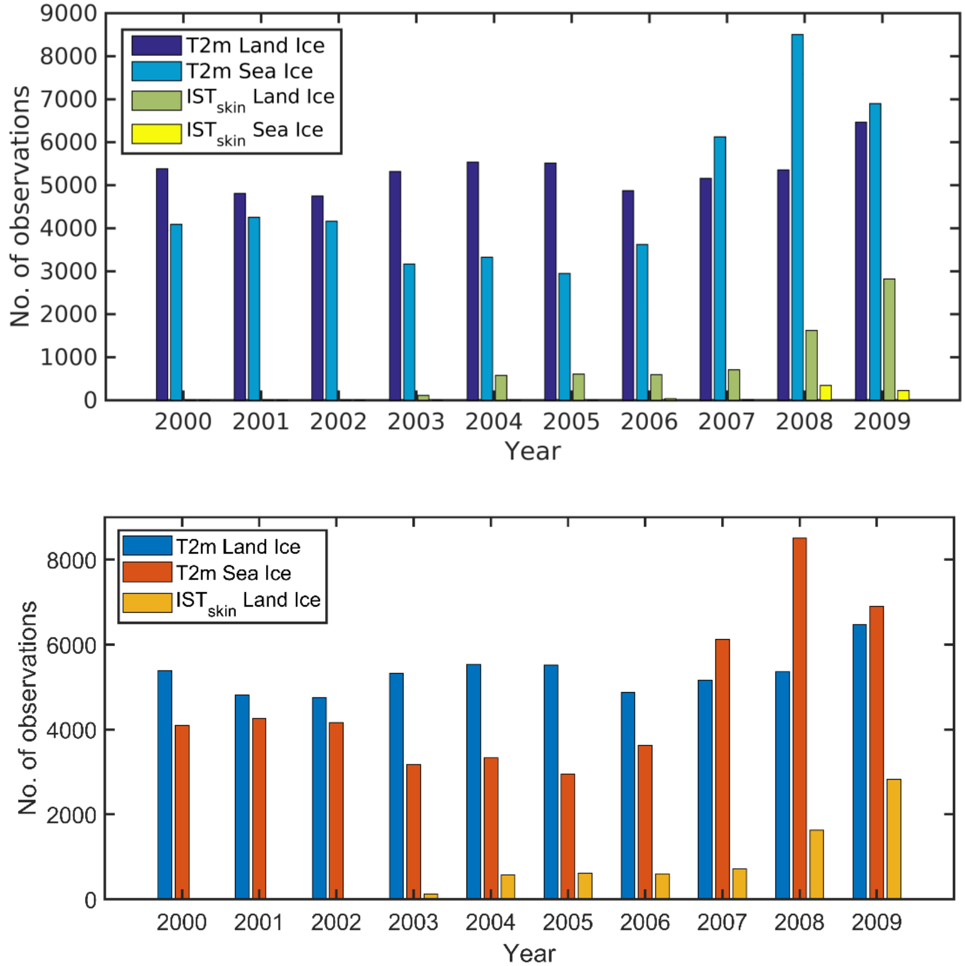


Figure 1: Total number of daily averaged in situ observations of T2m and IST_{skin} over Arctic land ice and sea ice per year covering the period 2000-2009.

Table 1: Overview of in situ observations used in this study, covering 2000-2009.

	No. of sites, (AWS, buoys or ships)	No. of days with observations	Surface Type	Observation Type	Temperature measurements
ACSYS	7	280	Sea ice	Buoy	T2m
ARM	2	2,846	Land snow	AWS	T2m, IST_{skin}
CRREL	10	1,031	Sea ice	Buoy	T2m
DAMOCLES	25	2,160	Sea ice	Buoy	T2m
ECMWF	196	27,235	Sea ice	Buoy	T2m~~, IST (11 buoys)~~
FRAMZY	11	251	Sea ice	Buoy	T2m
GC-NET	15	29,133	Land ice	AWS	T2m
POLARSTERN	1	189	Sea ice	Ship	T2m
PROMICE	8	2,685	Land ice	AWS	T2m, IST_{skin}

2.2 Satellite data

The satellite data used in this study is from the Arctic and Antarctic Ice Surface Temperatures from thermal Infrared satellite sensors (AASTI; Dybkjaer et al., 2018; Dybkjær et al., 2014; Høyer et al., 2019) data set, covering high latitude seas, sea ice, and ice cap clear sky surface temperatures based on satellite infrared measurements from the CLARA-A1 data set compiled by EUMETSAT's Climate Monitoring, Satellite Application Facility (Karlsson et al., 2013). The data set is based on one of the longest existing satellite records from the Advanced Very High Resolution Radiometer (AVHRR) instruments on board a long series of NOAA satellites. AASTI contains swath based (i.e. Level 2~~; (~~L2) ice surface skin temperature (IST_{skin_L2}) data processed and error corrected on the original Global Area Coverage (GAC) grid. The first version of the AASTI product, used in this study, is available from 2000 to 2009 in the original projection and resolution (L2), i.e. ~0.05 arc degree resolution and multiply daily coverage. Since 2000, seven different AVHRR instruments have been orbiting the globe, each 14 times per day, and thus providing approximately bi-hourly coverage of the Polar Regions (Figure 2). The number of operational satellites has increased from 2 to 6 from 2000 to 2009. The IST algorithm used to generate ~~in~~ the AASTI data set is based on thermal infrared brightness temperatures of AVHRR channel 4 (centre wavelength at ~11 microns) and 5 (centre wavelength at ~12 microns), and the satellite zenith angle. The algorithm is a split window algorithm, working within three temperature domains for each individual satellite (Key et al., 1997). The retrieval calibration of each domain has been done by relating modelled surface temperatures with modelled top-of-atmosphere brightness temperatures, determined by a radiative transfer model (Dybkjær et al., 2014). Cloud masking has been performed using the Polar Platform System (PPS) processing cloud processing software (Dybbroe et al., 2005a, 2005b).

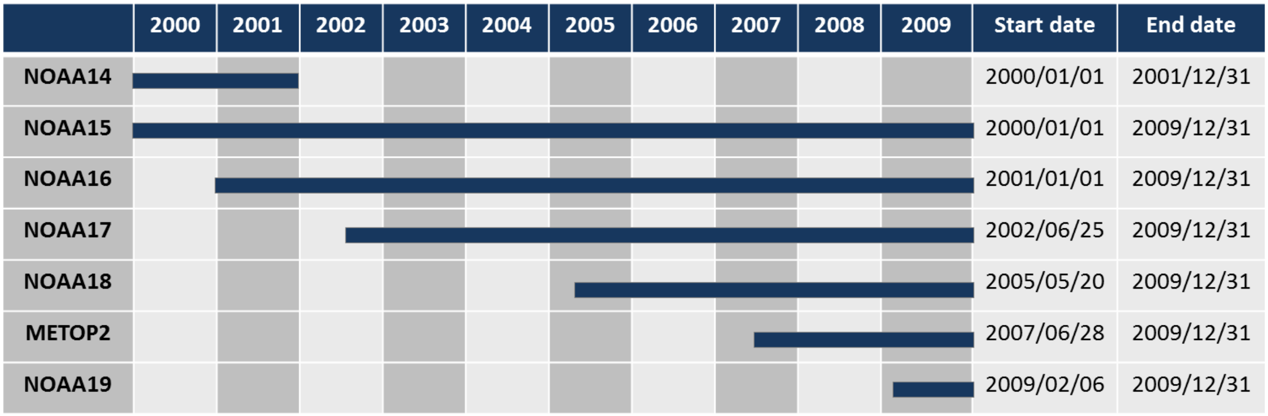


Figure 2: NOAA and Metop satellites carrying the AVHRR sensor, used for AASTI version 1.

As discussed in Merchant et al. (2017), satellite-based climate data records should include uncertainty estimates. The AASTI IST_{skin_L2} data come with uncertainties divided into three independent uncertainty components, each with different characteristics: the random uncertainty (μ_{rnd_L2}), a locally systematic uncertainty (μ_{local_L2}) and a large-scale systematic ("global") uncertainty (μ_{glob_L2}). These three components have been chosen since they behave differently when aggregating the observations in time or space (see Sect. 3.2). This uncertainty methodology has been developed within the SST community (Bulgin et al., 2016; Rayner et al., 2015) and will be followed here. The total uncertainty on the IST_{skin_L2}, μ_{total_L2} , is calculated by summing each component in quadrature (i.e., square root of sum of squares). Excluding the cloud mask uncertainty, grid-cell systematic uncertainties (μ_{glob_L2}) are set to a fixed value of 0.1°C to represent systematic uncertainties in the forward models (see e.g. Merchant et al., 1999; Merchant and Le Borgne, 2004). The AASTI IST_{skin_L2} data also come with a quality level (QL) from 1 (bad data) to 5 (best quality), with the addition of level 0 (no data) (GHRSST Science Team, 2010).

Here, we have aggregated the AASTI IST_{skin_L2} observations into 3 hourly and daily, gridded Level 3 (L3) averages (IST_{skin_L3}) of IST_{skin_L2} on a fixed 0.25 by 0.25 degrees regular geographical grid. The IST_{skin_L3} is calculated by averaging all available IST_{skin_L2} observations with a quality flag of 4 (good) or 5 (best) for a given date and within the 0.25 degree bin.~~.~~ This has been done to facilitate the development of the relationship model and to ease the user uptake. The data in the daily aggregated files contain mean surface temperature observations from 00 to 24 hours local solar time, ~~but also~~ 3-hourly bin averages of surface temperatures and also the number of observations in the eight time bins during each day. The 3-hourly numbers of observations ~~were~~ are used for estimating the satellite sampling throughout the day, and the 3-hourly temperature data to gain confidence in the daily cycle estimates (see quality checks below). ~~In the aggregation, all satellite observations with a quality flag of 4 (good) or 5 (best) were used.~~ Figure 3 shows the mean number of observations per day in each of the eight time intervals given in local time for the Arctic region. The variation in the coverage throughout the day is a combined effect of the satellite overpassing, performance of the cloud screening algorithm, and the cloud free conditions during the day. In addition, the fixed 0.25 degrees regular geographical grid results in a decreasing L3 bin area when approaching the

North Pole. The maximum in satellite coverage is generally seen around 80°N with a minimum at the North Pole. Cloud free conditions over the GrIS are primarily observed around noon and early afternoon.

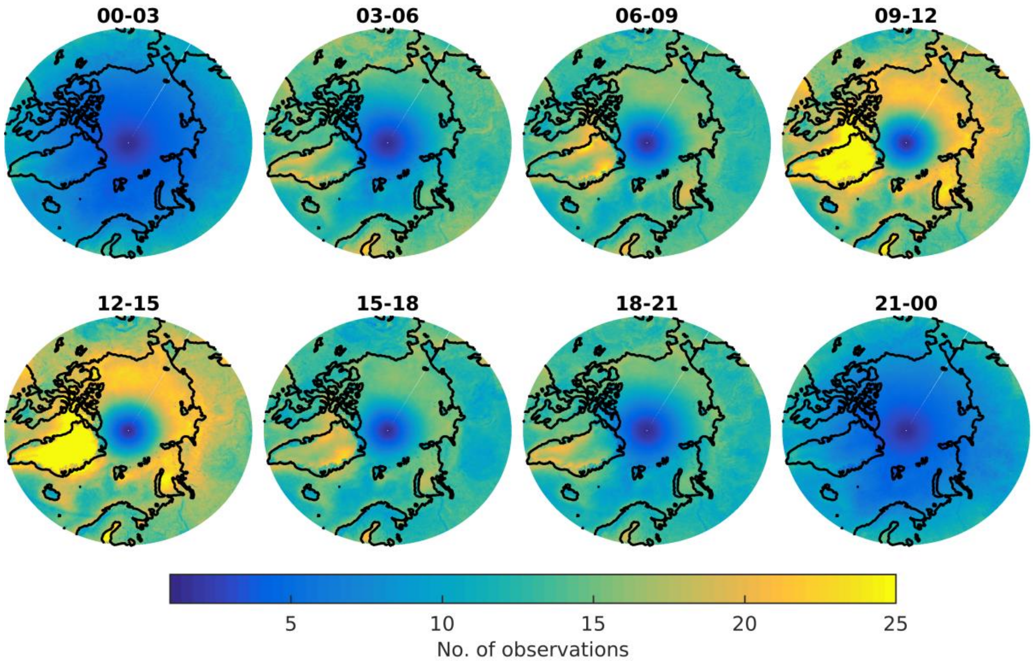


Figure 3: Mean number of observations per day in the L3 bins for each of the eight local solar time intervals, averaged for the period 2000-2009.

In order to best resolve the diurnal cycle with satellite information we require data during both night (between 18 and~~-~~ 6 local solar time) and day (between 6 and~~-~~ 18 local solar time) in order to calculate IST_{skin_L3}. ~~The IST_{skin_L3} is calculated by averaging all available IST_{skin_L2} observations for a given date.~~ A few more checks have been set up in order to minimize the temporal sampling errors and the effects of undetected clouds and outliers. Following Høyer et al. (2018), the IST_{skin_L3} is discarded if one of the following criteria is met:

- IST_{skin_L3} exceeds +5°C, indicating clear melting conditions or obviously wrong observations.
- The standard deviation of satellite IST_{skin_L2} during one day exceeds 7.07°C, corresponding to a sinusoidal daily cycle with a difference between day and night of 20°C.
- The difference between IST_{skin_L3} and the average of all available 3 h bin averages exceeds 10°C.
- IST_{skin_L3} is more than 10°C colder than the corresponding average of up to 24 neighbouring cloud free observations (in a 5 by 5 grid cell square) with the same surface type.

The criteria above have been derived from analysis and inspection of the satellite data and with considerations to the results presented in Nielsen-Englyst et al. (2019). The satellite-derived IST_{skin_L3} has seasonal differences in daily variability, with largest standard deviations during summer in Greenland and during winter for sea ice, where the freeze-up of sea ice causes higher variability along the sea ice margin (Fig. 4). The main uncertainty components of the IST_{skin_L3} estimates are erroneous cloud screening and the spatial variance of snow and ice surface emissivity, which are not accounted for in the retrieval algorithm. The presence of non-detected clouds will contribute to increased standard deviations and usually a cold IST_{skin_L3} bias, since the cloud tops and other atmospheric constituents generally are colder than the surface (Dybkjær et al., 2012).

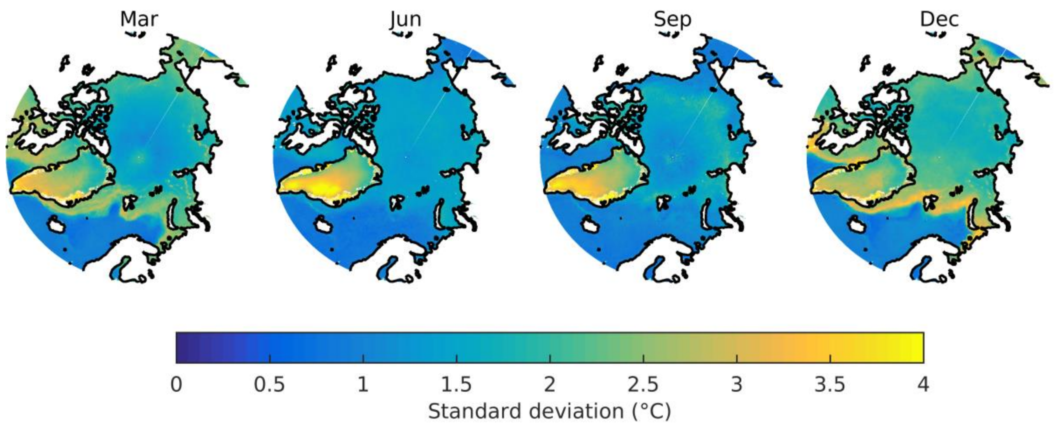


Figure 4: Standard deviations of daily satellite surface temperature observations for March, June, September and December, averaged for the years 2000-2009 (°C).

Additional satellite versus in situ differences arise when comparing satellite observations with pointwise ground measurements due to different spatial and temporal characteristics. To assess the magnitude of these effects, the IST_{skin_L3} data have been validated against in situ land ice temperatures from the PROMICE and ARM stations. Table 2 shows the validation results of IST_{skin_L3} against in situ skin temperatures (IST_{skin_insitu}) and in situ 2 meter air temperatures ($T2m_{insitu}$), respectively. The spatial and temporal sampling effects contribute to the overall uncertainty, but effects from erroneous cloud screening, algorithm simplifications, and uncertainties in the in situ observations are also included in the results. In general, IST_{skin_L3} correlates better with $T2m_{insitu}$ than with the IST_{skin_insitu}. Moreover, the IST_{skin_L3}-$T2m_{InSitu}$ difference shows smaller standard deviations than IST_{skin_L3}-IST_{skin_insitu}. However, as expected the biases and root mean squared differences (RMS) are larger for the IST_{skin_L3}-$T2m_{insitu}$ differences than for the IST_{skin_L3}-IST_{skin_insitu} differences. The reason is that the radiometric surface skin temperature can be significant different from the surface air temperature measurements (Adolph et al., 2018; Hall et al., 2008; Hudson and Brandt, 2005; Nielsen-Englyst et al., 2019; Vihma et al., 2008). On average, the skin temperature is colder than the air temperature, with the largest differences during clear-sky conditions and when the skin temperature is constrained by the melting point (melting snow has a maximum temperature of 0°C) (Nielsen-Englyst et al., 2019). The generally high correlations are dominated by the synoptic (2-5 days) and seasonal variations, which are pronounced in both IST and T2m.

Table 2. Validation of daily AASTI v.1 Level 3 IST (IST_{skin_L3}) against in situ IST_{skin} (IST_{skin_insitu}) and T2m observations ($T2m_{insitu}$). N: number of matchups, Corr: correlation, Std: standard deviation, and RMS: root mean square difference.

	IST_{skin_L3} - IST_{skin_insitu}	**IST_{skin_L3} – $T2m_{insitu}$**

Station	N	Corr	Bias	Std	RMS	corr	Bias	Std	RMS
ARM_Atq	1235	93.8	-2.47	3.69	4.44	93.7	-3.17	3.69	4.87
ARM_Bar	1594	94.1	-0.73	4.30	4.36	94.6	-1.14	4.02	3.86
PROMICE KAN-M	422	93.9	-3.65	3.37	4.96	94.6	-4.56	3.14	5.53
PROMICE KAN-U	239	93.9	-1.75	3.32	3.75	94.4	-3.39	3.17	4.64
PROMICE KPC-U	488	97.6	-1.31	2.62	2.92	98.2	-3.20	2.27	3.92
PROMICE NUK-U	296	77.7	-4.09	5.00	6.45	84.7	-7.19	4.01	8.23
PROMICE QAS-U	407	83.9	-1.65	4.20	4.51	86.3	-3.70	3.75	5.27
PROMICE SCO-U	403	91.5	-4.60	4.25	6.26	93.7	-7.55	3.75	8.43
PROMICE TAS-U	386	67.5	-1.03	5.43	5.52	79.5	-3.61	4.39	5.68
PROMICE UPE-U	125	88.2	-3.13	3.88	4.97	90.0	-5.49	3.50	6.50
All data	5595	92.9	-2.03	4.24	4.70	93.2	-3.36	4.12	5.32

3 Methods

3.1 Regression model

Nielsen-Englyst et al. (2019) analysed a large number of in situ stations with simultaneous T2m and IST_{skin} observations and showed that empirical relationships existed between T2m and IST_{skin}. It was also shown, however, that the relationships varied for ~~the~~ different regions. Based upon these results, it was decided to use a simple regression based method in this paper to derive the daily mean T2m from the satellite IST_{skin_L3} observations. Separate regression models have been derived for land ice and sea ice.

To test different types of regression models, the IST_{skin_L3} data have been matched up with in situ observations for each day (Høyer et al., 2018). This is done by requiring a distance to nearest in situ site of less than 0.5 degree (approximate 55 km in latitude). The matching threshold was chosen to ensure enough matchups to derive and validate the different regression models with sufficient statistical confidence. All in situ observations, described in Sect. 2.1., have been matched up with IST_{skin_L3} data, resulting in a total number of daily matchups of 65,810 from 275 different observation sites (see Table 1). These have been divided into two subsets: one for training and one for validation of the different regression models for land- and sea ice, respectively. This has been done while ensuring similar coverage of training and validation data over the two domains, which is shown in Fig. 5. The result is that 40% (13,792 ~~observations~~matchups) are used for testing the regression models (and generating the regression coefficients) and the remaining 60% (20,872 ~~observations~~matchups) are left for validation of the regression models over land ice. Over sea ice 48% (15,035 matchups~~observations~~) are used for testing and 52% (16,111 matchups) are left for validation.

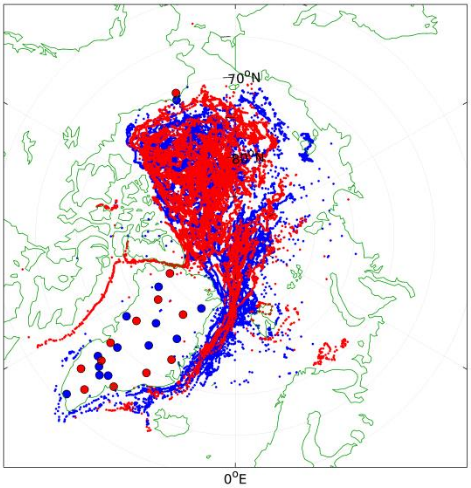


Figure 5: Positions of matchups on sea ice and land ice (red: training, blue: validation)

The regression model is based on multiple linear regression analysis using least squares (Menke, 1989). The multiple linear regression analysis equations can be written in matrix form,

$$\boldsymbol{d}^{obs} = \mathbf{G}\boldsymbol{m} + \boldsymbol{e} \quad (4)$$

$$\boldsymbol{d}^{pre} = \mathbf{G}\boldsymbol{m}, \quad (5)$$

where $\boldsymbol{d}^{obs}$ and $\boldsymbol{d}^{pre}$ are vectors containing the observed and modelled in situ air temperatures, respectively, $\mathbf{G}$ is a matrix containing the various predictors, $\boldsymbol{m}$ is a vector containing regression coefficients, and e is the fitting error.

The regression coefficients are found using damped least squares (Menke, 1989). The least squares method is used since the problem is generally over-determined, and the damping is added to limit effects of noisy data. The regression coefficients are thus given as:

$$\mathbf{G}^{-\mathbf{g}} = (\mathbf{G}^{\mathbf{T}}\mathbf{G} + \varepsilon^2\mathbf{I})^{-1}\mathbf{G}^{\mathbf{T}} \quad (6)$$

$$\boldsymbol{m} = \mathbf{G}^{-\mathbf{g}}\boldsymbol{d}^{obs}, \quad (7)$$

where $\mathbf{G}^{-\mathbf{g}}$ is called the generalized inverse, ε is a damping factor and **I** is an identity matrix (with ones in the diagonal and zeros elsewhere). The superscript operator T denotes transposing and -1 denotes inversion. We have tested a range of damping factors to assess the relation to the error coefficients. A damping factor ~~-~~of 0.2 was chosen to avoid overfitting noise in the data, while keeping the error coefficients low.

The choice of predictors is based on current knowledge of the parameters that influence the relationship between IST_{skin} and $T2m_{insitu}$ (Nielsen-Englyst et al., 2019), limited by the available satellite data. Nielsen-Englyst et al. (2019) showed that ~~on average $T2m_{insitu}$ is 0.65-2.65°C warmer than IST_{skin} with variations depending on region (lower ablation zone, upper-middle ablation zone, accumulation zone, seasonal snow cover and sea ice). T~~the T2m-Tskin difference varies over the ~~day and~~ season with smallest differences ~~around noon and early afternoon~~ during spring, fall and summer in non-melting conditions.

For that reason, we have also tested the effect of including a seasonal cycle as predictor. ~~Nielsen-Englyst et al. (2019) also found that at the observation sites located on the Arctic sea ice and snow covered regions of North Alaska the T2m$_{insitu}$-IST$_{skin}$ difference decreases almost linearly as a function of wind speed due to increased turbulent mixing of the air for higher wind speeds. Contrary, the maximum T2m$_{insitu}$-IST$_{skin}$ differences over the GrIS occur at wind speeds of about 5 m s^{-1}. This is also seen by Adolph et al. (2018) at Summit, GrIS and by Hudson and Brandt (2005) at the South Pole, and the feature is related to the pronounced katabatic winds in these regions. Furthermore, Nielsen-Englyst et al. (2019) found that the T2m$_{insitu}$-IST$_{skin}$ difference tends to decrease linearly as a function of the cloud cover fraction for all seasons and all regions. The reason for this is that clouds have a predominately warming effect on the skin temperature in the Arctic (Intrieri, 2002; Walsh and Chapman, 1998). Nielsen-Englyst et al. (2019) showed an almost linear relationship between the T2m$_{insitu}$-IST$_{skin}$ difference and the IST$_{skin}$, with larger differences for colder skin temperatures. Based on these findings we have calculated the correlations between satellite skin temperature (IST$_{skin_L3}$), in situ surface air temperatures (T2m$_{insitu}$), latitude (Lat), downward shortwave radiation (SWd) and not considering clouds (theoretical), and wind speed (WS) from ERA-Interim reanalysis. Since the cloud cover fraction and longwave radiation are unknown in this case, we have tested IST$_{skin_L3}$ as a predictor instead. The resulting correlations are shown in Table 3.~~

~~**Table 3 Correlations between satellite-measured IST$_{skin_L3}$, in situ measured T2m$_{insitu}$, latitude (Lat), theoretical downward shortwave radiation (SWd), and ERA-Interim wind speed (WS).**~~

		IST$_{skin_L3}$	T2m$_{insitu}$	Lat	SWd	WS
Land ice	IST$_{skin_L3}$	1.00	0.96	-0.22	0.72	-0.25
	T2m$_{insitu}$	0.96	1.00	-0.25	0.61	-0.28
	Lat	-0.22	-0.25	1.00	-0.05	0.10
	SWd	0.72	0.61	-0.05	1.00	-0.23
	WS	-0.25	-0.28	0.10	-0.23	1.00
Sea ice	IST$_{skin_L3}$	1.00	0.96	-0.07	0.79	-0.06
	T2m$_{insitu}$	0.96	1.00	-0.03	0.74	-0.07
	Lat	-0.07	-0.03	1.00	0.03	-0.04
	SWd	0.79	0.75	0.03	1.00	-0.21
	WS	-0.06	-0.07	-0.04	-0.21	1.00

~~The IST$_{skin_L3}$ and air temperatures are well correlated (above 90% correlation), and IST$_{skin_L3}$ also show correlation with the shortwave radiation. Part of the correlation between temperature and the theoretical shortwave radiation is expected to be due to correlation of a seasonal cycle in both signals, not necessarily indicating causality. Therefore, for the regression modelling, a seasonal cycle with fit of amplitude and phase was also tested.~~ A total of 5 regression models with different predictors have been tested (Høyer et al., 2018):

ÎST$_{skin}$: $$T2m_{sat} = \alpha_0 + \alpha_1 IST_{skin\_L3} \quad (8)$$

ÎST$_{skin}$SWd: $$T2m_{sat} = \alpha_0 + \alpha_1 IST_{skin\_L3} + \alpha_2 SWd \quad (9)$$

ÎST$_{skin}$WS: $$T2m_{sat} = \alpha_0 + \alpha_1 IST_{skin\_L3} + \alpha_2 WS \quad (10)$$

ÎST$_{skin}$Lat: $$T2m_{sat} = \alpha_0 + \alpha_1 IST_{skin\_L3} + \alpha_2 Lat \quad (11)$$

ÎST$_{skin}$Season: $$T2m_{sat} = \alpha_0 + \alpha_1 IST_{skin\_L3} + \alpha_2 \mathrm{COS}\left(\frac{t \cdot 2\pi}{1\,yr}\right) + \alpha_3 \mathrm{SIN}\left(\frac{t \cdot 2\pi}{1\,yr}\right) \quad (12)$$

The regression model in Eq. (8) is limited to an offset and a scaling of IST$_{skin_L3}$, where the latter term accounts for the synoptic and seasonal variations, which are the dominating factors in both IST and T2m variability. This part is thus included in all regression models tested. The ~~while all~~ other regression models also have a third predictor, which is included to examine how to best represent the residual variations in the T2m-IST difference. The model in Eq. (9) uses theoretical shortwave radiation, Eq. (10) uses the wind forcing, Eq. (11) uses latitude variation, and Eq. (12) uses a seasonal variation. In the regression model in Eq. (12), the seasonal variation is assumed to be the shape of a cosine function, $A \cdot \mathrm{COS}\left(\frac{t \cdot 2\pi}{1\,yr} - \varphi\right)$, where A is the amplitude, φ is the phase and t is time. Since $\mathrm{COS}(x_1 - x_2) = \mathrm{COS}(x_1)\mathrm{COS}(x_2) + \mathrm{SIN}(x_1)\mathrm{SIN}(x_2)$, the seasonal cycle can be rewritten to the form in Eq. (12) with $A = \sqrt{{\alpha_2}^2 + {\alpha_3}^2}$ and $\varphi = \mathrm{ARCTAN}\left(\frac{\alpha_3}{\alpha_2}\right)$.

The training data have been used to calculate the regression coefficients for each regression model covering land ice and sea ice, respectively. The ~~training data have been used to investigate the~~ performance of each regression model has been investigated using the training data and the results are shown in Table 4. The best ~~correlation~~ performance is found by using the regression model where T2m$_{sat}$ is predicted from IST$_{skin_L3}$ combined with a seasonal variation (ÎST$_{skin}$Season). This model predicts T2m$_{sat}$ better compared to the other regression models for both surface types, with correlations above 96 % and RMS values of 3.25-3.28°C against training data for both surface types (Table 4). In the following, we will use the regression model given in Eq. (12) with the seasonal term included and with separate regression coefficients for land ice and sea ice, respectively. The values are shown in Table 5. The phase corresponds to a maximum the 19th January and 12th February for land ice and sea ice, respectively. This is in agreement with Nielsen-Englyst et al. (2019) who found the strongest clear-sky inversion during the winter months (Dec-Feb) for all sites included in the analysis except from the ones located in the lower ablation zone (not included here), where pronounced surface melt takes place for long periods of time.

Table 4: Statistics on the relation between observed and modelled temperatures for the training data. N: number of matchups used for testing, Corr: correlation, RMS: root mean square difference. Since the training data are used for the regression, the bias is 0 and thus the standard deviation equals RMS.

	N	Corr (%)	RMS (°C)

Land ice	ÎST$_{skin}$	13792	95.7	3.51
	ÎST$_{skin}$SWd	13792	96.2	3.28
	ÎST$_{skin}$WS	13792	95.8	3.47
	ÎST$_{skin}$Lat	13792	95.8	3.48
	ÎST$_{skin}$Season	13792	96.3	3.28
Sea ice	ÎST$_{skin}$	15035	96.0	3.32
	ÎST$_{skin}$SWd	15035	96.0	3.32
	ÎST$_{skin}$WS	15035	96.0	3.32
	ÎST$_{skin}$Lat	15035	96.1	3.28
	ÎST$_{skin}$Season	15035	96.2	3.25

Table 5: Model regression coefficients for ÏST$_{skin}$Season.

	Offset, α_0 (°C)	IST$_{skin_L3}$ factor, α_1	Amplitude, A	Phase, φ
Land ice	4.20	1.06	2.26	-0.33
Sea ice	1.46	0.89	1.83	-0.75

3.2 Uncertainty estimates for T2m$_{sat}$

Uncertainty estimates on the derived T2m$_{sat}$ are crucial to facilitate the usage of the data set in modelling and for monitoring purposes. The uncertainty estimates of the satellite-derived T2m$_{sat}$ data follow the approach in Bulgin et al. (2016) and Rayner et al. (2015), which has also been used for the AASTI data. The uncertainty on a single T2m$_{sat}$ estimate is divided into random, locally correlated and systematic uncertainty components, with the total uncertainty μ_{total_t2m} given as the square root of the sum of the three squared components:

$$\mu_{total\_T2m} = \sqrt{\mu_{rnd\_T2m}{}^{2} + \mu_{local\_T2m}{}^{2} + \mu_{glob\_T2m}{}^{2}}$$

The random uncertainty component for the T2m$_{sat}$ belonging to a particular grid cell at a particular point in time is found by propagating the AASTI IST$_{skin_L3}$ random uncertainty through the regression model:

$$\mu_{rnd\_T2m} = \sqrt{\left(\alpha_1 \mu_{rnd\_L3}\right)^2},$$

with μ_{rnd_L3} given as the aggregated μ_{rnd_L2}:

$$\mu_{rnd\_L3} = \frac{\mu_{rnd\_L2}}{\sqrt{N}},$$

where N is the number of observations for each bin in the aggregation from L2 to L3. The $\sqrt{N}$ reduction applies because the random uncertainty of each L2 data point that goes into the L3 calculation is by definition independent from the other.

The L3 global uncertainty component does not average out in any aggregation and is thus transferred directly from the L2 uncertainty estimate and has been multiplied by α_1 to make up μ_{glob_T2m}:

$$\mu_{glob\_T2m} = \alpha_1 \mu_{glob\_L3} = \alpha_1 \cdot 0.1°\mathrm{C}$$

The μ_{local_T2m} contains both the local uncertainty component of L2, a sampling error μ_{lsamp_L3} related to sampling errors in space and time due to the aggregation, a relationship error, cloud mask uncertainty etc. When aggregating from L2 to daily L3, additional sources of uncertainty enter through the gridding process as IST_{skin_L3} can only be retrieved for clear-sky pixels. This introduces a temporal and spatial sampling uncertainty. If all our satellite observations were obtained during all-sky conditions we assume that the high polar temporal coverage is such that the temporal sampling uncertainty in the L3 files can be set to zero. However, this is not the case and using only clear-sky observations generally leads to a clear-sky bias in averaged IST_{skin} satellite observations when compared to in situ observations (Hall et al., 2012; Nielsen-Englyst et al., 2019; Rasmussen et al., 2018). The relationship error represents the standard deviation of the residuals calculated at in situ stations, where both skin and air temperatures are available, i.e. $T2m_{sat}$-$T2m_{insitu}$. Estimating all the different components that make up the μ_{local_T2m} is a very challenging task and is out of the scope of this paper. Instead, we estimate the μ_{local_T2m} component using a simple regression model fitted to the satellite derived T2m and in situ T2m differences. Separate models have been chosen for the land ice and sea ice, due to the differences in the error characteristics. The variables to include in the uncertainty regression models have been chosen from a careful examination of the matchup data set. For land ice and sea ice the most relevant variables were the IST_{skin_L3} itself and the number of 3 h time bins with observations in the L3, N_{bins}.

For land ice the regression model for μ_{local_T2m} is given as following:

$$\mu_{local\_T2m\_landice} = \beta_0 + \beta_1 IST_{skin\_L3} + \beta_2 N_{bins} ,$$

while the regression model for sea ice is given as:

$$\mu_{local\_T2m\_seaice} = \gamma_0 + \gamma_1 IST_{skin_{L3}} + \gamma_2 {IST_{skin_{L3}}}^2 + \gamma_3 N_{bins} .$$

The coefficients have been determined by fitting to the $T2m_{sat}$-$T2m_{insitu}$ standard deviations calculated for the training data with IST_{skin_L3} bin intervals of 2°C and N_{bins} interval of 1. The μ_{rnd_T2m} and μ_{glob_T2m} components have been removed from the standard deviations in each bin as well as an assumed in situ uncertainty of 0.1°C and an average sampling uncertainty of 0.5°C (Høyer et al., 2017a; Reeves Eyre and Zeng, 2017) before fitting the regression models. The optimal regression coefficients for each domain are listed in Table 6.

Table 6: Uncertainty model regression coefficients

Land ice	β_0 = 3.82°C	β_1 = -0.24	β_2 = -0.03	
Sea ice	γ_0 = 2.01°C	γ_1 = -0.06	γ_2 = -0.12	γ_3 = -0.001

4 Results

In Sect. 3.1 we selected the best (Eq. 12) of the 5 different algorithms and the derived coefficients (Table 4 and 5) to retrieve T2m from satellite surface temperature estimates. The dataset consists of daily estimates of mean air temperature on a 0.25 degree regular latitude-longitude grid, during the period 2000-2009 (Høyer et al., 2018; Kennedy et al., 2019). Days with clouds and few clear sky observations (as explained in Section 2.2) are not included in the dataset. However, for those days when the satellite derived T2m product is available, it provides an estimate of the daily averaged all-sky T2m (see discussion, Section 5). Each temperature estimate is associated with three components of uncertainty: random uncertainties on the 0.25 degree daily scale, synoptic scale correlated uncertainty and globally correlated uncertainty excluding uncertainties related to the masking of clouds. The three types of uncertainties are also gathered in a total uncertainty estimate. The land ice temperatures have been calculated for grid cells categorized as ice shelf by ETOPO1, averaged to the 0.25 degree grid (Amante and Eakins, 2009). Sea ice temperatures have been calculated for grid cells with sea ice concentrations above 30 %, according to OSISAF (Tonboe et al., 2016).

An evaluation of the product and the $T2m_{sat}$ regression model performance has been carried out by a comparison to the independent in situ data (i.e. validation subset described in Sect. 3.1). Figure 6 shows an example of daily near surface air temperatures on Jan 1st, 2008. Circles are in situ T2m measurements from coincidence independent AWSs and buoys. The overall model performance, when compared to all independent AWS and buoy observations, is summarized in Table 7. The satellite derived air temperatures are about 0.3°C warmer than measured in situ air temperature for both land ice and sea ice. The correlations are above 95 % for both surface types and the RMS is 3.47°C and 3.20°C for land and sea ice, respectively. Note that the uncertainty of the in situ data is also included in these RMS values.

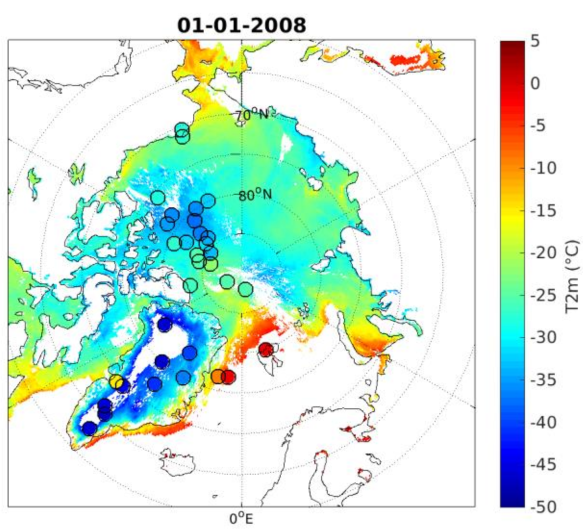


Figure 6: Daily mean air surface temperature over land ice and sea ice from January 1, 2008. Circles show in situ measurements.

Table 7: Statistics on the relation between satellite-derived and in situ measured temperatures for comparison with independent validation data. N: number of matchups used for validation, Corr: correlation, bias: $T2m_{sat} - T2m_{insitu}$ difference, Std: standard deviation, RMS: root mean square difference.

	N	Corr (%)	bias (°C)	Std (°C)	RMS (°C)
Land ice	20872	95.5	0.30	3.45	3.47
Sea ice	16111	96.5	0.35	3.18	3.20

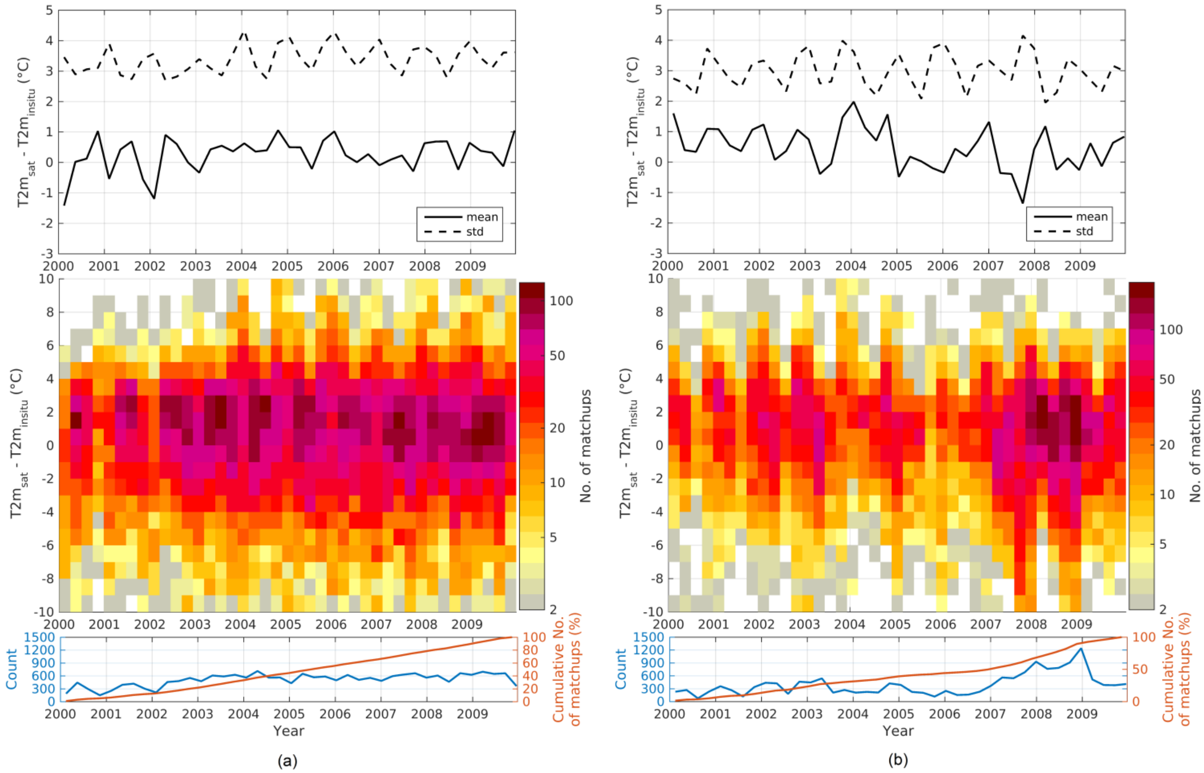


Figure 7: Estimated T2m minus observed T2m (bin size of 1°C) for the full time period (bin size of 90 days) for (a) land ice and (b) sea ice~~, respectively~~. The dashed lines are standard deviations while the solid lines are bias in the upper figure. The surface plots in the middle figures show the number of matchups in each bin, while the bottom plots show the number of matchups (blue) and the cumulative percentage of matchups (red) in each time bin.

Figure 7 shows the seasonal averaged independent validation statistics for land ice and sea ice, respectively. For both land ice and sea ice there are a seasonal dependency in standard deviation with largest values during winter and smallest during summer. This is probably explained by a better cloud screening performance during sunlit periods (Karlsson and Dybbroe, 2010). No significant seasonal cycle is seen in the mean bias, except for the sea ice region during 2000-2004, where there is a tendency to a warm bias during December and January.

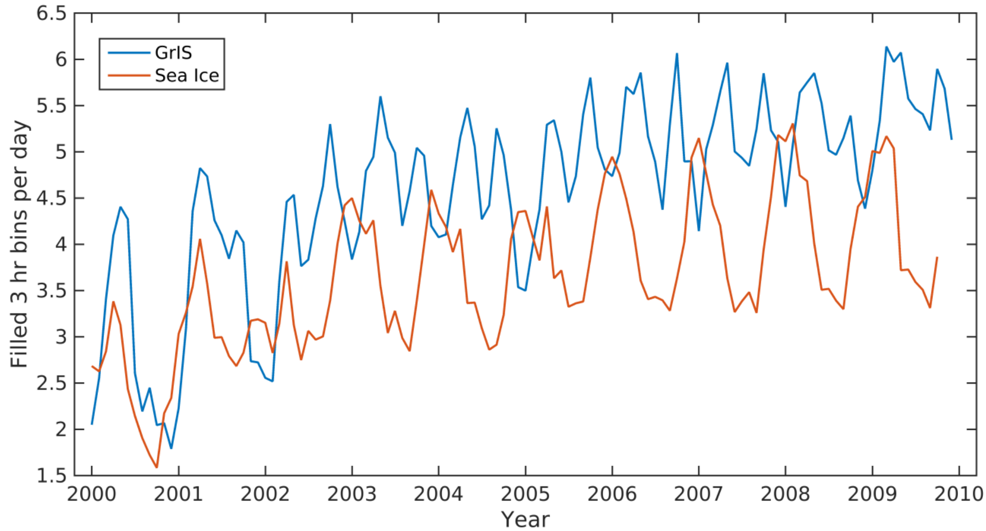


Figure 8: Average number of filled 3 h bins per day for the Greenland Ice Sheet and the Arctic Sea Ice, respectively.

As more satellite observations have become available over the time period a better coverage of the surface temperature is expected over time. Figure 8 shows the average number of filled 3 h bins per day for the GrIS and Arctic Sea Ice, 2000-2009. Both surface types show an increase in filled 3 h bins over time, with large seasonal variations. In most years sea ice has 1-1.5 filled bins per day more during winter than summer, due to a more extensive cloud cover over sea ice during summer. The GrIS typically has fewer filled bins per day during winter and summer, than spring and fall, which is also explained by differences in cloud coverage. Note that the increase in the average number of filled 3 h bins from 2000 to 2009 is not reflected in the performance of the T2m product (Figure 7).

Figure 9 shows $T2m_{sat}$-$T2m_{insitu}$ differences plotted as a function of AASTI L3 skin temperature for land ice and sea ice, respectively. Over land ice, the standard deviation decreases as a function of IST_{skin_L3}, while the bias is around zero for IST_{skin_L3} between -45°C and -10°C, and positive for higher temperatures and negative for lower temperatures. For sea ice the maximum standard deviation is found at skin temperatures of about -20°C, with smaller standard deviations for higher and lower IST_{skin_L3}. Positive biases are found for very cold skin temperatures (< -25°C) and for temperatures around the melting point (> -4°C), while the intermediate temperatures have a slightly negative bias. This effect is included in the uncertainty estimates as presented in Sect. 3.2, which include IST_{skin_L3} as a predictor for both land ice and sea ice.

Figure 10 shows the validation results of the estimated uncertainties, where the $T2m_{sat}$-$T2m_{insitu}$ difference is plotted against the theoretical total uncertainties as obtained in Sect. 3.2 for land ice and sea ice, respectively. The dashed lines represent the ideal uncertainty with the assumptions that the in situ observations have an uncertainty of 0.1°C and that the sampling uncertainty is 0.5°C. The estimated uncertainties show good agreement with the observed uncertainties for both land ice and sea ice, when the error bars follow the dashed line, which is the case here.

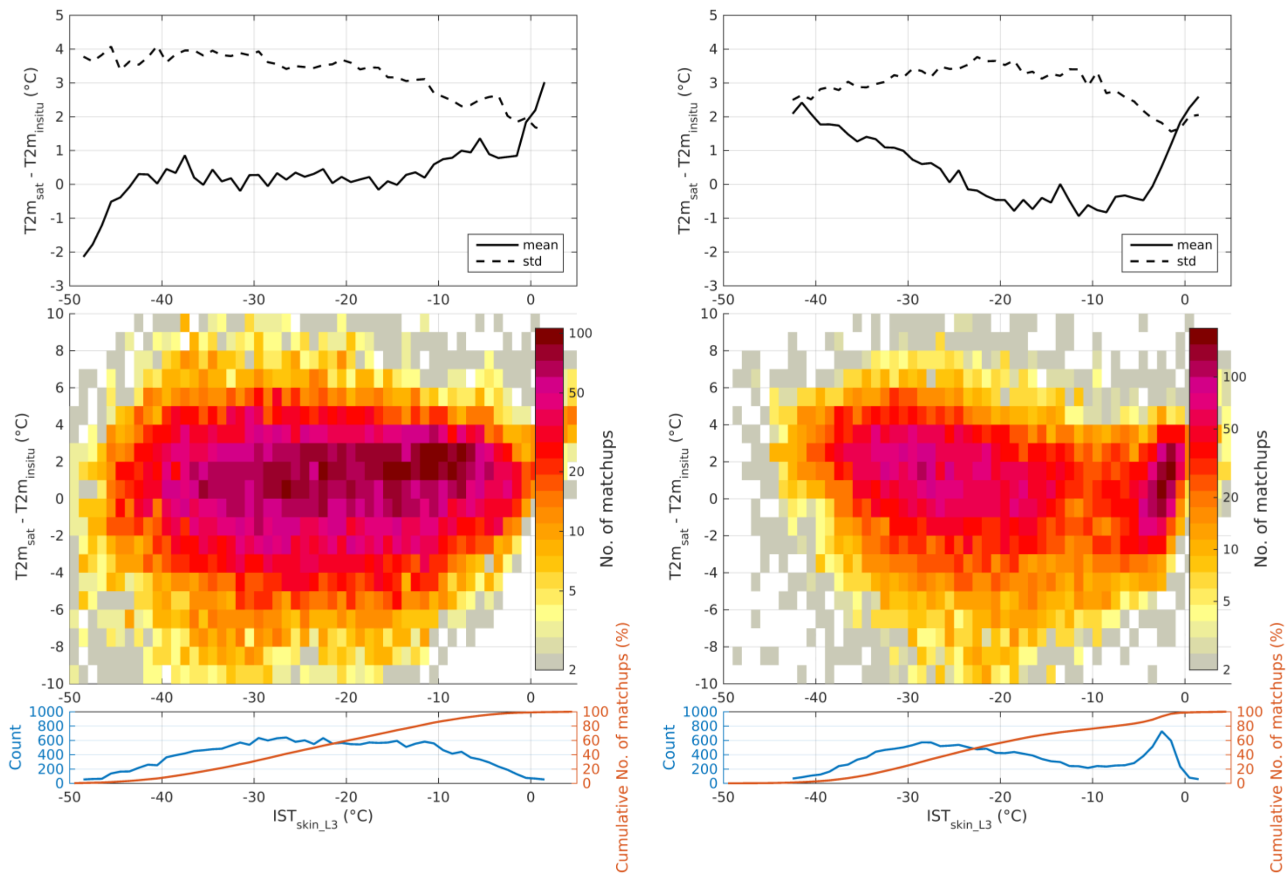


Figure 9: Estimated T2m minus observed T2m (bin size of 1°C) as a function of binned (bin size of 1°C) satellite IST_{skin_L3} for (a) land ice and (b) sea ice~~, respectively~~. The dashed lines are standard deviations while the solid lines are bias in the upper figure. The surface plots in the middle figures show the number of matchups in each bin while the bottom plots show the number of matchups (blue) and the cumulative percentage of matchups (red) in each IST_{skin_L3} bin.

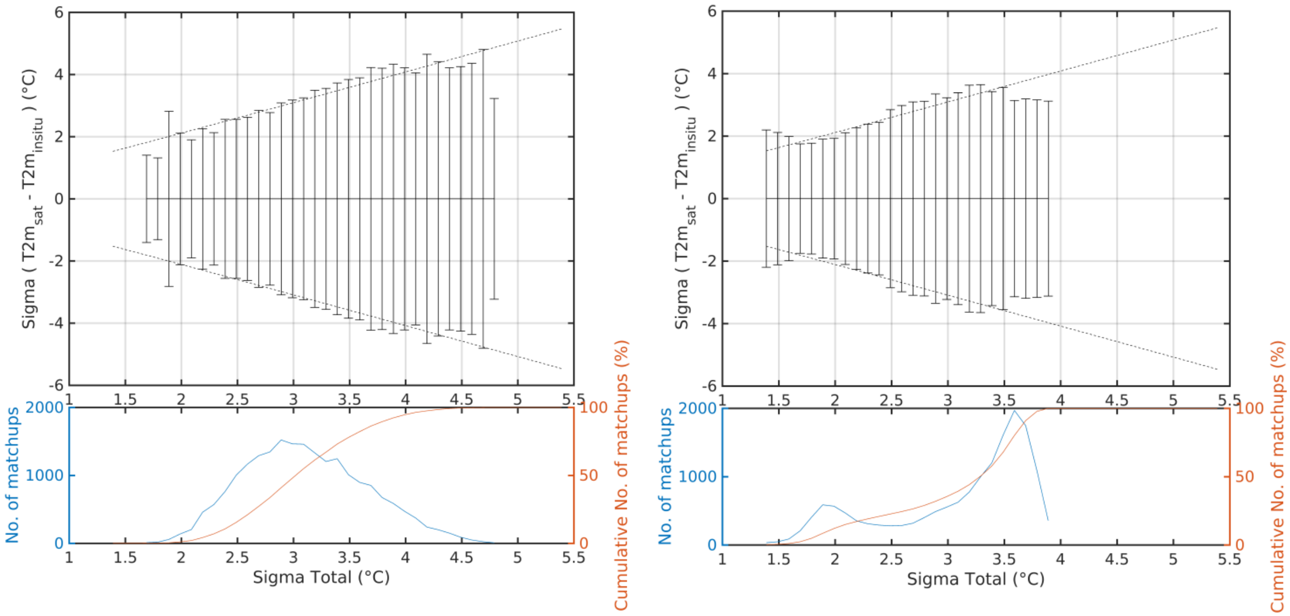


Figure 10: Satellite estimated T2m uncertainty validation with respect to independent in situ T2m for (a) land ice and (b) sea ice. Dashed lines show the modelled uncertainty accounting for uncertainties in the in situ T2m and the sampling error. Solid black lines show one standard deviation of the estimated minus in situ differences for each 0.1 °C bin. The bottom plots show the number of matchups (blue) and the cumulative percentage of matchups for each bin (red).

The performance of $T2m_{sat}$ has been compared to the performance of T2m from ECMWF's reanalysis ERA-Interim ($T2m_{ERA}$; Dee et al., 2011). Table 8 shows the performance of $T2m_{ERA}$ against the independent in situ T2m observations, which should be compared with the performance of the regression derived $T2m_{sat}$ as shown in Table 7. The comparison may not be truly independent as a number of stations and buoys have been assimilated into the ERA-Interim data product (Dee et al., 2011), which would favour the ERA-Interim in the comparison. Yet, the bias is significantly lower for $T2m_{sat}$ than for $T2m_{ERA}$, and the other validation parameters are similar, with slightly better correlation and standard deviation, but slightly worse RMS results for $T2m_{ERA}$.

Table 8: Statistics on the relation between ERA-Interim and in situ measured temperatures for independent test data. N: number of matchups used for validation, Corr: correlation, bias: $T2m_{sat}$ – $T2m_{insitu}$ difference, Std: standard deviation, RMS: root mean square difference.

	N	Corr (%)	Bias (°C)	Std (°C)	RMS (°C)
Land ice	20872	96.4	3.41	3.18	4.66
Sea ice	16111	96.9	1.14	3.02	3.22

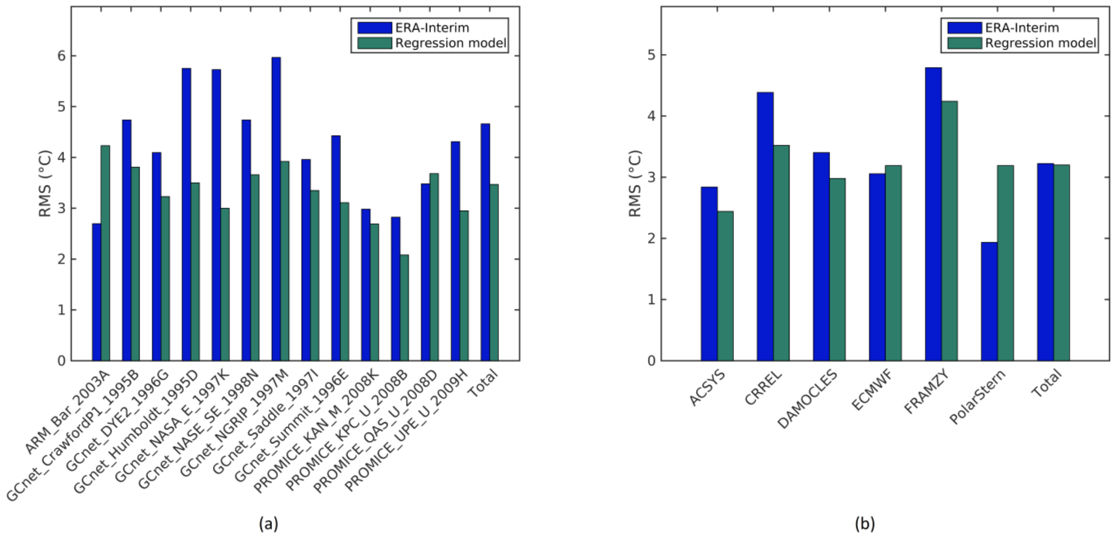


Figure 11: Root mean square (RMS) differences calculated for (a) land ice sites and (b) sea ice sites. Blue bars are RMS obtained by comparing in situ with ERA-Interim T2m, while green bars are in situ compared with the regression derived T2m. Only buoys with more than 200 observations are included. The last two bars listed as "total" are the RMS obtained by using all validation data.

Figure 11 gives an indication of the performance of $T2m_{sat}$ and $T2m_{ERA}$ at individual sites for each surface type. It shows the RMS difference between in situ measured T2m and (a) $T2m_{sat}$ and (b) $T2m_{ERA}$ for the independent test sites and for both surface types. Due to the large number of buoys these have been validated for each data source with all observations weighted equally. The last bars refer to the RMS obtained by validating all test sites in one long time series weighting all daily observations equally. The total $T2m_{sat}$ agrees better with in situ observations for both surface types compared to ERA-Interim. Over the GrIS $T2m_{sat}$ performs better than $T2m_{ERA}$, while ERA-Interim agrees better with in situ observations from the North Alaska site, Barrows. Over sea ice $T2m_{ERA}$ agrees better with in situ observations from ECMWF data stream and Polarstern. However, these may be assimilated into ERA-Interim. The independent in situ observations by ACSYS, CRREL, DAMOCLES and FRAMZY are better reproduced by the satellite-derived T2m. The errors in $T2m_{ERA}$ and $T2m_{sat}$ are expected to be independent and uncorrelated and a combination of the two data sets can therefore lead to an improved T2m estimate.

The monthly mean near surface air temperature estimates averaged over the GrIS have been shown in Fig. 12 for 2000-2009. The GrIS records a distinct annual cycle in near surface air temperature. The monthly mean air temperature typically reaches a maximum of -4°C during July and a minimum of about -28°C during winter. As is common for the Arctic environment, the temporal variability is largest during winter due to a more vigorous atmospheric circulation (Steffen, 1995).

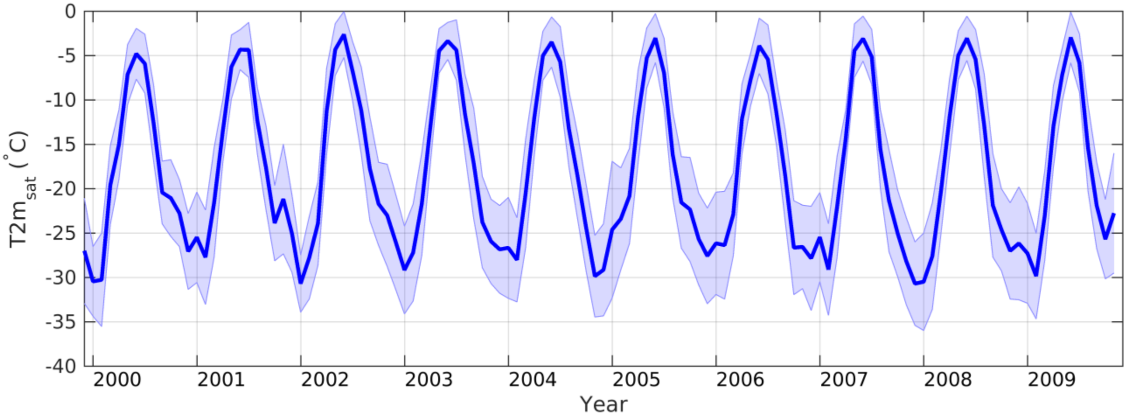


Figure 12: Monthly mean T2m$_{sat}$ for the Greenland Ice Sheet. The shading represents the variability.

The monthly mean T2m$_{sat}$ is shown in Fig. 13 for March, June, September and December averaged over the period 2000-2009. The interior and northern part of the GrIS is typically colder than other parts of the Arctic in all months, while the warmest regions are found along the sea ice marginal ice zone and the ablation zone of the GrIS. During summer little spatial variability in monthly mean T2m is found over the Arctic sea ice.

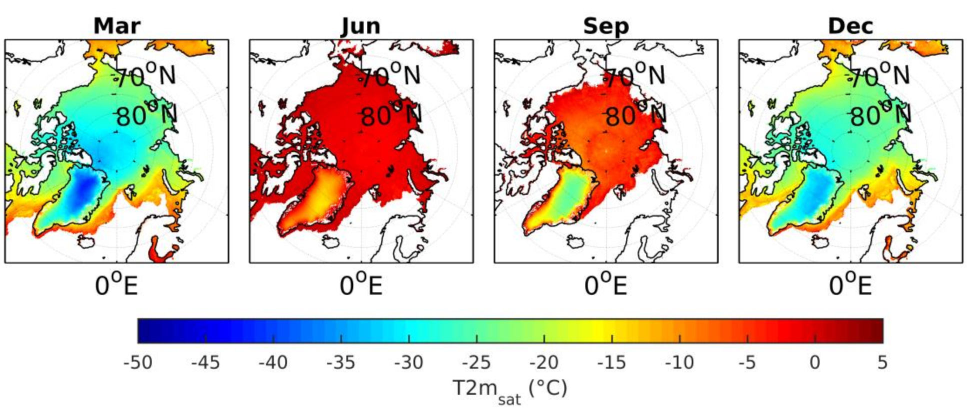


Figure 13: Monthly mean T2m$_{sat}$ during March, June, September and December, averaged for the period 2000-2009.

5 Discussion

Due to the limited number of in situ observations in the Arctic, and especially over sea ice, it is not a simple task to gather in situ observations for testing and validating the regression models. The sparse number of in situ observations over sea ice is the greatest challenge, as also discussed in Nielsen-Englyst et al., 2019. The lack of observations that represent all conditions and regions in the Arctic and the resulting matching threshold of 0.5 degree inflate the uncertainty in the pixel-to-point comparison, due to topography variations over the GrIS, and will therefore complicate the derivation and validation of the regression models.

As infrared satellites cannot measure the surface temperature during cloudy conditions a cold clear-sky bias is often observed in infrared satellite IST_{skin_L3} averages compared to all-sky temperature averages. When using satellite IST_{skin_L3} observations it is thus important to assess the clear-sky bias, which varies with different temporal averaging windows (Nielsen-Englyst et al., 2019). However, through the use of an empirical statistical method, which is trained against daily averaged in situ 2 m air temperatures (obtained both in clear sky and cloudy conditions), the conversion from IST_{skin_L3} to $T2m_{sat}$ removes the systematic IST_{skin_L3} clear-sky bias effects that may be present in the satellite data set. As a result, we obtain a $T2m_{sat}$ estimate which performs similar or better to the IST_{skin_L3} when compared against in situ observations.

For short-lasting (<24 hours) cloudy conditions the division into 3 h bin averages and the requirement of filled 3 h bins both during night (between 18 and~~-~~ 6 local solar time) and day (between 6 and ~~-~~18 local solar time) ensure that the diurnal cycle is best resolved despite the gaps with clouds. For long-lasting (>= 24 hours) cloudy conditions IST_{skin_L3} is not available and we do not retrieve $T2m_{sat}$ for these days. ~~and a~~A statistical technique or the use of atmospheric models and assimilation may be used to fill in the gaps. By using a statistical model to combine in situ observed and clear sky satellite derived T2m estimates (over land, lakes, ocean and ice), including uncertainty estimates, EUSTACE has provided a global and gap free daily analysis of surface air temperatures from 1850 to 2015 (Morice et al., 2019).

The product derived here shows an increasing coverage over the time period from 2000-2003 and a stable coverage for 2003-2009. The average daily coverage is 84% and 67% for land ice and sea ice, respectively, considering the stable 2003-2009 period and the 0.25 degree grid. When considering a 1 degree grid resolution, these numbers increase to 94% and 81%, respectively. The high percentages in coverage demonstrate that the gaps due to cloudy days are limited and that the data set contains a significant amount of information on the all-sky daily T2m even though it is based upon clear sky satellite observations.

Previous studies show a strong dependence of wind speed for both land ice and sea ice, but with different dependencies (Adolph et al., 2018; Hudson and Brandt, 2005; Miller et al., 2013; Nielsen-Englyst et al., 2019). However, the performance of the satellite derived T2m product did not improve much when including the wind speed information from ERA Interim (Table 4). The reason is that the quality of the ERA Interim wind speed is not adequate for use in the relationship model. Especially, the representation of katabatic winds in numerical weather prediction (NWP) models is a challenging task due to a high resolution needed in the vertical (Grisogono et al., 2007; Steeneveld, 2014; Weng and Taylor, 2003; Zilitinkevich et al., 2006), but also the processes of snow surface coupling, radiation and turbulent mixing are hampered by limited resolution, while their relative importance varies with wind speed (Sterk et al., 2013). More accurate information on the wind speed would very likely improve the performance of the regression model when including wind speed as predictor. In particular, the higher resolution NWP output may be very beneficial in the regions of the GrIS, where the local topography interacts with the wind through katabatic effects (DuVivier and Cassano, 2013; Oltmanns et al., 2015; Renfrew, 2004). At the time of the present work, the ERA5 analysis was not available (Copernicus Climate Change Service (C3S), 2017), but this may bring improvements in future work. Moreover, regional high resolution reanalysis products are currently being carried out within the Copernicus Arctic regional Reanalysis service C3S project (https://climate.copernicus.eu/copernicus-arctic-regional-reanalysis-service). It is likely that such products will provide winds that can be used within a relationship model.

The $T2m_{SAT}$ data set developed here only covers the Arctic, but the AASTI data set also covers the Antarctica. This implies that similar statistical methods can be derived for the Antarctic ice cap and sea ice. Preliminary investigations indicate that a T2m product can be derived for the ice caps with similar performance as for the GrIS, whereas the Southern Ocean sea ice is challenging due to very few in situ observations (Morice et al., 2012). For both Southern regions, more in situ observations are needed to repeat the work performed for the Arctic and to determine a reliable statistical model.

Including other available satellite products, such as Modis IST observations (Hall et al., 2004) or the (A)ATSR data set (Ghent et al., 2017) could improve upon the quality of the $T2m_{SAT}$ product. However, adding new data requires a detailed knowledge of the characteristics of the data set, such as sampling frequency and uncertainty of the IST observations. In addition, determination of the relationship model is needed again. At the same time, adding more satellite overpasses to the daily estimates may not improve the uncertainty of the products. This is evident when comparing Fig. 7 and 8 where the variation in the number of satellite observations during the record (Fig. 8) is not reflected in a similar variation in the performance of the product (Fig. 7). The uncertainty in the beginning of the record is comparable to the uncertainty in the end of the record, despite an almost doubling of the observed 3 hourly averages throughout the day.

The AASTI version builds upon the Clara version 1 data set from the CM-SAF. A version 2 of the data set is now available (Karlsson et al., 2017), facilitating the production of an AASTI version 2 data set that covers from 1982 up to present. With a consistency in the retrieval algorithm and data sets, it will be possible to use the relationship model to produce a satellite based climate data record of T2m from 1982 to today.

6 Conclusions

The air temperature over land ice and sea ice is an obvious indicator for Arctic climate change and it can easily be compared with climate change indicators from other regions. This study introduces a methodology for using satellite skin temperatures for estimating air temperatures, to compensate for the lack of in situ measurements, and as a supplement to reanalysis products. Daily near surface air temperatures (T2m) have been estimated based on daily clear sky satellite Level 3 (L3) observations of ice surface skin temperatures (IST_{skin_L3}) in the Arctic, using the Arctic and Antarctic ice Surface Temperatures from thermal Infrared satellite sensors (AASTI) reanalysis. A regression based method has been used and tuned against in situ observed T2m using IST_{skin_L3} observations covering both Arctic sea ice and land ice. In general, there is a good correlation between T2m and IST_{skin_L3}, due to the seasonal cycle in both IST and T2m. Different predictors have been tested to examine how to best capture the variability in the T2m-IST difference. These ~~As explaining factors~~predictors~~,~~ include latitude, theoretical downward shortwave radiation (not considering clouds), seasonal cycle, and wind speed (ERA-Interim reanalysis)~~ were tested. These factors~~ and were selected based on the current knowledge from the literature (Adolph et al., 2018; Hall et al., 2008; Hudson and Brandt, 2005; Nielsen-Englyst et al., 2019; Vihma and Pirazzini, 2005), limited by the available data. The seasonal cycle was introduced based upon the results from an analysis of in situ observations, where a seasonal cycle in the relationship between surface skin and near surface temperature was observed (Nielsen-Englyst et al., 2019). The ~~best~~ highest correlation and lowest RMS against the training data was found using a model where $T2m_{sat}$ is predicted from daily satellite IST_{skin_L3} combined with a seasonal variation assumed to have the shape of an annual harmonic. This model has been used to derive daily T2m on a 0.25 degree regular latitude-longitude grid from the clear sky AASTI IST_{skin_L3} ~~ice surface skin temperatures~~ over the Arctic during the period 2000-2009 (Kennedy et al., 2019), where different regression coefficients have been used for land ice and sea ice. Days with clouds or limited clear sky observations have been excluded from the analysis. Considering a 1 degree regular latitude-longitude grid, the average daily coverage of the $T2m_{sat}$ product is 94% over the GrIS and 81% for sea ice, considering the years 2003-2009. The days when the $T2m_{sat}$ is available the T2m estimate can be considered as a daily averaged all-sky T2m, since it has been tuned against all-sky in situ observations.

The estimated $T2m_{sat}$ data record has been validated against independent in situ measured 2 m air temperatures. The validation results indicate average correlations of 95.5% and 96.5% and average root mean square errors of 3.47°C and 3.20°C for land ice and sea ice, respectively. An uncertainty model has been developed and all daily $T2m_{sat}$ estimates come with a total uncertainty divided into a random, locally systematic and large-scale systematic uncertainty component. The total uncertainty of the satellite derived $T2m_{sat}$ shows good validation results when validated against independent in situ observations.

The satellite derived $T2m_{sat}$ product has been compared to ERA-Interim T2m estimates and has proven to validate similar or better compared with ERA-Interim estimates. The $T2m_{sat}$ product is independent of the quality of the NWP forecasts and thus represents an important ~~alternative~~ supplement to the model based T2m. The regression models presented here both work on satellite observations that are available from reprocessed records but opens up for a near real time estimation of T2m from satellites. The results obtained for the ice covered areas show that there is a large potential for using satellite observed surface temperatures for estimating near surface air temperatures. However, these estimates are not supposed to replace the already existing air temperature measurements, but rather to supplement these e.g. in areas where no in situ observations are currently available.

7. Data availability

The PROMICE data can be accessed through http://www.promice.dk (last access: 16 November 2018). The ARM data are available at https://www.archive.arm.gov/discovery/\#v/results/s/s::co (last access: 21 December 2018). GC-Net data can be found through doi:10.5067/6S7UHUH2K5RI (Kindig, 2010). Data from CRREL mass balance buoys are available from: http://imb-crrel-dartmouth.org (last access: 24 November 2016), while POLARSTERN data can be downloaded at https://dship.awi.de/Polarstern.html (last access: 24 November 2016. FRAMZY data are available from doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2002 (Brümmer et al., 2012b), http://dx.doi.org/doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2007 (Brümmer et al., 2011b), and http://dx.doi.org/doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2008 (Brümmer et al., 2011c), while ACSYS data are found here: DOI: 10.1594/WDCC/UNI_HH_MI_ACSYS2003. Damocles data can be found here: doi:10.1594/wdcc/uni_HH_MI_DAMOCLES2007 (Brümmer et al., 2011a). The traditional buoy and ship data obtained from ECMWF are distributed through the World Meteorological Organization's (WMO) Global Telecommunication System (GTS) and available for members at the ECMWF Meteorological Archival and Retrieval System (MARS). Finally, the AASTI IST$_{skin_L2}$ data are available from http://dx.doi.org/10.5285/60b820fa10804fca9c3f1ddfa5ef42a1~~doi:10.5285/60b820fa10804fca9c3f1ddfa5ef42a1~~ (Høyer et al., 2019). The derived surface air temperatures from satellite surface skin temperatures over ice can be downloaded from ~~doi:~~ http://dx.doi.org/10.5285/f883e197594f4fbaae6edebafb3fddb3 ~~10.5285/f883e197594f4fbaae6edebafb3fddb3~~ (Kennedy et al., 2019).

8 Author contribution

Pia Nielsen-Englyst, Kristine S. Madsen and Gorm Dybkjær compiled and quality checked the in situ data. Pia Nielsen-Englyst, Jacob L. Høyer and Kristine S. Madsen designed and developed the regression model and estimated uncertainties. Gorm Dybkjær, Jacob L. Høyer and Rasmus Tonboe developed the AASTI IST$_{skin_L2}$ data. Pia Nielsen-Englyst prepared the manuscript with contributions from all authors.

9 Competing interests

The authors declare that they have no conflict of interest.

10 Acknowledgements

This study was carried out as a part of the European Union Surface Temperatures for All Corners of Earth (EUSTACE), which is financed by the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171. The aim of EUSTACE is to provide a spatially complete daily field of air temperatures since 1850 by combining satellite and in situ observations. The author would also like to thank the data providers.

References

Ackerman, T. P. and Stokes, G. M.: The Atmospheric Radiation Measurement Program, Phys. Today, 56(1), 38–44, doi:10.1063/1.1554135, 2003.

Adolph, A. C., Albert, M. R. and Hall, D. K.: Near-surface temperature inversion during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures, The Cryosphere, 12(3), 907–920, doi:10.5194/tc-12-907-2018, 2018.

Ahlstrøm, A., van As, D., Citterio, M., Andersen, S., Fausto, R., Andersen, M., Forsberg, R., Stenseng, L., Lintz Christensen, E. and Kristensen, S. S.: A new Programme for Monitoring the Mass Loss of the Greenland Ice Sheet, Geol. Surv. Den. Greenl. Bull., 15, 61–64, 2008.

Amante, C. and Eakins, B. W.: ETOPO1 Global Relief Model converted to PanMap layer format, PANGAEA - Data Publl. Earth Environ. Sci., doi:10.1594/PANGAEA.769615, 2009.

van As, D.: Warming, glacier melt and surface energy budget from weather station observations in the Melville Bay region of northwest Greenland, J. Glaciol., 57(202), 208–220, doi:10.3189/002214311796405898, 2011.

Box, J. E., Colgan, W. T., Christensen, T. R., Schmidt, N. M., Lund, M., Parmentier, F.-J. W., Brown, R., Bhatt, U. S., Euskirchen, E. S., Romanovsky, V. E., Walsh, J. E., Overland, J. E., Wang, M., Corell, R. W., Meier, W. N., Wouters, B., Mernild, S., Mård, J., Pawlak, J. and Olsen, M. S.: Key indicators of Arctic climate change: 1971–2017, Environ. Res. Lett., 14(4), 045010, doi:10.1088/1748-9326/aafc1b, 2019.

Brümmer, B., Müller, G., Haller, M., Kriegsmann, A., Offermann, M. and Wetzel, C.: DAMOCLES 2007-2008 - Hamburg Arctic Ocean Buoy Drift Experiment: meteorological measurements of 16 autonomous drifting ice buoys, World Data Cent. Clim. WDCC, doi:10.1594/wdcc/uni_HH_MI_DAMOCLES2007, 2011a.

Brümmer, B., Müller, G., Lammert-Stockschläder, A., Jahnke-Bornemann, A. and Wetzel, C.: FRAMZY 2007 - Third Field Experiment on Fram Strait Cyclones and their Impact on Sea Ice: meteorological measurements of the research aircraft Falcon, 16 autonomous ice buoys and 13 autonomous water buoys, World Data Cent. Clim. WDCC, doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2007, 2011b.

Brümmer, B., Müller, G. and Wetzel, C.: FRAMZY 2008 - Fourth Field Experiment on Fram Strait Cyclones and their Impact on Sea Ice: meteorological measurements of 7 autonomous ice buoys, World Data Cent. Clim. WDCC, doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2008, 2011c.

Brümmer, B., Launiainen, J., Müller, G., Kirchgaessner, A. and Wetzel, C.: ACSYS 2003 - Arctic Atmospheric Boundary Layer and Sea Ice Interaction Study north of Spitsbergen: meteorological measurements of the research aircraft Falcon, 11 autonomous ice buoys and radiosoundings at the research vessels Aranda and Polarstern, , doi:10.1594/WDCC/UNI_HH_MI_ACSYS2003, 2012a.

Brümmer, B., Launiainen, J., Müller, G. and Wetzel, C.: FRAMZY 2002 - Second Field Experiment on Fram Strait Cyclones and their Impact on Sea Ice: meteorological measurements of the research aircraft Falcon, 15 autonomous ice buoys and radiosoundings at the research vessel Aranda, World Data Cent. Clim. WDCC, doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2002, 2012b.

Bulgin, C. E., Embury, O. and Merchant, C. J.: Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data, Remote Sens. Environ., 177, 287–294, doi:10.1016/j.rse.2016.02.021, 2016.

Cappelen, J.: Greenland—DMI Historical Climate Data Collection 1784–2015., DMI Rep. 16-04, Copenhagen, Denmark, Danish Meteorological Institute, Copenhagen, Denmark., 2016.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J. and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, pp. 1029–1136, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013.

Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), [online] Available from: https://cds.climate.copernicus.eu/cdsapp#!/home (Accessed 9 May 2019), 2017.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656), 553–597, doi:10.1002/qj.828, 2011.

DuVivier, A. K. and Cassano, J. J.: Evaluation of WRF Model Resolution on Simulated Mesoscale Winds and Surface Fluxes near Greenland, Mon. Weather Rev., 141(3), 941–963, doi:10.1175/MWR-D-12-00091.1, 2013.

Dybbroe, A., Karlsson, K.-G. and Thoss, A.: NWCSAF AVHRR Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling. Part I: Algorithm Description, J. Appl. Meteorol., 44(1), 39–54, doi:10.1175/JAM-2188.1, 2005a.

Dybbroe, A., Karlsson, K.-G. and Thoss, A.: NWCSAF AVHRR Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling. Part II: Tuning and Validation, J. Appl. Meteorol., 44(1), 55–71, doi:10.1175/JAM-2189.1, 2005b.

Dybkjær, G., Tonboe, R. and Høyer, J. L.: Arctic surface temperatures from Metop AVHRR compared to in situ ocean and land data, Ocean Sci., 8(6), 959–970, doi:10.5194/os-8-959-2012, 2012.

Dybkjær, G., Høyer, J. L., Tonboe, R. and Olsen, S. M.: Report on the documentation and description of the new Arctic Ocean dataset combining SST and IST., NACLIM Deliverable D32.28. [online] Available from: http://www.naclim.eu/ (Accessed 8 September 2018), 2014.

Dybkjaer, G., Eastwood, S., Borg, A. L., Høyer, J. L. and Tonboe, R.: Algorithm theoretical basis document (ATBD) for the OSI SAF Sea and Sea Ice Surface Temperature L2 processing chain. OSI205a and OSI205b, [online] Available from: http://osisaf.met.no/docs/osisaf_cdop2_ss2_pum_ice-conc_v1p4.pdf (Accessed 3 March 2018), 2018.

Ghent, D. J., Corlett, G. K., Göttsche, F.-M. and Remedios, J. J.: Global Land Surface Temperature From the Along-Track Scanning Radiometers: Global LST from the ATSRs, J. Geophys. Res. Atmospheres, 122(22), 12,167-12,193, doi:10.1002/2017JD027161, 2017.

GHRSST Science Team: The Recommended GHRSST Data Specification (GDS) 2.0, document revision 4, available from the GHRSST International Project Office, 2011, pp 123., 2010.

Good, E.: Daily minimum and maximum surface air temperatures from geostationary satellite data: Satellite min and max air temperatures, J. Geophys. Res. Atmospheres, 120(6), 2306–2324, doi:10.1002/2014JD022438, 2015.

Good, E. J., Ghent, D. J., Bulgin, C. E. and Remedios, J. J.: A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series, J. Geophys. Res. Atmospheres, 122(17), 9185–9210, doi:10.1002/2017JD026880, 2017.

Graversen, R. G., Mauritsen, T., Tjernström, M., Källén, E. and Svensson, G.: Vertical structure of recent Arctic warming, Nature, 451(7174), 53–56, doi:10.1038/nature06502, 2008.

Grisogono, B., Kraljević, L. and Jeričević, A.: The low-level katabatic jet height versus Monin-Obukhov height, Q. J. R. Meteorol. Soc., 133(629), 2133–2136, doi:10.1002/qj.190, 2007.

Hall, D., Box, J., Casey, K., Hook, S., Shuman, C. and Steffen, K.: Comparison of satellite-derived and in-situ observations of ice and snow surface temperatures over Greenland, Remote Sens. Environ., 112(10), 3739–3749, doi:10.1016/j.rse.2008.05.007, 2008.

Hall, D. K., Key, J. R., Case, K. A., Riggs, G. A. and Cavalieri, D. J.: Sea ice surface temperature product from MODIS, IEEE Trans. Geosci. Remote Sens., 42(5), 1076–1087, doi:10.1109/TGRS.2004.825587, 2004.

Hall, D. K., Comiso, J. C., DiGirolamo, N. E., Shuman, C. A., Key, J. R. and Koenig, L. S.: A Satellite-Derived Climate-Quality Data Record of the Clear-Sky Surface Temperature of the Greenland Ice Sheet, J. Clim., 25(14), 4785–4798, doi:10.1175/JCLI-D-11-00365.1, 2012.

Hansen, J., Ruedy, R., Sato, M. and Lo, K.: Global surface temperature change, Rev. Geophys., 48(4), doi:10.1029/2010RG000345, 2010.

Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in coupled models, Clim. Dyn., 21(3–4), 221–232, doi:10.1007/s00382-003-0332-6, 2003.

Høyer, J. L., Alerskans, E., Nielsen-Englyst, P., Thejll, P., Dybkjaer, G. and Tonboe, R.: Detailed investigation of the uncertainty budget for non-recoverable IST observations and their SI traceability (ESA Tech. Rep. FRM4STS OP-70), [online] Available from: http://www.frm4sts.org/wp-content/uploads/sites/3/2018/08/OP-70-FRM4STS_option3_report_v1-signed.pdf (Accessed 4 December 2018a), 2017.

Høyer, J. L., Lang, A. M., Tonboe, R., Eastwood, S., Wimmer, W. and Dybkjær, G.: Towards field inter-comparison experiment (FICE) for ice surface temperature (ESA Tech. Rep. FRM4STS OP-40)., [online] Available from: http://www.frm4sts.org/wp-content/uploads/sites/3/2017/12/OFE-OP-40-TR-5-V1-Iss-1-Ver-1-Signed.pdf (Accessed 2 November 2018b), 2017.

Høyer, J. L., Good, E., Nielsen-Englyst, P., Madsen, K. S., Woolway, I. and Kennedy, J.: Report on the relationship between satellite surface skin temperature and surface air temperature observations for oceans, land, sea ice, ice sheets, and lakes, Deliverable. [online] Available from: https://www.eustaceproject.org/eustace/static/media/uploads/d1.5_revised.pdf (Accessed 6 June 2018), 2018.

Høyer, J. L., Dybkjaer, G., Eastwood, S. and Madsen, K. S.: EUSTACE/AASTI: Global clear-sky ice surface temperature data from the AVHRR series on the satellite swath with estimates of uncertainty components, v1.1, 2000-2009, Centre for Environmental Data Analysis. [online] Available from: http://catalogue.ceda.ac.uk/uuid/60b820fa10804fca9c3f1ddfa5ef42a1, 2019.

Hudson, S. R. and Brandt, R. E.: A Look at the Surface-Based Temperature Inversion on the Antarctic Plateau, J. Clim., 18(11), 1673–1696, doi:10.1175/JCLI3360.1, 2005.

Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M. and Morice, C. P.: Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010:, J. Geophys. Res. Atmospheres, 117(D5), n/a-n/a, doi:10.1029/2011JD017139, 2012.

Karlsson, K.-G. and Dybbroe, A.: Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations, Atmospheric Chem. Phys., 10(4), 1789–1807, doi:10.5194/acp-10-1789-2010, 2010.

Karlsson, K.-G., Riihelä, A., Müller, R., Meirink, J. F., Sedlar, J., Stengel, M., Lockhoff, M., Trentmann, J., Kaspar, F., Hollmann, R. and Wolters, E.: CLARA-A1: a cloud, albedo, and radiation dataset from 28 yr of global AVHRR data, Atmospheric Chem. Phys., 13(10), 5351–5367, doi:10.5194/acp-13-5351-2013, 2013.

Karlsson, K.-G., Anttila, K., Trentmann, J., Stengel, M., Fokke Meirink, J., Devasthale, A., Hanschmann, T., Kothe, S., Jääskeläinen, E., Sedlar, J., Benas, N., van Zadelhoff, G.-J., Schlundt, C., Stein, D., Finkensieper, S., Håkansson, N. and Hollmann, R.: CLARA-A2: the second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data, Atmospheric Chem. Phys., 17(9), 5809–5828, doi:10.5194/acp-17-5809-2017, 2017.

Kennedy, J. J., Capponi, F., Ghent, D., Good, E. J., Høyer, J. L., Kent, E. C., Madsen, K. S., Mitchelson, J. R., Nielsen-Englyst, P. and Tonboe, R. T.: EUSTACE: Globally gridded clear-sky daily air temperature estimates from satellites with uncertainty estimates for land, ocean and ice, 1995-2016, , doi:10.5285/f883e197594f4fbaae6edebafb3fddb3, 2019.

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(2), 302–309, doi:10.1016/S0034-4257(97)89497-7, 1997.

Kindig, D.: Greenland Climate Network (GC-Net) Radiation for Arctic System Reanalysis, Version 1., Boulder Colo. USA NSIDC Natl. Snow Ice Data Cent., doi:10.5067/6S7UHUH2K5RI, 2010.

Knust, R.: Polar Research and Supply Vessel POLARSTERN operated by the Alfred-Wegener-Institute, J. Large-Scale Res. Facil. JLSRF, 3, doi:10.17815/jlsrf-3-163, 2017.

Koenig, L. S. and Hall, D. K.: Comparison of satellite, thermochron and air temperatures at Summit, Greenland, during the winter of 2008/09, J. Glaciol., 56(198), 735–741, doi:10.3189/002214310793146269, 2010.

Masson-Delmotte, V., Swingedouw, D., Landais, A., Seidenkrantz, M.-S., Gauthier, E., Bichet, V., Massa, C., Perren, B., Jomelli, V., Adalgeirsdottir, G., Hesselbjerg Christensen, J., Arneborg, J., Bhatt, U., Walker, D. A., Elberling, B., Gillet-Chaulet, F., Ritz, C., Gallée, H., van den Broeke, M., Fettweis, X., de Vernal, A. and Vinther, B.: Greenland climate change: from the past to the future: Greenland climate change, Wiley Interdiscip. Rev. Clim. Change, 3(5), 427–449, doi:10.1002/wcc.186, 2012.

Menke, W.: Geophysical Data Analysis: Discrete Inverse Theory, Elsevier, New York., 1989.

Merchant, C. J. and Le Borgne, P.: Retrieval of Sea Surface Temperature from Space, Based on Modeling of Infrared Radiative Transfer: Capabilities and Limitations, J. Atmospheric Ocean. Technol., 21(11), 1734–1746, doi:10.1175/JTECH1667.1, 2004.

Merchant, C. J., Harris, A. R., Murray, M. J. and Závody, A. M.: Toward the elimination of bias in satellite retrievals of sea surface temperature: 1. Theory, modeling and interalgorithm comparison, J. Geophys. Res. Oceans, 104(C10), 23565–23578, doi:10.1029/1999JC900105, 1999.

Merchant, C. J., Matthiesen, S., Rayner, N. A., Remedios, J. J., Jones, P. D., Olesen, F., Trewin, B., Thorne, P. W., Auchmann, R., Corlett, G. K., Guillevic, P. C. and Hulley, G. C.: The surface temperatures of Earth: steps towards integrated understanding of variability and change, Geosci. Instrum. Methods Data Syst., 2(2), 305–321, doi:10.5194/gi-2-305-2013, 2013.

Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter, M., Sathyendranath, S., Sandven, S., Sofieva, V. F. and Wagner, W.: Uncertainty information in climate data records from Earth observation, Earth Syst. Sci. Data, 9(2), 511–527, doi:10.5194/essd-9-511-2017, 2017.

Miller, N. B., Turner, D. D., Bennartz, R., Shupe, M. D., Kulie, M. S., Cadeddu, M. P. and Walden, V. P.: Surface-based inversions above central Greenland, J. Geophys. Res. Atmospheres, 118(2), 495–506, doi:10.1029/2012JD018867, 2013.

Morice, C. P., Kennedy, J. J., Rayner, N. A. and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set: THE HADCRUT4 DATASET, J. Geophys. Res. Atmospheres, 117(D8), doi:10.1029/2011JD017187, 2012.

Morice, C. P., Capponi, F., Kennedy, J. J., Killick, R. E., Lindgren, F., Mitchelson, J. R., Rayner, N. A. and Winn, J. P.: EUSTACE: Global daily air temperature combining surface and satellite data, with uncertainty estimates, for 1850-2015, v1.0, Centre for Environmental Data Analysis. [online] Available from: http://dx.doi.org/10.5285/468abcf18372425791a31d15a41348d9, 2019.

Nielsen-Englyst, P., Høyer, J. L., Madsen, K. S., Tonboe, R., Dybkjær, G. and Alerskans, E.: In situ observed relationships between snow and ice surface skin temperatures and 2 m air temperatures in the Arctic, The Cryosphere, 13(3), 1005–1024, doi:10.5194/tc-13-1005-2019, 2019.

Oltmanns, M., Straneo, F., Seo, H. and Moore, G. W. K.: The Role of Wave Dynamics and Small-Scale Topography for Downslope Wind Events in Southeast Greenland, J. Atmospheric Sci., 72(7), 2786–2805, doi:10.1175/JAS-D-14-0257.1, 2015.

Overland, J., Dunlea, E., Box, J. E., Corell, R., Forsius, M., Kattsov, V., Olsen, M. S., Pawlak, J., Reiersen, L.-O. and Wang, M.: The urgency of Arctic change, Polar Sci., doi:10.1016/j.polar.2018.11.008, 2018.

Perovich, D. K., Richter-Menge, J. A. and Polashenski, C. M.: Observing and understanding climate change: Monitoring the mass balance, motion, and thickness of Arctic sea ice, [online] Available from: http://imb-crrel-dartmouth.org (Accessed 24 November 2016), 2016.

Pielke, R. A., Davey, C. A., Niyogi, D., Fall, S., Steinweg-Woods, J., Hubbard, K., Lin, X., Cai, M., Lim, Y.-K., Li, H., Nielsen-Gammon, J., Gallo, K., Hale, R., Mahmood, R., Foster, S., McNider, R. T. and Blanken, P.: Unresolved issues with the assessment of multidecadal global land surface temperature trends, J. Geophys. Res., 112(D24), D24S08, doi:10.1029/2006JD008229, 2007.

Pithan, F. and Mauritsen, T.: Arctic amplification dominated by temperature feedbacks in contemporary climate models, Nat. Geosci., 7(3), 181–184, doi:10.1038/ngeo2071, 2014.

Rasmussen, T. A. S., Høyer, J. L., Ghent, D., Bulgin, C. E., Dybkjaer, G., Ribergaard, M. H., Nielsen-Englyst, P. and Madsen, K. S.: Impact of Assimilation of Sea-Ice Surface Temperatures on a Coupled Ocean and Sea-Ice Model, J. Geophys. Res. Oceans, 123(4), 2440–2460, doi:10.1002/2017JC013481, 2018.

Rayner, N., Good, S. and Block, T.: SST CCI Product User Guide, Project Document, SST_CCI-PUG-UKMO-201., [online] Available from: http://www.esa-sst-cci.org/PUG/pdf/SST_CCI-PUG-UKMO-201-Issue_1-signed.pdf (Accessed 15 March 2019), 2015.

Rayner, N. A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108(D14), doi:10.1029/2002JD002670, 2003.

Rayner, N. A., Auchmann, R., Bessembinder, J., Brönnimann, S., Brugnara, Y., Capponi, F., Carrea, L., Dodd, E. M. A., Ghent, D., Good, E., Høyer, J. L., Kennedy, J. J., Kent, E. C., Killick, R. E., van der Linden, P., Lindgren, F., Madsen, K. S., Merchant, C. J., Mitchelson, J. R., Morice, C. P., Nielsen-Englyst, P., Ortiz, P. F., Remedios, J. J., van der Schrier, G., Squintu, A. A., Stephens, A., Thorne, P. W., Tonboe, R. T., Trent, T., Veal, K. L., Waterfall, A., Winfield, K., Winn, J. and Woolway, R. I.: The EUSTACE project: delivering global, daily information on surface air temperature, Submitt. Bull. Am. Meteorol. Soc., 2019.

Reeves Eyre, J. E. J. and Zeng, X.: Evaluation of Greenland near surface air temperature datasets, The Cryosphere, 11(4), 1591–1605, doi:10.5194/tc-11-1591-2017, 2017.

Renfrew, I. A.: The dynamics of idealized katabatic flow over a moderate slope and ice shelf, Q. J. R. Meteorol. Soc., 130(598), 1023–1045, doi:10.1256/qj.03.24, 2004.

Richter-Menge, J. A., Perovich, D. K., Elder, B. C., Claffey, K., Rigor, I. and Ortmeyer, M.: Ice mass-balance buoys: a tool for measuring and attributing changes in the thickness of the Arctic sea-ice cover, Ann. Glaciol., 44, 205–210, doi:10.3189/172756406781811727, 2006.

Shuman, C. A., Steffen, K., Box, J. E. and Stearns, C. R.: A Dozen Years of Temperature Observations at the Summit: Central Greenland Automatic Weather Stations 1987–99, J. Appl. Meteorol., 40(4), 741–752, doi:10.1175/1520-0450(2001)040<0741:ADYOTO>2.0.CO;2, 2001.

Shuman, C. A., Hall, D. K., DiGirolamo, N. E., Mefford, T. K. and Schnaubelt, M. J.: Comparison of Near-Surface Air Temperatures and MODIS Ice-Surface Temperatures at Summit, Greenland (2008–13), J. Appl. Meteorol. Climatol., 53(9), 2171–2180, doi:10.1175/JAMC-D-14-0023.1, 2014.

Smith, T. M., Reynolds, R. W., Peterson, T. C. and Lawrimore, J.: Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006), J. Clim., 21(10), 2283–2296, doi:10.1175/2007JCLI2100.1, 2008.

Stamnes, K., Ellingson, R. G., Curry, J. A., Walsh, J. E. and Zak, B. D.: Review of Science Issues, Deployment Strategy, and Status for the ARM North Slope of Alaska–Adjacent Arctic Ocean Climate Research Site, J. Clim., 12(1), 46–63, doi:10.1175/1520-0442-12.1.46, 1999.

Steeneveld, G.-J.: Current challenges in understanding and forecasting stable boundary layers over land and ice, Front. Environ. Sci., 2, doi:10.3389/fenvs.2014.00041, 2014.

Steffen, K.: Surface energy exchange at the equilibrium line on the Greenland ice sheet during onset of melt, Ann. Glaciol., 21, 13–18, doi:10.3189/S0260305500015536, 1995.

Steffen, K. and Box, J.: Surface climatology of the Greenland Ice Sheet: Greenland Climate Network 1995-1999, J. Geophys. Res. Atmospheres, 106(D24), 33951–33964, doi:10.1029/2001JD900161, 2001.

Sterk, H. A. M., Steeneveld, G. J. and Holtslag, A. A. M.: The role of snow-surface coupling, radiation, and turbulent mixing in modeling a stable boundary layer over Arctic sea ice: MODELING A CLEAR-SKY ARCTIC SBL, J. Geophys. Res. Atmospheres, 118(3), 1199–1217, doi:10.1002/jgrd.50158, 2013.

Tonboe, R. T., Eastwood, S., Lavergne, T., Sørensen, A. M., Rathmann, N., Dybkjær, G., Pedersen, L. T., Høyer, J. L. and Kern, S.: The EUMETSAT sea ice concentration climate data record, The Cryosphere, 10(5), 2275–2290, doi:10.5194/tc-10-2275-2016, 2016.

Vihma, T. and Pirazzini, R.: On the Factors Controlling the Snow Surface and 2-m Air Temperatures Over the Arctic Sea Ice in Winter, Bound.-Layer Meteorol., 117(1), 73–90, doi:10.1007/s10546-004-5938-7, 2005.

Vihma, T., Jaagus, J., Jakobson, E. and Palo, T.: Meteorological conditions in the Arctic Ocean in spring and summer 2007 as recorded on the drifting ice station Tara, Geophys. Res. Lett., 35(18), doi:10.1029/2008GL034681, 2008.

Vose, R. S., Arndt, D., Banzon, V. F., Easterling, D. R., Gleason, B., Huang, B., Kearns, E., Lawrimore, J. H., Menne, M. J., Peterson, T. C., Reynolds, R. W., Smith, T. M., Williams, C. N. and Wuertz, D. B.: NOAA's Merged Land–Ocean Surface Temperature Analysis, Bull. Am. Meteorol. Soc., 93(11), 1677–1685, doi:10.1175/BAMS-D-11-00241.1, 2012.

Weng, W. and Taylor, P. A.: On Modelling the One-Dimensional Atmospheric Boundary Layer, Bound.-Layer Meteorol., 107(2), 371–400, doi:10.1023/A:1022126511654, 2003.

World Meteorological Organization: World Meteorological Organization (2014) Guide to Meteorological Instruments and Methods of Observation, [online] Available from: http://hdl.handle.net/11329/365 (Accessed 1 February 2018), 2014.

Zilitinkevich, S., Savijärvi, H., Baklanov, A., Grisogono, B. and Myrberg, K.: Forthcoming Meetings on Planetary Boundary-layer Theory, Modelling and Applications, Bound.-Layer Meteorol., 119(3), 591–593, doi:10.1007/s10546-006-9069-1, 2006.

~~Ackerman, T. P. and Stokes, G. M.: The Atmospheric Radiation Measurement Program, Phys. Today, 56(1), 38–44, doi:10.1063/1.1554135, 2003.~~

~~Adolph, A. C., Albert, M. R. and Hall, D. K.: Near-surface temperature inversion during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures, The Cryosphere, 12(3), 907–920, doi:10.5194/tc-12-907-2018, 2018.~~

~~Ahlstrøm, A., van As, D., Citterio, M., Andersen, S., Fausto, R., Andersen, M., Forsberg, R., Stenseng, L., Lintz Christensen, E. and Kristensen, S. S.: A new Programme for Monitoring the Mass Loss of the Greenland Ice Sheet, Geol. Surv. Den. Greenl. Bull., 15, 61–64, 2008.~~

~~Amante, C. and Eakins, B. W.: ETOPO1 Global Relief Model converted to PanMap layer format, PANGAEA - Data Publ. Earth Environ. Sci., doi:10.1594/PANGAEA.769615, 2009.~~

~~van As, D.: Warming, glacier melt and surface energy budget from weather station observations in the Melville Bay region of northwest Greenland, J. Glaciol., 57(202), 208–220, doi:10.3189/002214311796405898, 2011.~~

~~Box, J. E., Colgan, W. T., Christensen, T. R., Schmidt, N. M., Lund, M., Parmentier, F.-J. W., Brown, R., Bhatt, U. S., Euskirchen, E. S., Romanovsky, V. E., Walsh, J. E., Overland, J. E., Wang, M., Corell, R. W., Meier, W. N., Wouters, B., Mernild, S., Mård, J., Pawlak, J. and Olsen, M. S.: Key indicators of Arctic climate change: 1971–2017, Environ. Res. Lett., 14(4), 045010, doi:10.1088/1748-9326/aafc1b, 2019.~~

~~Brümmer, B., Müller, G., Haller, M., Kriegsmann, A., Offermann, M. and Wetzel, C.: DAMOCLES 2007-2008 - Hamburg Arctic Ocean Buoy Drift Experiment: meteorological measurements of 16 autonomous drifting ice buoys, World Data Cent. Clim. WDCC, doi:10.1594/wdcc/uni_HH_MI_DAMOCLES2007, 2011a.~~

~~Brümmer, B., Müller, G., Lammert-Stockschläder, A., Jahnke-Bornemann, A. and Wetzel, C.: FRAMZY 2007 - Third Field Experiment on Fram Strait Cyclones and their Impact on Sea Ice: meteorological measurements of the research aircraft Falcon, 16 autonomous ice buoys and 13 autonomous water buoys, World Data Cent. Clim. WDCC, doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2007, 2011b.~~

~~Brümmer, B., Müller, G. and Wetzel, C.: FRAMZY 2008 - Fourth Field Experiment on Fram Strait Cyclones and their Impact on Sea Ice: meteorological measurements of 7 autonomous ice buoys, World Data Cent. Clim. WDCC, doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2008, 2011c.~~

~~Brümmer, B., Launiainen, J., Müller, G., Kirchgaessner, A. and Wetzel, C.: ACSYS 2003 - Arctic Atmospheric Boundary Layer and Sea Ice Interaction Study north of Spitsbergen: meteorological measurements of the research aircraft Falcon, 11 autonomous ice buoys and radiosoundings at the research vessels Aranda and Polarstern, , doi:10.1594/WDCC/UNI_HH_MI_ACSYS2003, 2012a.~~

~~Brümmer, B., Launiainen, J., Müller, G. and Wetzel, C.: FRAMZY 2002 - Second Field Experiment on Fram Strait Cyclones and their Impact on Sea Ice: meteorological measurements of the research aircraft Falcon, 15 autonomous ice buoys and radiosoundings at the research vessel Aranda, World Data Cent. Clim. WDCC, doi:10.1594/WDCC/UNI_HH_MI_FRAMZY2002, 2012b.~~

~~Bulgin, C. E., Embury, O. and Merchant, C. J.: Sampling uncertainty in gridded sea surface temperature products and Advanced Very High Resolution Radiometer (AVHRR) Global Area Coverage (GAC) data, Remote Sens. Environ., 177, 287–294, doi:10.1016/j.rse.2016.02.021, 2016.~~

~~Cappelen, J.: Greenland - DMI Historical Climate Data Collection 1784–2015., DMI Rep. 16-04, Copenhagen, Denmark, Danish Meteorological Institute, Copenhagen, Denmark., 2016.~~

~~Copernicus Climate Change Service (C3S): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), [online] Available from: https://cds.climate.copernicus.eu/cdsapp#!/home (Accessed 9 May 2019), 2017.~~

~~Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656), 553–597, doi:10.1002/qj.828, 2011.~~

~~DuVivier, A. K. and Cassano, J. J.: Evaluation of WRF Model Resolution on Simulated Mesoscale Winds and Surface Fluxes near Greenland, Mon. Weather Rev., 141(3), 941–963, doi:10.1175/MWR-D-12-00091.1, 2013.~~

~~Dybbroe, A., Karlsson, K.-G. and Thoss, A.: NWCSAF AVHRR Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling. Part I: Algorithm Description, J. Appl. Meteorol., 44(1), 39–54, doi:10.1175/JAM-2188.1, 2005a.~~

~~Dybbroe, A., Karlsson, K. G. and Thoss, A.: NWCSAF AVHRR Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling. Part II: Tuning and Validation, J. Appl. Meteorol., 44(1), 55–71, doi:10.1175/JAM-2189.1, 2005b.~~

~~Dybkjær, G., Tonboe, R. and Høyer, J. L.: Arctic surface temperatures from Metop AVHRR compared to in situ ocean and land data, Ocean Sci., 8(6), 959–970, doi:10.5194/os-8-959-2012, 2012.~~

~~Dybkjær, G., Høyer, J. L., Tonboe, R. and Olsen, S. M.: Report on the documentation and description of the new Arctic Ocean dataset combining SST and IST., NACLIM Deliverable D32.28. [online] Available from: http://www.naclim.eu/ (Accessed 8 September 2018), 2014.~~

~~Dybkjaer, G., Eastwood, S., Borg, A. L., Høyer, J. L. and Tonboe, R.: Algorithm theoretical basis document (ATBD) for the OSI SAF Sea and Sea Ice Surface Temperature L2 processing chain. OSI205a and OSI205b, [online] Available from: http://osisaf.met.no/docs/osisaf_cdop2_ss2_pum_ice-conc_v1p4.pdf (Accessed 3 March 2018), 2018.~~

~~Ghent, D. J., Corlett, G. K., Göttsche, F.-M. and Remedios, J. J.: Global Land Surface Temperature From the Along-Track Scanning Radiometers: Global LST from the ATSRs, J. Geophys. Res. Atmospheres, 122(22), 12,167-12,193, doi:10.1002/2017JD027161, 2017.~~

~~GHRSST Science Team: The Recommended GHRSST Data Specification (GDS) 2.0, document revision 4, available from the GHRSST International Project Office, 2011, pp 123., 2010.~~

~~Good, E.: Daily minimum and maximum surface air temperatures from geostationary satellite data: Satellite min and max air temperatures, J. Geophys. Res. Atmospheres, 120(6), 2306–2324, doi:10.1002/2014JD022438, 2015.~~

~~Good, E. J., Ghent, D. J., Bulgin, C. E. and Remedios, J. J.: A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series, J. Geophys. Res. Atmospheres, 122(17), 9185–9210, doi:10.1002/2017JD026880, 2017.~~

~~Graversen, R. G., Mauritsen, T., Tjernström, M., Källén, E. and Svensson, G.: Vertical structure of recent Arctic warming, Nature, 451(7174), 53–56, doi:10.1038/nature06502, 2008.~~

~~Grisogono, B., Kraljević, L. and Jeričević, A.: The low-level katabatic jet height versus Monin-Obukhov height, Q. J. R. Meteorol. Soc., 133(629), 2133–2136, doi:10.1002/qj.190, 2007.~~

~~Hall, D., Box, J., Casey, K., Hook, S., Shuman, C. and Steffen, K.: Comparison of satellite-derived and in-situ observations of ice and snow surface temperatures over Greenland, Remote Sens. Environ., 112(10), 3739–3749, doi:10.1016/j.rse.2008.05.007, 2008.~~

~~Hall, D. K., Key, J. R., Case, K. A., Riggs, G. A. and Cavalieri, D. J.: Sea ice surface temperature product from MODIS, IEEE Trans. Geosci. Remote Sens., 42(5), 1076–1087, doi:10.1109/TGRS.2004.825587, 2004.~~

~~Hall, D. K., Comiso, J. C., DiGirolamo, N. E., Shuman, C. A., Key, J. R. and Koenig, L. S.: A Satellite-Derived Climate-Quality Data Record of the Clear-Sky Surface Temperature of the Greenland Ice Sheet, J. Clim., 25(14), 4785–4798, doi:10.1175/JCLI-D-11-00365.1, 2012.~~

~~Hansen, J., Ruedy, R., Sato, M. and Lo, K.: Global surface temperature change, Rev. Geophys., 48(4), doi:10.1029/2010RG000345, 2010.~~

~~Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in coupled models, Clim. Dyn., 21(3–4), 221–232, doi:10.1007/s00382-003-0332-6, 2003.~~

~~Høyer, J. L., Alerskans, E., Nielsen-Englyst, P., Thejll, P., Dybkjaer, G. and Tonboe, R.: Detailed investigation of the uncertainty budget for non-recoverable IST observations and their SI traceability (ESA Tech. Rep. FRM4STS OP-70), [online] Available from: http://www.frm4sts.org/wp-content/uploads/sites/3/2018/08/OP-70-FRM4STS_option3_report_v1-signed.pdf (Accessed 4 December 2018a), 2017.~~

~~Høyer, J. L., Lang, A. M., Tonboe, R., Eastwood, S., Wimmer, W. and Dybkjær, G.: Towards field inter-comparison experiment (FICE) for ice surface temperature (ESA Tech. Rep. FRM4STS OP-40)., [online] Available from: http://www.frm4sts.org/wp-content/uploads/sites/3/2017/12/OFE-OP-40-TR-5-V1-Iss-1-Ver-1-Signed.pdf (Accessed 2 November 2018b), 2017.~~

~~Høyer, J. L., Good, E., Nielsen-Englyst, P., Madsen, K. S., Woolway, I. and Kennedy, J.: Report on the relationship between satellite surface skin temperature and surface air temperature observations for oceans, land, sea ice, ice sheets, and lakes, Deliverable. [online] Available from: https://www.eustaceproject.org/eustace/static/media/uploads/d1.5_revised.pdf (Accessed 6 June 2018), 2018.~~

~~Høyer, J. L., Dybkjaer, G., Eastwood, S. and Madsen, K. S.: EUSTACE/AASTI: Global clear-sky ice surface temperature data from the AVHRR series on the satellite swath with estimates of uncertainty components, v1.1, 2000-2009, Centre for Environmental Data Analysis. [online] Available from: http://catalogue.ceda.ac.uk/uuid/60b820fa10804fca9c3f1ddfa5ef42a1, 2019.~~

~~Hudson, S. R. and Brandt, R. E.: A Look at the Surface-Based Temperature Inversion on the Antarctic Plateau, J. Clim., 18(11), 1673–1696, doi:10.1175/JCLI3360.1, 2005.~~

~~Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M. and Morice, C. P.: Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010:, J. Geophys. Res. Atmospheres, 117(D5), n/a-n/a, doi:10.1029/2011JD017139, 2012.~~

~~Karlsson, K.-G. and Dybbroe, A.: Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations, Atmospheric Chem. Phys., 10(4), 1789–1807, doi:10.5194/acp-10-1789-2010, 2010.~~

~~Karlsson, K.-G., Riihelä, A., Müller, R., Meirink, J. F., Sedlar, J., Stengel, M., Lockhoff, M., Trentmann, J., Kaspar, F., Hollmann, R. and Wolters, E.: CLARA-A1: a cloud, albedo, and radiation dataset from 28 yr of global AVHRR data, Atmospheric Chem. Phys., 13(10), 5351–5367, doi:10.5194/acp-13-5351-2013, 2013.~~

~~Karlsson, K.-G., Anttila, K., Trentmann, J., Stengel, M., Fokke Meirink, J., Devasthale, A., Hanschmann, T., Kothe, S., Jääskeläinen, E., Sedlar, J., Benas, N., van Zadelhoff, G.-J., Schlundt, C., Stein, D., Finkensieper, S., Håkansson, N. and Hollmann, R.: CLARA-A2: the second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data, Atmospheric Chem. Phys., 17(9), 5809–5828, doi:10.5194/acp-17-5809-2017, 2017.~~

~~Kennedy, J. J., Capponi, F., Ghent, D., Good, E. J., Høyer, J. L., Kent, E. C., Madsen, K. S., Mitchelson, J. R., Nielsen-Englyst, P. and Tonboe, R. T.: EUSTACE: Globally gridded clear-sky daily air temperature estimates from satellites with uncertainty estimates for land, ocean and ice, 1995-2016, , doi:10.5285/f883e197594f4fbaae6edebafb3fddb3, 2019.~~

~~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(2), 302–309, doi:10.1016/S0034-4257(97)89497-7, 1997.~~

~~Kindig, D.: Greenland Climate Network (GC-Net) Radiation for Arctic System Reanalysis, Version 1., Boulder Colo. USA NSIDC Natl. Snow Ice Data Cent., doi:10.5067/6S7UHUH2K5RI, 2010.~~

~~Knust, R.: Polar Research and Supply Vessel POLARSTERN operated by the Alfred-Wegener-Institute, J. Large-Scale Res. Facil. JLSRF, 3, doi:10.17815/jlsrf-3-163, 2017.~~

~~Koenig, L. S. and Hall, D. K.: Comparison of satellite, thermochron and air temperatures at Summit, Greenland, during the winter of 2008/09, J. Glaciol., 56(198), 735–741, doi:10.3189/002214310793146269, 2010.~~

~~Masson-Delmotte, V., Swingedouw, D., Landais, A., Seidenkrantz, M.-S., Gauthier, E., Bichet, V., Massa, C., Perren, B., Jomelli, V., Adalgeirsdottir, G., Hesselbjerg Christensen, J., Arneborg, J., Bhatt, U., Walker, D. A., Elberling, B., Gillet-Chaulet, F., Ritz, C., Gallée, H., van den Broeke, M., Fettweis, X., de Vernal, A. and Vinther, B.: Greenland climate change: from the past to the future: Greenland climate change, Wiley Interdiscip. Rev. Clim. Change, 3(5), 427–449, doi:10.1002/wcc.186, 2012.~~

~~Menke, W.: Geophysical Data Analysis: Discrete Inverse Theory, Elsevier, New York., 1989.~~

~~Merchant, C. J. and Le Borgne, P.: Retrieval of Sea Surface Temperature from Space, Based on Modeling of Infrared Radiative Transfer: Capabilities and Limitations, J. Atmospheric Ocean. Technol., 21(11), 1734–1746, doi:10.1175/JTECH1667.1, 2004.~~

~~Merchant, C. J., Harris, A. R., Murray, M. J. and Závody, A. M.: Toward the elimination of bias in satellite retrievals of sea surface temperature: 1. Theory, modeling and interalgorithm comparison, J. Geophys. Res. Oceans, 104(C10), 23565–23578, doi:10.1029/1999JC900105, 1999.~~

~~Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter, M., Sathyendranath, S., Sandven, S., Sofieva, V. F. and Wagner, W.: Uncertainty information in climate data records from Earth observation, Earth Syst. Sci. Data, 9(2), 511–527, doi:10.5194/essd-9-511-2017, 2017.~~

~~Miller, N. B., Turner, D. D., Bennartz, R., Shupe, M. D., Kulie, M. S., Cadeddu, M. P. and Walden, V. P.: Surface-based inversions above central Greenland, J. Geophys. Res. Atmospheres, 118(2), 495–506, doi:10.1029/2012JD018867, 2013.~~

~~Morice, C. P., Kennedy, J. J., Rayner, N. A. and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set: THE HADCRUT4 DATASET, J. Geophys. Res. Atmospheres, 117(D8), doi:10.1029/2011JD017187, 2012.~~

~~Morice, C. P., Capponi, F., Kennedy, J. J., Killick, R. E., Lindgren, F., Mitchelson, J. R., Rayner, N. A. and Winn, J. P.: EUSTACE: Global daily air temperature combining surface and satellite data, with uncertainty estimates, for 1850-2015, v1.0, Centre for Environmental Data Analysis. [online] Available from: http://dx.doi.org/10.5285/468abcf18372425791a31d15a41348d9, 2019.~~

~~Nielsen-Englyst, P., Høyer, J. L., Madsen, K. S., Tonboe, R., Dybkjær, G. and Alerskans, E.: In situ observed relationships between snow and ice surface skin temperatures and 2 m air temperatures in the Arctic, The Cryosphere, 13(3), 1005–1024, doi:10.5194/tc-13-1005-2019, 2019.~~

~~Oltmanns, M., Straneo, F., Seo, H. and Moore, G. W. K.: The Role of Wave Dynamics and Small-Scale Topography for Downslope Wind Events in Southeast Greenland, J. Atmospheric Sci., 72(7), 2786–2805, doi:10.1175/JAS-D-14-0257.1, 2015.~~

~~Overland, J., Dunlea, E., Box, J. E., Corell, R., Forsius, M., Kattsov, V., Olsen, M. S., Pawlak, J., Reiersen, L.-O. and Wang, M.: The urgency of Arctic change, Polar Sci., doi:10.1016/j.polar.2018.11.008, 2018.~~

~~Perovich, D. K., Richter-Menge, J. A. and Polashenski, C. M.: Observing and understanding climate change: Monitoring the mass balance, motion, and thickness of Arctic sea ice, [online] Available from: http://imb-crrel-dartmouth.org (Accessed 24 November 2016), 2016.~~

~~Pielke, R. A., Davey, C. A., Niyogi, D., Fall, S., Steinweg-Woods, J., Hubbard, K., Lin, X., Cai, M., Lim, Y.-K., Li, H., Nielsen-Gammon, J., Gallo, K., Hale, R., Mahmood, R., Foster, S., McNider, R. T. and Blanken, P.: Unresolved issues with the assessment of multidecadal global land surface temperature trends, J. Geophys. Res., 112(D24), D24S08, doi:10.1029/2006JD008229, 2007.~~

~~Pithan, F. and Mauritsen, T.: Arctic amplification dominated by temperature feedbacks in contemporary climate models, Nat. Geosci., 7(3), 181–184, doi:10.1038/ngeo2071, 2014.~~

~~Rasmussen, T. A. S., Høyer, J. L., Ghent, D., Bulgin, C. E., Dybkjaer, G., Ribergaard, M. H., Nielsen-Englyst, P. and Madsen, K. S.: Impact of Assimilation of Sea-Ice Surface Temperatures on a Coupled Ocean and Sea-Ice Model, J. Geophys. Res. Oceans, 123(4), 2440–2460, doi:10.1002/2017JC013481, 2018.~~

~~Rayner, N., Good, S. and Block, T.: SST CCI Product User Guide, Project Document, SST_CCI-PUG-UKMO-201., [online] Available from: http://www.esa-sst-cci.org/PUG/pdf/SST_CCI-PUG-UKMO-201-Issue_1-signed.pdf (Accessed 15 March 2019), 2015.~~

~~Rayner, N. A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108(D14), doi:10.1029/2002JD002670, 2003.~~

~~Rayner, N. A., Auchmann, R., Bessembinder, J., Brönnimann, S., Brugnara, Y., Capponi, F., Carrea, L., Dodd, E. M. A., Ghent, D., Good, E., Høyer, J. L., Kennedy, J. J., Kent, E. C., Killick, R. E., van der Linden, P., Lindgren, F., Madsen, K. S., Merchant, C. J., Mitchelson, J. R., Morice, C. P., Nielsen-Englyst, P., Ortiz, P. F., Remedios, J. J., van der Schrier, G., Squintu, A. A., Stephens, A., Thorne, P. W., Tonboe, R. T., Trent, T., Veal, K. L., Waterfall, A., Winfield, K., Winn, J. and Woolway, R. I.: The EUSTACE project: delivering global, daily information on surface air temperature, Submitt. Bull. Am. Meteorol. Soc., 2019.~~

~~Reeves Eyre, J. E. J. and Zeng, X.: Evaluation of Greenland near surface air temperature datasets, The Cryosphere, 11(4), 1591–1605, doi:10.5194/tc-11-1591-2017, 2017.~~

~~Renfrew, I. A.: The dynamics of idealized katabatic flow over a moderate slope and ice shelf, Q. J. R. Meteorol. Soc., 130(598), 1023–1045, doi:10.1256/qj.03.24, 2004.~~

~~Richter-Menge, J. A., Perovich, D. K., Elder, B. C., Claffey, K., Rigor, I. and Ortmeyer, M.: Ice mass-balance buoys: a tool for measuring and attributing changes in the thickness of the Arctic sea-ice cover, Ann. Glaciol., 44, 205–210, doi:10.3189/172756406781811727, 2006.~~

~~Shuman, C. A., Steffen, K., Box, J. E. and Stearns, C. R.: A Dozen Years of Temperature Observations at the Summit: Central Greenland Automatic Weather Stations 1987–99, J. Appl. Meteorol., 40(4), 741–752, doi:10.1175/1520-0450(2001)040<0741:ADYOTO>2.0.CO;2, 2001.~~

~~Shuman, C. A., Hall, D. K., DiGirolamo, N. E., Mefford, T. K. and Schnaubelt, M. J.: Comparison of Near-Surface Air Temperatures and MODIS Ice-Surface Temperatures at Summit, Greenland (2008–13), J. Appl. Meteorol. Climatol., 53(9), 2171–2180, doi:10.1175/JAMC-D-14-0023.1, 2014.~~

~~Smith, T. M., Reynolds, R. W., Peterson, T. C. and Lawrimore, J.: Improvements to NOAA's Historical Merged Land–Ocean Surface Temperature Analysis (1880–2006), J. Clim., 21(10), 2283–2296, doi:10.1175/2007JCLI2100.1, 2008.~~

~~Stamnes, K., Ellingson, R. G., Curry, J. A., Walsh, J. E. and Zak, B. D.: Review of Science Issues, Deployment Strategy, and Status for the ARM North Slope of Alaska Adjacent Arctic Ocean Climate Research Site, J. Clim., 12(1), 46–63, doi:10.1175/1520-0442-12.1.46, 1999.~~

~~Steeneveld, G.-J.: Current challenges in understanding and forecasting stable boundary layers over land and ice, Front. Environ. Sci., 2, doi:10.3389/fenvs.2014.00041, 2014.~~

~~Steffen, K.: Surface energy exchange at the equilibrium line on the Greenland ice sheet during onset of melt, Ann. Glaciol., 21, 13–18, doi:10.3189/S0260305500015536, 1995.~~

~~Steffen, K. and Box, J.: Surface climatology of the Greenland Ice Sheet: Greenland Climate Network 1995-1999, J. Geophys. Res. Atmospheres, 106(D24), 33951–33964, doi:10.1029/2001JD900161, 2001.~~

~~Sterk, H. A. M., Steeneveld, G. J. and Holtslag, A. A. M.: The role of snow-surface coupling, radiation, and turbulent mixing in modeling a stable boundary layer over Arctic sea ice: MODELING A CLEAR-SKY ARCTIC SBL, J. Geophys. Res. Atmospheres, 118(3), 1199–1217, doi:10.1002/jgrd.50158, 2013.~~

~~Tonboe, R. T., Eastwood, S., Lavergne, T., Sørensen, A. M., Rathmann, N., Dybkjær, G., Pedersen, L. T., Høyer, J. L. and Kern, S.: The EUMETSAT sea ice concentration climate data record, The Cryosphere, 10(5), 2275–2290, doi:10.5194/tc-10-2275-2016, 2016.~~

~~Vihma, T. and Pirazzini, R.: On the Factors Controlling the Snow Surface and 2-m Air Temperatures Over the Arctic Sea Ice in Winter, Bound.-Layer Meteorol., 117(1), 73–90, doi:10.1007/s10546-004-5938-7, 2005.~~

~~Vihma, T., Jaagus, J., Jakobson, E. and Palo, T.: Meteorological conditions in the Arctic Ocean in spring and summer 2007 as recorded on the drifting ice station Tara, Geophys. Res. Lett., 35(18), doi:10.1029/2008GL034681, 2008.~~

~~Vose, R. S., Arndt, D., Banzon, V. F., Easterling, D. R., Gleason, B., Huang, B., Kearns, E., Lawrimore, J. H., Menne, M. J., Peterson, T. C., Reynolds, R. W., Smith, T. M., Williams, C. N. and Wuertz, D. B.: NOAA's Merged Land–Ocean Surface Temperature Analysis, Bull. Am. Meteorol. Soc., 93(11), 1677–1685, doi:10.1175/BAMS-D-11-00241.1, 2012.~~

~~Weng, W. and Taylor, P. A.: On Modelling the One-Dimensional Atmospheric Boundary Layer, Bound.-Layer Meteorol., 107(2), 371–400, doi:10.1023/A:1022126511654, 2003.~~

~~Zilitinkevich, S., Savijärvi, H., Baklanov, A., Grisogono, B. and Myrberg, K.: Forthcoming Meetings on Planetary Boundary-layer Theory, Modelling and Applications, Bound.-Layer Meteorol., 119(3), 591–593, doi:10.1007/s10546-006-9069-1, 2006.~~