Interactive comment on "Near-surface thermal stratification during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures" *by* Alden C. Adolph et al.

Anonymous Referee #1

Received and published: 7 December 2017

General comments:

The authors present 40 days of near-surface air and surface skin temperature measurements at Summit station Greenland. They analyse the datasets with focus on identifying the timing and magnitude of near-surface thermal inversion. They compare their best measurements of surface skin temperature, from IR upward radiation, to the latest MODIS Land Surface Temperature (LST) and find smaller bias than in other studies validating that product.

The skin temperature is a major variable in the surface energy balance of snow and ice surfaces and has drawn little attention so far. Product such as MODIS LST is a valuable tool for quantifying that variable but needs ground validation before it is used for the whole Greenland ice sheet. The authors therefor address an important question in a well written study using adequate tools and methods.

However, the study suffers at the moment from the lack of temporal and spatial coverage. Indeed one can ask if these 40 days of observations are representative of the surface conditions on the ice sheet and adequate to validate the MODIS LST as a whole. Especially, it is unfair to compare a 40-days comparison to previous studies that all use multiyear or spatially distributed datasets. As far as I know, NOAA can provide suitable data (IR skin temperature) for a longer period at Summit and PROMICE stations for sites at lower elevation.

> *We agree that these 40 days of observations are not representative spatially or temporally of the Greenland Ice Sheet as a whole. However, we decline to include NOAA data or PROMICE data in this paper for the following reasons: Both of these datasets do measure upwelling longwave radiation, and the PROMICE datasets calculate a skin temperature by assuming blackbody radiation properties of snow and ice surfaces. However, it's a very wide band that is used (4.5 to 42 μm), whereas our IR sensor uses the band from 8-14μm to calculate temperature. MOD11 uses bands around 11μm and 12 μm to calculate skin temperature, which is much more similar to the sensor we use to measure IR skin temperature. Furthermore, the PROMICE stations are all at ice sheet margins. While it would be very interesting to use this data to compare to MOD/MYD11, our study is focused on the near-surface phenomenon in the dry snow zone. Comparison of skin temperature, air temperature and MODIS surface temperature at PROMICE stations at the ice sheet margins would make for an excellent future study. While our dataset may not be very long, we have strong confidence in the quality of our data, as the station was maintained fairly closely and was not left alone for more than 1 month. PROMICE stations are not as frequently maintained. Also, though our data is not representative of all radiative conditions due to summertime measurement in constant sunlight, there is still a large spread in the temperatures measured (from -30 C to -5 C).*

Additionally the study could also benefit from more in-depth definition and discussion of all the measurements (f.e. how is calculated MODIS LST, effect of surface emissivity and of the air below the sensor for IR radiation measurements *. . .*) and concepts (what is skin temperature, how is defined thermal inversion, which indicators and metrics are used to describe it, which are its drivers*. . .*) that are being used.

> *We have added further discussion of the topics you list throughout the manuscript.*

Eventually the use of unshielded thermochrons and thermocouples is highly sensitive to radiation (even more during midnight sun period) and should not be used without appropriate segregation of the erroneous periods or correction of radiative heating. At the end of the study, the authors even discard these measurements and only compare MODIS to IR surface radiation. I would recommend greatly minimizing the description of that data and any conclusion that derives from it (f.e. statement regarding inversion in the 5 cm above the surface).

> *We have decided to remove the thermochron and thermocouple data from the manuscript because the sensors were unshielded, and therefore we reconsidered the validity of the conclusions drawn from that data.*

Technical corrections:

l. 19: Define NOAA. Define again in main text.

We have added these definitions.

l.84: Please only mention the conclusions from Good (2016) relative to snowy surfaces.

We have limited the discussion of Good (2016) to only snow covered sites and seasons.

l. 100: “at but” word missing?

Sentence was editing and the word “at” was removed.

l. 105: It is said” Berkelhammer et al. (2016) discuss the impacts of the surface- based temperature inversions (with 2 m air temperature as the base)”. I can see in Berkelhammer et al. (2016) that they have measurements down to 20 cm. Please adjust or justify.

The measurements down to 20 cm are the isotope measurements. However, upon further review, they do measure surface temperature. It is unclear whether it is used or not in some of the calculations, but we have edited our description to reflect this new understanding: “Berkelhammer et al. (2016) discuss the impacts of the surface-based temperature inversions on boundary-layer dynamics, showing that the stability of the atmosphere prevents mixing and ultimately limits accumulation at Summit. These recent studies have investigated near-surface processes at Summit because of the importance of surface energy balance and turbulent snow/atmosphere exchange in climate monitoring and ultimately prediction of larger scale circulation and future change in ice mass balance. However, while some surface temperature measurements at Summit have been made (Berkelhammer et al. 2016), controls on surface temperature gradients in the lowest 2 meters of the atmosphere, which are most relevant for the remote sensing community and also have important implications for changing ice sheet dynamics, have not been explicitly studied at Summit, Greenland.”

l. 114: “There are”, there is

We edited this sentence to say “There are many different…”

l. 119: The equation points at the various variables that need to be calculated before the surface temperature can be determined. Please give a brief description of how it is being done for MODIS LST and how accurate is that calculation.

We have added more information about the MODIS LST. In lines 175-197, details of the MODIS surface temperature product are given.

l. 126: “its products are chosen as the remote sensing product” rephrase

Rewrote this to read: “The MODIS instrument produces widely-used land surface temperature (LST), which we will use as the remote sensing product for comparison in this work.”

l. 130: “A number . . .” here validation studies are described. It seems to be the same topic as the next paragraph so consider merging them. Also you discuss validation of MODIS products before you actually describe what these products are and how LST is being calculated. Consider rearranging paragraphs in this section.

The former Introduction, Background, and Methods sections have been rearranged with these suggestions (and those of the other reviewer) in mind.

l. 134: You mention the recurrent confusion between near-surface, 2 m air temperature and surface skin temperature. Please define your vocabulary at the beginning of the manuscript and then be specific every time temperature is mentioned. Here what does “ground surface temperature” refer to? Check other cases throughout the manuscript.

We have edited the sentence to clarify here that we were referring to both 2m and skin temperature (depending on what data the relevant study used).

l. 140: You mention a cold bias in Østby et al. (2014) but it is not reported in Table 1.

Thanks for pointing out the oversight. The bias is now reported in Table 1.

l. 144: “A bias in the data can obscure or alter trends within a dataset.” Very general statement, consider removing.

Statement has been removed.

l. 145: “Furthermore, it is possible that. . .” at this point, you have not exposed any theoretical (potentially from how MODIS LST is being calculated) or empirical reasons to think that one of the 2 m or skin temperature would match better than the other to the MODIS product. Either add a justification or reference for this hypothesis or move it to what is being interpreted from your validation.

We have edited the paragraph about the MODIS product to discuss the fact that it is indeed measuring the *skin temperature of the snow, which justifies this statement. This sentence was added: "*In the measurements of snow, the resulting temperature is representative of the top several microns of the surface at the snow/air interface because of the penetration depth of radiation at that wavelength, so it is indeed a skin temperature (Warren and Brandt, 2008)."

l. 148: What do you mean by "standard"?

We have removed that language.

l.154: "Both MOD/MYD11 and the preliminary version of the MOD29 special product were compared to our in situ data". If both have been compared and validated then both should be presented. The conclusion that a LST product developed for land performs better than a LST product developed for ice on the Greenland ice sheet is a very important conclusion. Quantifying and localizing the errors in MOD29 could be important for future studies. You can also validate only MOD/MYD11 but then don't mention that you validated MOD29 but not show the result.

Validation of MOD29 is indeed important; however, we have removed our mention of it, and discussion of this product is left to future studies.

l.156: In this paragraph, the authors should provide a clear and concise description on how LST (and all necessary variables such as water vapour, air temperature, emissivity, and cloud cover) are being calculated. It is important so that the reader can be reminded of the assumptions and uncertainties linked to the calculation and to which level of accuracy can be expected from the product.

We have added some details to this section as you suggest, especially in the uncertainties linked to the calculation. We feel that this section already has much more detail than other similar studies (Westermann et al. 2012, Ostby et al. 2014, Shuman et al. 2014, etc.) and the interested reader can seek further detail of the MODIS calculations in the cited literature.

l. 160: "feature" not clear what it refers to.

*We have rewritten this sentence to improve clarity. "*Calculation of surface emissivity is important in this product because MOD/MYD11 is a global product that estimates land surface temperature on all types of land cover types."

l. 161: "Over snow and ice, this. . ." Not very clear, please rephrase.

We have rewritten this sentence to improve clarity: "Because this study focuses on consistently snow-covered land, there was not significant variability in the emissivity; in band 32 the emissivity is 0.990 for each data point, and in band 31, the emissivity fluctuates between either 0.992 or 0.994."

l. 165: In this paragraph, please quantify the difference between C5 and C6. More especially, how much of the cold bias seen in validation studies can be explained by the defects of C5 products?

We conducted our comparison of in situ IR skin temperature to the C5 product to answer this question. We found a mean difference of 0.2°C between C5 and C6 in our study time and location, where the C6 temperatures are slightly higher than C5. This has been added to the manuscript and figures have been added to the supplement.

l. 171: "decrease in measured brightness temperatures" define brightness temperature or use vocabulary previously defined. Does this sentence imply that the corrections in C6 would lead to even colder bias if compared to the previously mentioned validation studies?

We have edited this section and added further discussion of C6 vs.C5 and figures to demonstrate the differences in the supplement.

l. 210: "the pixel that has the minimum distance" a pixel is an area so the station should be located within one at all time. Provide pixel size at some point in the manuscript.

The pixel is an area, but it is defined by central latitude/longitude point, so we use this minimum distance to find out which pixel our site is in. Pixel size is 1km x 1 km, and this information has been added to the manuscript.: Within each swath, we find the 1 km x 1 km square pixel that contains our measurement site by minimizing distance between pixel nadir point and our in-situ measurement site.

l. 213: "the nonsynchronicity may introduce some error to the comparison" It seems that the IR radiation comes is recorded at 5 min interval and 2 m air temperature at 1 min interval so this nonsynchronicity error could be removed by taking the measurements that are within few minutes of the MODIS acquisition time. Even if it is random noise, removing that error could potentially show better match of MODIS LST with ground measurements.

The IR skin temperature data is taken every 5 minutes, but what is saved in the datalogger is only the 30 *minute average (so that storage space could be minimized), so unfortunately we do not have 5 minute data for the IR skin temperature.*

l. 246 "Several different" redundant

Edited to read: "Several types of sensors…"

l. 247 "in order to compare this study to previous [. . .] studies" I assume the main goal was to validate MODIS, comparing to other validation studies comes after. Maybe rephrase.

This is a good point. As we have removed the thermochron and thermocouple data due to data quality concerns, this sentence is no longer in the manuscript.

l. 255 "In Koenig and Hall (2010) . . ." the following sentences should be in the methodology where the measurements and their expected limitations are being presented. Additionally Hall et al. (2015) seem to conclude that unshielded thermochrons are subject to measurement error. So why using/presenting that data at all? The use of unshielded thermocouple is subject to the same issue.

We have moved this section to the methods. We have decided not to include the thermocouple and thermochron data.

l. 270: "differences are much higher at lower wind speeds" Unfortunately, conditions favourable to inversions are also the one enhancing radiative heating of both thermochrons and thermocouple: low wind speeds at 2m imply even lower wind speeds at the surface and will hamper ventilation of the sensors leading to sustained radiation absorption even with low sunlight. Better information should be given to ascertain that this difference is due to inversion.

We decided not to include the thermocouple and thermochron data because we were unable to account for the issues associated with the unshielded temperature measurements and because their inclusion detracted from the central focus of the paper.

l. 274 "similar" quantify l.275 "larger" quantify

To address the previous 2 comments, the following edits were made: "While 2 m air temperature and IR skin temperature are similar during peak solar irradiance (Figure 5), with the mean difference in temperature being -0.32°C when incoming solar radiation is greater than 600 W m^{-2}. There is a larger difference between the two during the night-time, with 2 m air temperature higher than skin temperature by an average of 2.4°C when incoming radiation is less than 200 W m^{-2}."

l. 280 "increased discrepancy" seems redundant to what is said in the previous sentences. Also quantify here.

We have added the following quantification: "As the inversions appear diurnal in nature, the measurements are quite similar at higher temperatures (above -10°C, mean difference is -0.16°C), but at lower temperatures, there is increased discrepancy between 2 m temperature and snow skin temperature (below -20°C, mean difference is 3.5°C)."

l. 281 "most frequently" quantify.

We have added the following quantification: "There is a clear skew in the histogram, indicating that 2 m air temperature is most frequently higher than skin temperature (in 68% of measurements). This is true in both clear and cloudy sky conditions, where the percentage of measurements for which air temperature exceeds skin temperature is 70% in clear sky and 65% in cloudy sky."

l.289 The following paragraph brings in discussion about MODIS surface temperature when we are still in the ground measurement section. Consider moving to the next section.

We see your point, but this is not a comparison of our data to MODIS temperatures, so we think it still fits best within this section.

l. 294 "Previous studies acknowledge that near-surface stratification may be part of the cause of the discrepancy" this is very important and has not been mentioned clearly in the introduction. How did these studies arrive to these conclusions?

More information on this has been added in the introduction: Shuman et al. 2014 acknowledge that differences between 2 m air temperature and skin temperature caused by inversions could cause bias in their comparison to MODIS, but at the time there was insufficient data to suggest whether inversions would persist in central Greenland and in the very near surface.

l. 311 "In the summer. . ." and following. Can you be more quantitative, f.e. binning observation into night and day time and showing that the mean differences in each group are statistically different. Maybe also find a cut off value in radiation below which you have significantly stronger inversion.

We have added more quantitative analysis as follows: "The mean differences reported above as -0.16±0.88°C *when temperatures are above -10°C and as 3.5±2.4°C when temperatures are below -20°C. A paired t-test shows that these means are not equal to one another with a p-value of less than 0.001."*

l. 319 You are just repeating what is said about Good (2016) in the introduction.

We have removed the discussion of Good (2016) as it did not add anything new.

l. 322 not clear what "median difference" with a +/- sign refers to.

Upon further reading of Good (2016), I do not think the comparison to the "median difference" was relevant.

l. 332 Can you specify what is the average solar radiation and zenith angle during that time.

This is no longer relevant as we have removed the thermochron data from the study because we decided that the measurements were flawed since the thermochrons were not shielded.

l. 335 Since you aim at validating MODIS LST, better or worse match with it is not a good reason for discarding ground measurements. Consider quantifying the error of the thermochrons using IR temperature instead. If these measurements were not considered reliable enough to validate MODIS LST, why are they being used to quantify the thermal inversion earlier in the study?

We have removed the thermochron data from the study because we decided that the measurements were flawed since the thermochrons were not shielded.

l. 335 please give mean bias along with RMSE

We have removed the thermochron data from the study because we decided that the measurements were flawed since the thermochrons were not shielded.

l. 350 please give mean bias along with RMSE

The mean bias was already reported here. "This is also evident in Figure 10, where MOD/MYD11 products combine to yield and RMSE of 1.6°C (n=374) when compared with IR skin temperature, and there is a mean bias of 0.7±1.4°C."

l. 356 please state at which viewing angle the error is maximal and quantify that maximum error

We have indicated the range of viewing angles and the slope of the fit line. To indicate the maximum error and the angle at which it occurs may be misleading as there are some points that appear as relative outliers. Linear regression provides a better way to consider these errors.

l. 360 Why not comparing MODIS to the thermocouple data that you also presented earlier? If, as for the thermochrons it is considered as erroneous measurements it should also be stated and in that case it shouldn't be used for *quantifying* near surface inversion earlier in the study. Also consider comparing MODIS with the 2 m temperature over the study period and see if you find same bias as previous validation studies using 2 m temperature.

We have removed thermochron and thermocouple data from the manuscript. We have added a comparison of 2m temperature to MODIS temperature. The new paragraph is as follows: "While we do not believe that 2 m air temperature is a good proxy for skin temperature, for demonstration purposes, we have compared the 2 m air temperature measurements to the MOD/MYD11 product in Figure 10b. In doing so, we find an RMSE of 3.1°C and a mean bias of 1.9±2.5°C (n=374). This results in a similar RMSE to Shuman et al. (2014) of 3.5°C, though the mean bias of our comparison is slightly less than their bias was at 3°C. This comparison further illustrates the importance of using skin temperatures in MODIS validation studies. Shuman et al. (2014) were unable to conclusively say that any of their bias was a result of using 2 m air temperature instead of skin temperature, and in fact they did not think it was likely that any inversion effects would cause the gradually increasing bias with decreasing temperature because there was insufficient research on the presence of near-surface inversions in the dry snow zone in Greenland. The comparison of Figure 10a and 10b shows that at least in the summer, inversions were likely to have played a large role in their 2014 results."

l.362 Along with the present discussion, it would be important to analyse the performance of MOD35 using the MMCR (false positive, false negative) to draw more broad conclusions (also using findings from Østby et al. (2014)) about when and where MODIS LST might be wrong because of cloud cover.

We have added an analysis of false positives, true positives, false negatives, and true negatives into our discussion of cloud masking: "In comparing the MMCR data to the MOD/MYD11, we find that of the 1059 times that the site was within the field of view of the satellites in June and July of 2015, there were 585 instances when both MMCR and MODIS detected cloud cover, 288 instances when both MMCR and MODIS indicated clear sky. This indicates 82% agreement. There were 86 false negatives (where MMCR indicates clouds and MODIS does not) and 100 false positives (where MMCR indicates clear sky, and MODIS indicates clouds). Østby et al. (2014) also use in-situ cloud data to filter out MODIS surface temperatures that are impacted by the presence of clouds in their study in Svalbard. They found an overall false negative *rate of 17%, whereas our false negative rate was 8%. Their work shows that the MOD35 cloud mask performs more poorly in the winter than in the summer, so the difference in false negatives is likely due to more favourable conditions for effective cloud masking due to constant sunlight during our measurement period."*

l. 381 "5 cm nearest to the..." this conclusion comes from potentially heated thermo- couple. Needs to be mentioned.

Thermocouple data have been removed.

l. 387 "the lower RMSE is likely . . ." here you make a hypothesis when you could actually show it. Is your RMSE greater when comparing MODIS to 2 m temperature? Please discuss mean bias along with RMSE.

*We have added the comparison of MODIS to 2 m temperature and included the data in the paper. In the conclusion, we have edited this sentence as follows: "*The lower RMSE and mean bias is likely a result of measuring the skin temperature using an IR instrument directly (instead of using 2 m air temperature, which resulted in an RMSE of 3.1°C and a mean bias of 1.9°C)."

Interactive comment on “Near-surface thermal stratification during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures” *by* Alden C. Adolph et al.

Anonymous Referee #2

Received and published: 12 December 2017

General Comments:

The authors present results from a new field campaign near Summit Station, Greenland conducted between 8th June and 18th July 2015. This campaign data is used to investigate near-surface temperature inversions at the site and validate MODIS MOD/MYD11 Collection 6 products. Overall this is an interesting paper which presents original results from a short field campaign as well as from validation of MODIS collection 6 products over Greenland. The manuscript is generally well-written and the methods used are appropriate.

However, I feel that there are issues which need to be addressed by the authors in redrafting this manuscript. There are also sections of this manuscript which would benefit from a tightening of prose and improvement of structure. Both are elaborated on in the specific comments.

Specific Comments:

It was hard to ascertain the new and original contributions that this manuscript provides to current scientific understanding at first. Outlining these in the introduction or similar would aid reader understanding.

> *We have edited the final paragraph of the introduction to make our contributions more clear. The final paragraph of the introduction now reads as follows: "We use our original dataset to determine how summertime meteorological conditions impact near-surface inversions (beneath 2 m height) on the ice sheet at Summit. Furthermore, we provide a validation of MODIS land surface temperatures, and show that the use of 2 m air temperature for MODIS validation is not recommended due to the presence of near-surface inversions. Lastly, we use in situ cloud data to show that the accuracy of the MODIS surface temperature product could be improved through stricter cloud masking."*

Line 14: suggest “can be assessed” or similar rather than “are assessed” as satellite derived temperatures over land are not in common use for this yet.

> *We have edited this sentence as you suggest.*

Line 53-54: The section starting “however, satellite remote sensing” should mention that the focus of this manuscript is on thermal infrared remote sensing, rather than microwave remote sensing, of surface temperatures. It should also be mentioned that thermal infrared remote sensing observations are affected by cloud cover. The sentence currently gives the impression that the satellite derived surface temperatures are spatially and temporally complete.

> *We have added this information as you suggest. This section now reads as follows: "In addition, thermal infrared satellite remote sensing provides the opportunity to collect surface temperature with large spatial coverage and sub-daily to weekly temporal resolution, depending on cloud conditions. In this study, we will focus on the MODerate-resolution Imaging Spectroradiometer (MODIS) thermal infrared land surface temperature (LST) product."*

Line 60-62: “inversions. . . which may cause a disparity between the ‘surface’ temperature at 2m and the actual skin temperature of the snow surface”. This sentence currently gives the impression that there is uncertainty about whether inversions cause a disparity between skin and 2 m air temperatures over snow and ice. Yet, the references cited in Section 2.1 note both the presence of inversions in these areas and their effect on temperature stratification. In a related issue, the manuscript could be read as suggesting that the issue of inversions (or other causes of disparity between 2m and skin temperatures) in validation of surface temperatures from thermal infrared remote sensing has not previously been considered. However, various previous studies have noted the issue of the difference between skin and 2 m air temperature over snow and ice surfaces in relation to remote sensing of surface temperatures and their validation. The effect of inversions on the difference between skin and 2 m air temperature over snow and ice surfaces has been noted by e.g. Comiso et al, 2003; Yu et al, 1995. There are also other reasons why differences may be seen between these two temperatures over cryospheric surfaces under clear sky conditions e.g. ice crystal precipitation (Yu et al, 1995) and latent heat effects (Wiese et al, 2015). As a result it is recommended that validation of satellite surface temperatures over land areas is done with situ surface temperatures (ideally from ground-based radiance measurements) if possible (Guillevic et al, 2017). In light of these previous studies and recommendations, this sentence at least should be rewritten and references included in the manuscript to give credit to previous work looking at the issue of the difference between skin and 2 m air temperature over snow and ice surfaces in relation to validation of remotely sensed surface temperatures.

Thank you for sharing these references. We absolutely want to give credit to these authors and have included this information in the introduction, and in section 3.2.1.

Lines 66-73: Does question b) refer to the specific in situ sensors you use, or to the use of in situ skin temperature data for validation in preference to air temperatures? I could not find these questions referred back to in the results or conclusions. Would suggest removing these or referring back to them later in the manuscript.

I have replaced these questions with statements about the main contributions of our work (relevant to comment above).

Section 2: The content of the background section was informative and interesting. However, in combination with the introduction and methods I found the structure of the manuscript a little hard to follow here. This section would benefit from a rewrite, or a restructuring. Some suggestions follow. The content of the surface temperature inversion section would provide a nice introduction to the issues raised in lines 55-65. The Remote Sensing of Surface Temperatures section mostly focuses on MODIS rather than remote sensing of surface temperatures more generally. The content relating to MODIS products could be moved to Section 3.2 or the section could be renamed.

We have restructured the Introduction/Background/Methods sections to improve the flow and clarity. The background on inversions was moved into the introduction, and the background on remote sensing was moved into the introduction and methods sections.

Line 215 to 220: these metrics are in very common use so the equations do not need stating unless there are notable differences from how they are commonly applied

We have removed the equations from the manuscript.

Line 241: There is very little discussion relating to figure 3. Reconsider including this figure or provide more discussion of the results shown.

We have added more discussion of this figure in the first paragraph of section 3.1. We believe it provides a nice context for the comparisons that follow.

Section 4.1: I found this section quite hard to read and understand. Please rewrite or restructure. Also, please include some values for the differences (bias, RMSE, etc.) noted between the in situ sensor surface temperatures to provide context to previous study comparisons (lines 304-329).

With the removal of the thermochron and thermocouple data, we have restructured this section and included bias and RMSE between 2m air temperature and IR skin temperature.

Lines 246-260: Given the issues noted with the thermochrons, please reconsider including these or provide some comment on whether these sensors are suitable for measuring surface temperature. If the authors decide to retain the thermochron analysis, in Figure 4 the difference between the thermochrons and the other sensors in- creases noticeably after day of year 167 but before they are buried by snow. Please provide some comment on this.

We have decided to remove the analysis with the thermochrons from the paper because the sensors were unshielded and so the quality of the data was uncertain.

Lines 266-267: "The measurements show". Ambiguous as to whether this refers to the measurements affected most by solar heating or other observations.

No long relevant as we have decided not to include these measurements in our manuscript.

Lines 304-318: Did Hall et al. 2008 use the same infrared radiometer as in the study detailed in the manuscript? Table 1 suggests this study only looked at 2 m air temperature.

Hall et al (2008) states that 'surface temperatures were derived from two positions on an AWS mast' and also that Everest 4000 4ZL TIR sensor that was placed 50 cm above the surface to record the surface temperature assuming a snow emissivity of 0.99. They did not use the IR data for MODIS validation.

Line 317-318: Sentence is ambiguous as to whether the need for future studies refers to the results of your study specifically, or in general.

Sentence has been edited as follows: "Future studies beyond our analysis that incorporate all seasons are needed to investigate this discrepancy and determine conditions under which 2 m air temperature is, or is not, *a good proxy for snow skin temperature."*

Lines 323-324: Is there publicly available snow depth information at these sites to address this question if it is not included in the paper?

Using satellite imagery, I was able to determine that these sites are not continuously snow covered in the summer. As a result, we are not including comparisons to their summer results.

Lines 324-329: Why was the thermocouple data not compared to MODIS? If there is a reason this should be stated.

The thermocouple data was not compared to MODIS because it is not a skin temperature, as MODIS is. Furthermore, issues of thermocouple heating have caused us to remove this dataset from the manuscript.

Lines 353-359: Do the authors have any thoughts on what else could be causing the remaining differences? If so please include this.

We believe that a stricter cloud mask would improve the difference, which we address in the next section. Furthermore, non-synchronicity between ground based measurements and MODIS measurements may be to blame for random noise. We have included the following sentence: "We believe that the differences may be due to insufficient cloud masking and perhaps to imperfect synchronicity of measurements, where in situ skin measurements represent an average of 30 minutes but the MODIS measurement represents a shorter time window."

Lines 366-368: I think the authors say "improving the cloud mask" when they mean increasing cloud masking strictness which may improve the product but also overflag cloud so that there is loss of un-cloud contaminated data ("reduce the amount of measurements available"). Data loss due to cloud masking, assuming that the pixels re- moved are genuinely cloud contaminated and the cloud masking is therefore accurate, is not a problem as these pixels will not contain sensible infrared surface temperature estimates. The issue is when there is significant over-flagging of cloudy pixels, leading to loss of non-cloud contaminated data, due to an increase in cloud masking strictness. If so these sentences (and lines 395-396) should be re-written.

Yes, you are correct, and we have edited our text to improve clarity. "In determining the strictness of the cloud mask used, there is a trade-off due to the need to mask out all cloud contaminated pixels but not overflag data, which results in the generation of false positives and removes pixels that were in fact clear."

Line 384: Do not use the word "correct" here as this suggests that MODIS is perfectly accurate, when actually MODIS data (and indeed any observation) will not measure the true (generally unknown) value of the surface temperature. There are always biases and uncertainties when measuring.

This is a good point. The wording has been edited as follows: "indicate that the MODIS data has only a very slight cold bias (-0.7°C)"

Lines 390-391: Sentence unclear in meaning. Are the authors suggesting using MODIS data and in situ 2 m air temperature to study inversions? If so, are the biases between the MODIS and in situ skin temperatures understood adequately to allow such a study? Lines 353-359 suggest not.

This is indeed what we were suggesting. We think that with further work, these biases would be adequately understood to allow such a study. The text has been edited as follows: "Furthermore, the validation presented in this study of the strong correlation between MODIS surface temperature and snow skin temperature in the summer lays a groundwork for inversions to be studied more extensively in locations where 2 m air temperature is currently measured."

Table 1: This table could do with a little restructuring and/or reduction of text as it is currently a little difficult to read and understand. Also, if this is for studies over land ice only please include this information in the caption.

The table has been edited to reduce text and hopefully improve clarity. The caption has been edited to clarify that this data is for "snow-covered regions."

Technical Corrections:

Line 47: "for understanding ice sheet..." rather than "of understanding ice sheet..."?

Correction made.

Line 53: remove extra space in "Fausto et al., 2012) ; however"

Correction made.

Line 534: remove curly brackets

Correction made.

Figure 1: The location dot is a little small. The north arrow is also a little difficult to see.

Correction made.

Table 1: missing “n=” on last row.

Correction made.

References:

Comiso, J. C. (2003). Warming trends in the Arctic from clear sky satellite observations. Journal of Climate, 16(21), 3498-3510.

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J.A., Remedios, J., Román, M. & Camacho, F. (2017). Land Surface Temperature Product Validation Best Practice Protocol. Ver- sion 1.0. In P. Guillevic, F. Göttsche, J. Nickeson & M. Román (Eds.), Best Practice for Satellite-Derived Land Product Validation (p. 60): Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/lst.001. https://lpvs.gsfc.nasa.gov/PDF/CEOS_LST_PROTOCOL_Oct2017_v1.0.0.pdf

Overland, J. E., & Guest, P. S. (1991). The Arctic snow and air temperature budget over sea ice during winter. Journal of Geophysical Research: Oceans (1978–2012), 96(C3), 4651-4662.

Wiese, M., Griewank, P., & Notz, D. (2015). On the thermodynamics of melting sea ice versus melting freshwater ice. Annals of Glaciology, 56(69), 191.

Yu, Y., Rothrock, D. A., & Lindsay, R. W. (1995). Accuracy of sea ice temperature derived from the advanced very high resolution radiometer. Journal of Geophysical Research: Oceans (1978–2012), 100(C3), 4525-4532.

All of these references have been added except for Wiese et al., which was a very interesting paper, but did not seem directly relevant to our study.

Near-surface ~~thermal stratification~~temperature inversion during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures

Alden C. Adolph[1,2], Mary R. Albert[1], Dorothy K. Hall[3,4]

[1]Thayer School of Engineering, Dartmouth College, Hanover, NH, USA
[2]Department of Physics, St. Olaf College, Northfield, MN, USA
[3]Earth System Science Interdisciplinary Center / University of Maryland, College Park, MD, USA
[4]Cryospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence to: Alden C. Adolph (Adolph1@stolaf.edu)

Abstract

As rapid warming of the Arctic occurs, it is imperative that climate indicators such as temperature be monitored over large areas to understand and predict the effects of climate changes. Temperatures are traditionally tracked using in situ 2 m air temperatures~~, but in remote locations where few ground-based measurements exist, such as on the Greenland Ice Sheet, temperatures over large areas are~~ and can also be assessed using remote sensing techniques. Remote sensing is especially valuable over the Greenland Ice Sheet where few ground based air temperature measurements exist. Because of the presence of surface-based temperature inversions in ice-covered areas, differences between 2 m air temperature and the temperature of the actual snow surface (referred to as “skin” temperature) can be significant and are particularly relevant when considering validation and application of remote sensing temperature data. We present results from a field campaign extending from 8 June through 18 July 2015, near Summit Station in Greenland to study surface temperature using the following measurements: skin temperature measured by an infrared (IR) sensor, ~~thermochrons, and thermocouples;~~ 2 m air temperature measured by a National Oceanic and Atmospheric Administration (NOAA) meteorological station; and a MODerate-resolution Imaging Spectroradiometer (MODIS) surface temperature product. Our data indicate that 2 m air temperature is often significantly higher than snow skin temperature measured in~~-~~ situ, and this finding may account for apparent biases in previous~~ surface temperature~~ studies of MODIS products that used 2 m air temperature for validation. This inversion is present during summer months when incoming solar radiation and wind speed are both low. As compared to our in~~-~~ situ IR skin temperature measurements, after additional cloud masking, the MOD/MYD11 Collection 6 surface-temperature standard product has an RMSE of 1.0°C and a mean bias of 0.4°C, spanning a range of temperatures from -35°C to -5°C. For our study area and time series, MODIS surface temperature products agree with skin surface temperatures better than previous studies indicated, especially at temperatures below -20°C where other studies found a significant cold bias. ~~The~~ We show that the apparent “cold bias” present in ~~others’ comparison~~other comparisons of 2 m air temperature and MODIS surface temperature ~~is perhaps~~may be a result of the near-surface temperature inversion~~ that our data demonstrate.~~. Further investigation of how in~~-~~ situ IR skin

temperatures compare to MODIS surface temperature at lower temperatures (below -35°C) is warranted to determine if thiswhether a cold bias does indeed existexists for those temperatures.

1. Introduction

The Arctic is experiencing warming at a more rapid rate than the rest of the world (Stocker, 2014), but the impacts of this increased temperature extend beyond the polar region. Declining sea ice extent and retreat of glaciers contribute to a powerful ice-albedo feedback that leads toresults in further warming on a large scale. This increased warming leads to declining mass balance of the Greenland Ice Sheet, contributing to global sea level rise. Quantifying current and future ice sheet mass balance remains an active area of research (e.g. Rignot et al., 2011; Rae et al., 2012; Vernon et al., 2013) and is critical to improving projections of sea level rise. Declining Greenland Ice Sheet mass balance is driven in part by changes in surface energy balance, which drives surface temperature and surface melt (Box, 2013; van den Broeke et al., 2016). Furthermore, the energy balance at the snow surface controls the interactions between the snow surface and the atmospheric surface layer. The net surface energy balance is defined by the net shortwave and longwave radiation, as well as sensible and latent heat fluxes, and heat flux from the underlying snow and ice. The net radiation at the surface affects the stability of the near-surface atmosphere and the extent to which turbulent heat exchange occurs between the snow surface and the lower atmosphere, impacting both local and regional circulation and climate.

Surface temperature is a critical component offor understanding ice sheet mass balance and for tracking changes in surface energy balance, however making accurate measurements of surface temperature across the vast expanse of the Greenland Ice Sheet over a long period of time is challenging (Reeves Eyre and Zeng, 2017). The installation of automatic weather stations (AWS) across the ice sheet has begun to provide point meteorological data at many locations through programs such as Greenland Climate Network (GC-Net) (e.g. Steffen et al., 1996; Steffen and Box, 2001; Shuman et al., 2001) and the Programme for Monitoring of the Greenland Ice Sheet (PROMICE)), which monitors both skin and air temperatures (e.g. Ahlstrøm, et al, 2008; van As et al., 2011; Fausto et al., 2012) ; however,). In addition, thermal infrared satellite remote sensing provides the opportunity to collect surface temperature with large spatial coverage and sub-daily to weekly temporal resolution, depending on cloud conditions. In this study, we will focus on the MODerate-resolution Imaging Spectroradiometer (MODIS) thermal infrared land surface temperature (LST) product.

"Surface" temperatures in climatological studies often refer to 2 m air temperature (Hudson and Brandt, 2005) as it is a standard measurement at meteorological stations around the globe; however, remotely-sensed surface temperatures from satellite-borne sensors in the cryosphere aremeasure the radiometric surface temperature, which is the actual "skin" temperature of the surface at the snow/air interface (Warren and Brandt, 2008). In the polar regions, the high albedo of snow in the visible part of the spectrum, and high emissivity of snow at longer wavelengths often leads to the phenomenon of inversions, where temperature increases with altitude. While often studied on the scale of tens of meters to kilometres above the snow surface (e.g. Philpot and Zillman, 1970; Reeh, 1989; Kahl, 1990), these temperature gradients have been shown to persist within the lowest two meters above the snow surface (Hudson and Brandt, 2005), which may cause a disparity between the "surface" temperature at 2 m and the actual skin temperature of the snow surface. In validation studies or use of remotely sensed temperatures, this distinction is important. Additionally, these temperature gradients resulting from changes in net radiation have important implications for understanding turbulent exchange between the snow and the atmosphere, which ultimately affects larger scale circulation.

In the summer of 2015, we conducted a field campaign near Summit Station, Greenland to investigate several methods of determining skin and near-surface air temperatures including use of data from the MODerate resolution Imaging Spectroradiometer (MODIS). We use these data to answer the following questions: a) How do summertime meteorological conditions impact near-surface inversions (beneath 2 m height) on the ice sheet at Summit, Greenland? b) How do MODIS surface temperature products compare to in-situ measurements of temperature, and which "surface" temperature measurements are appropriate for direct comparison with MODIS? c) Can the accuracy of MODIS algorithms to calculate surface temperature be improved through better cloud-masking?

2. Background

2.1. Surface-Based Temperature Inversions

means relatively little solar radiation is absorbed, and high emissivity of snow at longer wavelengths as compared to the emissivity of the atmosphere often leads to temperatures at the surface that are lower than the air above, a phenomenon called an inversion. The presence of surface-based inversions in the hundreds of meters of the lower atmosphere in the polar regions has long been established (Sverdrup, 1926) as a feature that results from low absorption of solar radiationand is defined by snow and measuring temperature at two different heights to determine the high emissivity of snow as compared to magnitude of the atmosphere.temperature difference over the relevant height difference. Inversions have been characterized in Greenland and the wider Arctic (Reeh, 1989; Kahl, 1990, Overland and Guest, 1991) as well as in Antarctica (Philpot and Zillman, 1970). Conditions tothat cause inversions are most frequently met in winter when incoming radiation is low. Surface-based inversions have typically been studied with 2m air temperature as the "base" of the inversion and the height of the inversion extending hundreds of meters or more into the atmosphere or higher. However, work by Hudson and Brandt (2005) demonstrated the presence of a surface-based temperature inversion below 2 m in the winter of 2001 at South Pole in Antarctica, showing that the largest temperature gradient was in the 20 cm nearest to the snow surface. However, Good (2016) presents measurements of skin temperature and 2 m air temperature from sites around the globe, and findfinds that at their polar sites, during snow-covered seasons in fall, winter, and spring, these two temperatures generally agree well, with the caveatscaveat that there is a reduced amplitude of diurnal cycle temperatures at 2 m and that the agreement is worse during summer due to solar insolation. Unlike in Hudson and Brandt (2005) and this study, data presented in Good (2016) are not from continuously snow-covered sites. .

Hall et al. (2008) analysed 2 m air temperature data and skin temperature data from across Greenland and discussed conditions that lead to near-surface thermal stratification over snow-covered areas. Incoming solar irradiance and wind speed are two major controls on thermal stratification. Temperature inversions occur when the incoming solar irradiance is small (i.e. during night) and the snow surface emits longwave radiation; the net radiation at the surface is negative, causing heat transport from the air to the snow surface. The opposite phenomenon of temperature lapse can occur when there is significant incoming solar irradiance resulting in net positive radiation at the surface, with higher temperatures closer to the ground surface and upward heat transport from the snow surface to the air. ~~Strong winds~~Winds can serve to neutralize these temperature gradients by mixing air masses.

In recent years, studies have been conducted on surface energy balance and near-surface processes in Greenland (e.g. Miller et al., 2013; 2015; 2017; Berkelhammer et al., 2016) and Antarctica (e.g. van As et al., 2005; van den Broeke et al., 2006; Kuipers Munneke et al., 2012). At our study site ~~in particular~~ at Summit, Greenland, Miller et al. (2013) studied the inversions over two years ~~at~~ but consider the 2 m air temperature to be the base of these inversions, and they did not investigate the surface processes beneath 2 m height. They find that inversions are prevalent in winter months and are less intense during summer months and that the presence of clouds results in weaker inversions. In Miller et al. (2015) the impact of clouds on the surface energy budget at Summit is further investigated, and the warming effect of clouds on 2 m air temperatures is shown in all seasons. Details of the Summit, Greenland surface energy balance are extensively documented in Miller et al. (2017). Berkelhammer et al. (2016) discuss the impacts of the surface-based temperature inversions ~~(with 2 m air temperature as the base)~~ on boundary-layer dynamics, showing that the stability of the atmosphere prevents mixing and ultimately limits accumulation at Summit. These recent studies have investigated near-surface processes ~~in the atmosphere above 2 m~~ at Summit because of the importance of surface energy balance and turbulent snow/atmosphere exchange in climate monitoring and ultimately prediction of larger scale circulation and future change in ice mass balance. ~~However,~~Though some surface temperature measurements at Summit have been made (Berkelhammer et al. 2016), controls on surface temperature gradients in the lowest 2 meters of the atmosphere, which are most relevant for the remote sensing community and also have important implications for changing ice sheet dynamics, have not been ~~definitively~~explicitly studied at Summit, Greenland.

~~2.2. Remote Sensing of Surface Temperature~~

~~There are a number of different remote sensing instruments that measure radiance in the thermal infrared part of the electromagnetic spectrum in order to determine surface~~In remote sensing validation studies or use of remotely sensed temperatures, this distinction between 2 m air temperature and skin temperature is important and has been demonstrated in polar regions. In work using satellite data to study warming trends in the Arctic, Comiso (2003) presents a dataset from an Arctic sea ice study showing correlation between 2 m air temperature and skin temperature that had been averaged monthly. Over sea ice, there was an average offset of 0.34°C between air and skin temperature (a temperature lapse), but the author indicates that similar data from Greenland show a negative offset, perhaps due to inversions that are not well understood. Indeed, best practices for thermal remote sensing validation indicate that ground-based radiance measurements that yield a skin temperature provide the best validation of remote sensing land surface temperature products (Guillevic et al., 2017). Because these data have not always been available, previous studies have used a variety of measurement types for remote sensing surface temperature validation.

temperature, including the Advanced Very High Resolution Radiometer (AVHRR), the Advanced Thermal Emission and Reflection Radiometer (ASTER), the Enhanced Thematic Mapper Plus (ETM+), and the MODIS. The theoretical basis for determining temperature of a snow surface based on measured thermal infrared radiance is described by Hook et al. (2007) and Hall et al. (2008) as follows:

$$Ls_\lambda = \left[\epsilon_\lambda L_{bb,\lambda}(T) + (1-\epsilon_\lambda) L_{sky,\lambda}\right]\tau_\lambda + L_{atm,\lambda}$$

where Ls_λ is the radiance measured by the sensor on a given satellite, ϵ_λ is the surface emissivity at a given wavelength, $L_{bb,\lambda}(T)$ is the spectral radiance from a black body as a function of temperature, $L_{sky,\lambda}$ is the spectral downwelling radiance from the atmosphere on the surface, τ_λ is the spectral transmittance through the atmosphere, and $L_{atm,\lambda}$ is the spectral radiance upwelling from atmospheric emission and scattering. If emissivity, sky radiance, transmittance, and path radiance are known, surface temperature can be determined through measurements of the radiance at the sensor.

The MODIS instrument produces widely-used land surface temperature (LST), and its products are chosen as the remote sensing product for comparison in this work. This instrument, aboard the Terra and Aqua satellites, has been collecting radiance data from 24 February 2000 to present. The surface temperature products of the Greenland Ice Sheet are used as a baseline to investigate future warming trends (e.g. Hall et al. 2012), to monitor melt events on the ice sheet (Hall et al., 2013), and as input for surface mass balance or snowpack modeling (Fréville et al., 2014; Shamir and Georgakakos, 2014; Navari et al., 2016). A number of validation studies present results acquired over various time scales and in different locations to determine the accuracy of the MODIS surface temperature products in the cryosphere (Hall et al., 2004, 2008; Koenig and Hall, 2010; Westermann et al., 2012; Hachem et al., 2012; Shuman et al., 2014; Østby et al., 2014; Shamir and Georgakakos, 2014; Hall et al., 2015; Williamson et al., 2017). Table 1 provides summary statistics related to the results of many of these validation studies and is discussed in further detail in the discussion section. Overall, a negative bias is present in nearly all validation studies, where the MODIS surface temperature is lesslower than the measured ground surface temperatureskin or 2 m air temperatures, and this bias is particularly prevalent at temperatures below -20°C.

Some studies (e.g., Hall et al., 2004; Hall et al.,, 2008; Shuman et al., 2014) use 2 m air temperature to validate the MODIS surface temperature products, which may be part of the reason for the biases that are consistently present. Shuman et al. 2014 acknowledge that differences between 2 m air temperature and skin temperature caused by inversions could cause bias in their comparison to MODIS, but at the time there was insufficient data to suggest whether inversions would persist in central Greenland and in the very near-surface. Other studies use thermochrons, either shielded (e.g., Hall et al., 2015) or during darkness (Koenig and Hall, 2010). However, Westermann et al. (2012) and Østby et al. (2014) both use pyrometers to measure thermal longwave radiation and estimate surface (skin) temperature, and these studies also find a cold bias in the MODIS surface temperatures. Østby et al. (2014) indicate that this bias is present at lower temperatures during the winter (and that there is a slight warm bias in the MODIS temperatures during summer), whereas Westermann et al. (2012) show a cold bias at higher temperatures. Identifying if and when this bias is indeed present is critical to the use of the MODIS surface temperature products over the ice sheet. A bias in the data can obscure or alter trends within a dataset. Furthermore, it is possible We hypothesize that a cold bias between 2 m air temperature and skin surface temperature could be indicative of physical processes of temperature inversion and not any issue of MODIS data validityinstrument calibration, and coupled datasets can be used to further develop our understanding of temperature processes in polar regions.

There are two standard MODIS surface temperature products that may be used to study Greenland surface temperature: the MOD/MYD11 Collection 6 product and the MOD/MYD29 Collection 6 product, where MOD refers to the Terra MODIS product and MYD refers to the Aqua MODIS product. The MOD/MYD11 product was developed as a land surface temperature product (Wan and Dozier, 1996; Wan, 2008, 2014). MOD/MYD29 was developed as an ice surface temperature product (Key and Haefliger, 1992; Key et al., 1997; Hall et al., 2004 and 2012), and while it is typically not available on land, it will be available as a special product over the Greenland Ice Sheet after further development. Both MOD/MYD11 and the preliminary version of the MOD29 special product were compared to our in situ data, and MOD/MYD11 provided a better match to the data, so we use MOD/MYD11 in the analysis.

The MOD/MYD11 method of surface temperature determination uses radiance in MODIS bands 31 and 32, which correspond to 11μm and 12 μm, respectively. The algorithm used to estimate temperature is referred to as a "split window" technique because the differences between the 11μm and 12 μm bands are used to account for atmospheric effects on the measured radiance. MOD/MYD11 estimates an emissivity value based on land cover, presence of water vapor, and estimated air temperature near the surface using other MODIS bands. This feature exists because MOD/MYD11 is a global product that estimates land surface temperature on all types of land cover types. Over snow and ice, this presents very little actual variability; in all of the data we used, the emissivity in band 32 was 0.990, and in band 31, the emissivity fluctuates between either 0.992 or 0.994. For cloud masking, MOD/MYD11 uses MOD/MYD35, the standard MODIS cloud mask product. This product gives a probability that a pixel is clear. MOD/MYD11 masks out anything below 95% probability of a clear pixel.

Previous MODIS surface temperature validation studies have used Collection 5 (C5) products; Collection 6 (C6) products became available in 2014. Improvements were made in the C6 MODIS product, most notably to rectify degradation in the calibration of the Terra sensors that was apparent in C5; however the sensor degradation was largely in the visible part of the spectrum and not in the thermal infrared part of the spectrum used to calculate surface temperature (Lyapustin et al., 2014; Polashenski et al., 2015; Casey et al., 2017). MOD/MYD11 C6 benefits from improved stability of emissivity values and improved algorithms to account for viewing angle over its C5 counterpart (Wan, 2014). Additionally, in C6, the calibration of bands 31 and 32 (used in surface temperature calculation) is improved, resulting in a decrease in measured brightness temperatures. Furthermore, cloud mask algorithms are improved in C6 (Riggs et al., 2017).

3In the summer of 2015, we conducted a field campaign near Summit Station, Greenland to measure skin and near-surface air temperature to study near-surface thermal stratification and determine its impact in validation of the MODIS land surface temperature product. We use our original dataset to determine how summertime meteorological conditions impact near-surface inversions (beneath 2 m height) on the ice sheet at Summit. Furthermore, we provide a validation of MODIS land surface temperatures, and show that the use of 2 m air temperature for MODIS validation is not recommended due to the presence of near-surface inversions. Lastly, we use in situ cloud data to show that the accuracy of the MODIS surface temperature product could be improved through stricter cloud masking.

2. Methods

~~3~~2.1. In~~-~~ situ Measurements

To characterize ~~temperatures in the lower 2 m of the atmosphere and on the snow surface skin~~snow skin temperature, an autonomous ~~temperature~~ measurement station was installed approximately 10 km NNW of Summit, Greenland (indicated on a map in Figure 1) at an undisturbed ~~snow~~ site for 40 days between June 8~~, 2015~~ and July 18, 2015. ~~The following measurements were made at the station with the sensors indicated:~~

~~1. Snow surface skin temperature using~~A Campbell Scientific/Apogee Precision Infrared (IR) Radiometer [Model: SI-111]

~~2. Snow surface skin temperature using two iButton thermochron sensors [Model: DS 1922L,~~ was used ~~in Koenig and Hall (2010)]~~

~~3. Temperature above the snow surface at 5cm height and within the snow at the following depths: 0cm, 5cm, 10cm, 15cm, 25cm, and 50cm using type T thermocouples~~

~~A schematic of the measurement set up is shown in Figure 2. For all measurements, temperatures are measured every 5 minutes, then averaged and recorded in 30-minute intervals. The thermochron sensors were placed on the snow surface with a string tied around the circumference of the sensor and attached~~ to ~~a stake in the snow. The thermochron sensors were a silver color, and they were not shielded. The sensors were occasionally buried by falling or drifting/blowing snow. From June 8 to 25 the station was visited and maintained every 2-3 days; between June 26 and July 18, the station was unmaintained, and the thermochron data for that period are not included in the analysis. Thermochron sensors were factory calibrated within a few months of deployment.~~

~~To~~ measure ~~the snow temperature with depth and in the air above the snow surface, type T thermocouple wires were fed through hollow white delrin rods approximately 0.5 cm in diameter and 30 cm in length, and the delrin rods were mounted into a central PVC pipe that was then buried in the snow so that the measurements were at the depths as described above. The ends of the thermocouple wires were stripped approximately 0.5 cm from the end and twisted several times with pliers; they were not coated with additional weather-proofing. The thermocouple measurements were calibrated against the Campbell Scientific SI-111 several weeks before deployment. Although measurements at all depths were collected, the focus in this current investigation uses only the temperatures measured at 5 cm height above the snow surface.~~

~~The Campbell Scientific SI-111 Precision Infrared Radiometer~~skin temperature of the snow. The instrument covers the wavelength range from 8 to 14 µm. It has a stated absolute accuracy of ±0.5°C from -40°C to -10°C, and ±0.2°C from -10°C to 65°C. The sensor was factory calibrated within several months of its deployment. The sensor was mounted on a horizontal rod extending approximately 60 cm out from the supporting tripod, and the sensor was approximately 60 cm from the surface, pointed directly downward. The field of view of the sensor is 22° half angle, so the legs of the tripod did not affect the measurements. Figure 2 shows an image of the sensor setup.

3.2. MODIS Products

The high latitude location of Summit, Greenland puts it within the field of view of the MODIS instruments on Terra and Aqua multiple times each day. To compare in-situ measurements to the temporally coincident MODIS collections, we use swath-level products whose file names contain the UTC time of collection within ±5 minutes. Within each swath, we select the pixel that has the minimum distance from the latitude and longitude coordinates of our in-situ measurement site. Comparisons between temperatures from the MODIS product and the in-situ measurements that are within 30 minutes of one another are used in the analysis. As skin and near-surface air temperatures can fluctuate within a span of 30 minutes, the non-synchronicity may introduce some error to the comparison, but errors should be random and non-systematic as 30-minute windows of both increasing and decreasing temperature are included in the analysis.

In comparisons of MODIS data to in situ measurements, the bias and root mean square error (RMSE) are calculated as follows:

$$Bias = \frac{1}{n}\sum_{i=1}^{n} y_i - x_i$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}$$

where n is the number of observations in the dataset, and x and y are the two datasets being compared. Unless otherwise noted, all errors are reported as a single standard deviation.

3.3. Summit Meteorological Monitoring

Summit Station was the location of the Greenland Ice Sheet Program 2 (GISP2) deep core site and has operated continuously as a year-round station for nearly a decade. NOAAThe National Atmospheric and Oceanic Administration (NOAA) has operated a meteorological station at Summit, measuring the 2 m air temperature using a shielded Logan PT139 sensor. Additionally, wind speed and incoming solar radiation data were also measured as part of the NOAA station data (NOAA ESRL Global Monitoring Division, 2017). The data provided by NOAA and used in this paper have a one minute temporal frequency, and we take a 30 minute average of the data so that the 2 m air temperature is comparable to the IR skin temperature measurements. Further details of the 2 m air measurements are outlined in Shuman et al. (2014). Additionally, through the Integrated Characterization of Energy, Clouds, Atmospheric state, and Precipitation at Summit (ICECAPS) project, a number of instruments to monitor cloud, atmosphere, and precipitation were installed at Summit in 2010. One of these instruments is the millimetre wavelength cloud radar (MMCR), a custom-built Doppler 35 GHz radar that was built in-house and measures reflectivity, mean Doppler velocity, Doppler spectra, and Doppler spectrum width (data available at http://www.archive.arm.gov). More information about the MMCR can be found in Moran et al. (1998). We use MMCR data in this study to detect the presence of clouds and determine the accuracy of the MODIS cloud mask, again employing the higher temporal frequency measurements and calculating 30-minute averages so that the data are comparable to our in-situ skin temperature measurements.

2.2 Remote Sensing of Surface Temperature with MODIS

There are many different remote sensing instruments that measure radiance in the thermal infrared part of the electromagnetic spectrum to determine skin temperature, including the Advanced Very High Resolution Radiometer (AVHRR), the Advanced Thermal Emission and Reflection Radiometer (ASTER), the Enhanced Thematic Mapper Plus (ETM+), and the MODIS. The theoretical basis for determining temperature of a snow surface based on measured thermal infrared radiance is described by Hook et al. (2007) and Hall et al. (2008) as follows:

$$Ls_{\lambda} = [\epsilon_{\lambda} L_{bb,\lambda}(T) + (1 - \epsilon_{\lambda}) L_{sky,\lambda}]\tau_{\lambda} + L_{atm,\lambda}$$

where Ls_{λ} is the radiance measured by the sensor on a given satellite, ε_{λ} is the surface emissivity at a given wavelength, $L_{bb,\lambda}(T)$ is the spectral radiance from a black body as a function of temperature, $L_{sky,\lambda}$ is the spectral downwelling radiance from the atmosphere on the surface, τ_{λ} is the spectral transmittance through the atmosphere, and $L_{atm,\lambda}$ is the spectral radiance upwelling from atmospheric emission and scattering. If emissivity, sky radiance, transmittance, and path radiance are known, surface temperature can be determined through measurements of the radiance at the sensor. In the measurements of snow, the resulting temperature is representative of the top several microns of the surface at the snow/air interface because of the penetration depth of radiation at the wavelengths used, so it is indeed a skin temperature (Warren and Brandt, 2008).

The MODIS instrument produces widely-used land surface temperature (LST), which we use as the remote sensing product in this work. This instrument, aboard the Terra and Aqua satellites, has been collecting radiance data from 24 February 2000 to present. The surface temperature products of the Greenland Ice Sheet are used as a baseline to investigate surface temperature trends (e.g. Hall et al. 2012), to monitor melt events on the ice sheet (Hall et al., 2013), and as input for surface mass balance or snowpack modeling (Fréville et al., 2014; Shamir and Georgakakos, 2014; Navari et al., 2016). In this study, we use the MOD/MYD11 Collection 6 product, where MOD refers to the Terra MODIS product and MYD refers to the Aqua MODIS product. This product has a pixel size of 1km x 1 km.

The MOD/MYD11 algorithm was developed to map land surface temperature (Wan and Dozier, 1996; Wan, 2008, 2014) using radiance in MODIS bands 31 and 32, which correspond to a center wavelength of 11µm and 12 µm, respectively. The algorithm used to estimate temperature is referred to as a “split window" technique because the differences between the 11µm and 12 µm bands are used to account for atmospheric effects on the measured radiance. MOD/MYD11 estimates an emissivity value based on land cover (assessed from bands 3-7, 13, and 16-19), presence of water vapor and estimated air temperature atmospheric profiles using MODIS sounding channels (Wan and Dozier, 1996). Emissivity can vary widely because MOD/MYD11 is a global product that estimates land surface temperature on all types of land cover types. Because this study focuses on consistently snow-covered land, there was not significant variability in the emissivity; in band 32 the emissivity is 0.990 for each data point, and in band 31, the emissivity fluctuates between either 0.992 or 0.994 as determined from MOD/MYD11. For cloud masking, MOD/MYD11 uses MOD/MYD35, the standard MODIS cloud mask product which uses data from multiple MODIS bands for cloud detection. This product gives a probability that a pixel is clear. MOD/MYD11 masks out anything below 95% probability of a clear pixel. The accuracy of the MOD/MYD11 product is limited by the uncertainties of radiative modelling, the uncertainty of absorption and scattering coefficients of aerosols and water vapor, and the atmospheric profiles of temperature and water vapor (Wan and Dozier, 1996). For surfaces with a known emissivity, the accuracy of the MOD/MYD11 is within 1°C (Wan, 1999). For further information on the MOD/MYD11 algorithm and associated uncertainties, consult Wan and Dozier (1996) and Wan (1999, 2008, 2014).

Previous MODIS surface temperature validation studies have used Collection 5 (C5) products; Collection 6 (C6) products started to become available in 2014. Improvements were made in the C6 MODIS algorithms, most notably to rectify degradation of some sensors on the Terra satellite. However the sensor degradation was largely affecting bands in the visible part of the spectrum and not in the thermal infrared part of the spectrum used to calculate surface temperature (Lyapustin et al., 2014; Polashenski et al., 2015; Casey et al., 2017). MOD/MYD11 C6 benefits from improved stability of emissivity values and improved algorithms to account for viewing angle over its C5 counterpart (Wan, 2014). Additionally, in C6, the calibration of bands 31 and 32 (used in surface temperature calculation) is improved. Supplemental Figure 1 shows comparisons of C5 and C6 data at our study site over the time period of interest. On average, C6 results in temperatures 0.2°C higher than C5. The temperatures differences are larger at higher temperatures. Finally, the cloud mask algorithms are improved in C6 (Riggs et al., 2017), resulting in a less strict cloud mask over Greenland.

The high latitude location of Summit, Greenland puts it within the field of view of the MODIS instruments on Terra and Aqua multiple times each day.

4To compare in situ measurements to the temporally coincident MODIS collections, we use swath-level products whose file names contain the UTC time of collection within ±5 minutes. Within each swath, we find the 1 km x 1 km square pixel in which our measurement site is located by minimizing distance between pixel nadir point and our in situ measurement site. Comparisons between temperatures from the MODIS products and the in situ measurements that are within 30 minutes of one another are used in the analysis. As skin and near-surface air temperatures can fluctuate within a span of 30 minutes, the non-synchronicity may introduce some error in the comparisons, but errors should be random and non-systematic, as 30-minute windows of both increasing and decreasing temperature are included in the analysis.

3. Results and Discussion

3.1. Near-Surface Temperature Measurements

The IR skin temperature measurements ~~and the thermocouple measurements of temperature~~ operated continuously ~~without interruption~~ during the 40-day campaign. ~~The thermochron dataset ends on~~The station was visited several times between June 8 and June 25, ~~because the thermochrons were not maintained after that date and subsequently became buried in the snow.~~ though no maintenance was required, and then left unmaintained for the remainder of the measurement period. A

Formatted: Font: Not Bold

time series of the IR skin temperature is presented in Figure 3. The snow skin temperature measured using the IR sensor varied between approximately -34°C and -2°C during the measurement period.

4 Gray vertical bars in the figure indicate the presence of clouds as detected by the MMCR radar, and while the diurnal cycles are clear throughout the time series, there is more high frequency fluctuation in **.1. Near-Surface Temperature Measurements**

Several different types of sensors were used to measure snow skin and near-surface air temperature during this field campaign in order to compare this study to previous MODIS surface temperature validation studies that used these measurement methods. Figure 4 shows a time series of four different temperature measurements forcloudy periods.

Our IR skin temperature measurements are compared in a subset of the study period. The diurnal cycle of temperature is present in all temperature signals despite continuous solar illumination due to the changing zenith angle of the sun throughout the day. The difference between the thermochron skin temperature measurement and other near-surface air and skin temperatures is illustratedtime series to the 2 m air temperature measurements at Summit Station in Figure 4. Because the thermochron has a mass of several grams and is silver in color, its lower albedo and thermal mass results in heating during peak solar hours. In fact, the thermochron temperature often reports above freezing skin temperatures at timesThis time window shows a clear sky period when we are certain that no surface melting was occurring. Because the thermochron was not shielded and has a different albedo than the snow, the thermochron did not provide an accurate skin temperature record when subject to solar illumination. In Koenig and Hall (2010), temperatures were monitored in the winter during polar night when there was not an issue of solar illumination. Hall et al. (2015) used thermochrons for measurement in March and April in Barrow, Alaska, when there was some sunlight during part of the day. In the Hall et al. (2015) study, both shielded and unshielded thermochrons were used at each site to study and account for issues of solar heating. The shielded thermochrons provided a better match to MODIS surface temperatures than did the non-shielded thermochrons during sunlightdiurnal cycles are clear and conditions.

Measured IR skin temperature is shown in time series with 5 cm thermocouple temperature and 2 m air temperature for a subset of the study period (approximately 5 days) in Figure 5 to illustrate details of the temperature time series. Temperatures measured 5 cm above the surface produce higher values during the peak sunlight periods of the day than do the 2 m air temperature and the IR skin temperature. Also, their temperatures are not as low as the IR skin temperature at night. The mid-day difference is likely due to heating of the small amount of exposed thermocouple wire. The wire is silver, and though it has a very small mass, it has some potential to absorb solar radiation and heat up during peak solar irradiance. The measurements show that thermal stratification causes a difference between snow skin temperature and 5 cm air temperature. When considering only those periods when incoming solar radiation is less than 300 W m^{-2} (to eliminate solar heating effects), near-surface thermal stratification causes temperature differences between the IR skin temperature and the 5 cm thermocouple air temperature of up to 6.5 °C. for inversion are most favorable. These differences are much higher at lower wind speeds; a stronger wind shear allows the system to overcome the stability in temperature and promotes heat flux from the air to the snow

~~surface. Weaker winds cannot overcome the temperature stability so the temperature differences persist. (see supplemental Figure S1).~~

Thermal stratification in the lowest several meters of the atmosphere is ~~most~~ prominently seen in the difference between 2 m air temperature and IR skin temperature~~. While~~ (Figure 4). 2 m air temperature and IR skin temperature are similar during peak solar irradiance ~~(Figure 5), there~~, with the mean difference in temperature equal to -0.32°C when incoming solar radiation is greater than 600 W m^{-2}. There is a larger difference between the two during the night-time, with 2 m air temperature ~~much~~ higher than skin temperature~~.~~ by an average of 2.4°C when incoming radiation is less than 200 W m^{-2}. This ~~is caused by~~ near-surface ~~inversions~~inversion is due to low incoming solar radiation and emission of longwave radiation from the snow surface during the night. This stable condition prevents turbulent heat exchange and allows the inversion to persist. Figure ~~6a~~5a shows a direct comparison between the 2 m air temperature measured at the NOAA weather station at Summit and the in~~-~~ situ IR skin temperature measured 10km NNW of Summit. ~~The~~As the inversions appear diurnal in nature, the measurements are quite similar at higher temperatures (above -10°C, mean difference is -0.16°C), but at lower temperatures, there is increased discrepancy between 2 m temperature and snow skin temperature~~.~~ (below -20°C, mean difference is 3.5°C). Figure ~~6b~~5b shows a histogram of the differences between the same 2 m air temperature and IR skin temperature. There is a clear skew in the histogram, indicating that 2 m air temperature is most frequently higher than skin temperature~~,~~ (in 68% of measurements). This is true in both clear and cloudy sky conditions~~.~~, where the percentage of measurements for which air temperature exceeds skin temperature is 70% in clear sky conditions and 65% in cloudy sky conditions.

Figure ~~7~~6 shows the magnitude of the temperature difference between 2 m and snow skin temperature as a function of concurrent wind speed, with the color of the marker indicating the concurrent incoming solar radiation. It is clear that increasing wind speed serves to reduce any temperature gradient in the lower meters of the atmosphere, and that at peak solar radiation, there are no inversions present. These differences are much higher at lower wind speeds; a stronger wind shear allows the system to overcome the stability in temperature and promotes heat flux from the air to the snow surface. Weaker winds cannot overcome the temperature stability so the temperature differences persist. ~~Specifically~~Specifically, for the data presented here, at incoming solar radiation above 600 W m^{-2} or wind speeds greater than approximately 7 m s^{-1}, there were ~~not~~no inversions greater than 2°C in the 2 m above the snow surface.

The presence of this near-surface thermal inversion is of particular interest in the context of previous MODIS surface temperature comparison studies. Several studies have used 2 m air temperature to compare to MODIS surface temperature products (Hall et al., 2004, 2008; Shuman et al., 2014). These studies consistently report a "cold bias" in the MODIS surface temperatures (see Table 1), where MODIS surface temperature is lower than concurrently measured 2 m air temperature. In Shuman et al. (2014), a comparison of MOD29 to 2 m air temperature results in a cold bias of approximately 3°C, and the authors note that the disagreement was larger for lower temperatures. Previous studies acknowledge that near-surface stratification may be part of the cause of the discrepancy, but also highlight other potential causes such as issues of calibration of the MODIS instruments at very low ~~(<~~~(less than approximately -20°C) temperatures (Wenny et al., 2012; Xiong et al., 2015), errors in cloud masking, and potential atmospheric interference. The data presented in Figure ~~6~~5 show that near-surface thermal stratification may play quite a large role in the discrepancies found between MOD29 and 2 m air temperatures (see Figure 1 of Shuman et al. (2014)). Inversions, which are present during periods of lower incoming solar radiation, and thus frequently lower temperature, result in offsets between skin and 2 m air temperature. Because the MODIS products ~~are a~~provide skin temperature (Warren and Brandt, 2008), the difference seen in Shuman et al. (2014) between 2 m air temperature and MODIS temperature at these lower temperatures could in fact be a signature of inversions, which the authors indeed acknowledge but did not have the data to explore. Comparisons of 2 m air temperature to MODIS surface temperature ~~allows~~allow us to see how potentially pervasive these inversions could be, though further measurements are needed to determine their presence in non-summer seasons.

Hall et al. (2008) present a figure (their Figure 2) similar to our Figure ~~6a~~5a, in which measured IR skin temperature is plotted vs. 2 m air temperature measured at Summit Station in Greenland from 2000 to 2001. However, they found a consistent offset between 2 m air temperature and skin temperature (of approximately 1°C), a trend that does not vary with temperature. In contrast, our measurements show that the offset is larger at lower temperatures than at higher temperatures and has a much larger magnitude than 1°C; inversions up to 12°C were measured in our data (Figure ~~6c).~~5c). The mean differences are reported above as -0.16±0.88°C when temperatures are above -10°C and as 3.5±2.4°C when temperatures are below -20°C. A paired t-test shows that these means are not equal to one another with a p-value of less than 0.001. In the summer, inversions are present only when solar radiation is low, and therefore temperatures are typically low, so discrepancies between 2 m air temperature and skin temperature only occur during periods of high solar zenith angle. During day time in summer, when there is more incoming radiation and temperatures are typically higher, there is good agreement between measured 2 m air temperature and skin temperature. Because the Hall et al (2008) data span a longer time scale over all seasons, it is possible that the seasonality effects of studying only summer are the root of the differences in our results. However, because inversions are known to be more persistent in the winter than in the summer, we might expect that the trend of larger offsets at lower temperatures would be even more pronounced when all seasons are included. Future studies~~ ~~, beyond our analysis here, that incorporate all seasons are needed to investigate this discrepancy and determine~~ seasons and~~ conditions under which 2 m air temperature is, or is not, a good proxy for snow skin temperature.

~~Good (2016) presents results from a study of atmospheric temperatures over many different types of land cover, comparing 2 m air temperature and skin temperature measured from an infrared radiation pyrometer, similar to the instrument used in this study. At polar sites, they find that 2 m air temperature has a reduced diurnal amplitude as compared to skin temperature, but that the two temperatures are generally in good agreement (median differences of ±1.1°C) except in the summer. However, it is unclear if these sites are snow covered in the summer, which may explain why their results differ from ours during this period. Because of the potential issues associated with using 2 m air temperature as a proxy for snow skin temperature, we elect not to compare this to MODIS temperature products. In the following sections, we do compare thermochron data to MOD/MYD11 products, with a distinction between night data and all data because of the issues of heating during peak solar irradiance. We also compare all IR skin temperature to MOD/MYD11-derived surface temperature using swath products when we can match the times of the MODIS and in-situ derived temperatures.~~

~~**4.2. In-situ Temperature Comparisons to MODIS Temperature Products**~~

~~**4.2.1. Thermochron Temperature Comparison**~~

~~It is useful to extract the thermochron data each day from 21:30 UTC - 7:15 UTC (spanning the time around the largest solar zenith angles and therefore lower incoming solar radiation) to compare this "night" data to MOD/MYD11 swath temperatures. A comparison of all the thermochron data and only "night" data is shown in Figure 8. Using only night data results in a more favorable comparison to MOD/MYD11 data than using data from the full diurnal cycle. However, even when only the night data are used, the agreement is only fair (RMSE = 4.7 °C). This agreement does appear to be somewhat better during times of higher wind speed (see Supplemental Figure S2). It is likely that during this time of year, there is still too much incoming solar radiation even during high solar zenith angle to use unshielded thermochrons for accurate skin temperature measurement.~~

~~**4.2.2.**~~

3.2. In situ Temperature Comparisons to MODIS Surface Temperature Products

3.2.1. IR Skin Temperature Comparison

Figure ~~9~~7 shows a time series of a subset of the measurement period with the 30-minute IR skin temperature measurements overlain with the MOD/MYD11 ~~surface temperature product.~~LSTs. MOD/MYD11 does not provide a surface temperature when the cloud mask indicates that there are clouds present, which is why there are some gaps in the data (i.e. at day 186/187). Most of the time series shown in Figure ~~9~~7 is during a consistently cloudless period. Terra (MOD) passes over Summit several times in the latter half of the day as temperatures are dropping. Aqua (MYD) passes over Summit as temperatures are typically increasing within the diurnal cycle. The algorithm to calculate temperature from measured radiance is the same in the two different satellites. Figure ~~9~~7 shows that there is generally good agreement between IR skin temperature and both MOD11 and MYD11 products. This is also evident in Figure ~~10~~8a, where MOD/MYD11 products combine to yield and RMSE of 1.6°C (n=374) when compared with IR skin temperature, and there is a mean bias of ~~0.7±1.4°C.~~ -0.7±1.4°C. Separate results for Terra and Aqua are not significantly different (see Table 1). In contrast to the results from Shuman et al. (2014), there does not seem to be an increase in the difference between MODIS surface temperature and in-situ temperature as temperatures decrease.

~~To investigate the root of discrepancies between MODIS surface temperature and IR skin temperature, we consider the sensitivity of the difference between MOD/MYD11 surface temperature and in-situ~~While we do not believe that 2 m air temperature is a good proxy for skin temperature, for demonstration purposes, we have compared the 2 m air temperature measurements to the MOD/MYD11 product in Figure 8b. In doing so, we find an RMSE of 3.1°C and a mean bias of 1.9±2.5°C (n=374). This comparison results in a similar RMSE to Shuman et al. (2014) of 3.5°C, though the mean bias of our comparison is slightly less than their bias was at 3°C. This comparison further illustrates the importance of using skin temperatures in MODIS validation studies. Shuman et al. (2014) were unable to conclusively say that any of their bias was a result of using 2 m air temperature instead of skin temperature, and in fact they did not think it was likely that any inversion effects would cause the gradually increasing bias with decreasing temperature because there was insufficient research on the presence of near-surface inversions in the dry snow zone in Greenland. The comparison of Figure 8a and 8b shows that at least in the summer, inversions were likely to have played a large role in their 2014 results.

As compared to other MODIS validation studies, these results indicate a closer match between in situ measurements and MODIS temperature products as indicated by smaller RMSE and mean bias (see Table 1). While the length of our study is short in comparison to many of the other works referenced, the use of a different in situ sensor is likely a key factor, and there is still a significant range of temperatures captured within our study. In comparing our results to other studies, it is also important to consider that we are using a Collection 6 product, which has seen improvements from previous versions. The Collection 5 cloud mask was more conservative over the Greenland Ice Sheet than is the Collection 6 cloud mask. If we consider only swaths that are considered cloud free by both C5 and C6 (n=341) and compare the MODIS surface temperature to our IR skin temperature, we find that the C6 performs slightly better than C5 with a lower RMSE (C6: 1.44°C, C5: 1.57°C) and lower mean bias (C6: -0.70°C, C5: -0.93°C). The comparisons are shown graphically in Figure S2 of the supplement.

However, there are still some differences between IR skin temperature and MODIS surface temperature in our validation study. To investigate the root of these discrepancies, we consider the sensitivity of the difference between MOD/MYD11 surface temperature and in situ skin temperature as a function of the following parameters: IR skin temperature, solar zenith angle, and sensor viewing angle. These results are presented in Figure ~~11~~9. The only significant relationship is between temperature difference and MODIS sensor view angle ($p = 0.0029$). ~~This means~~The viewing angle varies between 0° and 66°, and the slope of the trend (-0.01°C/°) indicates that at larger viewing angles, there is a larger difference between the MODIS surface temperature and our measured IR skin temperature, but it does not explain much of the variance, as the R^2 value is only 0.02. There is not a significant trend with temperature or with solar zenith angle. As these variables do not explain much of the difference, we believe that the differences may be due to insufficient cloud masking (discussed in the following section) and perhaps to imperfect synchronicity of measurements, where in situ skin measurements represent an average of 30 minutes but the MODIS measurement represents a shorter time window. Yu et al. (1995) suggest that ice crystal precipitation present during inversions may also cause differences between in situ and satellite skin temperatures, though they caused a warm bias rather than a cold bias.

~~4~~3.2.~~4~~2. Using In~~-Situ~~ situ Cloud Data to Improve MODIS Surface Temperature

Using the ~~millimeter cloud radar (~~MMCR~~)~~ data from Summit, we identify periods when there were clouds present above Summit Station. While our IR skin temperature measurements were 10km away, we believe that this is still a relatively good proxy for cloudiness, as we resample the data to cover a 30 minute window, so we feel it is more reflective of a larger area. Figure ~~12~~10 shows ~~the~~a comparison of IR skin temperature to the MOD/MYD11 reduced data, when cloud-affected pixels are removed~~, for MOD/MYD11~~. There is an improvement in the RMSE of the data comparison when the cloud-affected data are removed~~. In determining the strictness of the cloud mask used, there is a trade-off between the number of data points~~

available and the accuracy of the data retrieved. While improving the cloud mask would improve the data product, it would reduce the amount of measurements available. Østby et al. (2014) also use in- (from 1.6°C to 1.0°C) and the mean bias is also reduced from -0.7°C to -0.4°C. In determining the strictness of the cloud mask used, there is a trade-off due to the need to mask out all cloud contaminated pixels but not overflag data, which results in the generation of false positives and removes pixels that were in fact clear. In comparing the MMCR data to the MOD/MYD11, we find that of the 1059 times that the site was within the field of view of the satellites in June and July of 2015, there were 585 instances when both MMCR and MODIS detected cloud cover, 288 instances when both MMCR and MODIS indicated clear sky. This indicates 82% agreement. There were 86 false negatives (where MMCR indicates clouds and MODIS does not) and 100 false positives (where MMCR indicates clear sky, and MODIS indicates clouds). Østby et al. (2014) also use in situ cloud data to filter out MODIS surface temperatures that are impacted by the presence of clouds in their study in Svalbard. They found an overall false negative rate of 17%, whereas our false negative rate was 8%. Their work shows that the MOD35 cloud mask performs more poorly in the winter than in the summer, so perhaps our results from June and July actually showcase athe difference in false negatives is likely due to more favorable favourable conditions for effective cloud masking due to constant sunlight during our measurement period.

54. Conclusions

Data collected during a 40-day field campaign at Summit, Greenland in June and July of 2015 are used to improve understanding of near-surface temperature on an ice sheet, particularly with respect to MODIS land surface temperature retrieval products. In our comparison of different types of temperature measurement, we find that thermochrons and thermocouple wires, used to measure skin and near-surface air temperature during periods of polar day, can heat up, which may lead to erroneous temperature measurement, and that the thermochrons heated more than did the thermocouples. We alsoLST retrieval products. We find that at Summit, 2 m air temperature is often significantly higher than skin temperature during the summer months, particularly at periods of low incoming solar radiation and low wind speed. This near-surface inversion is even present in the 5 cm nearest to the snow surface. This result is important because previous studies that have used 2 m air temperature to validate MODIS surface temperature products have concluded that there was a cold bias in the MODIS data, but our results indicate that the MODIS data may indeed be correct,has only a very slight cold bias (-0.7°C), and the 2 m air temperature is simply not always reflectivenecessarily representative of skin temperature. Indeed, it is because of the differences between 2 m air temperature and MODIS temperature that we began to see the pervasiveness of the inversion. We do find that there is a slight cold bias in the MOD/MYD11 surface temperature products as compared to in- situ IR skin temperature, but it is not as large as previous studies have reported, and the RMSE is 1.6°C. The lower RMSE isand mean bias are likely a result of measuring the skin temperature using an IR instrument directly (instead of using 2 m air temperature)., which resulted in an RMSE of 3.1°C and a mean bias of 1.9°C). During our study period, we measured temperatures down to approximately -30°C. In the future, we plan to extend studies of this type to longer spans of time to determine if these results also are representative of lower temperatures. Furthermore, the validation presented in this study of the strong correlation between MODIS surface temperature and snow skin temperature in the summer ~~would allow~~lays a groundwork for inversions to be studied more extensively in locations where 2 m air temperature is currently measured. Finally, by using in situ cloud radar data, we confirm, as has been noted in previous studies, that the MODIS cloud mask did not remove all cloud-obscured data from the dataset. When we remove data that were cloud-obscured using the MMCR, the RMSE of MOD/MYD11 improves to 1.0°C. This indicates that ~~improved~~stricter cloud-masking in the MODIS surface temperature products could improve the accuracy of the data collected~~, although it would reduce the total amount of surface temperature measurements available~~.

Acknowledgements

~~Thank you to~~We thank the National Science Foundation and Vasilii Petrenko for the opportunity to conduct this field work, and ~~to~~ Polar Field Services and staff at Summit Station for logistic support. Data including 2 m air temperature, wind speed, and irradiance were provided by NOAA's Earth System Research Laboratory Global Monitoring Division. Cloud radar data were provided by the ICECAPS program. This work was funded by NSF-GRFP 2014186404 and NSF-1506155. Dorothy Hall was funded by NASA-NNX16AP80A.

References

Ahlstrøm, A.P. and PROMICE project team: A new programme for monitoring the mass loss of the Greenland ice sheet, In Bennike O. and Higgins A.K. eds. Review of survey activities 2007, GEUS, Copenhagen (Geological Survey of Denmark and Greenland Bulletin 15), 2008.

Berkelhammer, M., Noone, D. C., Steen-Larsen, H. C., Bailey, A., Cox, C. J., O'Neill, M. S., Schneider, D., Steffen, K., and White, J. W. C.: Surface-atmosphere decoupling limits accumulation at Summit, Greenland, Science Advances, 2, doi:10.1126/sciadv.1501704, 2016.

Box, J. E.: Greenland ice sheet surface mass balance reconstruction. Part II: surface mass balance (1840-2010), Journal of Climate, 26, 6974-6989, doi: 10.1175/JCLI-D-12-00518.1, 2013.

Casey, K. A., Polashenski, C. M., Chen, J., ~~and~~& Tedesco, M.: Impact of MODIS sensor calibration updates on Greenland ~~ice sheet~~Ice Sheet surface reflectance and albedo trends~~,~~. The Cryosphere~~ Discussions~~, 11(4), 1781, 2017~~, 1–24,~~ . https://doi~~:~~.org/10.5194/tc-~~2017-38,~~ 11-1781-2017.

Comiso, J. C.: Warming trends in the Arctic from clear sky satellite observations. Journal of Climate, 16(21), 3498-3510, 2003.

Fausto, R., van As, D. and Ahlstrøm, A.: Assessing the accuracy of Greenland ice sheet ice ablation measurements by pressure transducer, Journal of Glaciology, 58, 1144-1150, doi: 10.3189/2012JoG12J075, 2012.

Fréville, H., Brun, E., Picard, G., Tatarinova, N., Arnaud, L., Lanconelli, C., Reijmer, C., and van den Broeke, M.: Using MODIS land surface temperatures and the Crocus snow model to understand the warm bias of ERA-Interim reanalyses at the surface in Antarctica, The Cryosphere, 8, 1361–1373, doi:10.5194/tc-8-1361-2014, 2014.

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: Font: +Headings (Times New Roman), English (United States)

Formatted: Font: +Headings (Times New Roman), English (United States)

Formatted: Font: +Headings (Times New Roman), English (United States)

Formatted: Font: +Headings (Times New Roman), English (United States)

Good, E. J.: An in situ-based analysis of the relationship between land surface “skin" and screen-level air temperatures, Journal of Geophysical Research: Atmospheres, 121, 8801–8819, doi:10.1002/2016JD025318, 2016.

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J.A., Remedios, J., Román, M. and Camacho, F.: Land Surface Temperature Product Validation Best Practice Protocol. Version 1.0. In P. Guillevic, F. Göttsche, J. Nickeson and M. Román (Eds.), Best Practice for Satellite-Derived Land Product Validation (p. 60): Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/lst.001. https://lpvs.gsfc.nasa.gov/PDF/CEOS_LST_PROTOCOL_Oct2017_v1.0.0.pdf

Hachem, S., Duguay, C. R., and Allard, M.: Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain, The Cryosphere, 6, 51–69, doi:10.5194/tc-6-51-2012, 2012.

Hall, D. K., Key, J. R., Casey, K. A., Riggs, G. A., and Cavalieri, D. J.: Sea ice surface temperature product from MODIS, IEEE transactions on geoscience and remote sensing, 42, 1076–1087, doi:10.1109/TGRS.2004.825587, 2004.

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

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, Journal of Climate, 25, 4785–4798, doi:10.1175/JCLI-D-11-00365.1, 2012.

Hall, D. K., Comiso, J. C., DiGirolamo, N. E., Shuman, C. A., Box, J. E., and Koenig, L. S.: Variability in the surface temperature and melt extent of the Greenland ice sheet from MODIS, Geophysical Research Letters, 40, 2114–2120, doi:10.1002/grl.50240, 2013.

Hall, D. K., Nghiem, S. V., Rigor, I. G., and Miller, J. A.: Uncertainties of temperature measurements on snow-covered land and sea ice from in situ and MODIS data during BROMEX, Journal of Applied Meteorology and Climatology, 54, 966–978, doi:10.1175/JAMC-D-14-0175.1, 2015.

Hook, S. J., Vaughan, R. G., Tonooka, H., and Schladow, S. G.: Absolute Radiometric In-Flight Validation of Mid Infrared and Thermal Infrared Data From ASTER and MODIS on the Terra Spacecraft Using the Lake Tahoe, CA/NV, USA, Automated Validation Site, IEEE Transactions on Geoscience and Remote Sensing, 45, 1798–1807, doi:10.1109/TGRS.2007.894564, 2007.

Hudson, S. R. and Brandt, R. E.: A look at the surface-based temperature inversion on the Antarctic Plateau, Journal of Climate, 18, 1673–1696, doi:10.1175/JCLI3360.1, 2005.

Key, J. and Haefliger, M.: Arctic ice surface temperature retrieval from AVHRR thermal channels, Journal of Geophysical Research: Atmospheres, 97, 5885–5893, doi:10.1029/92JD00348, 1992.

Key, J. R., Collins, J. B., Fowler, C., and Stone, R. S.: High-latitude surface temperature estimates from thermal satellite data, Remote Sensing of Environment, 61, 302–309, doi:10.1016/S0034-4257(97)89497-7, 1997.

Koenig, L. S. and Hall, D. K.: Comparison of satellite, thermochron and air temperatures at Summit, Greenland, during the winter of 2008/09, Journal of Glaciology, 56, 735–741, 2010.

Kuipers Munneke, P., van den Broeke, M. R., King, J. C., Gray, T., and Reijmer, C. H.: Near-surface climate and surface energy budget of Larsen C ice shelf, Antarctic Peninsula, The Cryosphere, 6, 353-363, doi:10.5194/tc-6-353-2012, 2012.

Lyapustin, A., Wang, Y., Xiong, X., Meister, G., Platnick, S., Levy, R., Franz, B., Korkin, S., Hilker, T., Tucker, J., et al.: Scientific impact of MODIS C5 calibration degradation and C6+ improvements, Atmospheric Measurement Techniques, 7, 4353–4365, doi:10.5194/amt-7-4353-2014, 2014.

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, Journal of Geophysical Research: Atmospheres, 118, 495–506, doi:10.1029/2012JD018867, 2013.

Miller, N. B., Shupe, M. D., Cox, C. J., Walden, V. P., Turner, D. D., and Steffen, K.: Cloud Radiative Forcing at Summit, Greenland, Journal of Climate, 28, 6267–6280, doi:10.1175/JCLI-D-15-0076.1, 2015.

Miller, N. B., Shupe, M. D., Cox, C. J., Noone, D., Persson, P. O. G., and Steffen, K.: Surface energy budget responses to radiative forcing at Summit, Greenland, The Cryosphere, 11, 497–516, doi:10.5194/tc-11-497-2017, 2017.

Moran, K. P., Martner, B. E., Post, M., Kropfli, R. A., Welsh, D. C., and Widener, K. B.: An unattended cloud-profiling radar for use in climate research, Bulletin of the American Meteorological Society, 79, 443–455, doi:0.1175/1520-0477(1998)079<0443:AUCPRF>2.0.CO;2, 1998.

Navari, M., Margulis, S., Bateni, S., Tedesco, M., Alexander, P., and Fettweis, X.: Feasibility of improving a priori regional climate model estimates of Greenland ice sheet surface mass loss through assimilation of measured ice surface temperatures, The Cryosphere, 10, 103–120, doi:10.5194/tc-10-103-2016, 2016.

NOAA ESRL Global Monitoring Division. Meteorology measurements at Summit Station, Greenland. Compiled by the Observatory Operations Group. National Oceanic and Atmospheric Administration (NOAA), Earth System Research Laboratory (ESRL), Global Monitoring Division (GMD): Boulder, Colorado, USA, 2017.

Østby, T. I., Schuler, T. V., and Westermann, S.: Severe cloud contamination of MODIS Land Surface Temperatures over an Arctic ice cap, Svalbard, Remote Sensing of Environment, 142, 95–102, doi:10.1016/j.rse.2013.11.005, 2014.

Overland, J. E., and Guest, P. S.: The Arctic snow and air temperature budget over sea ice during winter. Journal of Geophysical Research: Oceans (1978–2012), 96(C3), 4651-4662. 1991.

Polashenski, C. M., Dibb, J. E., Flanner, M. G., Chen, J. Y., Courville, Z. R., Lai, A. M., Schauer, J. J., Shafer, M. M., and Bergin, M.: Neither dust nor black carbon causing apparent albedo decline in Greenland's dry snow zone: Implications for MODIS C5 surface reflectance, Geophysical Research Letters, 42, 9319–9327, doi:10.1002/2015GL065912, 2015.

Rae, J., Aðalgeirsdóttir, G., Edwards, T., Fettweis, X., Gregory, J., Hewitt, H., Lowe, J., Lucas-Picher, P., Mottram, R., Payne, A., et al.: Greenland ice sheet surface mass balance: evaluating simulations and making projections with regional climate models, The Cryosphere, 6, 1275–1294, doi:10.5194/tc-6-1275-2012, 2012.

Reeh, N.: Parameterization of melt rate and surface temperature on the Greenland ice sheet, Polarforschung, 59, 113–128, 1989.

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

Riggs, G. A. and Hall, D. K.: MODIS sea ice products user guide to collection 6, National Snow and Ice Data Center, University of Colorado, Boulder, CO, USA, 2015., Hall, D. K., & Román, M. O. (2017). Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records. *Earth System Science Data*, *9*(2), 765.

Rignot, E., Velicogna, I., van den Broeke, M. R., Monaghan, A., and Lenaerts, J. T.: Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise, Geophysical Research Letters, 38, doi:10.1029/2011GL046583, 2011.

Shamir, E. and Georgakakos, K. P.: MODIS Land Surface Temperature as an index of surface air temperature for operational snowpack estimation, Remote Sensing of Environment, 152, 83–98, doi:10.1016/j.rse.2014.06.001, 2014.

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, Journal of Applied Meteorology, 40, 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), Journal of Applied Meteorology and Climatology, 53, 2171–2180, doi:10.1175/JAMC-D-14-0023.1, 2014.

Steffen, K. and Box, J.: Surface climatology of the Greenland ice sheet: Greenland Climate Network 1995–1999, Journal of Geophysical Research: Atmospheres, 106, 33 951–964, doi:10.1029/2001JD900161, 2001.

Steffen, K., Box, J., and Abdalati, W.: Greenland climate network: GC-Net, US Army Cold Regions Research and Engineering (CRREL), CRREL Special Report, pp. 98–103, 1996.

Stocker, T. F.: Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2014.

Sverdrup, H. U.: Scientific work of the Maud Expedition, 1922-1925, The Scientific Monthly, 22, 400–410, 1926.

Toller, G., Xiong, X., Sun, J., Wenny, B. N., Geng, X., Kuyper, J., Angal, A., Chen, H., Madhavan, S., and Wu, A.: Terra and Aqua moderate resolution imaging spectroradiometer collection 6 level 1B algorithm, Journal of Applied Remote Sensing, 7, 073 557–073 557, doi:10.1117/1.JRS.7.073557, 2013.

Van As, D., Van Den Broeke, M., and Van De Wal, R., Daily cycle of the surface layer and energy balance on the high Antarctic Plateau, Antarctic Science, *17*(1), 121-133, doi: 10.1017/S095410200500252X, 2005

Van As D., Fausto, R.S. and PROMICE Project Team: Programme for monitoring of the Greenland Ice Sheet (PROMICE): first temperature and ablation records, In Bennike O., Garde A.A. and Watt W.S. eds., Review of survey activities 2010. GEUS, Copenhagen, 73–76 (Geological Survey of Denmark and Greenland Bulletin 23), 2011.

Van Den Broeke, M., Reijmer, C., Van As, D., and Boot, W., Daily cycle of the surface energy balance in Antarctica and the influence of clouds, *International Journal of Climatology*, *26*(12), 1587-1605, doi: 10.1002/joc.1323, 2006.

Formatted: English (United States)

Formatted: English (United States)

Formatted: Character scale: 105%

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg,W. J., van Meijgaard, E., and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, The Cryosphere, 10, 1933–1946, doi:10.5194/tc-10-1933-2016, 2016.

Van Tricht, K., Lhermitte, S., Lenaerts, J. T., Gorodetskaya, I. V., L'Ecuyer, T. S., Noël, B., van den Broeke, M. R., Turner, D. D., and Van Lipzig, N. P.: Clouds enhance Greenland ice sheet meltwater runoff, Nature communications, 7, doi:10.1038/ncomms10266, 2016.

Vernon, C. L., Bamber, J., Box, J., Van den Broeke, M., Fettweis, X., Hanna, E., and Huybrechts, P.: Surface mass balance model intercomparison for the Greenland ice sheet, Cryosphere (The), 7, 599–614, doi:10.5194/tc-7-599-2013, 2013.

Wan, Z: MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD) Version 3.3, https://modis.gsfc.nasa.gov/data/atbd/atbd_mod11.pdf, 1999.

Wan, Z.: New refinements and validation of the MODIS land-surface temperature/emissivity products, Remote sensing of Environment, 112, 59–74, doi:10.1016/j.rse.2006.06.026, 2008.

Wan, Z.: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product, Remote sensing of Environment, 140, 36–45, doi:10.1016/j.rse.2013.08.027, 2014.

Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature from space, IEEE Transactions on geoscience and remote sensing, 34, 892–905, doi:10.1109/36.508406, 1996.

Warren, S. G. and Brandt, R. E.: Optical constants of ice from the ultraviolet to the microwave: A revised compilation, Journal of Geophysical Research, 113, doi:10.1029/2007jd009744, 2008.

Wenny, B., Xiong, X., and Madhavan, S.: Evaluation of Terra and Aqua MODIS thermal emissive band calibration consistency, in: SPIE Remote Sensing, pp. 853 317–853 317, International Society for Optics and Photonics, doi:10.1117/12.974230, 2012.

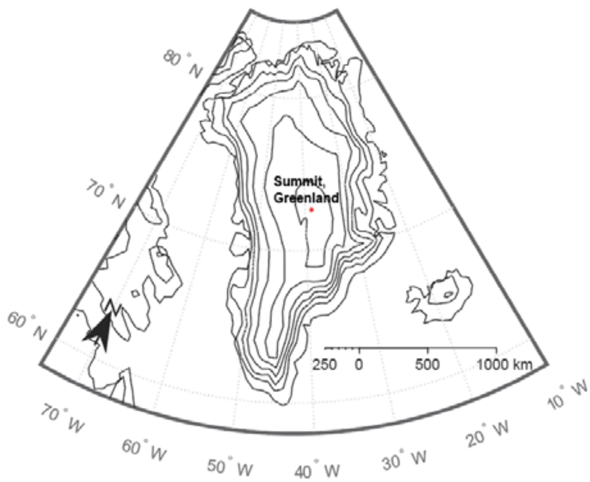
Westermann, S., Langer, M., and Boike, J.: Systematic bias of average winter-time land surface temperatures inferred from ~~{~~MODIS~~}~~ at a site on Svalbard, Norway, Remote Sensing of Environment, 118, 162 – 167, doi:http://dx.doi.org/10.1016/j.rse.2011.10.025, 2012.

Williamson, S. N., Hik, D. S., Gamon, J. A., Jarosch, A. H., Anslow, F. S., Clarke, G. K., and Rupp, T. S.: Spring and summer monthly MODIS LST is inherently biased compared to air temperature in snow covered sub-Arctic mountains, Remote Sensing of Environment, 189, 14–24, doi:10.1016/j.rse.2016.11.009, 2017.

Yu, Y., Rothrock, D. A., and Lindsay, R. W.: Accuracy of sea ice temperature derived from the advanced very high resolution radiometer, Journal of Geophysical Research: Oceans (1978–2012), 100(C3), 4525-4532, 1995.

Xiong, X.,Wu, A.,Wenny, B. N., Madhavan, S.,Wang, Z., Li, Y., Chen, N., Barnes,W. L., and Salomonson, V. V.: Terra and Aqua MODIS Thermal Emissive Bands On-Orbit Calibration and Performance, IEEE Transactions on Geoscience and Remote Sensing, 53, 5709–5721, doi:10.1109/TGRS.2015.2428198, 2015.

Formatted: Character scale: 105%


80 ° N
70 ° N
60 ° N
Summit,
Greenland
250 0 500 1000 km
N
70 ° W
60 ° W
50 ° W
40 ° W
30 ° W
20 ° W
10 ° W

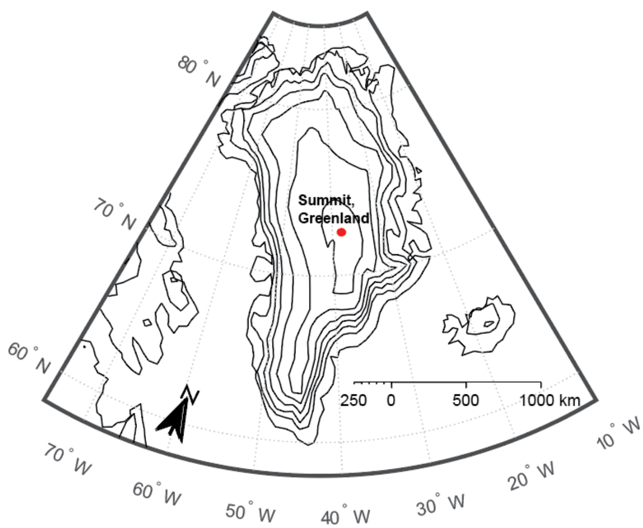

80 ° N
70 ° N
60 ° N
Summit,
Greenland
250 0 500 1000 km
N
70 ° W
60 ° W
50 ° W
40 ° W
30 ° W
20 ° W
10 ° W

Figure 1: Map indicating the location of Summit, Greenland, the study site for remote sensing and in- situ temperature comparisons. Contour lines represent elevation change of 500m500 m. Latitude and longitude coordinates for the measurement site are 72.65923°N, 38.57067°W.

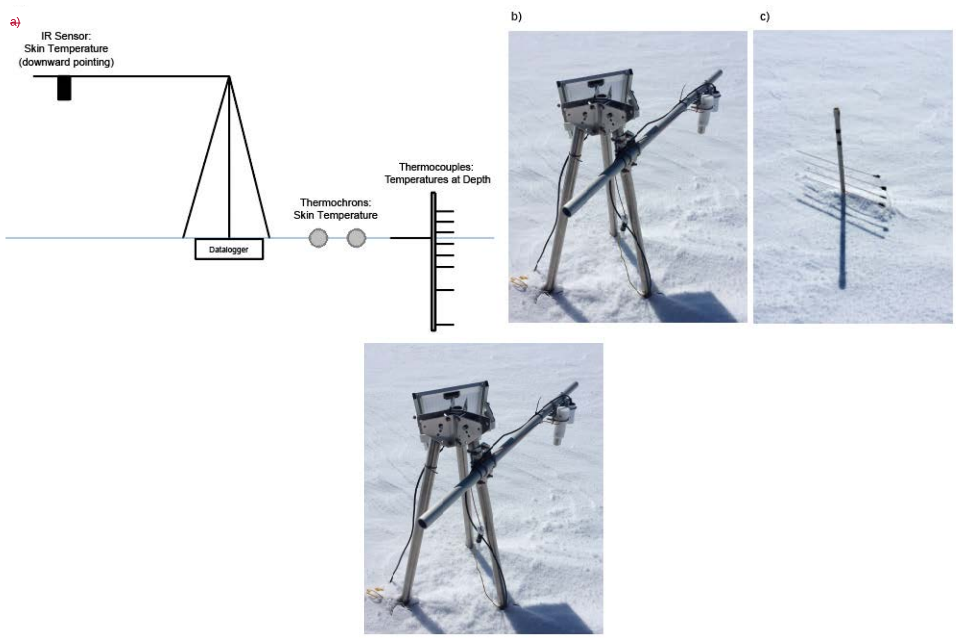


Figure 2: a) Schematic of the types of measurements that were made at the remote station near Summit, Greenland. Measurements included IR skin temperature, thermochron-measured skin temperature, and thermocouple-measured snow temperature with depth, b) Image of the IR skin temperature sensor and tripod set up, and c) Image of the thermocouple wire set up to measure temperature at fixed heights above the snow and depths within the snow. .

~~Temperature (°C)~~

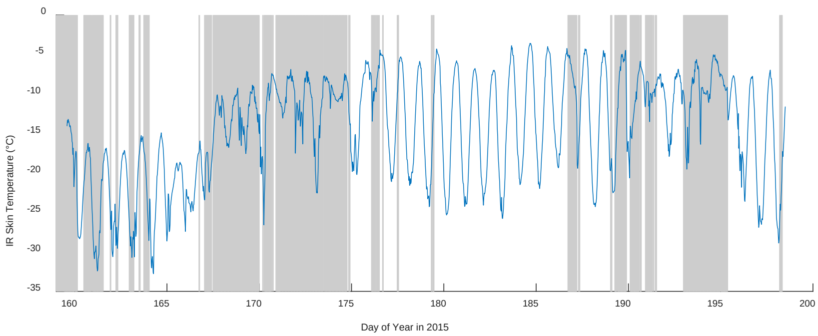
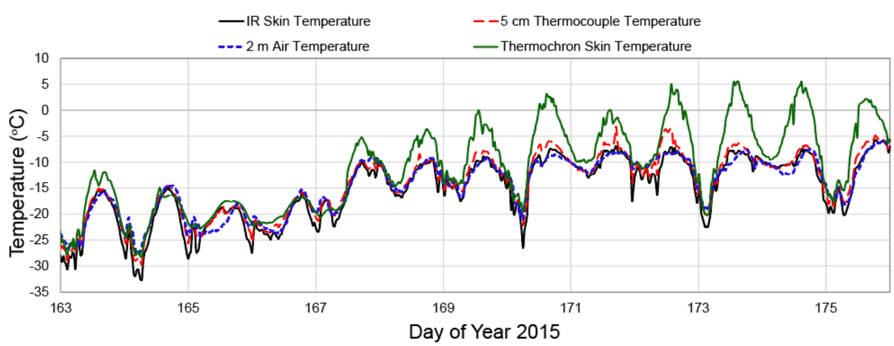


Figure 3: Time series of skin temperature at Summit, Greenland measured with SI-111 IR thermometer (blue). Grey bars indicate presence of clouds as detected by a millimeter cloud radar at Summit Station.

Formatted: Centered

Formatted: Justified



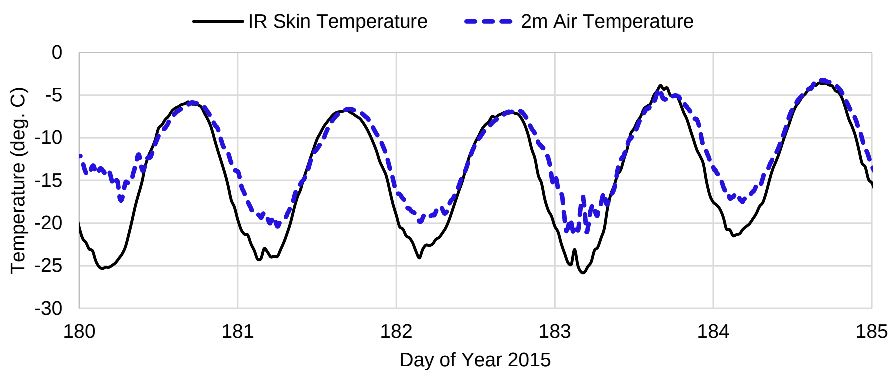


Figure 4: Time series of IR skin temperature, thermochron skin temperature, 2 m air temperature, and 5 cm thermocouple temperature for the duration of the thermochron measurements. In direct sunlight, thermochrons record higher temperatures than IR sl nperatures, 5cm air temperature and 2 m air temperature.

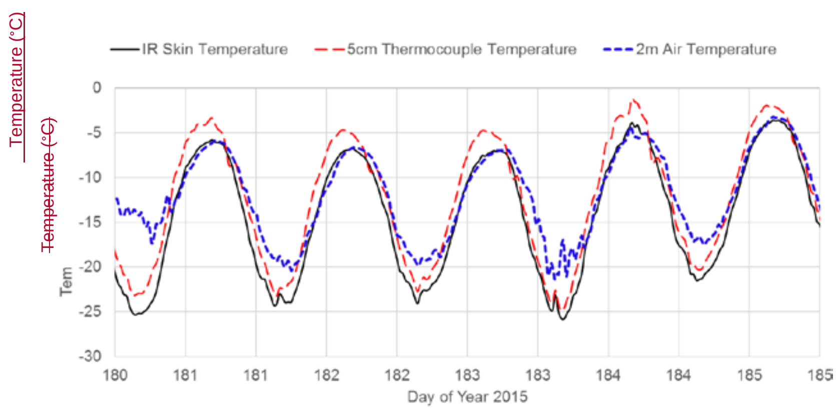


Figure 5: Time series of IR skin temperature, 2 m air temperature, and 5 cm thermocouple temperature during a clear sky period near Summit, Greenland.

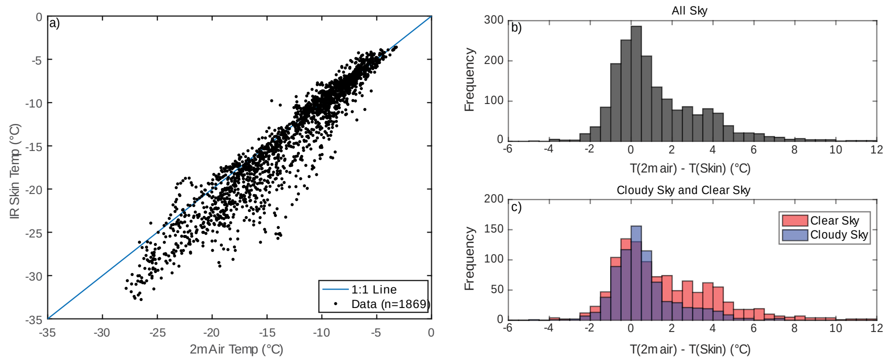


Figure ~~6~~5: a) Comparison of 2 m air temperature to IR skin temperature at Summit, Greenland during June and July 2015. The difference between air and skin temperature is largest at lower temperatures. b) Histogram of the difference between 2 m air temperature and IR skin temperature during the study period in June and July of 2015 at Summit, Greenland during all sky conditions and c) clear sky and cloudy sky conditions (as detected by MMCR data) separated. The difference is skewed to positive temperature differences indicating higher air temperatures than skin temperatures.

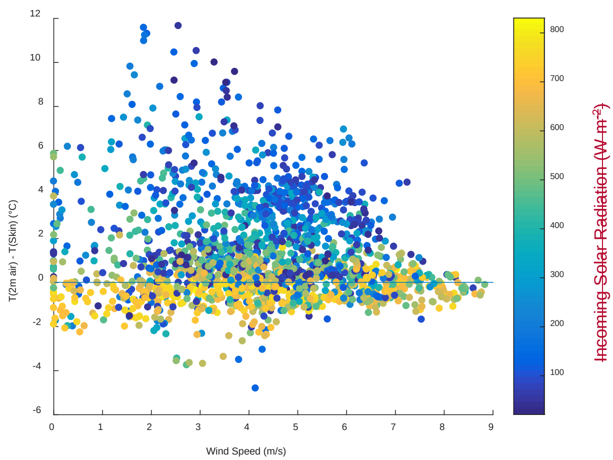


Figure ~~7~~6: Difference between 2m air temperature and IR skin temperature showing the presence of strong surface-based inversions at low wind speeds and low values of incoming solar radiation (indicated by the marker colour).

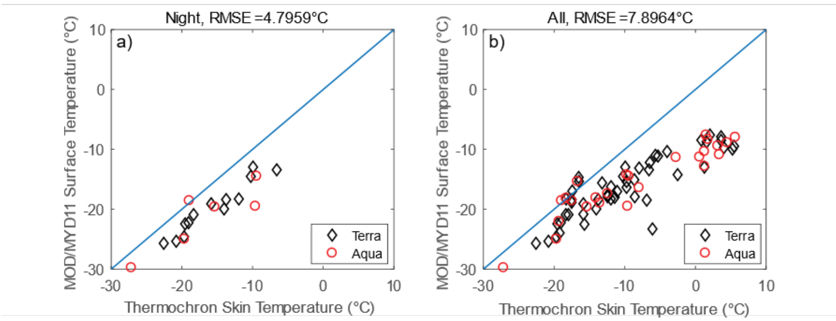


Figure 8: Comparison of thermochron skin temperatures to MOD/MYD11 C6 surface temperature product a) during the night and b) for all available data. Agreement improves for night-time measurements because thermochrons are not heated by peak solar radiation, but there is considerable spread in the data.

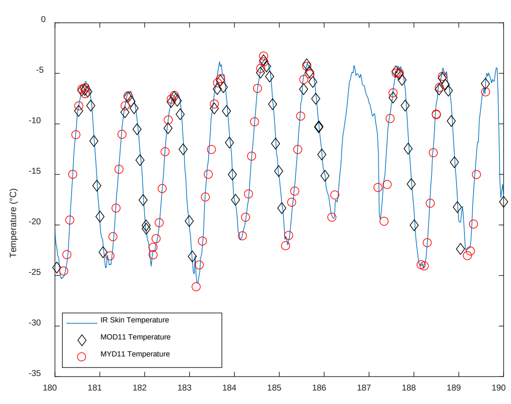


Figure 9

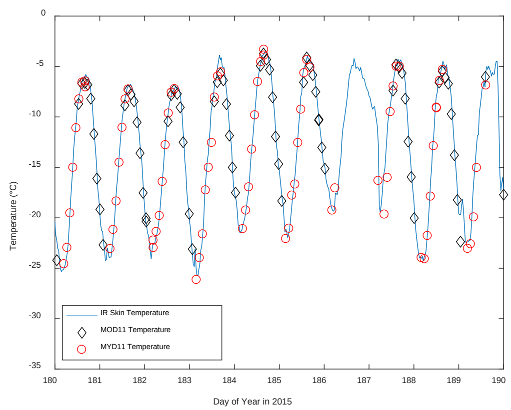


Figure 7: Time series as shown in Figure 3 with only a temporal subset of data presented to clearly show the diurnal cycle of temperature during fairly clear conditions. Note that the MOD/MYD11 product shows good agreement with IR skin temperature throughout the diurnal cycle.

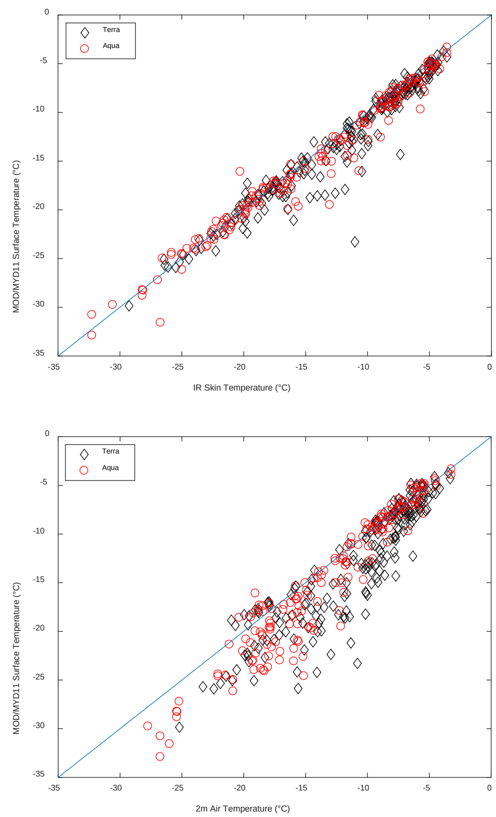


Figure ~~10:~~8: a) Direct comparison of in~~-~~situ IR skin temperature data with MOD/MYD11 C6 surface temperatures. Agreement between satellite and ground-based measurements is quite good (RMSE = 1.6°C, n=374), and there is not a noticeable difference between the performance of the MOD11 and MYD11 temperature products, on the Terra and Aqua satellites, respectively. b) Direct comparison of 2 m air temperature with MOD/MYD C6 surface temperatures. This is illustrative of bias that may be inferred if 2 m air temperature is used in validation studies when inversions are present. RMSE = 3.1°C.

Formatted: Centered

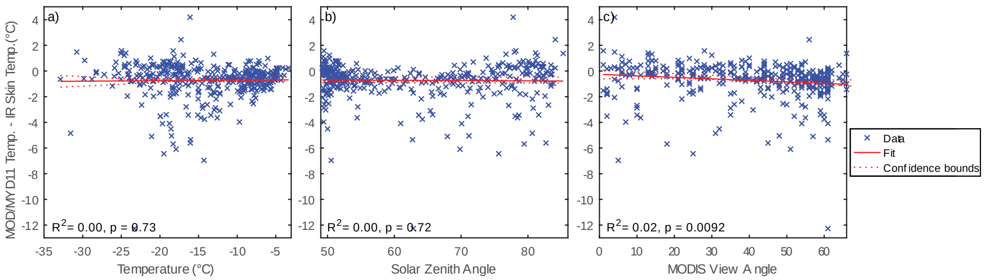


Figure ~~11~~9: Difference in temperature measured from MOD/MYD11 and in-situ IR skin temperature measurements as a function of a) IR skin temperature, b) solar zenith angle, and c) MODIS viewing angle. The only significant relationship is that the temperature difference is sensitive to the MODIS viewing angle. While the relationship is statistically significant, it is not a strong control on the temperature difference.

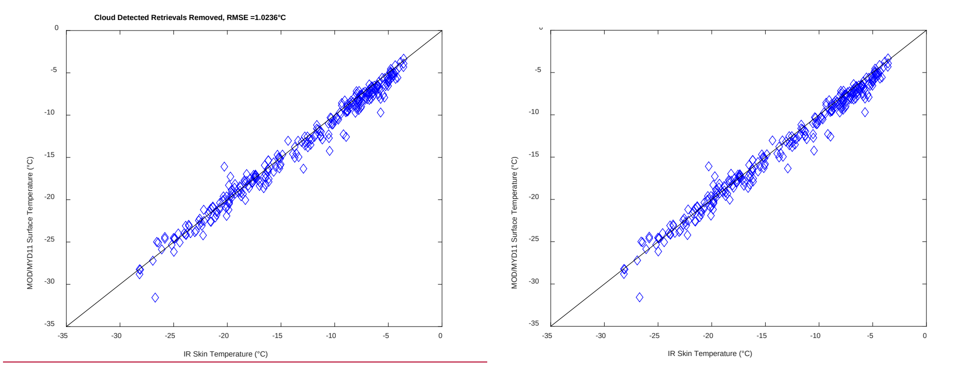


Figure ~~12~~10: Comparison of MOD/MYD11 to in-situ IR skin temperature after cloud-affected data are removed. The RMSE is 1.0°C and the mean bias is -0.4°C.

Table 1: Summary statistics from recent literature comparing MODIS surface temperature products to in- situ surface temperature measurements in snow-covered regions.

Study	Location	Temperature Range	Temperature Measurement	MODIS Product	MODIS Collection	Sample Number	RMSE	Bias
Hall et al. 2004	South Pole	-70º°C to -20º°C	2m Air Temperature	29	MOD/MYD29 Collection 4	255	1.7º°C (n=255)	-1.2º°C
Hall et al. 2008	15 Greenland AWS	-40º°C to 0º°C	2m Air Temperature (only during neutral thermal stratification)2m Air	11	MOD/MYD11 Collection 4	48	2.1º°C (n=48)	-0.3º°C
Koenig and Hall 2008	Summit, Greenland	-41º°C to -20º°C	Thermochron Skin Temperature	11	MOD/MYD11 Collection 5	62	3.1º°C (n=62)	-3.4º°C
		-60º°C to -20º°C	2m Air Temperature	11	MOD/MYD11 Collection 5	259	4.1º°C (n=250)	-5.5º°C
Westermann et al. 2012	Ny Alesund, Svalbard	-40º°C to 0º°C	IR Skin Temperature	11	MOD/MYD11 Collection 55			~-3º°C
Shuman et al. 2014	Summit, Greenland	-60º°C to 0º°C	2m Air Temperature	29	MOD29 (Special Greenland Product) Collection 5	2536 2270	All: 5.3º°C (n=2536) Filtered: 3.5º°C (n=2270)	~-3º°C
OtsbyØstby et al. 2014	Svalbard	-45º°C to 0º°C	IR Skin Temperature	11	5	3941 1065	MOD/MYD11 Collection 5All: 5.3°C Filtered: 3.0°C	All: 5.-3º.3°C (n=3941) Filtered: 3.-0º.3°C (n=3941)
Hall et al. 2014	Barrow, Alaska (tundra site)	-42º°C to -20º°C	Thermochron Skin Temperature	11	MOD11 Collection 5	-69		-2.3±3.9º°C (n = 69)
				11	MYD11 Collection 5	84		0.6±2.0º°C (n = 84)

Inserted Cells

Inserted Cells

Inserted Cells

Inserted Cells

Inserted Cells

Deleted Cells

Formatted: English (United Kingdom)

Inserted Cells

Inserted Cells

This Study	Summit, Greenland	-30º°C to 0º°C	IR Skin Temperature	11	MOD/MYD11 Collection 6 (C6)6	374 288	All: 1.6º°C (n=374) Cloud Filter: 1.0º°C (n=288)	All: -0.7±1.4º°C Cloud Filter: -0.4±0.9º°C
				MOD11 C6	6	1.8ºC (n=207)	1.8°C	-0.8±1.6º°C
				MYD11 C6	6	1.4ºC (167)	1.4°C	-0.6±1.3º°C

Inserted Cells

Inserted Cells

Inserted Cells