Thank you for your constructive comments. We have addressed your general comments in a bulk response and your specific comments line by line below. Comments are indicated by boldface and italicized text; our responses in normal text and preceded by **[Author response]**.

Review of "Evaluating sources of an apparent cold bias in MODIS land surface temperatures in the St. Elias Mountains, Yukon, Canada"

Overview:

This manuscript describes analysis of in-situ air temperature data relative to the MODIS LST product (MYD21) with supplemental data from ASTER, Landsat, and additional meteorological measurements at two sites in the St. Elias Mountains. The central focus of the manuscript is to determine the causes of difference between in-situ air temperature measurements and MODIS LST data (called the "MODIS offset"). Three main hypotheses are explored: 1) The large footprint of the MODIS pixels results in the offset; 2) The lack of constraint in surface emissivity results in the offset; 3) Near-surface temperature inversions lead to this offset. The work demonstrates that the first two hypothesis are unlikely to play a large role in the offset, and that the near-surface inversions are likely a reason why there is a difference between in-situ air temperature and MODIS LSTs. While this has been shown in previous studies, the authors indicate that this is the first time that inversions have been identified as a cause of this offset in this type of environment. The manuscript makes a compelling case that the system studied is an important one given the role of Alaskan glaciers in sea level rise and the fact that the surface temperatures in this region are understudied. Understanding how representative MODIS LSTs (and other remote sensing surface temperature products) are of actual surface conditions is critical to effectively monitoring this region.

Overall, this is important work and a useful contribution. I think that the work can be improved by increased clarity and discussion about the fact that a comparison of 2m air temperature and MODIS LST is not a 1:1 comparison. I see that it is addressed in the text, but I believe it needs to be a bit more explicit. For example: in Section 4.4, I think the language of "corrected" MODIS LST is a bit misleading because there isn't a true error in MODIS LSTs identified; it is simply determined that it is not equivalent to 2m air temperature. The linear regression implemented is a tool to adjust the LST to an air temperature. Furthermore, in the abstract and conclusion section, I think it is important to clarify that this analysis doesn't rule out the possibility that there may be measurement errors in the MODIS algorithm for determining surface temperature because there was not a direct comparison to in-situ skin temperature in this study. While convincing evidence is provided that inversions are likely a source of difference between MODIS LST and in-situ 2m air temp, it is still possible that MODIS LSTs have error relative to in-situ skin temp too. Lastly, in the title, I would be hesitant to call this a cold bias without explicitly stating that the "bias" is relative to air temperature measurements.

I have several specific comments to improve clarity and increase detail about methodology, but the manuscript is overall well-written, the figures are clear, and the methods are sound. This is a very rich dataset, and the analysis and results provide an important contribution to our understanding of near surface temperatures in understudied snow and ice covered regions.

[Author response]
We understand that clarity regarding the difference between 2 m air temperature and MODIS LST is essential, and we will address this more explicitly. For example, see our edited statement in response to comments on Line 41. We will also add statements to the abstract and conclusion clarifying that this study does not rule out errors in the MODIS algorithm. Finally, we will change our title to "Offset of MODIS land surface temperatures from in situ measurements in the Upper Kaskawulsh Glacier region (St. Elias

mountains) indicates near-surface temperature inversions" and insert the following statement preceding our goal to clarify the distinction between air and surface temperatures.

Edited Line 99: Lastly, it may be that the LST offset does not arise during the calculation of LSTs at all, but is a real physical temperature difference between the surface and air due to the development of a near-surface temperature inversion. Although MODIS LSTs can be a useful complement to in situ air temperatures, the two cannot be directly compared and physical differences between the two must be accounted for when using them together.

Specific Comments:

Lines 4-5: In the statement that MODIS LSTs are offset from AWS data, is this referring to prior studies or the current work? Please clarify.

[Author response]

This statement is in reference to prior studies.

Updated text to read: However, MODIS LSTs in the St. Elias Mountains have been found in prior studies to show an offset from available weather station measurements, the source of which is unknown.

Lines 15-17: I agree that understanding near-surface physical properties is critical to convert MODIS LSTs from a surface temperature measurement to an air temperature measurement, but this is different from improving the accuracy of the MODIS LST, which this work doesn't directly address because it compares in-situ air temp (and not in-situ skin temp) to MODIS LSTs. Please revise this statement accordingly.

[Author response]

Updated text to read: These results demonstrate that efforts to derive an air temperature measurement from MODIS LST data should focus on understanding near-surface physical properties rather than refining the MODIS sensor or LST algorithm.

Line 41: The statement that all in-situ temperatures are air temperatures is key here. I think it requires a bit of further discussion. Given the Guillevic et al. (2017) report, validation should be done with skin temperature wherever possible. I understand that in-situ skin temperature data is not available in this case, and I believe comparison to in-situ air temperature is a worthy endeavor, but I think it should be explicitly addressed here (or elsewhere) that this is different from a standard validation of the MODIS LST product.

[Author response]

Updated text to read: Remote sensing temperatures include the final surface temperature product, as well as "brightness temperature", or the surface temperature of a perfect blackbody emitter under the same conditions. In contrast, temperatures measured in situ are directly measured using instruments onsite, and can be measured for both the earth's surface and the air above it. Surface temperatures measured in situ provide important validation for remote sensing surface temperatures such as MODIS LSTs. However, because of a lack of in situ surface temperature data in our study region, unless otherwise stated, all in situ temperatures used here refer to the air ~2m above the land surface. Our study is therefore not a standard validation of the MODIS LST product, but rather an evaluation of its use in conjunction with in situ air temperatures to characterize the near-surface temperature conditions of the St. Elias region.

Table 1: Consider adding to this table further information about the footprint of each sensor, the specific products/instruments used, and the uncertainty associated with each measurement (if

available).

[Author response]

We will update Table 1 to contain the below information:

	Name	Measurement technique	Footprint	Product/instrument	Uncertainty
Air temperatures	Divide AWS	In situ	Point measurement	Campbell 107F	± 0.2°C
				HOBO S-THB-M008 12-bit sensor	± 0.21°C
	Eclipse AWS	In situ	Point measurement		
	iButton	In situ	Point measurement	Maxim Integrated iButton Data Logger DS1922L	± 0.5°C
Surface temperatures	MODIS LST	Remote sensing	1 km	MYD21	
	ASTER surface temperature	Remote sensing	90 m	AST_08	
	MODIS BT	Remote sensing	1 km	MODTBGA_006	
	ASTER BT	Remote sensing	90 m	Calculated from ASTL1T following Ndossi and Avdan (2016)	
	Landsat BT	Remote sensing	Resampled from 100 m to 30 m (Landsat 8)	LC08	
			Resampled from 120 m to 30 m (Landsat 5,7)	LE07 LT05	

Lines 118-122: In the window when the two sensors at Divide overlap, which dataset is used?

In the window where the two sensors at Divide overlap, the HOBO S-THB-M008 12-bit sensor was used because it provides contemporary solar radiation, relative humidity, wind speed, and pressure data.

Lines 123-125: Was the container containing the iButton sensor ventilated? Was it a light color to limit absorbed solar radiation? I know that there can be issues with iButton sensors heating up during periods of high incoming solar radiation.

[Author response]

Available temperature data at Eclipse are lower quality than at Divide, with limited temporal coverage and sensors not up to World Meteorological Organization standards. We therefore focus on data from Divide, but include available data from Eclipse with the caveat that results are less robust. Temperatures at Eclipse were obtained from an AWS from 2005-2007, and a Maxim Integrated iButton Data Logger DS1922L (±0.5°C) from 21 May 2016 to 17 May 2017, both located on or near a bedrock outcrop ~3 km from the site of an ice core drilled at Eclipse in 2016 (Fig. 3). The AWS recorded hourly averages of 5 minute sampling intervals using digital sensors housed in a passively vented radiation shield at a height of approximately 2 m (Williamson et al., 2020). The iButton recorded temperatures at 3-hour intervals and was placed inside an unvented clear plastic container shielded with rocks. Because data is so limited at Eclipse, we combine the AWS and iButton datasets for maximum coverage at the site. We refer to both the Divide AWS and the combined Eclipse iButton and AWS data as "AWS" for the remainder of this paper.

Lines 125-126: It looks in Figure 3 like the Eclipse Weather Station and Eclipse iButton datasets do not overlap, so it's not clear to me how it was determined that the records were consistent. Please clarify and provide data if needed, perhaps in a supplement.

[Author response]

We did not perform a robust test of the datasets' consistency, instead choosing to focus on data from Divide, which is more abundant. Around 88% of the temperature data used in this study came from Divide. Additionally our examination of other meteorological variables and our surface energy balance calculations are all performed with data from Divide. However, we still included what data we had from Eclipse to supplement the results at Divide. See edited statement above in response to comment on Lines 123-125.

Lines 135-143: I think that choosing to look at a nearby MODIS pixel that does not include the darker nunatak surfaces is probably a good idea, but the implications of this choice should be further explored. Perhaps the air temperature above the dark surface actually is higher than the air temperature above the nearby ice/snow. Comparing the MODIS pixels and providing that information in Table 2 is great, and I think that in the discussion of results, the manuscript should come back to this and address what the implications of this choice are on the results.

[Author response]

Updated text to read: Our goal is to determine the dominant source of the offset in MODIS LSTs at glaciated sites in the St. Elias. Because the Eclipse and Divide AWS are located on nunataks, we test for the LST offset using MODIS data encompassing adjacent ice core sites ~3 km from each AWS location, thereby excluding the dark nunatak surface from the MODIS pixel and focusing on the ice surface (Fig. 2). We compute the difference in MODIS LST between the ice core site grid cell (containing only ice) and the AWS site grid cell (containing ice and rock) to determine whether the inclusion of the nunatak has a discernible effect on the MODIS LST.

To be inserted in results section (and add figure plotting temperature differences between AWS and ice core sites): MODIS data at the Divide AWS nunatak and adjacent ice core site have a median temperature difference of 0.86°C and interquartile range of 1.97°C. The difference between the two sites shows greater variability in the fall (IQR = 3.21) and winter (IQR = 3.98) than in the spring (IQR = 0.70) and summer (IQR = 1.80), with the ice core site tending to be slightly colder in the winter (median temperature difference of -0.49°C), but warmer in the spring (median = 0.95°C), summer (median = 1.33°C) and fall (median = 0.28°C). Temperature differences between the Divide AWS and ice core site are summarized in Table 2.

Note: Table 2 will be edited to show differences between the ice core and AWS sites as medians rather than means so as to be more directly comparable with median LST offsets from AWS temperatures.

To be inserted in the discussion section: MODIS LSTs at the ice core site do not tend to be colder than at the AWS site except during the winter. The inclusion of the warmer nunatak surface in the MODIS grid cell at the AWS site fails to provide a compelling explanation for the colder wintertime LSTs at the ice core site, given that more of the rock surface would likely have snow cover during the winter. The colder wintertime

LSTs at the ice core site may contribute to the MODIS LST offset from in situ temperature measurements examined in this study. However, this contribution is too small (median = -0.49°C) to explain the magnitude of the MODIS LST offset at the Divide ice core site (median = -8.40°C). In the spring, summer, and fall, the LSTs at the ice core site tend to be slightly warmer than at the AWS site. Results here may therefore underestimate the magnitude of the MODIS LST offset from AWS temperatures in these seasons.

Line 145: I see that later in the manuscript there is a justification provided for choosing to only consider data from the time window from 11:00 – 1:30. I think it would be appropriate to address this here in the methods.

[Author response]

The comment provided in the results section (lines 326-328) serves to point out that we don't capture diurnal cycles in our data because of the limited time period of data acquisition; however this is a result of, rather than a reason for, limiting data acquisition to the period from 12:00-13:00. MODIS overpass times are all from approximately 11:00-13:30. We further narrow our data acquisition to the hours for which the viewing angle of our site is < 30°, which are from 12:00-13:00.

Updated text to read: Temperature differences between the Divide AWS and ice core site are summarized in Table 2. MODIS LST data were obtained for the period 2000-2020 (https://lpdaacsvc.cr.usgs.gov/appeears/) for dates with minimal cloud cover between the hours of 12:00 and 13:00 (local solar time), when viewing angle is less than 30°, to mitigate the effect of viewing angle on temperature and emissivity. At Divide, 742 MODIS images spanning 2002-2020 were analyzed. Seasonally, 203 images were acquired in spring (MAM), 169 in summer (JJA), 188 in fall (SON), and 182 in winter (DJF). The average time between scenes at Divide was 9 days after filtering. At Eclipse, 100 MODIS images were analyzed: 87 spanning June 2005 through June 2007 and 13 spanning November 2016 through February 1017. Each MODIS image was paired with the closest hourly measurement available in the AWS data.

Lines 145-147: Just to clarify, the 742 images at Divide span 20 years of data, and the 100 images at Eclipse span ~2 years of data. Is that correct? Please state that clearly here. Also, please specify how the MODIS images and AWS data are paired in time. Is it simply the closest hourly measurement that is paired with the MODIS temperature, or is it some kind of average of multiple AWS measurements?

[Author response]

Addressed above

Section 2.3: Please clarify if only ASTER and MODIS images from the 11-1:30 window that had paired AWS data are used in this comparison.

[Author response]

Updated text to read: To test if the LST offset is a result of the MODIS sensor's large footprint, we calculate the offset of both ASTER (90 m footprint) and MODIS (1 km footprint) surface temperatures from AWS measurements and then compare the magnitude of the offsets. We use only ASTER and MODIS images from the 12:00-13:00 window that had paired AWS data in this comparison.

Lines 168-169: Please explain how Landsat images are "examined for cloud cover." Is there a particular algorithm used, and are particular thresholds implemented?

[Author response]

Landsat imagery was visually examined for cloud cover. Only images with no visibly identifiable clouds were analyzed.

Updated text to read: Landsat top of atmosphere brightness temperature imagery (https://earthexplorer.usgs.gov/) was visually examined for cloud cover, and cloud-free pixels were extracted for analysis using QGIS.

Equations 2+3: I believe it might be appropriate to use ≈ instead of = in these expressions.

[Author response]

Updated text to read:

$$E_N \approx E_S\downarrow + E_S\uparrow + E_L\downarrow + E_L\uparrow$$

$$T_S \approx \left(\frac{E_L\downarrow + E_S\downarrow(1-\alpha)}{\epsilon_S\sigma}\right)^{0.25}$$

Line 200: Provide a bit more information about the atmospheric emissivity from ERA5.

[Author response]

Updated text to read: We calculate downward longwave radiation as follows, using 2 m air temperature (T_a) from Divide and atmospheric emissivity (ε_a) derived from the ERA5 reanalysis longwave radiation product. We use only the derived emissivity from the ERA5 product, rather than the total downward radiation in order to use measured values (in situ 2 m air temperature) where possible. ERA5 outputs have a spatial resolution of 31 km; data span 2002-2019 every six hours (Hersbach et al., 2020). Atmospheric emissivity increases with increasing surface vapor pressure (Staley and Jurica, 1971). Our atmospheric emissivity values ranged from ~0.48 to 1. Atmospheric emissivity measured over the Sierra Nevada (Spain) from 2005-2011 ranged from ~0.4-1 (Herrero and Polo, 2012).

Line 202: Please provide a bit of description of the surface types that are present at the two sites. Is surface melt common? This doesn't necessarily need to be right here, but the point about albedo raised the question of if this is a reasonable assumption or not.

[Author response]

Surface melt is present but limited at both sites, which are situated in the accumulation zone. Surface melt at these sites does not result in standing surface water, but rather saturates or percolates below the surface, limiting its effect on surface albedo. Observed early melt season (May/June) surface conditions were a fairly soft and flat snow surface with no sastrugi, drifting, or other surface features.

Section 3.1 (and elsewhere): When comparing different seasons, I think that the results of the Wilcoxon rank sum tests can be referenced in the tables and don't need to be repeated in each paragraph. The many parentheticals make it a bit challenging to read and the statistics are clearly presented in the tables. Maybe the journal or author preferences dictate that it should also be in each paragraph, but I think it interferes a bit with clarity.

[Author response]

We agree that the parentheticals interfere with clarity and will remove them.

Figure 4: 1:1 plots of in-situ air temperature vs. MODIS LST are very instructive, and should potentially be included in addition to these box plots.

[Author response]

We will include 1:1 plots of in-situ air temperature vs. MODIS LST

Lines 247-248: Please indicate if it was tested that the data are normally distributed prior to linear regression, and normalize data if necessary.

[Author response]

Edit line 174: We compare differences between AWS and MODIS LST data to wind speed and solar radiation data obtained from the Divide AWS. We transform LST offsets, wind speed, and solar radiation data to approximately normal distributions using a box-cox transformation and normalize each dataset around zero. We then perform linear regressions on LST offsets vs. wind speed, LST offsets vs. solar radiation, and LST offsets vs. wind speed under low (<400 Wm^{-2}) levels of solar radiation.

Regressions on normalized data:

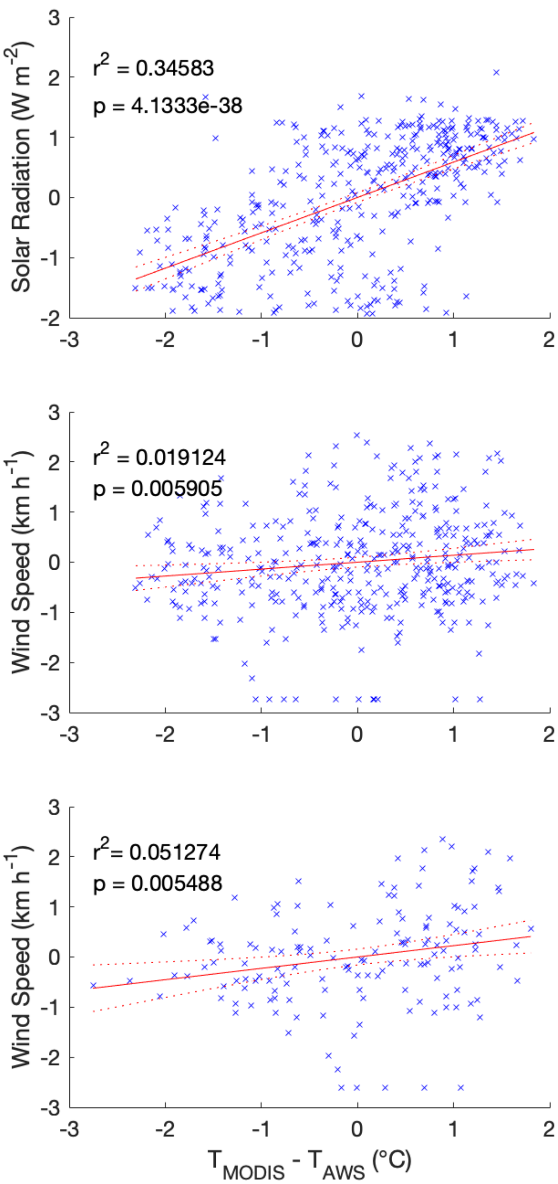

Figure 8: Adding a horizontal line at 0 C would be helpful in figures a and b; Is this data for all 20 years when Divide AWS data are available?

[Author response]

This is data through 2019, as we did not have ERA5 data for 2020. We will add a horizontal line at 0°C.

Figure 9: Adding a horizontal line at 0 C would be helpful in figures a and b

[Author response]

We will add a horizontal line at 0°C

Lines 287-290: Is it possible that the lack of difference between offsets in MODIS brightness and surface temperatures in the fall and winter is also because of increased inversions relative to other seasons?

[Author response]

Yes! We chose to focus in this section (4.2) on the implications for emissivity and keep the discussion of inversions contained in section 4.3 for clarity. We will add an explicit statement commenting on the lack of difference in MODIS brightness and surface temperature offsets in fall and winter in relation to inversions either here or after our discussion of inversions below.

Lines 294-295: I would potentially expect a wider variability in surface conditions (and thus emissivity) in the summer, but this may be site specific. Perhaps explain more why winter could result in more rapidly changing emissivity.

[Author response]

Updated text to read: Emissivity increases with surface melt, and decreases with increasing particle size and density, which can occur due to either packing or sintering of grains as the snow surface evolves following a snowfall event (Salisbury et al., 1994). At Divide, summertime emissivity changes are likely dominated by alteration of the surface snow by melt, while wintertime emissivity changes are likely dominated by snow surface evolution following snowfall events, which occur more frequently in the winter. The relative magnitude of summer and winter emissivity changes are unknown and may result in the seasonal difference in outcome of the LST algorithm.

Lines 304-305: The analyses exploring links between accumulation and the LST offset seem to have covered a wide array of options. Was "days since last accumulation" one of the ways this data was analyzed? I couldn't quite tell from the descriptions and that might be a reasonable metric to consider.

[Author response]

Yes, it was. We've rephrased the statement below to mention this clearly.

Updated text to read: However, given the low temporal resolution of the MODIS data relative to the accumulation record (1 image per day vs. 1 sample per hour), we found no relationship either between the LST offset and individual snowfall events or between the LST offset and the total accumulation each month, the percent of days with accumulation each month, or the mean days between accumulation each month. We also found no relationship between the LST offset and days since last accumulation.

Lines 314-315: Why used different emissivity values for the different sites, or is this a typo?

[Author response]

This is a typo. The different emissivities used are end-member snow emissivities both at Divide, not at Divide and Eclipse.

Line 321: Is it a discrepancy because it is not accounted for given the assumptions used in the T_surface calculations presented here?

[Author response]

Updated text to read: The smaller magnitude of surface-air temperature offsets at Summit, Greenland and the South Pole relative to our study sites may be due to a stronger influence of turbulent fluxes at Summit and the South Pole or to variations in albedo, as both turbulent fluxes and surface albedo can be strong controls on surface energy balance (Braithwaite and Olesen, 1990; Oerlemans, 1991; Ebrahimi and

Marshall, 2016).

Lines 323-338: Consider also comparing to 2m near surface inversions investigated in Nielsen-Englyst et al. (2019).

[Author response]

Updated text to read: In comparing the magnitude of the summer LST offset here (JJA Mdn = 0.98°C), to prior studies, the offsets presented here are smaller than previously observed summer MODIS LST offsets in the St. Elias (5–7°C, Williamson et al. 2017). However, these prior LSTs were daily averages of maximum and minimum values, with most of the offset being attributed to the inclusion of minimum LSTs (Williamson et al., 2017). In contrast, this study uses a single daily LST value and coincident AWS measurements acquired between 11:00 a.m. and 1:30 p.m, when surface and air temperatures are near their maximum, thereby eliminating the effects of any diurnal cycle on observed LST offsets. Our modeled temperature inversions show a diurnal cycle, which is more dramatic in the summer than the winter because of the greater difference between incoming solar radiation during the day and night, and is likely responsible for the higher magnitude of the previously observed summer LST offsets (Fig. 9; Tables 6, 7). A comparison of surface and air temperatures measured in situ over 29 Arctic sites also shows a diurnal cycle that is most pronounced in spring and absent during the winter and polar night (Nielsen-Englyst et al., 2019). The magnitude of the summer LST offset at Eclipse and Divide is in close agreement with temperature inversions observed at Summit, Greenland, where 2 m air and surface temperatures have been contemporaneously measured in situ. During June–July 2015, Summit surface temperatures were 0.32 to 2.4°C lower than 2 m air temperatures (Adolph et al., 2018). At three northern Alaska sites, summer clear-sky surface temperatures were higher than corresponding 2 m air temperatures (Barrow and Atqasuk in 2010, and Olitok Point in 2014; (Good, 2016)). In contrast to sites in Greenland and the St. Elias, these northern Alaskan sites are characterized by seasonal snow cover. Sites with seasonal snow cover present challenges for interpretation because they experience surface melt and a drastic change in surface type over the course of the melt season. Across glaciated areas, sites in the accumulation zone have been found to have the weakest near-surface inversions during the summer, while sites in the ablation zone have been found to have the strongest near-surface inversions during the summer, likely because of the change in surface type with over the melt season (Nielsen-Englyst et al., 2019).

Line 342-343: Perhaps I'm missing something, but why wouldn't winter surface temperatures also be high then?

[Author response]

Updated text to read: The dip in modeled summer temperature inversions (Fig. 8) is the result of our 0°C surface temperature cap, which is a simplistic numerical correction for unrealistically high summer surface temperatures above the melting point. Because the 0°C cap is applied after the calculation of surface temperatures and does not address the mechanisms of inversion development, the distinction between capped temperatures slightly over 0°C and uncapped temperatures slightly below 0°C is somewhat arbitrary. We therefore focus on the magnitudes and seasonal patterns of calculated inversions during summer and winter rather than during the shoulder seasons where the temperature cap likely biases our results.

Lines 370-375: Please explain in a bit more detail how the data are processed before going into the linear regression. I suggest framing the linear fit with the actual variables desired (instead of generic x and y). The corrected MODIS temperatures have a low mean error because of how they are calculated, but how does the standard deviation compare in magnitude to the interannual variability? Is a better fit to the data achieved if it is not averaged annually but instead each paired data point (of 2 m air temp and MODIS

LST) is part of the regression?

[Author response]

We chose not to fit the data using each paired point because our specific goal here was to see if MODIS LSTs could be used to help interpret paleo records on annual timescales.

To be inserted in methods:

2.6 MODIS LSTs and melt

To evaluate whether MODIS LSTs can be used in conjunction with in situ air temperatures to examine the conditions associated with surface melt, we compare interannual trends between the two and reconcile the difference between MODIS LSTs and AWS temperatures using a simple linear regression. We are interested in interannual trends since interpretation of paleo records often occurs on interannual timescales. We therefore calculate the mean annual value for both MODIS LSTs and AWS temperatures. Because we calculate annual means rather than examine individual MODIS LST and AWS temperature pairs, we use all available MODIS LSTs and AWS temperatures, rather than only the subset of dates for which we have both. We fit a linear model to mean annual MODIS LSTs and AWS temperatures using the MATLAB function fitlm(), taking the AWS temperature to be the response variable. We then use the coefficients from this linear fit model to generate a set of converted mean annual MODIS LSTs ($LST_{converted} = -3.35 + 0.49LST$). The RMSE of our linear fit is 1.88°C, and the interannual variability spans a range of 5.33°C.

Minor Comments:

Line 42: Consider starting a new paragraph with the sentence that begins "Instrumentation…"

[Author response]

Alternatively, it may make sense to start a new paragraph with the sentence that begins "MODIS LSTs are a valuable tool…". Either way, we agree that breaking this into separate paragraphs improves organization.

Line 68: Consider phrasing as "apparent MODIS LST offset" here and elsewhere?

[Author response]

Alternatively, it may improve clarity to use the term "MODIS LST offset" throughout with an explicit statement that surface and air temperatures are different quantities. "Apparent MODIS LST offset" implies that the offset isn't real. If interpreted to mean that MODIS LSTs are not in fact biased, this is true, because surface and air temperatures are different quantities. However, if interpreted to mean that MODIS LSTs and air temperatures should be the same, this is false. Regardless of which phrasing is used, consistency throughout is important for clarity, and that we will revise.

Line 86: Consider adding a definition of a nunatak

[Author response]

Updated text to read: (note the many ridges and nunataks, or exposed areas of rock, shown in Fig. 2),

Lines 88-89: Consider reversing sentence that begins with "Here…" so that it discusses inversions first, which were just being discussed. I found the sentence a bit confusing on the first read because I was unsure which hypotheses were being referenced.

[Author response]
Updated text to read: Here, we evaluate the plausibility of near-surface temperature inversions in the St. Elias and test alternative hypotheses to explain the offset in MODIS LSTs in the region.

Line 221 (and elsewhere): forgotten degrees C units on reported median.
[Author response]
Updated text to read: Mdn = -2:90°C (also edited elsewhere)

Line 222: "MODIS temperature difference data spaces >10 degrees C" – is this saying the maximum offset is greater than 10 degrees C? Please clarify.
[Author response]
No, this is referring to the variability of MODIS offsets in each season.

Updated text to read: In all seasons, observed MODIS offsets vary by more than 10°C, with the range of winter values being greatest at 35.56°C at Divide and 25.13°C at Eclipse.

Line 249: Perhaps an "overwhelming majority" or some other language would be better suited to describe 95%.
[Author response]
Updated text to read: An overwhelming majority (95%) of MODIS LST offsets…

Line 256: "modeled temperature difference data is greater than 60 degrees C" – does this mean that the modeled inversions are sometimes up to 60 degrees? Please clarify.
[Author response]
Updated text to read: In the winter, modeled inversion strength varies by up to 60°C.

Lines 275-280: Since the manuscript already revealed that the brightness temp and surface temp do show the same pattern, going through the hypothetical scenarios in the way they are framed confused me a bit. Consider rephrasing to make it clear that one of them we already know to be what was observed.
[Author response]
Updated text to read: We find similar seasonal distributions of offset from AWS temperatures in MODIS LSTs and MODIS brightness temperatures, suggesting that the preferential fall and winter offset is not introduced by the conversion from brightness temperature to surface temperature or the emissivity values used in this conversion (Fig. 5). Moreover, Landsat brightness temperatures also show a pattern of greater offset from AWS temperatures in the fall and winter. The observed apparent cold bias in MODIS LSTs is therefore not unique to the MYD21 product or even the MODIS sensor. Unfortunately, due to the limited availability of ASTER data, too few images exist to examine any seasonal pattern.

Citations:

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. Version 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

Nielsen-Englyst, P., Høyer, J. L., Madsen, K. S., Tonboe, R., Dybkjær, G., & Alerskans, E. (2019). In situ observed relationships between snow and ice surface skin temperatures and 2 m air

temperatures in the Arctic. The Cryosphere, 13(3), 1005-1024.

References:

Adolph, Alden C., Mary R. Albert, and Dorothy K. Hall. 2018. "Near-Surface Temperature Inversion during Summer at Summit, Greenland, and Its Relation to MODIS-Derived Surface Temperatures." *The Cryosphere* 12: 907–20. https://doi.org/10.5194/tc-12-907-2018.

Braithwaite, Roger J., and Ole ? Olesen. 1990. "A Simple Energy-Balance Model to Calculate Ice Ablation at the Margin of the Greenland Ice Sheet." *Journal of Glaciology* 36 (123): 222–28. https://doi.org/10.3189/S0022143000009473.

Ebrahimi, Samaneh, and Shawn J. Marshall. 2016. "Surface Energy Balance Sensitivity to Meteorological Variability on Haig Glacier, Canadian Rocky Mountains." *Cryosphere* 10 (6): 2799–2819. https://doi.org/10.5194/TC-10-2799-2016.

Good, Elizabeth Jane. 2016. "An in Situ-Based Analysis of the Relationship between Land Surface Skin and Screen-Level Air Temperatures." *Journal of Geophysical Research: Atmospheres* 121 (15): 8801–19. https://doi.org/10.1002/2016JD025318.

Herrero, J., and M. J. Polo. 2012. "Hydrology and Earth System Sciences Parameterization of Atmospheric Longwave Emissivity in a Mountainous Site for All Sky Conditions." *Hydrol. Earth Syst. Sci* 16: 3139–47. https://doi.org/10.5194/hess-16-3139-2012.

Hersbach, Hans, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, et al. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146 (730): 1999–2049. https://doi.org/10.1002/QJ.3803.

Nielsen-Englyst, Pia, Jacob L. Høyer, Kristine S. Madsen, Rasmus Tonboe, Gorm Dybkjær, and Emy Alerskans. 2019. "In Situ Observed Relationships between Snow and Ice Surface Skin Temperatures and 2 m Air Temperatures in the Arctic." *Cryosphere* 13 (3): 1005–24. https://doi.org/10.5194/TC-13-1005-2019.

Oerlemans, J. 2016. "The Mass Balance of the Greenland Ice Sheet: Sensitivity to Climate Change as Revealed by Energy-Balance Modelling:" *Http://Dx.Doi.Org.Wv-o-Ursus-Proxy02.Ursus.Maine.Edu/10.1177/095968369100100106* 1 (1): 40–49. https://doi.org/10.1177/095968369100100106.

Salisbury, John W., Dana M. D'Aria, and Andrew Wald. 1994. "Measurements of Thermal Infrared Spectral Reflectance of Frost, Snow, and Ice." *Journal of Geophysical Research: Solid Earth* 99 (B12): 24235–40. https://doi.org/10.1029/94JB00579.

Staley, D. O., and G. M. Jurica. 1972. "Effective Atmospheric Emissivity under Clear Skies." *Journal of Applied Meteorology and Climatology* 11 (2): 349–56.

Williamson, Scott N., David S. Hik, John A. Gamon, Alexander H. Jarosch, Faron S. Anslow, Garry K. C. Clarke, and T. Scott Rupp. 2017. "Spring and Summer Monthly MODIS LST Is Inherently Biased Compared to Air Temperature in Snow Covered Sub-Arctic Mountains." *Remote Sensing of Environment* 189 (February): 14–24. https://doi.org/10.1016/J.RSE.2016.11.009.