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

Overview

The manuscript describes a thorough comparison between in-situ 2m air temperature data and a MODIS LST product for two sites in the St Elias mountains. The authors use additional data from ASTER, Landsat, ERA-5 and AWSs to further support their conclusions.

The main aim of the study is to determine the cause of the difference between 2-m air temperature measurements and the MODIS LST. Three possible causes are explored: the large MODIS footprint, errors in the surface emissivity, and near-surface temperature inversion. The study finds that the latter is the most likely cause of the temperature offset. The manuscript gives insight into how well MODIS LST represents the actual surface conditions in the remote St Elias mountains, which is important for future monitoring.

The manuscript is generally well-written and the presented results are interesting. I have two areas of concern, but most of my comments are minor:

1. I think you should be a bit clearer about the fact that you can't directly compare the 2m air temperature and the land surface temperature. In some sections, you talk about "correcting" the MODIS LST (e.g. Figure 10) – it is not necessarily that the MODIS data is biased, it is just because you are comparing two different things. For the same reason, I would also be careful calling it a "bias" in the title.

[Author response]
We agree that this language is confusing and will remove the usage of "bias" and "correct" in reference to the MODIS offset and conversion to air temperatures. We will also 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.

2. You mention that both AWS are situated on nunataks, but you don't really go into detail on the effect of this. If they are placed on nunataks, and not on glacier ice, could this not also be causing part of the bias? See also my specific comments for L 136-138 and L 270.

[Author response]
See response to comment on L136-138.

Specific comments

L 4-5: is this referring to previous work over St Elias? Or to the current study?

[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.

L 39: “"brightness temperature", an intermediate temperature product used to produce the final surface temperature.” - I would explain what this is here, not just call it an intermediate product

[Author response]

Updated text to read: Remote sensing temperatures include the final surface temperature product, as well as "brightness temperature", or the temperature of a perfect blackbody emitter under the same conditions.

Table 1: can you add a bit more info about the different data sources here? Resolution (temporal and spatial) and maybe uncertainty.

[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	

L 109: What do you mean with “more influential”? In terms of current sea level rise?

[Author response]

Updated text to read: Importantly, Alaska contains much more glacial ice than the European Alps, and therefore has a larger potential contribution to global sea level rise.

L 111: you do not define “Divide Icefield” as “Divide” until later in the text

[Author response]

Edited L108: The AWS record from Icefield Divide (hereafter referred to as "Divide") is, to our knowledge, the longest such record from a glaciated high alpine area outside the European Alps.

L 125-126: How do you know the datasets are consistent, when the time periods do not overlap? Please clarify.

[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. However, we still included what data we had from Eclipse to supplement the results at Divide.

Updated text to read: We combine the Eclipse AWS and iButton datasets for maximum coverage at the site.

Figure 2: Where is the location of the iButton?

[Author response]

The iButton was located at the same site as the AWS. We will add this to the legend and caption.

Table 2: change "Ice core site – AWS site" to "MODIS Ice Core Site – MODIS AWS site", to clarify it is not in situ observations.

[Author response]

Edited table heading: "MODIS Ice Core Site – MODIS AWS site"

L136-138: I am not sure I follow this. You are comparing two MODIS pixels – one with only ice, and one with ice and a nunatak, to find out the difference in temperature between the nunatak surface and the ice surface? If so, this is interesting, but should be clarified and mentioned in the discussion. In addition, I would guess that the difference between the AWS and the ice covered ground is bigger than found in this comparison, since both MODIS pixels does contain some glaciated area.

[Author response]

Yes, here we are comparing two MODIS pixels – one with only ice, and one with ice and a nunatak. However, our purpose is not really to find out the difference in temperature between the nunatak surface and the ice surface, but rather simply to see if there is a discernible difference. We are interested in the ice surface, so if there is a difference in the MODIS pixel with only ice vs. ice and rock, we want to mitigate the effect of the rock in our subsequent analysis.

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.

Insert 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 that 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.

L 144: why only between 11 and 1:30?
[Author response]
MODIS overpass times are all from approximately 11:00-13:30. We choose to 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, 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.

L 145-147: Do I understand correctly, that the 700+ images at Divide span 20 years, and the 100 images at Eclipse span ~2 years of data?
[Author response]
See edited statement for L144 above

L 199: How do you get the downward radiation for Eclipse/iButton?
[Author response]
We only apply the simple energy balance model at Divide because of the lack of data at Eclipse. L339-0340 contains a typo including Eclipse and is edited below:

Results from the simple energy balance model predict no summertime inversion at all, with surface temperatures being a median of 0.77°C higher than 2m air temperatures using an emissivity value of ε_S = 0.95 and 0.75°C higher than 2m air temperatures using an emissivity value of ε_S = 0.99.

L 200: can you provide a bit more info about the ERA-5 product you use?
[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).

Page 10, 11 and others: consider your number of significant digits – are your results really that accurate? I would stick to 1-2 significant digits. Also in e.g. L 255: if it is a simple model, it probably does not have an accuracy of 3 significant digits.

[Author response]

Significant digits will be reduced to 2 throughout

Figure 6: this is at Divide?

[Author response]

Yes, this is at Divide

Edited figure title: Comparison of the MODIS LST offset (MODIS-AWS) with measured solar radiation and wind speed at Divide.

L 270: Why do you compare MODIS, Landsat and Aster over the ice core location and not the AWS location (if I understand figure 2 correctly). If you are investigating the cause of the difference in AWS and measured LST, it would make more sense to look at the AWS location – especially since the AWS is on a nunatak, you would be able to better investigate the effect of this.

[Author response]

We compare MODIS, Landsat and ASTER over the ice core location because our motivation is to use remote sensing methods to obtain temperature records over large ice-covered areas and we therefore do not want to focus on just the nunatak. See response to comment on L136-138. We initially tested the footprint and emissivity hypotheses using all three sensors. Although we used AWS temperatures to calculate the offset of each remote sensing product, these tests essentially involved comparing remote sensing products against one another. When we began to address the possibility of a near-surface temperature inversion, we then incorporated MODIS LST data from the AWS site because it is a more direct comparison between surface and air temperatures at the same site. We did not feel it necessary to re-test the footprint and emissivity hypotheses with remote sensing data over the AWS site because: a) the temperatures between the sites are very similar and produce the same seasonal distribution of MODIS LST offsets, and b) whether we use the ice core site or the AWS site to calculate offsets for each sensor, it is the relationship between sensors that is pertinent for the footprint and emissivity hypotheses, and as we had already eliminated those hypotheses, we do not investigate them in further detail as we do near-surface temperature inversions.

L 302: How is the DIVIDE snowfall record measured? Maybe give some information about this in the data section.

[Author response]

To be inserted in section 2.1 Study sites and in situ data:

The Divide snow accumulation record was obtained using a Campbell Scientific SR50 ultrasonic snow depth sounder instrument. The instrument provided twice-daily readings of its distance from the snow surface at the Icefield Discovery Camp during the period spanning 2003-2012, corrected for the variability in speed of sound with air temperature.

L 314-315: Why are you using different emissivities for the two sites?

[Author response]

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

Figure 10: What happened in 2020? The AWS temperature is much lower than the MODIS temperature.

[Author response]

Our data for 2020 were incomplete at the time of acquisition, only running through June, so the warmest months were not included. For clarity, 2020 will be omitted.

References:

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.

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