Thank you for your very thorough review of our paper. 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].

**This manuscript should be rejected for publication.**

**General comments:**

*The major finding seems to be the difference in air temperature and glacier temperature is the result of
an inversion. Surface temperatures of snow and ice are usually colder than air temperatures. This is
particularly true during the melt season and is a pretty well established fact. I am unsure what this
manuscript offers that is not already present in the literature.*

*The aims and goals should be refined and better described in the manuscript. The organisation of the
manuscript requires substantial revision.*

*No lapse rates are reported, which is the typical way to identify an inversion.*

*Editing for clarity is required. Details need to be added to the many vague statements in the
manuscript.*

*Typically these sorts of studies use orders of magnitude more data than what appears here. An
argument needs to be presented that the small data set is adequate. Static (or literature) values need to
be replaced with measurements where possible.*

*The placement of figures in the narrative is disjointed and not logical. Many results are being
presented in the discussion.*

*Many references concerning the relationship between energy balance and glacier mass balance are
missing.*

*There is substantially more meteorological data available than what has been used in this analysis.*

[Author response]

The contribution of our work is providing a comparison of air and surface temperatures in an
understudied region containing a large amount of glaciated area. Although it is known that surface
temperatures of snow and ice are often colder than air temperatures, there are few studies that examine
this phenomenon in alpine terrain or address its relevance to monitoring alpine glaciers and interpreting
alpine paleoclimate records. Additionally, our data suggest that near-surface inversions are more
pronounced in the winter, not during the melt season, and thus warrant further examination. As the
importance of regional variability becomes ever clearer from a glacier mass balance perspective, the
limitations of current temperature monitoring in the St. Elias mountains require attention. Our focus here
is on MODIS LSTs as a tool to understand a vulnerable and rapidly changing region, rather than on the St.
Elias mountains as a case study to examine MODIS LSTs and air temperatures. We have framed our goal
to address three specific hypotheses (footprint, emissivity, inversion) about the source of the offset
between MODIS LSTs and in situ air temperatures in the St. Elias mountains. Each hypothesis is
addressed in its own set of methods/results/discussion subsections. We have reorganized portions of the
manuscript for clarity. Specifically, we have ensured that each subsection in the results and discussion is
preceded by a corresponding subsection in methods, and rearranged the figures accordingly. We will add
statements making the link between our three hypotheses and monitoring temperature/interpreting paleoclimate records more explicit, as this appears to be the most unclear aspect of our goal and motivations based on comments above.

Regarding clarity and vague statements, please see our responses to comments below, specifically, L21, L29, L96, L130, L367. Please also see responses to comments below for added references. We’ve used measurements wherever possible; measurements are extremely limited in this region, which is a large motivation for the study. Regarding lapse rates, although lapse rates are typically used to report large-scale inversions, the same is not true for near-surface inversions, which is what we are specifically dealing with here. We will add a statement clarifying that our use of the term “inversion” specifically refers to near-surface inversions throughout the paper.

We use the longest meteorological record ever compiled for the St. Elias for our analysis, and indeed one of the longest such records when considering alpine regions globally (please see response to comment on Line 108 below). Williamson et al. (2020) summarize the meteorological stations in the St. Elias region. Out of 16 stations, 7 are located on glaciated terrain. The others are on talus, alpine tundra, grass or gravel. Six out of the 7 stations on glaciated terrain are located above 2500 m. None of the non-glaciated station sites are above 2500 m. Four of the 6 stations on glaciated terrain above 2500 m have records less than or equal to two years in length. The remaining two are the stations we use at Divide, giving a combined record of 20 years. Although the Eclipse record is too short to produce a robust dataset on its own, we include data from Eclipse in our analysis because of its proximity to Divide. We do not include the other three (Ogilvie Glacier, King Col, and Prospector-Russel Col) because their records are likewise too short to produce robust datasets.

**Specific comments:**

**Abstract:**

**MODIS LST can also be sparse or absent**

[Author response]

When we say that in situ measurements are sparse or absent in the St. Elias, we mean that there are very few measurements distributed across a large area. Although a dataset of MODIS LSTs may not have very good temporal resolution (since the surface is often obscured by cloud cover), they do provide measurements in space that realistically cannot be obtained using in situ methods because of the challenges associated with installing and maintaining equipment in the remote mountainous environment.

**MODIS LSTs are offset… which each LST measurement, average LST, minimum, maximum…**

[Author response]

This is referring to each individual LST measurement as compared to the closest hourly weather station measurement.

**Footprint usually refers to swath width, or some derivative. Is it the grid cell size you are referring to?**

[Author response]

Yes, we are referring to the grid cell size.

**Snow emissivity is >0.8 and can be close to being a blackbody, so it is intuitive that brightness temperature would also contain bias.**

[Author response]
It may be intuitive, but we consider this important to explicitly address and report since the intuitive answer to a question may not always be the right one.

**Line 21:** …with far reaching impacts. This is the kind of statement that the manuscript is peppered with and is virtually meaningless: please revise, here and throughout.

[Author response]
Updated text to read: In recent decades, the high latitudes (>60°) have warmed at a more rapid rate than the rest of the planet, with impacts extending to distant lower latitude regions (Winton, 2006; Serreze and Barry, 2011; You et al., 2021).

We have edited the statement to be more specific in referring to impacts in lower latitude regions, as the point of this statement is that high latitude temperature changes are relevant to processes across other parts of the globe. We believe this is pertinent context for studying temperature changes in the high latitudes.

**Line 23:** reduced the Earth’s albedo, further accelerating warming. Please provide credible references for this statement. Most studies do exactly what you are doing which is confusing correlation and causation. Perhaps as the temperature increases more snow is melted, and the newly exposed area provides a negligible amount of atmospheric warming. For context read: [https://www.nature.com/articles/s41598-018-27348-7](https://www.nature.com/articles/s41598-018-27348-7). Alternatively, the reduction in snow and ice causes a warming, but the amount of increase in temperature cannot be disentangled from warm air advection. Alternatively, the snow albedo feedback melts glaciers pretty efficiently.

[Author response]
Updated text to read: In particular, the loss of Arctic glaciers has reduced the Earth’s albedo (which can further accelerate warming) and contributed to global sea level rise (Budyko, 1969; Lian and Cess, 1977; Serreze and Barry, 2011; Zemp et al., 2019; Hugonnet et al., 2021).

**Line 23-24:** As written this statement is not correct. Hugonnet et al.(2021) didn’t analyse albedo, nor was it mentioned in Zemp.

[Author response]
See above for added references.

**Line 25:** Some ink should probably be spilled on your geographical definition of the Arctic. From a climatology point of view (i.e., Arctic Amplification) Arctic is defined as north of the Arctic Circle.

[Author response]
To avoid misunderstandings with the use of “Arctic”, we will change our language to “high-latitude”, and define this explicitly as above 60°N.

**Line 29:** I don’t know how many crucial and criticals I have seen to this point. The writing will pack more punch if these types of words are used substantially less often.

[Author response]
Updated text to read:

Line 28: Additionally, Alaskan glaciers are losing mass at some of the highest rates globally (~66.7 Gt yr⁻¹), and therefore remain pertinent to projections of global sea level rise (Hugonnet et al., 2021). The greater North Pacific cordillera contains over 40 mm of global sea level rise potential in a combination of large icefields and small alpine glaciers, making widespread monitoring of glacier mass in the region a worthwhile endeavor (Farinotti et al., 2019).
Remote sensing temperature products are especially useful for relating glacier behavior and mass balance to climatological changes in rugged alpine regions where glaciers tend to be at higher elevations than most nearby weather stations.

We recommend continued work to understand near-surface thermal processes in these complex regions, including obtaining in situ air and surface temperatures to validate these results.

Despite some uncertainty about the exact mechanism for the MODIS offset, and the lack of an accurate physical correction, MODIS LSTs can still shed light on the important question of surface melt and mass balance in the North Pacific…

This work provides a critical step forward in using remote sensing imagery to expand in situ records and thus provide insight into past and present temperature changes in the St. Elias Mountains and broader North Pacific region.

**Line 32: controlled by atmospheric warming: not necessarily true, these might simply be correlated.**

Updated text to read: Glacier mass changes are driven by changes in the surface energy balance; in effect, glacier mass loss is largely associated with atmospheric warming, which relates to surface energy balance through its influence on downward longward radiation and sensible heat transfer (Cuffey and Paterson 2010).

**Line 33: continued -> projected?**

Updated text to read: In order to better predict the impacts of continued atmospheric warming, we need to monitor temperature change and glacier response.

**Line 34: delete “to be able”**

See above response to comment on Line 33

**Line 38: What does “Remote sensing temperatures include the final surface temperature” mean?**

Updated text to read: A variety of temperature products can be obtained using remote sensing methods, including the final surface temperature product, as well as "brightness temperature", or the temperature of a perfect blackbody emitter under the same conditions.

**Line 44: high temporal resolution and long temporal record; they provide two decades… what resolution, which decades? Always provide dates, rates, numbers, values, colours, weights, dimensions, etc. when describing quantitative subjects.**

Updated text to read: MODIS LSTs are a valuable tool for monitoring climate in remote regions because they provide more than two decades (2000-present) of near-daily imagery under clear-sky conditions.
It is important that low and high elevation sites are in contact with different air masses because this means data from a low elevation site cannot be used to represent a high elevation site and vice versa. The moisture content is not the key piece here and is removed in the edited statement below. Updated text to read: Lower elevation sites are in contact with different air masses and are sensitive to different sources of variability than their high elevation counterparts, so data from these stations are not necessarily representative of climatic behavior at glaciated alpine sites (McConnell, 2019).


[Author response]
Updated text to read: Modeling studies (Chen et al., 2003; Giorgi et al., 1997) predict that warming rates increase with elevation. Although not a universal phenomenon (Ohmura, 2012), elevation-enhanced warming has been observed in a number of alpine mountain ranges including the St. Elias and greater North Pacific cordillera (high elevation sectors of Alaska and parts of the Yukon and British Columbia; Williamson et al., 2020; Diaz et al., 2014; Pepin et al., 2015; Rangwala and Miller, 2012).

Line 71: “surface itself” should be replaced with details like where the photons are being emitted e.g., from the top x nm of the snow and ice, etc.

[Author response]
We use the term “surface itself” because “surface” is the standard term used in the literature and is sufficient information to contrast with the air. Something like the top x nm of the snow and ice would be more important language if our point were to contrast the surface with deeper snow/ice.

Line 75: This paper is relevant here: https://journals.ametsoc.org/view/journals/clim/26/5/jcli-d-12-00250.1.xml. There is probably only a very minor contamination issue.

[Author response]
Updated text to read: Without accurate cloud masking, apparent cold biases in MODIS LSTs have been previously observed at Summit, Greenland in both summer (3°C; Koenig and Hall, 2010) and winter (5°C; Shuman et al., 2014). However, the cloud mask has since been updated to address this problem (Yao et al., 2020). Additionally, previous work in the St. Elias mountains indicates that MODIS LSTs over warm (>0°C) surfaces are an average of <2°C (Williamson et al., 2013).

Line 77: Summit should have Greenland appended to it, here and elsewhere, when referring to the summit of GIS.

[Author response]
We will add “Greenland” to all usages of “Summit”.

Line 80: More detail is required here: There is more forcing that downwelling solar. Air parcel advection plays a role. And why does it have to be balanced- the temperature might be changing? Provide rationale.

[Author response]
Updated text to read: Near-surface temperature inversions occur when the surface is colder than the air directly above it and develop over glaciated regions when heat transfer from the surface to the air occurs as a result of an energy imbalance at the surface-air interface (Adolph et al., 2018). Such energy imbalances can occur under low incoming solar radiation, when upward longward radiation emitted by the earth’s surface may exceed downwelling energy fluxes (Adolph et al., 2018). Snow surfaces often have a high emissivity (0.949-0.997 in the 10.5-12.5 μm range; Hori et al. 2006) relative to the
atmosphere, which has been observed to be as low as ~0.4, depending on water vapor content (Herrero and Polo 2012). This difference in emissivities requires the snow surface to cool relative to the air above as it equilibrates (Hudson and Brandt, 2005).

Line 82: efficient emitter than the atmosphere - implies the atm has a lower emissivity than snow surface. Provide details. Atmospheric emissivity is mainly dependent on water vapour concentration.  
[Author response]  
See edited statement above for line 80

Line 92: pixel is a picture element of a computer screen, where the minimum resolution is set by the screen parameters. Using pixel to describe a remote sensing array element or grid cell is common usage, but not technically correct.  
[Author response]  
We will replace all uses of “pixel” with “grid cell”.

Updated text to read: The heterogeneity of the St. Elias’ environment (surface type, elevation, aspect, incline, wind scouring, shading) may not be well represented by the average temperature value of a MODIS grid cell.

Line 96: How exactly would “disparate changes in emissivity” lead to a bias? Provide details.  
[Author response]  
Updated text to read: Therefore, the icefields undergo disparate changes in emissivity over hours to days, meaning that identifying a single representative emissivity value is challenging. Employing too high an emissivity value in the calculation of LST would result in too low a surface temperature.

[Author response]  
Global
Plotted above are all of the GHCN weather stations located higher than 2500 m a.s.l. with records longer than 20 years (the length of the Divide record). The Antarctic stations are not relevant here, as we focus on alpine areas and not ice sheet plateaus. Many stations in the western U.S. (and likely elsewhere) are likewise irrelevant, as they are not located in glaciated terrain (see station sites plotted atop RGI 6.0 glaciers by region below). Even without removing the Antarctic and non-glaciated locations, the AWS record at Divide is the only high-elevation weather station record of its length in Alaska, as well as northern Canada and the Russian Arctic.

Western Canada/US

Asia

Europe
The AWS record from Divide is, to our knowledge, the longest such record from a glaciated high alpine area in Alaska and the surrounding area.

*Figure 3: Landsat has different sensors (MSS, TM, ETM+, etc.) so either break these up in the figure or identify differences in the text/caption, or both.*

[Author response]
We used LS8 OLI-TIRS, LS7 ETM+, and LS5 TM. We will break these up in the figure/text/caption.

*Line 119: Is air temp. samples on the hour of hourly averages of sub-hour measurements? MODIS LSTs are essentially samples.*

[Author response]
In situ temperatures at Divide were obtained from two adjacent AWS located on small nunataks, the first of which used a Campbell 107F temperature probe (±0.2°C) housed inside a solar radiation shield, which recorded hourly readings from 2002-2015. The second AWS was located ~300 m from the first, and recorded hourly temperatures with a HOBO S-THB-M008 12-bit sensor (±0.21°C) housed inside a solar radiation shield from 2009-present (Fig. 3). Both sensors at Divide were located approximately 2m above the surface. The height of sensors above the surface changed with snow accumulation; however, accumulation on nunataks at Divide is typically limited by intense wind scouring so the sensor height above the surface remains relatively constant over time. Both sensors collected temperature data as hourly averages of 5 minute sampling intervals (Williamson et al., 2020).

**Line 122: Not correct. As snow level changes the Divide sensor’s height above the surface will change. It is possible that it also gets buried in some of the winter months.**

[Author response]
See response to comment on Line 119 above.

**Line 124: “plastic container”? Provide details. Was this vented passively? Exposed to direct sunlight?**

[Author response]
Updated text to read: 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 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 subhourly 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 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.

**Line 125: We combine the Eclipse AWS and iButton datasets… Why? Is this a valid method? Provide sensitivity analysis.**

[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 Line 124.

**Line 126: consistent. - define, preferably statistically.**

[Author response]
See above

**Line 130: “employ an improved method” provide details and why relevant here.**

[Author response]
Updated text to read: The MOD21 and MYD21 (together referred to as MxD21) products dynamically retrieve emissivity values for each grid cell, rather than assigning them based on land cover as was done
for the MxD11 products previously examined (Williamson et al., 2017; McConnell, 2019), and have been shown to correct for MxD11 apparent cold biases over barren, but not glaciated, surfaces (Hulley, 2017; Li et al., 2020; Yao et al., 2020). This change in emissivity assignment is relevant for clarification that the MODIS LSTs we are using are not exactly the same product as those used in prior studies.

**Line 135 (and below): It appears results are provided before methods have been described. It is not clear what is being compared. Is it daily averages of air temperature? Have temperatures (air and MODIS) been temporally matched?**

**[Author response]**

Table 2 has been moved to the results section.

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. MODIS LST data were obtained for the period 2000-2020 (https://lpdaacsvc.cr.usgs.gov/appeears/) for dates with minimal cloud cover and a viewing angles < 30°, to mitigate the effect of viewing angle on temperature and emissivity. At Divide, 742 MODIS images taken between 11:00 a.m. and 1:30 p.m. 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 taken between 11:00 a.m. and 1:30 p.m. were analyzed. Seasonally, 25 images were acquired in spring, 24 in summer, 29 in fall, and 22 in winter. The average time between scenes at Eclipse was ~43 days after filtering. MODIS LSTs were subtracted from the nearest hourly in situ air temperature measurement to calculate their offset from in situ temperatures. A small number of summer MODIS LST offset results were skewed by air temperatures well above 0°C (30 dates with air temperature > 4°C, 5 dates with air temperature > 8°C), as the snow surface cannot warm above freezing without melting. Removing these dates reduced the temporal coverage of the summer MODIS LST offset data, but had no effect on the seasonal distribution of the offset.

Move to results section: MODIS data at the Divide AWS nunatak and adjacent ice core site has a mean temperature difference of 0.27°C and standard deviation of 2.20°C. The difference between the two sites shows greater variability in the fall (std = 140 2.64) and winter (std = 2.95) than in the spring (std = 1.44) and summer (std = 0.77), with the ice core site tending to be slightly colder (mean winter temperature difference of -0.80°C). This may be due to the inclusion of the warmer nunatak surface in the MODIS pixel at the AWS site. Temperature differences between the Divide AWS and ice core site are summarized in Table 2.

**Line 141: “This may be due to the inclusion of the warmer nunatak surface” - this is testable by comparing time series from grid cells which contain less (or none) exposed rock.**

**[Author response]**

Yes, in lines 135-142 we are comparing two MODIS grid cells – one with only ice, and one with ice and a nunatak. Our purpose is to see if there is a discernible difference between the two. 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. We find that there is a difference, and the
statement “this may be due to the inclusion of the warmer nunatak surface” is a possible explanation. We reorganize this (see above response to comment on Line 135) to separate our methods and results more clearly.

**Line 144: What is the rationale for using only <30 degrees view angle? Is there a sensitivity analysis or a citation to confirm this?**

[Author response]
The coefficients in the MODIS LST algorithm vary with viewing angle; the RMSE between calculated and modeled surface brightness temperatures (which are used in MODIS LST calculations) increases exponentially with view angle (Hulley et al., 2016). The RMSE is <1°C for viewing angles <60°; we cap our viewing angles at 30° to be conservative.

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 (Hulley et al., 2016). 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.

**Line 145: This temporal subset will sample somewhere below the maximum daily temperature. This also seems to be a very small amount of data from what should be available from a 20 year time series, from two sensors and multiple daily overpasses.**

[Author response]
We chose this temporal subset because that is when we could obtain the most MODIS overpasses with low viewing angles that could be matched with close in situ measurements. We will include a comment that this is sampling below the maximum daily temperature. Unfortunately data acquisition is severely limited by cloud cover, which is why a 20 year timeseries yields such limited data.

**Line 146: “The average time between scenes” describe what this means and why it is important - as written I have no idea what it means.**

[Author response]
Updated text to read: The mean time elapsed between two consecutive analyzed images at Divide was nine days.

**Line 151: Removing these data?**

[Author response]
Updated text to read: Removing these data reduced the temporal coverage of the summer MODIS LST offset data, but had no effect on the seasonal distribution of the apparent bias.

**Line 160: Under development as of when?**

[Author response]
Updated text to read: Landsat surface temperatures remain under development (as of July 2021) and were therefore not included in this study.
**Line 167: TOA Tb is not really a useful metric to compare to surface temperature.**

*Author response*
We included the Landsat TOA brightness temperature because it was the closest available product used in the calculation of surface temperatures. However, if its inclusion is more confusing than useful, we can remove it.

**Line 170-175: Provide bounding values for “low” and “high”.**

*Author response*
Updated text to read: To test whether the MODIS LST offset reflects pervasive near-surface temperature inversions, we examine whether the offset is more pronounced under conditions that facilitate near-surface inversions, namely low levels of incoming solar radiation and low wind speeds. At Summit, Greenland, no inversions greater than 2°C were observed in the 2 m above the snow surface when incoming solar radiation was above 600 W m\(^{-2}\) or wind speed was greater than approximately 7 m s\(^{-1}\) (Adolph et al., 2018). Over 22 sites in Greenland, maximum temperature inversions were observed at wind speeds of 5 m s\(^{-1}\).

**Line 176: “would” -> could.**

*Author response*
Updated text to read: To test if a near-surface temperature inversion could occur under surface conditions at Divide and Eclipse, we compare the MODIS LST offset to the offset from AWS temperatures of surface temperatures calculated using the following simple energy balance model. The net surface energy balance (E\(_N\)) can be expressed by…

**Line 176: What does “physically plausible under surface conditions” mean exactly?**

*Author response*
See edit for line 176 above

**Line 177: “theoretical model of temperature inversions. To” - Provide details and a space after the period.**

*Author response*
See edit for line 176 above

**Line 185: Typically the terms you avoid are small compared to the dominant terms you include. Provide a range of values for all the terms. This will allow the reader to evaluate the effect of removing some terms.**

*Author response*
Updated text to read: We ignore E\(_G\) because it is often small relative to both radiative and turbulent fluxes, and several studies (e.g. Brock and Arnold, 2000; Hock and Noetzli, 1997; Favier et al., 2004) have validated energy models in which it is omitted (Hock and Holmgren, 1996; Pellicciotti et al., 2009, Yang et al., 2021). Subsurface energy fluxes have been found to represent 1-2% of the total heat flux on glacier surfaces (Giesen et al., 2008; Yang et al., 2011; Zhang et al., 2017).

**Line 188: It rained at the summit of Greenland Ice Sheet this year, so probably better to rephrase this sentence.**

*Author response*
Part of the value we see in this study stems from the fact that the St. Elias Icefields are vastly different than Greenland, so our focus here is on observations specifically in the St. Elias.
Author response

Why assume $E_{\text{SN}}=0$, when it will most certainly not be, either seasonally or annually?

We assume $E_{\text{SN}}=0$ because our aim here is simply to get an estimate of whether near-surface inversions are possible at our study site. At a site high in the accumulation zone with relatively limited melt, $E_{\text{SN}}=0$ is an adequate approximation for this purpose. The implication of this assumption is that our modeled surface temperature may be underestimated in the summer and overestimated in the winter, meaning that wintertime inversions may be even more pronounced than our model results suggest.

Line 200: Provide range of values for atmospheric emissivity.

Updated text to read: We calculate downward longwave radiation as follows, using 2 m air temperature ($T_a$) from Divide and atmospheric emissivity ($\varepsilon_a$) 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 200: ERA 5 Land produces a downwelling longwave variable. Why wasn’t this incorporated into the analysis?

We chose not to incorporate the ERA5 downwelling longwave variable because we wanted to use in situ data wherever possible in the analysis.

Equation 3: Provide more information about how this equation was derived. And why use a literature value for albedo? There is considerable variation, spatially and temporally, in albedo. Why not use the coincident MODIS albedo?

Updated text to read: $E_L^{\uparrow}$ is the energy emitted by the earth’s surface and can be described by:

$$E_L^{\uparrow} = \sigma T_S^4$$

where $\varepsilon$ is surface emissivity, $\sigma$ is the Stefan-Boltzmann constant, and $T_S$ is surface temperature (Cuffey and Paterson, 2010). Expressing $E_L^{\uparrow}$ in terms of its components...

We used the mean measured albedo value from Divide in the interest of incorporating as much in situ data as possible into our analyses. Prior work in the St. Elias (Williamson et al., 2016) has demonstrated issues with MODIS albedo values arising from confusion between snow and cloud cover. We therefore avoided using the MODIS albedo product to eliminate this unnecessary source of uncertainty.

Line 203: MODIS provides emissivity values. What are these for the given days sampled in this study? What are the seasonal ranges of snow emissivity?

The emissivity values for the days sampled in this study range from 0.930 to 0.988. In spring (MAM) they range from 0.944 to 0.988. In summer (JJA) they range from 0.930 to 0.986. In fall (SON) they range from 0.942 to 0.988. In winter (DJF) they range from 0.942 to 0.986. Boxplots by season of the MODIS
emissivity values from bands 29, 31 and 32 are shown below. The range of emissivity values is similar in all seasons, so we consider distinguishing by season unnecessary for our simple model. The distribution of emissivity values is also skewed toward higher values, so we consider the 0.95 value from Hori et al. (2006) a reasonable choice for our lower emissivity bound.

**Line 208: Differences between median values? I am unsure what “Median differences” is.**

[Author response]
We first calculated the difference for each paired AWS measurement and MODIS LST and then took the median of these values.

Statement to be added to methods sections (around lines 145-150): We calculate the difference between each paired AWS measurement and MODIS LST to produce a timeseries of the offset between the two. We report the median values of the resultant timeseries for each season below.

Updated text to read: In comparing MODIS LSTs with AWS temperatures at Divide and Eclipse, we find the MODIS LST offset to be greatest during the fall and winter (Table 3). We report a warmer surface as a positive difference and a colder surface as a negative difference. The difference between AWS temperatures and MODIS LSTs at Divide is larger in the fall (Mdn = -4.43°C) and winter (Mdn = -8.40°C) than in the spring…

**Line 209: Which is warmer, surface or air? Not clear.**

[Author response]
See edited statement for line 208 above

**Line 210: Are these distributions normally distributed? There are tests to determine this.**

[Author response]
The distributions are not normal, which is why we used the Wilcoxon rank sum test instead of a two-sided T test.

*I gather that seasonal averages use all of the data from 2000-2020. Are air temperature and surface temperature changing at the same rate? Are inversions weakening over time? Are rates of temperature change similar between seasons? Is there a monotonic trend in emissivity? All of these things will influence your results.*

[Author response]
Yes, seasonal averages use all of the data from 2000-2020. There are no clear seasonal trends in air or surface temperature or their rates of change over this time period. Air and surface temperatures change at similar rates over this period. There is also no clear trend in the LST offset or monotonic trend in emissivity over this period. The most important consideration for our analysis of seasonal differences in the LST offset is whether the relationship between air and surface temperatures has exhibited different changes in each season over the period from 2000-2020, and we do not see this.
Line 223: Temperature has not been measured to the precision being reported.
[Author response]
MODIS temperatures are measured to the 0.01 and temperatures from the Divide AWS are measured to the 0.001.

Line 247: R^2 =0.02 is statistically significant? How big was this data set?
[Author response]
The dataset had 395 pairs. See the response to comment on Figure 7 below for more information. Despite a statistically significant r^2 between the temperature difference and wind speed, we find the correlation weak to the point of being negligible.

Figure 5: I am not sure what the point of this figure is? The two MODIS thermal bands will differentially absorb in the atmosphere, which is the basis for the split window LST algorithm. To work out the atm. emissivity, atm. column water vapour is required.
[Author response]
The purpose of Figure 5 is to illustrate that the MODIS brightness temperatures show the same pattern of offset from in situ measurements (greater in fall and winter) as the LSTs do. Therefore, this seasonal offset doesn’t come from the introduction of emissivity during the LST calculations.

**Figure 6: Are these data temporally matched? It must be sampled data because a daily average of 1000 w/m^2 is not feasible.**

[Author response]
Yes, the data are temporally matched. Wind speeds were taken for the dates/times that we had already calculated the MODIS LST offset from in situ measurements.

**Line 263: “averaging temperature” - means what?**

[Author response]
Updated text to read: The LST offset in ASTER data indicates that MODIS LSTs do not display an offset from AWS temperatures simply because they are mean temperatures over square kilometer grid cells rather than point measurements.

**Line 266: Air temperature scales over 100s of km, so not surprising.**

[Author response]
Perhaps unsurprising, but still worth noting, especially given the variability of the terrain in this area.

**Figure 7: I am skeptical about the magnitude of the p-values reported here. These should be checked.**

[Author response]
We checked the p-values and re-ran the correlations using different methods to transform the dataset to normal (log, exponent, boxcox). The boxcox method provided the best transform and the resulting correlations are shown below.
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.

**Figure 9:** Earlier in the methods you said the time of MODIS capture was between 11AM and 1:30PM. Why the different diurnal time range here? Same issue with Table 6.

[Author response]
We use a wider diurnal time range in Figure 9 and Table 6 to illustrate why our MODIS LST offsets (calculated using individual time-matched data points) is smaller than previously studied LST offsets (calculated using daily averages). Figure 9 and Table 6 deal with inversions calculated from ERA 5 and Divide AWS data, not MODIS data.

**Line 289:** “suggesting that emissivity values during these seasons may contribute to the offset” - how exactly?

[Author response]
See response to comment on line 96 above

If there is a trend in cloud cover change then both downwelling shortwave and longwave radiation will be altered over the course of the study period. This could add a substantial amount of error to the results. This needs to be analysed.

[Author response]
All of our data are from days with cloud-free MODIS imagery, so the impact of a trend in cloud cover change would be a change in the distribution of our sampling dates. Distribution of our sampling dates is shown below by year and overall.
Emissivity values may be especially poorly known under winter conditions because of rapidly changing snow surface characteristics during and following frequent snowfall events, resulting in the seasonal difference in outcome of the LST algorithm as seen at Divide.

Williamson et al. (2020) put inversion level at approximately 1200 masl. The two stations used here are 1000 to 2000 m above this level, and are not situated in valleys where cold air drains and collects.

All uses of “inversion” in this study refer to near-surface inversions within 2 m of the snow surface, rather than large-scale inversions such as the ones mentioned in Williamson et al. (2020). We will insert a statement to clarify our usage of the term.

See response to comment on Line 314 above. Radiosonde and re-analysis data have been used to study large-scale inversions, but we are focused on near-surface energy balance processes, which is why we use a simple energy balance model.
Line 367: “Surface melt is primarily driven by high air temperature” - what is high? And melt is correlated with air temperature. There are many examples in the literature of melt rate being influenced by short and longwave radiation.

[Author response]
Updated text to read: Surface melt correlates with air temperatures, largely because of increased longwave atmospheric radiation (an important source of energy for melt) with higher temperatures (Ohmura, 2001; Cuffey and Paterson, 2010). Higher air temperatures tend to occur under cloudy conditions when no MODIS imagery is available (Walsh and Chapman, 1998). MODIS LSTs may therefore be inadequate for examining temperature conditions associated with individual extreme melt events.

Line 369-370: MODIS albedo correlates very well to glacier mass balance. There are many examples of this to be found in the literature. MODIS can’t measure albedo under cloud cover. I am not sure the statement presented in the manuscript is correct.

[Author response]
Although MODIS albedo may correlate well to glacier mass balance overall, it may not capture individual extreme melt events. See response to comment on Line 367 above.

Figure 10: Corrected is the wrong word. LST and air temperature are not the same thing and should display offsets. These offsets are important for understanding the energy transfer between surface and atmosphere. If the goal is to produce air temperature fields originating from MODIS LST, then ‘converted’ instead of ‘corrected’ might be a better option. Further, there are many examples of methods to convert LST in the literature, most of which do not appear in the manuscript. The AWS data from 2020 is suspiciously cold.

[Author response]
We agree that the use of “corrected” is misleading. We will change all usages of “corrected” to “converted”. 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.

Edited (Line 370): However, interannual trends in MODIS LSTs agree well with those in AWS temperatures ($r^2 = 0.23$ and $p < 0.05$; Fig. 10). Various methods have been used to convert MODIS LSTs to air temperatures, including advanced statistical and modeling frameworks (e.g. Hengl et al., 2012; Benali et al., 2012; Emamifar et al., 2013; Wenbin et al., 2013; Janatian et al., 2017; Zhang et al., 2016; Hooker et al., 2018; Zhang et al., 2018; Zhang et al., 2021) Here we apply a simple linear regression ($y = 3.35 + 0.49x$) to show that the difference between mean annual MODIS LSTs and AWS temperatures (mean error of $0.00 ± 1.77 \degree$C) can be reconciled, effectively converting MODIS LSTs to air temperatures and enabling their use for both qualitative and quantitative applications related to glacier melt and mass balance on annual timescales.

Line 376: Snow and ice melts when its temperature reaches $0\degree$C not when the air temperature above it reaches $0\degree$C. So the rationale here needs to be revisited.

[Author response]
Snow and ice do indeed melt when their temperature reaches $0\degree$C; however, the rate of snow and ice melt does also exhibit a close relationship with air temperatures (Ohmura, 2001). Numerous studies have used this relationship to reconstruct air temperatures using the presence and thickness of melt layers observed in the snowpack or ice core (e.g. Abram et al., 2013; Alley and Anandakrishnan, 1995; Das and Alley, 2008; Fisher et al., 1995; Herron et al., 1981; Winski et al., 2018).

References:


https://doi.org/10.1038/s41561-019-0300-3.


https://doi.org/10.1038/293389a0.


https://doi.org/10.1007/S00704-012-0687-X.

https://doi.org/10.3189/172756409787769555.


https://doi.org/10.1007/s10584-012-0419-3.

https://doi.org/10.1016/j.gloplacha.2011.03.004.


