

The Cryosphere Discuss., referee comment RC1
<https://doi.org/10.5194/tc-2022-113-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on tc-2022-113

Sam Herreid (Referee)

Referee comment on "High-resolution debris-cover mapping using UAV-derived thermal imagery: limits and opportunities" by Deniz Tobias Gök et al., The Cryosphere Discuss., <https://doi.org/10.5194/tc-2022-113-RC1>, 2022

"High-resolution debris cover mapping using UAV-derived thermal imagery: limits and opportunities" by Gök et al. uses thermal data collected during eight repeat UAV flights throughout one day of a portion of Tsijiore-Nouve Glacier in Switzerland to evaluate estimates of debris cover thickness. The authors use two established methods to estimate debris cover, but move research on this subject forward by their thorough consideration of thermal data drift and offset, the continued advancement of using thermal cameras on UAVs and the repetition of flights over the course of one diurnal cycle. The authors do a nice job evaluating the debris thickness estimates against field measurements, however, I think a key deficiency in the results are debris thickness difference maps showing the variability of a quantity, which, if estimated perfectly, should be a grid of zeros. Even integrating to show the change (i.e. error) in total debris volume for this study domain might, for example, help future erosion rate studies looking to use debris thickness changes as a way to quantify bedrock erosion to understand the magnitude of error in these methods. Below are mostly minor, but some more major in line comments.

ABSTRACT

L10: be explicit: e.g. supraglacial debris surface temperature or LST over a debris cover

L11: can you say at x cm resolution rather than high?

L16: Can you make here in the abstract a statement on which method you found to be most successful / efficient if the RMSE is essentially the same for both? This could be formulated as a recommendation for future studies.

L16: This isn't exactly true that diurnal variability in LST controls the relationship between LST and debris thickness. Maybe you mean to say the success of a rational function to express the relationship varies predictably with time of day? Maybe also conditionally with clouds cover and precipitation?

L20 This last sentence is unclear, do you mean independent measurements of LST or a better camera on your UAV? And also a little bit of a downer. Glaciologists can't practically afford a cooled microbolometer and they certainly aren't going on UAVs anytime soon. I would either make a more achievable suggestion or try a more positive closing sentence.

INTRODUCTION

L23: Add citation(s) to first sentence, Scherler et al, 2018: Herreid and Pellicciotti, 2020

L23: Add citation supporting "debris cover is generally rather thin, usually less than a meter"

L24: I would change "profound" to something like "exponentially compounding influence ...where melt rates are accelerated..." or at least " it can have a profound"

L26: by "heavily" do you mean spatially or in thickness?

L29: I think it is established that the advancing you cite here is occurring more because of elevation and precipitation rates, not as a function of debris cover.

L33: inappropriate location for citations

L39: Cite the recent studies

L39: There are practical reasons why there are more studies on extent change rather than thickness change, it's an easier problem with a better control on errors, it's intermediate studies like this one that may slowly help add the z component to change analyses.

L43: The list followed by "vary rapidly" reads poorly, maybe you mean the abundance of

these features can increase or decrease rapidly or the abundance can vary dramatically between different glaciers or different parts of one glacier?

L45-46: what does "distribution" constitute beyond knowledge of "extent and thickness"? Also a continuous model of debris thickness is more likely, not comprehensive observations.

L49: Incorrect citation to Ostrem 1959, this paper established the relation between debris thickness and melt rate.

L49: In (3) explain where the debris thickness comes from, a residual from the methods you describe.

L52: I think you might mean "fitting ... to" rather than "extrapolating ... using"

L57: What do you mean by recent technological advances? Uncooled microbolometers have been around for quite a while. One limitation may be that it's a single sensor imaging a single wavelength window rather than a split window approach some thermal satellite sensors use to solve for and remove atmospheric attenuation.

L60: What are the opportunities and limits? What is the research question? And what is the desired resolution, acquisition frequency and look angle? Some idea of these are needed to motivate this statement.

L60: I think it's more "approaches of acquiring a thermal image" rather than "applications"

L61-62: the physics and methodologies seem the same to me whether the thermal imagery is acquired via oblique imagery or an airborne sensor. I would say what remains elusive is mapping high resolution thermal for entire glaciers or several glaciers. A UAV is a good candidate for achieving this, although still a step beyond the scope of this paper.

L64 and throughout: LST not "LST's"

STUDY AREA

L70: coordinate units

MATERIALS AND METHODS

L102: I understand this, but maybe say point measurements are ambiguous or difficult to make in an unbiased / representative way.

L103: how does this boundary look in your debris thickness estimates? Does this mean there were areas of >30 cm debris, just not sampled?

Section 3.2: I think the authors did a really nice job here illustrating a problem that future glaciologists might encounter and giving a solution to it. Is the spline, rather than a median value shift, capturing spatial variability in drift? I don't think you said yet in the paper how long a flight path took to fly, do we assume Fig. 2 is roughly one snapshot in time? Was this approach applied for each of the flights? Was drift and offset roughly similar for each flight? It seems like bare ice temperature in (c) is still not 0C? I trust that the air temperature during the survey was well above freezing and indicative of bare ice being at the pressure melting point, but it might be worth noting that there's nothing initially alarming about ice being less than 0C. Consider changing the section title to something like thermal image drift and offset correction.

Section 3.3: The thermal images were stitched and then orthorectified or orthorectified and then stitched? From Fig. 1 I'm guessing the number of frames is around 600 for each flight, can you give the number here. You say that you identified the GCPs in the thermal images but there are only 6 GCPs so most thermal images have no reference, this makes me guess that you stitched the thermal images first. How did you manage overlap? Did you take a mean or trim one image? It's clear from the section above you applied some corrections for drift and offset, but still I'm guessing there are residuals and the temperature may be slightly different for some of the overlap just due to temperature change during the gap time between image acquisition paths.

L150: watch section numbering

L166: you are making the assumption that atmospheric attenuation is linear, but still I think I agree that this is a negligible term in this study.

L171: Convective heat transfer was more 'deemed not present' than included

L173, Eq. 2: All fluxes have units of Wm^{-2} so ΔS must as well, how is this a rate of change of heat, where is the time derivative coming from?

L175: What exactly is the ground in this context?

L185: with a larger

L219: stored heat flux

L221: where did d in the second term come from?

L229: Fig. 7 suggests this null area of unphysical SEBM results is nontrivial, this seems like a very notable detraction, a lot of future analyses that will use debris thickness maps will need to be continuous. The lack of a meaningful solution in areas with debris also raises concern about the area where there is a solution.

L230: I think you mean training

RESULTS

L241: How many scenes make up the corner of (b) that is incorrect? The geometry of the erroneously cold area seems odd. Do you think the $>0.5\text{C}$ data is reliable nearby? I'm curious why the bare ice / ice cliff geometry isn't consistent between images? That would be a data quality red flag if stable thermal features aren't captured in a repeatable fashion.

L247: Is $0/360$ degrees north?

L280: Fig. 5. The plot of time of LST_{max} is really clever! But how is it a function of terrain aspect? Wouldn't it be a combination of aspect and the thermal inertia of the debris? Isn't \bar{T}_d just $LST/2$ per definition in this paper $(LST + 0)/2$ (L211)? Does this quantity really have more meaning than LST alone?

L289: Somewhere can you add a histogram of debris thickness please? Fig. 6(e,f) suggest

many of the measurements were 5cm or less. L101 said the mean was 9cm but maybe also state the median? Or perhaps a different drawing order in the fig or different data visualization if there are too many lines to show the whole dataset. I'm just concerned that the methods used here are better suited (meaning more stable for method evaluation) to slightly thicker debris.

L290: methods not results

L305: I don't get this, are you saying G is actually large for thin debris or just a very subtle max over a time series of low values? G might be sensitive to inaccurate LST over thin debris cover because the ice should keep the numerator of G low or near zero, but a thin debris thickness denominator might make G sensitive to LST error. Debris thickness measurements can also be quite sensitive over thin debris covers, where getting the local mean thickness wrong by 1 cm could be a high percentage error, while a 3 cm error over a 50 cm debris cover will propagate less error through these sorts of calculations.

L316: I don't find this sufficient. Reality doesn't return an undefined solution, I think if the model fails and you understand why, you should write a piecewise function that either returns continuous results or at least sets constraints on where the method is applicable.

L313: Fig, 7: If the debris thickness estimates were continuous over the same domain the histograms would be informative, but with data voids they aren't very comparative. I would like to see difference maps, with respect to perhaps one reference time that agreed particularly well with measurements. The main challenge in this problem is taking variable input datasets and returning a constant value. Based on the images alone in Fig. 7 it looks like a general trend is preserved but the pixel to pixel debris thickness will be shown to vary quite a bit unrealistically over hours and I also see what look like artifacts in the data. Likely the same artifacts that are discussed in the text but I would be hesitant to keep them for use in the later analysis of this paper. It's possible the quality of thermal sensor that can fit in a DJI UAV isn't quite high enough to return clean data.

L317-318: I would like to read more quantitative results than "relatively consistent in time"

Line 320: Line 101 states the mean debris thickness is 9 cm so a RMSE of 6 to 8 cm means nearly 100% error for the average case.

L3320-325: Fig, 8. I think this is just about how well debris thickness can be predicted from mostly or exclusively thermal data, and I think one should be careful from trying to see what we all wish it would show. "Correlate well with observations even if they do not follow the 1:1 line" sounds like seeing what you want to see, a successful predictive model really should be clustered around the 1:1 line.

L328: I'm not sure it's so clear about the thin debris, if you change the plot limits to 5 cm I think it might look like random scatter. The distance away from 1:1 might be less, but with respect to the magnitude of the debris thickness there may not be clear evidence of predictive capability. This tight cluster of thin debris cover near the origin likely also weights the RMSE favorably. If you consider only debris greater than 5 cm thick, for example, the RMSE will likely increase.

L335: Are you computing the standard deviation from 7 or 8 measurements? Is that enough?

L347: How did you divide the dataset?

DISCUSSION

L370: comma after pattern

LL371: I think you need to show how well the model results are consistent in time as well as approach the testing dataset to make a statement about method viability.

L405: Fig. 12: x-axis units

L445: this could be the piecewise condition mentioned above

L450-452: I think this might be true with a more advanced methodology, but the results from this study show, what look like, large changes in debris thickness over a day when the quantity should be found as static. Mapping high contrast features like ice cliffs from high resolution thermal data has already been shown to be useful, but studies relating changes in debris thickness to erosion rates or surface properties will need quite high confidence debris maps, since the changes are likely to be on the order of cm per decade.

L460: the high emissivity of rocks limit reflected radiation making the viewing angle less of a factor, but still can be accounted for. The angularity of rock clasts make even a nadir look-angle non-normal to most surfaces.

L462: I would add battery limitations on range as well.

L475: What about a more expensive camera as well? I assume higher end FLIR uncooled microbolometers have a higher accuracy even in less ideal conditions.