Comment on tc-2021-94

Evan Miles (Referee)

The study by Van Tricht et al presents a set of experiments to find the best image processing approach to resolve glacier surface mass balance from high-quality multi-year UAV measurements. The study is generally well implemented and seems to produce an approach to derive reasonable distributed SMB estimates based on the correspondence with stake measurements, suggesting filtering of thickness and velocity gradients, as well as direct flux divergence values, over a multiple-ice-thickness length scale. The study also highlights the importance of high-accuracy ice thickness data.

This is a welcome contribution to the field, as implementations of the continuity approach have come increasingly into vogue over the past few years, and UAV surveys have made high-precision topographic data a routine aspect of glacier monitoring. This study particularly offers the potential of deriving spatially-distributed flux divergence values, rather than based on glacier segments or for a single point. It is especially nice that authors have made use of a great set of field data to evaluate their model well; this is an important and challenging step . The manuscript is nicely written and the work is presented well. I do have a number of comments for the authors, including a few more substantial methodological concerns which the authors should consider, particularly with regards to the choice of filters and assumptions/effects thereof.

General comments:

It is not apparent that the authors have evaluated whether the filtering has significantly impacted the total net volume change (or net flux divergence) – is mass conserved after all the filtering? Essentially, neither the variable box average filter or the exponential decay filter are conservative – they change the local and non-local mean values. I think it is likely that this does not result in a significant difference in the total volume loss or total flux divergence, but the authors should evaluate whether the difference falls within the uncertainty of the underlying data, or whether the filtering has internally broken mass conservation.

I found it strange for all values to be in ice equivalent! This avoids the difficult topic of

density (less difficult in the ablation area) but I think it would be better to make an assumption, present your results in the standard unit, and then briefly mention this problem in the discussion (future work/opportunities/needs).

One of the filtering steps (for the dH/dt data) aims to correct apparent surface changes due to advection of surface features. I would recommend that the authors examine the flow corrections of e.g. Brun et al (2018) which are a more appropriate approach to resolve this problem, but requires switching from an Eulerian to Lagrangian frame, which is anyhow more appropriate for comparing to a stake.

For both velocity and thinning datasets, the authors have discarded data outside the glacier boundary, but they should instead use these data to empirically assess the quality of their results. Similarly, the combination of the ice thickness datasets is a bit awkward; it would make more sense to me to integrate all available bed elevation observations to produce an optimal ice thickness dataset. However, this may not be feasible for multiple reasons, and would require reprocessing nearly all datasets, so I can understand that the authors may wish to avoid that course.

The justification of an exponential decay filter is not particularly strong, and I am concerned that the optimal distance results may be sensitive to the type of filter used. It would be worthwhile to consider other approaches or justify this choice (and its implicit assumption).

Minor comments:

L26-L27. suggest '…several times the local ice thickness, *accomplished in this study* using an exponential decay filter.' There are several ways to consider the larger scale variations in stresses; your study nicely points to the exponential decay filter as one good approach.

L42. I'd suggest spelling out interpolation here and in the conclusions. Not so many letters but improves readability.

L51-53. Suggest citing Berthier et al (2012) or similar studies here. Also a key point is that even if one corrects for ice flow, assumptions regarding ice density are needed to assess glacier mass balance, which are not the same assumptions as for the glacier scale (e.g. Huss, 2013).

L60. There are many other relevant references here, and you can't be exhaustive, but see also Brun et al., (2018) and Wagnon et al (2020).

L64. I'd be more explicit that many studies reduce the problem by using flux gates. I suppose this is a one-dimensional reduction and indeed smoothing, but it's also an entirely different kind of data processing.

L152. Is this the difference in SMB values or the rate?

L171-183. The disagreement of the two thickness datasets is interesting and not terribly surprising. However, I am surprised that the authors have relied on the two modelled datasets (based on distinct field measurements and necessarily extrapolated using physical principles) rather than integrating the underlying field measurements of bed elevations in an updated ice thickness estimate. Surely this would result in a better consensus than the mean of two model results? This would allow you to also integrate your 2020 measurement(s).

L188. These values differ by 18-24%

L255. Did you look at your off-glacier dH measurements (bias and standard deviation) at all? As with velocity, these provide an important measurement of error.

L260. Yes, these undulations are quite annoying from a geodetic perspective (and they mostly cancel one another out) but I find it strange to filter them out, when you could instead flow-correct the DEM (or underlying point cloud) using your velocity data (see Brun et al, 2018). An important test for this filtering is whether volume change is conserved before and after – do the volume-change measurements before and after the filtering equate? They are likely to be very close, but bear in mind that you are eliminating real volume changes; this is a good reason to consider warping your DEM to correct for glacier flow.

L277. Why is 25 in the denominator here? Should this be delta-x?

L294. Please indicate what ImGRAFT settings you used, which correlator, etc, just for replicability.

L303. Are the glacier outlines from a standard source (which?) or did you digitize them yourself?

L303. It's fine to remove them for your filtering, but the off-glacier velocity vectors give you an important metric of velocity uncertainty!

L304. Do you median filter the u and v components or the speed (and direction)?

L357. These studies don't really 'resample' to a lower resolution, they are instead resolving the flux across a gate or set of gates.

L367. The choice of an exponential decay filter is interesting! I'm not fully convinced that an exponential is the best choice but neither is it a bad filter – have you tested or considered other filters for this? Unlike others (Gaussian, Weibull, mean/median), the exponential decay filter considers the closest observations to be the best estimate, giving the local value the highest weight. Do you have a reason to assume this to be the case? For instance, a block mean/median (as used for the dh/dt) would consider all nearby measurements to be a reasonable estimate. I see the exponential decay for a perturbation of ice thickness, but this is really the inverse process, no? I mean, the noise in the flux divergence map (which should be shown somewhere, by the way!) is in part due to ice thickness/velocity errors and in part due to the far-field longitudinal stresses. Did you test or consider other filtering approaches? For instance a Gaussian filter exhibits a similar distance decay, but considers any of the nearby estimates to be reasonable and weights them more or less equally.

L400. A concern here is that this filtering can internally 'break' conservation of mass. That is to say that the net flux divergence over your survey zone should equate to the flux into your survey zone; this should be the case with your gridded flux divergences, but after this filtering (which is not mass-conserving!) the net flux divergence (the integral of all pixel flux divergences) can change. Have you checked this? It is likely that the difference is within the uncertainty bounds of the flux into your lower domain, but this is a key problem with applying such filtering, and should be mentioned.

L403-407. The justification for applying both filters is not entirely clear to me from the text here. I see (later) that it does appear to improve the MAE results, but the text here could make the case more clearly for this.

L446. Looks like the comma wasn't meant to be here

L451. The exception of the debris-melt-reduction is the terminal ice cliff, as identified by e.g. Immerzeel et al (2014).

L453-456. It is certainly possible that this pattern is due to avalanches, but it's worth nothing as well that this is probably the least-constrained part of your survey area; do you have GCPs or GVPs that high?

L467. 'are occurring' -> 'are evident'

L490. 'Constant box filter' -> not clear what you mean here. Do you mean the average value of a box (e.g. altitudinal or longitudinal bin)?

L499. What is the physical meaning of k=3? Three times the local surface velocity? The spatial pattern looks nice, but in my mind it would be much better to flow-correct your DEM as in Brun et al (2018). This of course represents a shift from an Eulerian to Lagrangian frame.

Figures 5 and 7. dh/dt has units of m per year.

Figure 8. The black dashed line for VM is missing on the right panel. These would represent the ideal cross section to assess whether your filtering has conserved mass (flux through cross section = integral of flux divergence).

L540-542. I'd avoid the speculation about a slowdown; you would need very different data and methods to assess this meaningfully!

L553. Section reference to 4.3 within 4.3

L605. This shows a problem of the filtering used here – the possible local ablation increase, if present in the dH/dt, has been filtered out.

Figure 12. Very nice summary; the processing has produced some very nice results in terms of reproducing the stake measurements!

L627. Nice to see this lateral heterogeneity highlighted, and very nice to see the local deviations from a linear ablation gradient. Can you indicate the standard deviation of SMB for elevation bins (for example) based on your data?

L638. See e.g. van Woerkom et al, 2019.

L648-655. This is a really nice result, because it highlights an area where the assumption b=b_s breaks down. Do you have any measurements to isolate the surface and subsurface melt in these areas?

L667. I suggest making a note that the perturbation details will follow

L670. Is this MAE computed for all years' data or a single year?

L672-674. I'd appreciate a few more details about the implementation of these random perturbations. I can't tell if this is the equivalent of 'random Gaussian fields' or something else. Are the perturbation magnitudes also random (per realization? Per lump?).

L679. Do the perturbations of F, thickness, and velocity also follow the same 'patch' approach?

Figure 14. It is clear that the thickness perturbations have had the greatest effect, but these were also the greatest perturbation (30% vs 10% for velocity), due to the uncertainty of the underlying dataset(s). As such, these results do not really correspond to 'sensitivity' but somewhat more the 'uncertainty' attributable to an individual dataset. I.e. if the thickness data had a 10% uncertainty, how would its associated errors compare to the velocity-associated errors?

L705. My only criticism here is that only the exponential decay has been tested, and I am not sure that it is necessarily better or conclusive, so I would suggest softening the language around 'necessary' to instead indicate that the exponential filter produces closer correspondence to stake measurements.

L738. I'm missing a discussion section considering implications/recommendations/future work? Density considerations? How important are the lateral variations in SMB (how much impact do they have on mass loss overall?) etc.

Brun, F., Wagnon, P., Berthier, E., Shea, J. M., Immerzeel, W. W., Kraaijenbrink, D. A., … Arnaud, Y. (2018). Ice cliff contribution to the tongue-wide ablation of Changri Nup Glacier, Nepal, central Himalaya. *The Cryosphere*, *12*, 3439–3457. https://doi.org/10.5194/tc-12-3439-2018

Van Woerkom, T., Steiner, J. F., Kraaijenbrink, P. D. A., Miles, E. S., & Immerzeel, W. W. (2019). Sediment supply from lateral moraines to a debris-covered glacier in the Himalaya. *Earth Surface Dynamics*, *7*(2), 411–427. https://doi.org/10.5194/esurf-7-411-2019

WAGNON, P., BRUN, F., KHADKA, A., BERTHIER, E., Dibas SHRESTHA, VINCENT, C., … JOMELLI, V. (2020). Reanalysing the glaciological mass balance series of Mera Glacier (Nepal, Central Himalaya) using geodetic mass balance, (2013), 1–9. Retrieved from Under review