THE OHIO STATE UNIVERSITY

Byrd Polar & Climate Research Center

Glacier Dynamics Research Group

Scott Hall Room 123
1090 Carmack Road
Columbus, OH 43210

chudley.1@osu.edu

3 May 2022

Dear Dr Berthier,

RE: Comments on tc-2022-33

We have received 2 reviews on our manuscript tc-2022-33 ***Empirical correction of systematic orthorectification error in Sentinel-2 velocity fields for Greenlandic outlet glaciers***, both of which found our work methodologically sound and of relevance to the wider community. We are grateful to Drs Mouginot and Altena for their constructive comments.

We have adapted the manuscript and introduced new text and figures in light of the comments, which we describe in detail below. In addition, we have made minor textual changes to the manuscript that were not directly requested by reviewers, in order to improve the clarity of the manuscript in light of the reviewer recommendations, or to fix minor errors in the text. We have attached a revised manuscript with changes tracked. Unless otherwise specified, line numbers in our response to reviewers refer to change-tracked version of the manuscript.

Thank you for your consideration of our revised manuscript, which we hope is now acceptable for publication.

Sincerely,

Tom Chudley

On behalf of all co-authors

Review 1: Jeremie Mouginot

Chudley et al. presents a study on the correction of the orthorectification error in surface ice velocity generated from Sentinel-2 acquisitions in Greenland. The results are sounds, clear and well described in the manuscript. Having a methodology to use the potential of Sentinel-2 images to track the dynamic evolution of Greenland's glaciers is a significant advance and will be useful for many other studies. The paper focuses mainly on the methodological aspects of remote sensing and as such could have its place in a journal more related to this type of subject. Nevertheless I believe that the topic will be of interest to the glaciological community and so could be published in The Cryosphere. Therefore, I recommend publication and have only minor comments below.

We are grateful to Dr Mouginot for his helpful and supportive review. Below, we have outlined how we have revised the manuscript in light of his comments and recommendations.

Comments:

An important point of the study to justify the processing Sentinel-2 images from different orbits (cross-track pairs) is the increase in the number of measurements that are useful for capturing rapid changes in velocity and producing dense time series. I think that the advantage of such an approach over considering only observations of similar orbits (repeat-track pairs) should be reinforced. In short, I think showing time series with and without the cross-track pairs could be useful to highlight the advantage over using repeat-track pairs.

We have included a new figure (figure S4, see below) to address this comment and that of the minor comment addressing P12L250 by Reviewer #2. that separates cross-track and repeat-track pairs. Figure S[X]a shows this for all pairs in 2019 at Jakobshavn site a, whilst Figure S[X]b shows the same for only for short baselines (<10 days). This highlights the particular utility in creating dense time series consisting of short-term velocity pairs. We have highlighted this point and associated figure in the discussion at line 313-315.

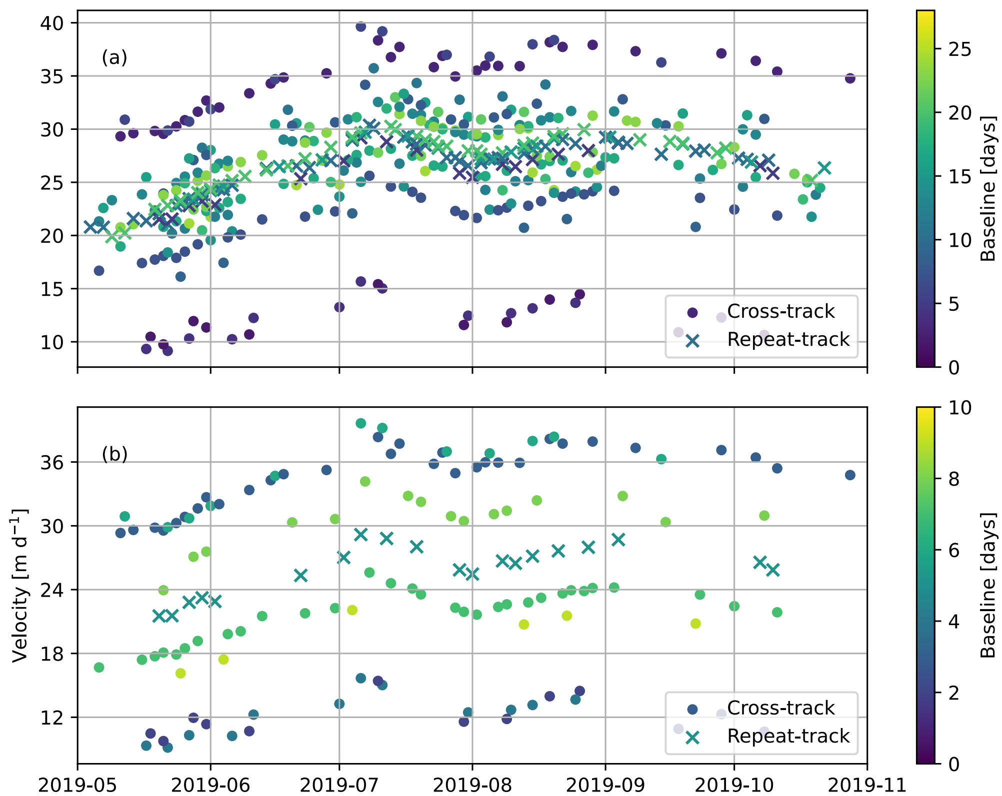

Figure S4: (a) Time-series of uncorrected observations at Jakobshavn site a, coloured by temporal baseline and with same-track pairs marked as crosses. (b) Same as a, but showing only temporal baselines less than ten days.

Similarly, the processing of Sentinel-2 data does not seem to provide additional information to that obtained by Sentinel-1 (Fig. 7). Would it be possible to find an example where S2 would fill a gap with respect to the already published time series in Sentinel-1 MEaSUREs (or Landsat-8 ITS_LIVE) ?

We have updated figure S5 (figure S4 in previous draft) to additionally include Sentinel-1 MEaSUREs data showing GP regression applied to all datasets over the period of velocity slowdown at Store Glacier in 2018. This shows that the Sentinel-2 dataset is better able to resolve this period of rapid slowdown, which takes place over a period of ~2 weeks, than the comparative datasets, which cannot resolve the slowdown to less than one month. We have updated the text where we refer to fig. S5 (L298, L316) to better reflect this.

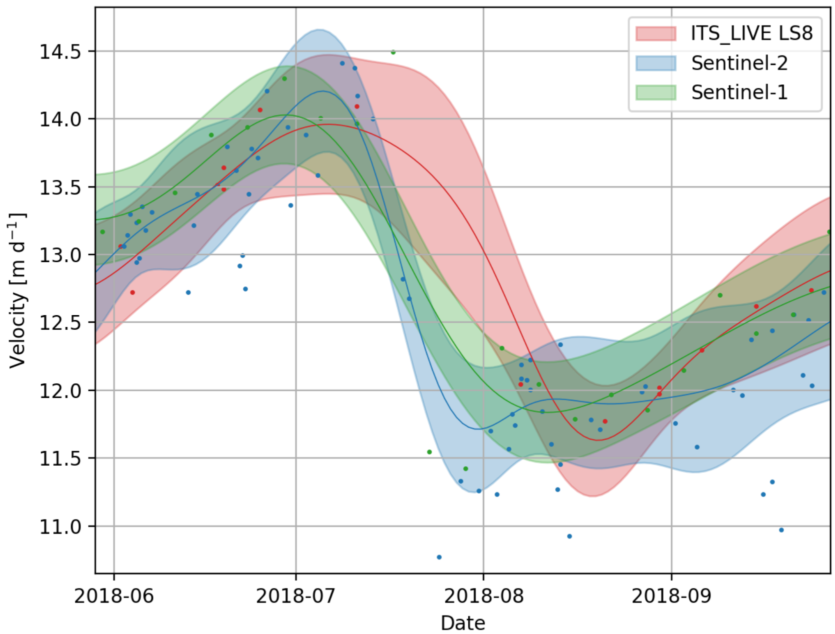

Figure S5: Summer slowdown at Store Glacier site (a) in 2018, highlighting the inability to capture the rapid slowdown in sparse Landsat 8 and Sentinel-1 derived data. Points mark individual velocity measurements, solid lines are Gaussian Process models, and shaded regions mark the 2-sigma confidence interval.

Gaussian process regression could be further described. Although a short description of the kernels used is given, it seems difficult with the details provided to reproduce the results obtained in Figure 7.

We have modified and further added to the description with the exact kernel functions used to aid the reader in replicating the processing (paragraph beginning L199). Although the explanation is short, it is not deceptively so: the scikit-learn library is both high-level and user-friendly. The core processing chain can be performed in six lines of code and an applicable example of the Rasmussen and Williams (20006) implementation for seasonally variable time-series datasets is outlined in a short tutorial [link] in the documentation.

"We performed GP regression using the GaussianProcessRegressor() implementation in the Python scikit-learn library. We model ice velocity as the sum of two kernels (covariance functions) representing seasonal and short-term variability respectively. Our implementation follows Rasmussen and Williams (2006), ignoring long-term variability over the five-year period we assess. Seasonal variability is incorporated as the product of an exponential sine squared kernel (ExpSineSquared()) and a radial basis function kernel (RBF()). The exponential sine squared kernel has a fixed periodicity of one year and the length-scale as a free parameter. The radial basis function kernel has a free length-scale. We implement short-medium term variability as a rational quadratic kernel (RationalQuadratic()), with free length-scale and alpha (controlling the diffusivity of the length-scale) parameters. Finally, we incorporate velocity error estimates (section 2.1.3) into modelling directly via the alpha parameter of the GaussianProcessRegressor()."

In Fig.6 : The name “Jacobshavn” is not consistent with the text (Jakobshvan)

This has been corrected in the new figure.

Review 2: Bas Altena

The authors present an implementation/ workflow description, to use displacement products from different orbital tracks of Sentinel-2. They are not alone, as such efforts are becoming popular in other fields of remote sensing of the cryosphere as well e.g.: [Lavergne et al. 2021]. For this work the authors have opted for a more methodological approach, instead of a paper looking at the physical drivers, which fit into a journal such as TC. Seen in this light the work does not go deeper than a desciption, at several places the reader is left in the dark why certain steps are taken. The work would become more interesting when such heuristics and rationale are explained or appropriate earlier work is mentioned as implementation argument.

We are grateful to Dr Alterna for his helpful and supportive review. Below, we have outlined how we have revised the manuscript in light of his comments and recommendations.

Major Comments

Looking more detailed to the work, the advancement made by the authors is less clear to me. Though it is a valid implementation, through setting of a hypothesis and testing through emperical relations, it reads as an isolated peace of work. As such relations for Sentinel-2 (and Landsat 8) have already been set out by [Kääb et al. 2016] and [Altena & Kääb, 2017]. A brief description is given about a vector projection method as presented in [Altena & Kääb, 2017]. Though this is a stripped down version, for large scale processing pipelines, as demonstrated in this work. While the major work presented in [Altena & Kääb, 2017] deals with a framework of harmonization of different orbits and elevation offset. Which later is extended to changing topography over time [Altena et al. 2019]. The limitations of the projection method are correct (which is a stripped version), but these do not apply to the core framework. Hence, the presented work is halfway the work presented previously by others. If one assumes stable topography and a perfect co-registration (which is done by the authors), all cross-orbit displacements observing a region on the ground are related. Why are they then isolated? It does not look like an improvement...?

We are grateful for this clarification of the relationship between this and the extensive work Dr Altena has done in the realm of Sentinel-2 correction. Our work follows on from the stripped-down method presented in Altena & Kääb (2017) in that it is simple and suitable for application to large-scale processing pipelines. In fact, as described in the text, our method *requires* a large-scale application in order to have enough data to reliably reconstruct the error. We have modified the text introducing Dr Altena's work in order to better reflect this (beginning L76):

"Finally, recent work has begun to develop frameworks that harmonise orbit and elevation offsets in Sentinel-2 glacier velocity datasets that do not require any prior knowledge of Δh (Altena & Kääb, 2017; Altena et al. 2019). However, the complexity of these methods necessitate simplified pipelines for operational use and bulk processing (Altena & Kääb, 2017). The simplified method presented by Altena & Kääb (2017) takes advantage of the fact that, if glacier flow is known a priori..."

Furthermore, we have rewritten the concluding sentence of this paragraph to better clarify the primary step forward we wished to make in our methodology (i.e. the limitation of the stripped-down projection method) (beginning L90):

"…it would be desirable to develop of method that, like Altena and Kääb (2017), remains simple, computationally efficient, and does not require any prior knowledge of Δh, but is also able to produce a geographically continuous record of velocity that is not spatially limited by the relationship between satellite geometry and flow direction."

In general context is missing, as other efforts like presented in e.g.: [Rosenau et al. 2015] are not mentioned. Results of the change in temporal baselines is mentioned and a decision is taken, but an assessment is missing, while [Millan et al. 2019] do show interesting results on this aspect. Hence, why is such knowledge not taken into account, and do the authors branch off?

We are grateful for the identification of highly relevant papers that were not included in our original submission. We have now discussed Rosenau et al. (2015), and their own approach to addressing orthorectification error, in the introduction (beginning L65):

Rosenau et al. (2015) provide their own improved orthorectification for Landsat imagery by orthorectifying L1G data to the ~1" ASTER Global Digital Elevation Model (GDEM) V2, rather than the ~30" Global 30 Arc Second Elevation (GLOTOPO30) dataset that was standard for L1C products at the time. However, the non-orthorectified product (L1B) for Sentinel-2 is not made available, meaning that this approach is not viable.

Please see our responses below to line comments P8L176 and P9L205 for a more detailed address of the temporal baseline and the relevance of the work of Millan et al. 2019.

As a final note, this review needs to be done on paper, but if done in speech, it would have been on a friendly tone. I do think the authors present nice work, and this work should be seen as work in progress, like any scientific enedavour. My comments are encouragements, with the aim to set this work to a higher level. A potential is present within this effort, but not exploited to its fullest.

We are grateful to Dr Altena for his helpful and thorough review and of course take his comments in the spirit with which they were intended.

Minor Comments

- most people scan a paper by reading the abstract first, but context is skewed here. Why would people use cross-track data?

We have added an additional sentence to the abstract (L11-13), highlighting the motivated for cross-track image pairs:

"As a result, most standard processing chains ignore cross-track pairs, which limits the opportunity to fully benefit from Sentinel-2's high-frequency observations during periods of intermittent coverage or for rapid dynamic events."

Typically, this is an exotic way of processing, thus why the authors put much emphasis on stating the presence of enormous errors in ESA products is a bit strange. It misses the overview, as this work is a nice contribution, but has limited impact.

The consideration of the errors introduced by PlanetDEM-90 is, to us, one of the key conclusions of the paper. Major effort has been put forward, indeed primarily by Dr Alterna, into developing methods of correcting errors introduced by this DEM into cross-track velocity pairs from Sentinel-2. All of these methods, in some way, work around the fact that the DEM is not publicly available. Here, we show strong evidence that the underlying DEM for Greenland is in fact available in an alternative public dataset, opening up a clear avenue to work towards analytical solutions to correct orthorectification errors in the 2016-2021 velocity data. As ESA has not announced plans for a reprocessing of the 2016-2021 Sentinel-2 data to the updated Copernicus DEM in a way comparable to the Landsat Collection-1/2 products, this is a key finding in providing these analytical solutions for this period in the future.

In response to this and a comment P17L346 below, we have updated the conclusion (beginning L369) to better reflect the key parallel impacts we consider our work to make: (i) a simple and efficient empirical method for correcting Sentinel-2 glacier velocity fields in large-scale datasets; and (ii) advances in determining the underlying DEM sources for 2015-2021 Sentinel-2 imagery over Greenland, and the opportunities this raises for analytical solutions.

- Gaussian Processes are popularized in our field by [Huggonet et al. 2021], hence this is an argument to use this approach. Now justification is missing.

Thank you for identifying this – we have further referenced Huggonet et al. in the methods section (L193).

- Why are these specific glaciers used, if cross-orbits are of interest, Northern Greenland is very interesting, since overpasses amost occur every day, see Fig.1 [Altena et al. 2019].

The study glaciers we have chosen are well-studied outlet glaciers that are significant in terms of contribution to GrIS discharge, are of ongoing interest to the glaciological community, and represent a diverse range of seasonal behaviours. Interestingly, we chose glaciers outside Northern Greenland for precisely the reason mentioned here: Northern Greenland glaciers already have a dense temporal coverage even without cross-track pairs. Focussing on glaciers further south allows us to examine the viability of filling in coverage with cross-track pairs even where coverage is non-optimal.

Typos and Details

Since I am not a native speaker, I am not able to give any feedback on typos, nonetheless some details might be improved:

p1 l29 Landsat8 is now a fleet together with Landsat9, having same orbit repeat cycle of 8 days

Fixed – added "*now 8 days from 2022 onwards with the addition of Landsat 9*" (L31-32)

p1 l29 Sentinel-2A and Sentinel-2B, need capital letters

Fixed (L31)

p2 l41 it might be a cultural thing, but please do use words instead of newly introduced abbreviations, especially as this is abbreviation is later changed into something else, and not used anymore. If you like abbreviations, than help the reader a bit and include an appendix with a nomenclature.

We have removed the ‘relative orbit’ abbreviation from the paper (L43, and throughout).

p2 l56 with the advent of so many same orbit repeat aquisitions, orthoretification errors are not a "significant" issue. The cross-orbit velocities are a "nice-to-have". Please do not oversell.

The use of ‘significant here refers to the scale of the DEM errors relative to Landsat 8, rather than a reflection on the relative importance of the problem. We have rephrased to be more specific: *“large vertical DEM errors (10s-100s metres)”* (L56)

p3 l67 "is not freely available" -> "is a commercial product"

Fixed (L73)

p3 l81 such a methodology is presented in the [Altena and Kääb] study, so what is the unique contribution...?

We have rewritten this section (L90-93) to better highlight the distinct features of our method relative to the Altena & Kääb study, as discussed in our response to the first major comment.

p4 l93 please consider using a logarithmic colorbar, as most (~95%) of the figures are dark blue now...

The use of a logarithmic scale for velocities would indeed by better at highlighting variation at the low end of velocities but it also acts to reduce the colour scale at the upper end, which is where the key variations detectable by this method exist (in particular, it would dampen the visible variation in figure 5). In the interests of keeping the colour scale consistent throughout the paper, we have chosen to retain a linear colour scale.

p6 l117 AOI, why is this acronym introduced, if it is only used here?

We retain this acronym as (i) it is commonly used across the literature; (ii) although limited in scope, we use it a total of six times in a relatively short period, which would be unwieldy to read otherwise; (iii) Our use of the phrase in the figure 2 flowchart, where “Area of Interest” would be hard to fit in to the box.

We have, however, noticed that mistakenly introduced the phrase twice in the first submission. We have fixed this, and now only introduce it once on L121.

p6 l130 here the newly introduced RO acronym is again replaced, what is the use?

We have now removed the RO acronym (see above). Cross-track and repeat-track are not replacing the ‘relative orbit’ term directly: instead, they refer to configurations of relative orbits for the respective velocity pairs. One solution to maintain an internal consistency may be to refer to these as *cross-relative orbit* and *repeat-relative orbit* respectively, but this is unwieldy. In referring to relative orbits, we are staying consistent to ESA/Copernicus programme terminology; in using cross-track and repeat-track, we are reverting to phrases commonly used in the literature. We think this is a suitable compromise.

p6 l130 the orbital number is even in the filename, as is also the case for Landsat, so this line is obsolete.

For simplicity, we have changed this sentence to "*Velocity fields are grouped by the relative orbits of their respective source image pairs"* (L142)

p7 l145 where does this 5x heuristic threshold comes from?

This threshold was identified based on manual assessment of a range of options. Below 5 pairs, the quality of the correction fields is notably degraded (as we are calculating the median from a low number of observations, this is perhaps unsurprising). We now better highlight this when the concept is first introduced (L150-251).

P7 l148 why does the assumption of stable geometry still hold, even for an highly dynamic outlet as Sermeq Kujalleq (Jakobshavn Isbræ) [Joughin et al. 2020 & Riel et al. 2021]?

Our justification of this assumption is outlined in the paragraph beginning L154: ultimately, the initial value of Δh is so large that, even using worst-case figures for a highly dynamic glacier, the maximum displacement error contributed by this assumption over our study period is at most on the order of ~1 m and thus subsumed by other error terms.

However, this does raise the point that this assumption will not hold for the new Copernicus DEM, which will have lower initial Δh values. We have added new discussion of this in section 4.3 (L361-363):

"Additionally, the lower initial Δh value will likely invalidate the assumption we make in this study that further elevation change will have a negligible impact on our assumption of stable geometry."

p7 l157 the description of vertical DEM error is to broad, since it is not random but has mostly a systematic effect.

We have rewritten this sentence in order to better communicate our intent, which is that our correction scheme is designed to correct for surface elevation change which we assume is largely negligible outside of the ice sheet. We update the sentence to read as such (L169-171):

"Displacements are corrected only over ice […], as our correction scheme is designed to correct for surface elevation change which we assume is largely negligible outside of the ice sheet boundaries."

p7 l162 it is not clear to me if stable ground is used, to co-register the imagery? please write it down, or give a motivation why not

We have made this further explicit in our methods section (L136-137):

"…and applied the relative sensor model bias compensation module as co-registrations. For co-registration, we define the off-ice region using the Greenland Ice Mapping Project (GIMP) ice mask"

p8 l166 why are velocities used here, while for orthorectification the temporal baseline is almost irrelevant. Please use appropriate units, i.e.: meter.

Error in absolute displacement should be broadly consistent across all velocity fields. However, the same displacement has a much large impact on shorter timescales, and hence error will be strongly dependent on the temporal baseline (see e.g. Figure [S3] – and also Millan et al. 2019 Fig. 4). We discuss this in the paper section 3.2 (L272-274). Ultimately, it is velocity error that is important to consider when differencing and filtering the fields, and also an inputs to our Gaussian processing pipeline. Reporting errors in velocity in further consistent with previous work (e.g. Millan et al. 2019).

p8 l169 this is a strange formulation, but why is this classical statistic used and not robust measures like "median of absolute difference" (MAD)? Also, the registration of Sentinel-2 has a flight line dependent positioning error (see fig.15 in [Kääb et al. 2016]), hence treating U and V as uncorrelated delutes the effectiveness of this threshold (see fig.2 in [Altena et al. 2021]).

We use standard deviation to be consistent with our comparative products (in particular, ITS_LIVE). The use of standard deviation is also consistent with previous work (e.g. Millan et al. 2019). The method outlined in Altena et al. 2021 provides an interesting perspective on calculating an isotropic uncertainty, and we look forward to begin implementing it in future pipelines upon the publication of the final work.

p8 l174 "errors" > "deviations/differences"

Clarified to RMSE error estimate (L188)

p8 l176 please justify your decision, and have a look at [Millan et al. 2019]

We mistakenly referred to Figure S1 here and not Figure S3, which shows the clear outlier of 2-day fields vs 3+ to properly justify our decision – we have corrected this in the text (L190). We think that the recommendation to consult Millan et al. 2019 is referring us specifically to section 3.3.6, which suggests that a good rule of thumb for detecting velocity changes is that 2σ must be <10% of the flow speed (σ is calculated as the standard deviation in ice-free areas). Our errors of ~1 m/day (Fig S3) suggest that 3-day+ fields are appropriate for identifying change at rapidly regions of fast-flowing outlet glaciers. Please note also that our use of Gaussian Processing allows for further management of uncertainty in the stacked time series, allowing useful information to be contributed even from fields with higher uncertainties.

p8 l180 "estimate" > "distill" and what is "true velocity", and are you able to back this claim, please rephrase

We retain estimate but replace ‘true’ with ‘continuous’ (L193).

p9 l205 please see [Millan et al. 2019]

We think that this comment is suggesting that we could/should use >30 day baselines in order to be able to better detect change over slower-flowing sectors of our study AOIs. Whilst this is true, we do not do so for three reasons: (i) Our focus here is to show the validity of our method in detecting rapid variation (~days-weeks) in faster-flowing outlets, (ii) long baselines (on the order of hundreds of days) may not ultimately be useful in extracting useful information about the seasonal dynamics of the Greenland Ice Sheet, and multidecadal change has already been assessed (e.g. Tedstone et al. 2015); and (iii) Processing the pairs is time- and resource-intensive and, based on our focus outlined in points (i) and (ii), there are ultimately practical decisions to make as to how many pairs is desirable to process. Our focus here is on extracting a dense high-temporal-resolution dataset, so for us the thirty day maximum baseline suits our needs.

P9 l207 why are velocity units used as a threshold for variable temporal base line data?

See our response above to P8L166.

p9 l210 please indicate the flight direction of the Sentinel-2 satellite, similar to inSAR maps.

We provide this information within Figure S2. We consider it more useful as a separate map as we can also provide information about the location of the orbit relative to the AOI. We now include a reference to Figure S2 in the caption of Figure 3 (L226)

p10 l223 please give an indication of the intersection angle, how much is the base-to-height or its angle

External imaging geomety could be useful in determining terrain-induced errors. However, since we are restricted to using orthorectified imagery provided in the tiling scheme, imaging geometries such as base to height ratio are less useful for determining errors. Indeed, one of the main objectives of the presented method is to derive an empirical terrain correction that is scaleable to large datasets and independent of specific pair geometries.

p10 l228 why so much hypothesis, while cross-track photogrammetry is around for some time?

We remove 'hypothetically' from this sentence, which implied that it was speculation rather than an expected result (L243)

p11 l231 "true" > "reference"?

Fixed (L246)

p12 l250 the figure can be improved, mayby make most stuff black and let the red points change in colour for different time intervals? As Sermeq Kujalleq (Jakobshavn Isbræ) seems to have several consistent off-sets. Where do they come from, please describ it in the text, this makes it more informative to the reader. Is it also interesting to look at the flow direction? So assumptions on this aspect might propagate into this variance? Maybe scatter plots are not the best to use, thus connsider using line plots to draw the temporal baselines as e.g.: fig.4 in [Charrier et al. 2022]

Thank you for providing feedback to improve this figure. Previously, including colour scale information on the points in the main figure created a visually noisy and difficult to interpret dataset. We have instead included a new figure (figure S4, see below) to address both this comment and that of the first comment of Reviewer #1. In it, we visualise the Jakobshavn data coloured by the temporal baseline of the scene, showing that the consistent offsets relate to the temporal baseline of the velocity pairs, as absolute displacement errors will have a higher relative effect on short-baseline fields. We note this and include a reference to the figure at L273 in Section 3.2 of the text.

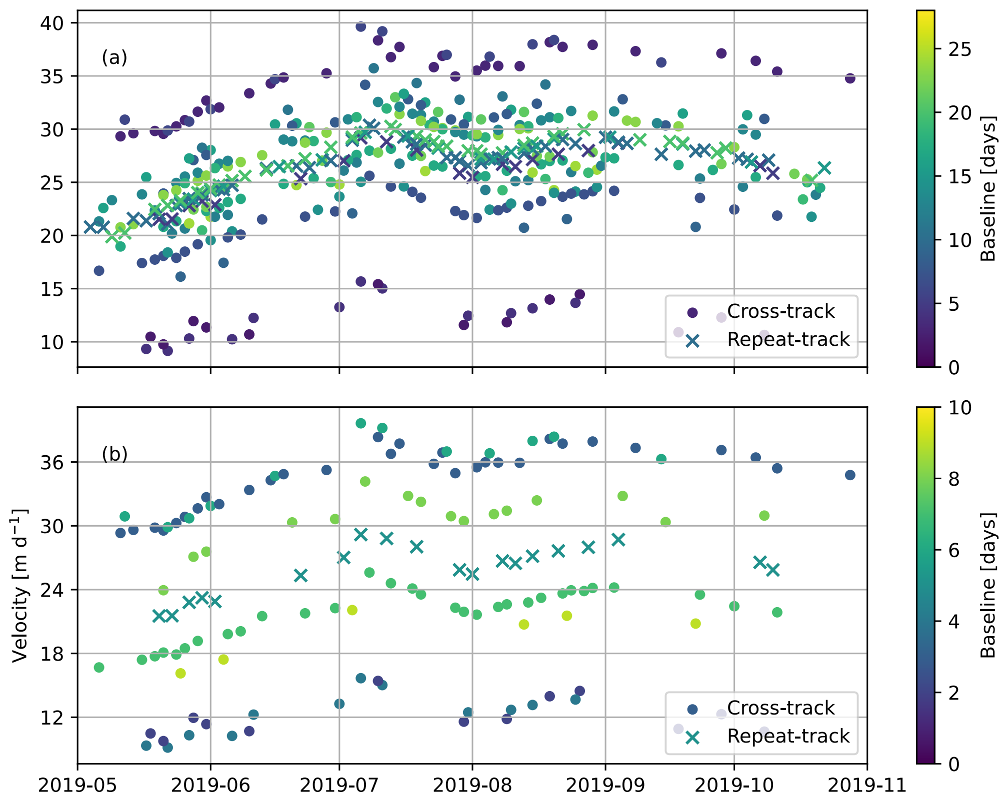

Figure S4: (a) Time-series of uncorrected observations at Jakobshavn site a, coloured by temporal baseline and with same-track pairs marked as crosses. (b) Same as a, but showing only temporal baselines less than ten days.

p12 l254 why not compare against off-glacier stable terrain?

Our purpose here is to compare our derived data against comparative time-series datasets. Our comparison against the uncertainty comes in the form of our error estimates.

p12 l261 this method does not resolve this issue of dynamic thinning, so why so specific about the cause?

Here we simply intended to refer to the total Δh, which, in these contexts, largely occurs from dynamic thinning. We agree that this unclear, and have removed the reference to dynamic thinning from the text (L276).

p13 l264 "high uncertainty of optical feature tracking", where does this loose claim come from?

We agree with this assessment and have removed it from the final text (L279)

p13 l265 please rephrase the GP sentence

In combination with comments above, we have removed ‘true’ in favour of ‘continuous’ (L280).

P13 l271 this is interesting, why is that? Please give more depth to the subject

We agree that this is an interesting feature. We address this in the Discussion section in paragraph beginning L323. The cause is unclear, but we consider it likely related to underlying vertical uncertainty in the Sentinel-1 correction DEM at the specific AOI we sample here.

p17 l346 very vague conclussion, but this can be improved if more in depth analysis are done.

Based on this and above comments, we have modified the conclusion to better highlight the two parallel contributions we consider out study to make: (i) a simple and efficient empirical method for correcting Sentinel-2 glacier velocity fields in large-scale datasets; and (ii) advances in determining the underlying DEM sources for 2015-2021 Sentinel-2 imagery over Greenland, and the opportunities this raises for analytical solutions (beginning L369).

Empirical correction of systematic orthorectification error in Sentinel-2 velocity fields for Greenlandic outlet glaciers

Thomas R. Chudley[1], Ian M. Howat[1,2], Bidhyananda Yadav[1], Myoung-Jong Noh[1]

[1]Byrd Polar and Climate Research Center, Ohio State University, Columbus, OH, USA
[2]School of Earth Sciences, Ohio State University, Columbus, OH, USA

Correspondence to: Tom Chudley (chudley.1@osu.edu)

Abstract. By utilising imagery from overlapping orbits, the Sentinel-2 programme offers high-frequency observations of high-latitude environments well in excess of its 5-day repeat rate, which is valuable for obtaining large-scale records of rapid environmental change. However, the production of glacier velocity datasets from optical feature tracking of Sentinel-2 imagery is limited by the orthorectification error in ESA products, which introduces significant systematic errors (on the order of tens of metres) into displacement fields produced from cross-track image pairs. As a result, most standard processing chains ignore cross-track pairs, which limits the opportunity to fully benefit from Sentinel-2's high-frequency observations during periods of intermittent coverage or for rapid dynamic events. Here, we use temporally complete glacier velocity datasets to empirically reconstruct systematic error, allowing for the corrected velocity datasets to be produced for four key fast-flowing marine-terminating outlets across the Greenland Ice Sheet between 2017 – 2021. We show that corrected data agrees well with comparison velocity datasets derived from optical (Landsat 8) and synthetic aperture radar (Sentinel-1) data. The density of available velocity pairs produces a noisier dataset than for these comparative records, but a best-fit velocity reconstructed by time-series modelling can identify periods of rapid change (e.g., summer slowdowns), even where gaps exist in other datasets. We use the empirical error maps to identify that the commercial DEM used to orthorectify Sentinel-2 scenes over Greenland between 2017 – 2021 likely shares data sources with freely available public DEMs, opening avenues for the analytical correction of Sentinel-2 glacier velocity fields in the future.

1 Introduction

Continuous glacier velocity datasets derived from medium-resolution satellite programmes have become increasingly available in recent years, forming a key part of investigations into ice discharge (King *et al.* 2018; Gardner *et al.* 2018; Mankoff *et al.* 2020), glacier dynamics (Poinar and Andrews, 2021; Dehecq *et al.* 2019), and characterisation of seasonal glacier behaviour (Vijay *et al.* 2021; Moon *et al.* 2014). Globally comprehensive scene-pair velocity fields from medium-resolution satellite data are available from both optical feature-tracking and SAR speckle-tracking techniques: e.g. for Landsat 8 optical data, the ITS_LIVE programme (Gardner et al., 2018, 2021); and for Sentinel-1 SAR data, the MEaSUREs and PROMIICE programmes for Greenland (Joughin, 2021a, Solgaard *et al.* 2021) and the RETREAT programme for glaciers and ice caps

(Friedl *et al.* 2021). The Sentinel-2 mission holds further promise in deriving glacier velocities from optical imagery compared to Landsat 8, offering an improved repeat time of 5 days (with both Sentinel-2A and -2B) compared to 16 days (now 8 days from 2022 onwards with the addition of Landsat 9), and a resolution of 10 m in the visible and near-infrared portion of the electromagnetic spectrum compared to 30 m (15 m panchromatic) for Landsat 8. This dense temporal coverage increases the chances of finding cloud-free image pairs, particularly when making use of cross-track imagery at high latitudes. However, as yet, the use of Sentinel-2 velocity fields - particularly in the form of large-scale public datasets - are limited.

A particular problem for Sentinel-2 feature tracking is the presence of systematic orthorectification errors. These are lateral off-nadir offsets in orthorectified satellite imagery resulting from vertical differences between the Digital Elevation Model (DEM) surface used to orthorectify the imagery and the true surface at the time of acquisition. Over solid bedrock, these offsets occur due to DEM errors, but the issue is exacerbated in glacial environments, where significant (10s of metres or more) real elevation change may occur between the DEM and image acquisition times due to changes in ice surface elevation (hereafter *Δh*) resulting from, mainly, sustained flow acceleration, increased surface melt and, subsequently, rapid ice thinning (e.g. King et al. 2020). When tracking displacement between two optical scenes from the same orbital path (in Sentinel mission terminology, the 'relative orbit'), orthorectification errors will be the same across the two images and will be eliminated in the final displacement map. However, in scene pairs from different orbits, a systematic error will be present as the vector sum of the two orthorectification errors. Sentinel-2 is particularly vulnerable to orthorectification error, suffering from an order-of-magnitude greater terrain bias than Landsat 8 (Altena and Kääb, 2017). This is in part due to the wide viewing angle of Sentinel-2 compared to Landsat 8: with Sentinel-2's swath width of 290 km, a vertical DEM error of *Δh* can result in a worst-case offset of ~*Δh*/5.4 at the maximum off-nadir distances, whilst for Landsat 8's swath width of 185 km this is only ~*Δh*/7.8 (Kääb *et al.* 2016). However, in the L1C and L2A data provided by ESA, the large errors are also related to the DEM chosen to orthorectify Sentinel-2 imagery. Until the 23rd August 2021 (30th March 2021 for Europe and Africa), the commercial PlanetDEM 90 m global elevation model (https://planetobserver.com/global-elevation-data/) was used to orthorectify Sentinel-2 data. Little public information exists as to which data sources were used to construct the PlanetDEM outside of the Shuttle Radar Topography Mission (SRTM) acquisition zone. However, Kääb *et al.* (2016) suggest that high-latitude source datasets are shared with the Viewfinder Panorama 3" DEM (de Ferranti, 2014), which use, among other sources, 20th century topographic maps to reconstruct high-latitude ice topography (J. de Ferranti, *pers. comm.*). This decades-old source data could explain the large vertical DEM errors (10s-100s metres) compared to Landsat 8, which, in Collection-2 processing, uses more recent region-specific elevation models at high latitudes (Franks *et al.* 2020), such as the ArcticDEM, Greenland Ice Mapping Project (GIMP) DEM, and Alaskan National Elevation Dataset. As a result, orthorectification errors remain a significant issue for producing consistent Sentinel-2 glacier velocity fields. Kääb *et al.* (2016) recommend that feature tracking using cross-track image pairs should be performed only for ice displacements that are at least one order of magnitude larger than the expected orthorectification error, whilst Nagy *et al.* (2019) recommend not using cross-track pairs at all. This limits the benefits of Sentinel-2's dense data coverage, reducing the number of available image pairs to only those with temporal baselines of 5 (10, 15 etc.) days. Being able to remove or account for the off-nadir orthorectification error is highly desirable to unlock the full potential of dense Sentinel-2 temporal coverage at high latitudes.

A range of solutions have been implemented to account for orthorectification errors in medium-resolution optical satellite imagery. Rosenau *et al.* (2015) provide their own improved orthorectification for Landsat imagery by orthorectifying L1G data to the ~1" ASTER Global Digital Elevation Model (GDEM) V2, rather than the ~30" Global 30 Arc Second Elevation (GLOTOPO30) dataset that was standard for L1C products at the time. However, the non-orthorectified product (L1B) for Sentinel-2 is not made available, meaning that this approach is not viable. With access to the PlanetDEM 90, Ressl and Pfeifer (2018) were able to generate a predicted offset field for Sentinel-2 relative orbits over Austria by using the PlanetDEM, the known orbits of Sentinel-2, and a reference DEM (the AustriaDEM) as ground truth. Here, rays were projected from the satellite orbital path to the reference DEM and intersected with the PlanetDEM to derive the off-nadir offset. However, this method not only requires access to the PlanetDEM (which is a commercial product) but also a reference elevation model that is accurate at the time of image acquisition, which is challenging for glaciated regions where, in places, surface elevations are changing significantly on an interannual timescale.

Finally, recent work has begun to develop frameworks that harmonise orbit and elevation offsets in Sentinel-2 glacier velocity datasets that do not require any prior knowledge of Δh (Altena & Kääb, 2017, Altena *et al.* 2019). However, the complexity of these methods necessitate simplified pipelines for operational use and bulk processing (Altena & Kääb, 2017). The simplified method presented by Altena & Kääb (2017) takes advantage of the fact that, if glacier flow is known a priori, it is possible to map the offset onto this flow direction as the offset vector always occurs perpendicular to the flight path (or, for dual orbits, along the epipolar line between the two satellite locations). This is an elegant solution that can be implemented in normal image matching pipelines without requiring elevation data, but comes with two primary limitations. The first is the assumption that flow direction is stable over time, which is justified on sub-decadal timescales for ice streams and glaciers that are not undergoing significant changes in their geometries, such as occurs during surges. However, the second limitation of this method is that comprehensive coverage is restricted by two criteria: (i) when the flow direction bearing is in the same direction as the epipolar line, the displacement will be mapped to infinity; and (ii) when displacement is within the measurement error, the same effect can occur. As a result, the authors filter velocities where the flow direction is within 20˚ of the epipolar line, and where displacement is >2.5 times the matching accuracy. Hence, the final corrected velocity field is discontinuous, dependening on the relationship between satellite geometry and surface topography. To produce continuous and dense velocity fields taking advantage of the entire Sentinel-2 record, it would be desirable to develop a method that, like Altena & Kääb (2017), remains simple, computationally efficient, and does not require any prior knowledge of Δh, but is also able to produce a geographically complete record of velocity that is not spatially limited by the relationship between satellite geometry and flow direction.

Here, we take advantage of five years of Sentinel-2 imagery to generate empirical corrections for systematic orthorectification error in ice surface velocity fields at four key marine-terminating outlet glaciers around the Greenland Ice Sheet. We describe the process by which we produce dense and continuous velocity datasets from 2017 – 2021, before validating our results by comparing them to publicly available velocity datasets at four key outlet glaciers.

2 Methods

We produce and present velocity data for four major marine-terminating glaciers: Sermeq Kujalleq (Store Glacier), Sermeq Kujalleq (Jakobshavn Isbræ), Helheim Glacier, and Kangerlussuaq Glacier (fig. 1). As two of these glaciers share a Greenlandic name, we hereafter refer to them by their alternative names used in scientific literature (Store Glacier and Jakobshavn Isbræ).

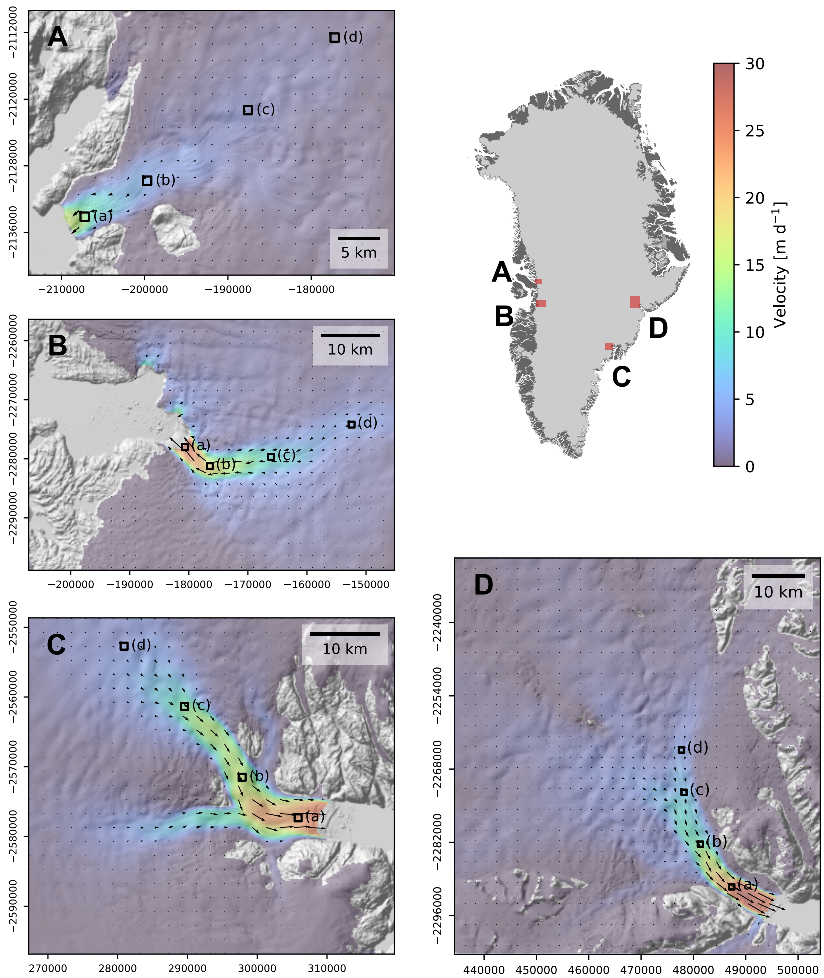

Figure 1: Reference velocity fields (median velocity of repeat-orbit pairs; see section 2.1.2) for the four marine-terminating glaciers presented in this paper, ordered anti-clockwise from North: (a) Store Glacier; (b) Jakobshavn Isbræ; (c) Helheim Glacier; and (d) Kangerlussuaq Glacier. 1 × 1 km sample sites a-d are marked for each glacier - see Table S2 for precise coordinates. Backgrounds are GIMP DEM hillshades (Howat, 2017); coordinates in NSIDC Polar Stereographic North. Inset: location in Greenland of all outlet glacier AOIs (red).

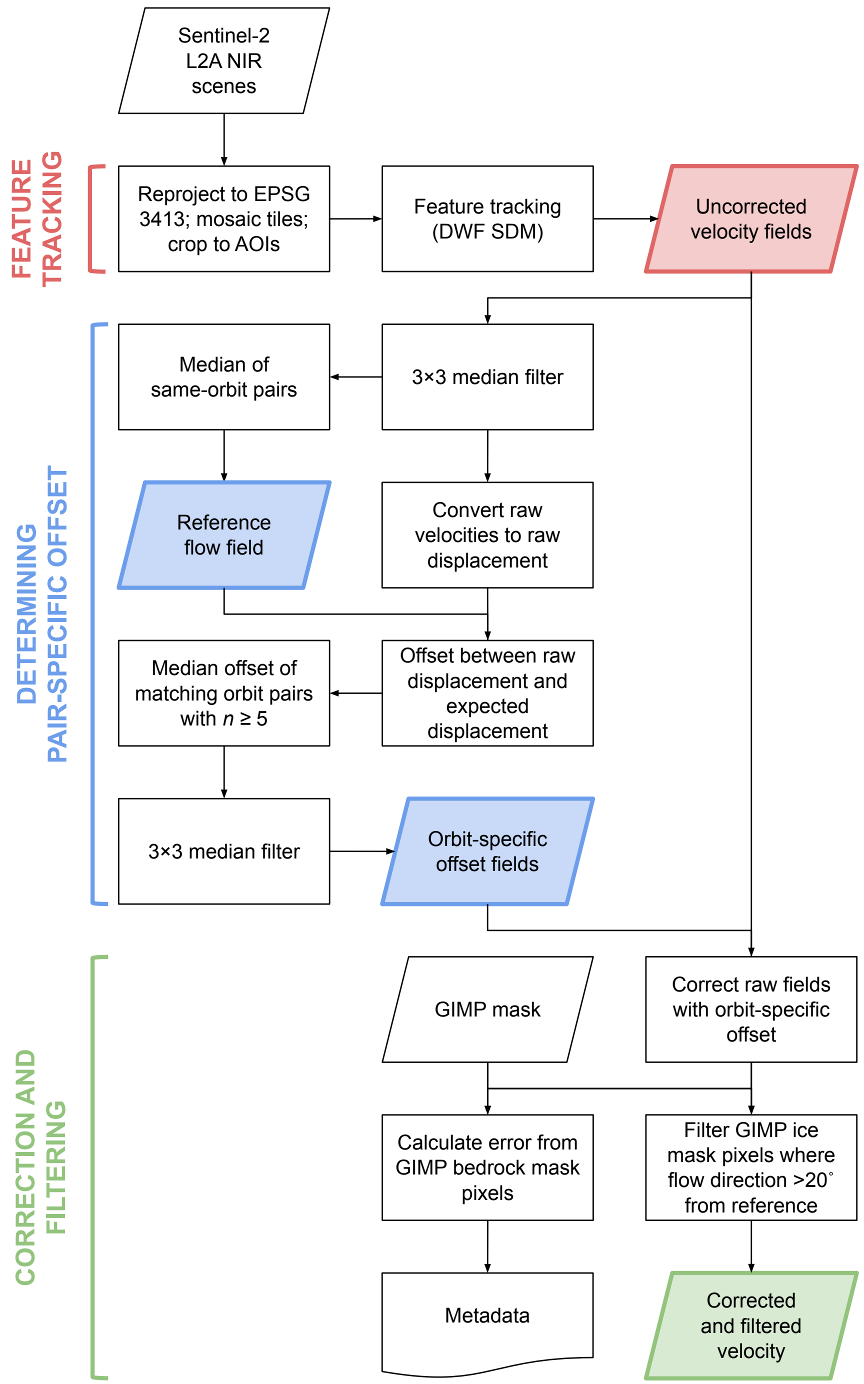

Figure 2: Workflow for deriving corrected and filtered glacier velocity fields from Sentinel-2 Level-2A data.

As orthorectification error is a function of satellite geometry, it should result in a consistent offset (in units of absolute displacement rather than velocity) across all velocity fields generated from the same relative orbit pairs. We assume that other errors in the velocity field (e.g. image matching error, coregistration error) are (a) random; and (b) do not correlate with specific relative orbit pairings. Hence, by drawing from a large base dataset, we can infer the average orthorectification error over the study period for specific relative orbit pairs by measuring the average offset between (i) the ice displacement measured from Sentinel-2 scenes and (ii) the expected displacement from a reference velocity field. We refer to this difference as the orbit-pair offset field, and subtract it from measured fields to generate final corrected fields (fig. 2).

2.1 Data production

2.1.1 Velocity field production

Data are produced at four marine-terminating glaciers, matching the spatial extent of the areas of interest (AOIs) in the MEaSUREs Selected Glacier Site Velocity Maps from Optical Images collection (Howat, 2020). We downloaded all Bottom-Of-Atmosphere corrected Level-2A (L2A) Sentinel-2 data with a cloud cover < 50% from the Amazon Web Services (AWS) Registry of Open Data using the `sat-search` STAC API (https://github.com/sat-utils/sat-search). Sentinel-2 data is staged using a pre-defined tiling scheme based on the Military Grid Reference System (MGRS). The pre-processing pipeline includes: subsetting Sentinel-2 tiles for each glacier bounding polygon AOI; reprojecting the raster to a common WGS 84 / National Snow and Ice Data Centre (NSIDC) Sea Ice Polar Stereographic North (EPSG:3413) projected coordinate system; mosaicking the adjacent overlapping tiles; and finally clipping the mosaicked raster to the glacier Area of Interest (AOI). We only considered scenes until the 23rd August 2021, when the L1C orthorectification process switched to a new geolocation procedure and underlying DEM (section 4.3).

Velocity fields at 100 m resolution were produced using feature tracking methods, performed using a Directional Weighted Filtering (DWF) algorithm based on the Surface Extraction from TIN-based Search-space Minimization (SETSM) approach (Noh and Howat, 2019). The method was originally developed for precisely estimating the surface displacement map (SDM) by compensating relative sensor model biases (minimising co-registration errors) and removing orthorectification errors caused by height changes through true DEMs. For the purpose of this research, we modified the algorithm to use orthorectified images directly, and applied the relative sensor model bias compensation module as co-registrations. For co-registration, we define the off-ice region using the Greenland Ice Mapping Project (GIMP) ice mask (Howat *et al.* 2014; Howat, 2017). The SDM processing is fully automated except in using an a priori, or seed, velocity field to specify maximum displacements for determining the initial resolution in the coarse-to-fine processing scheme (Noh and Howat, 2019). Here, we used InSAR-derived velocity fields between 2016 and 2017 (Joughin, 2021b) as the seed.

2.1.2 Estimating orbit-pair offsets

Velocity fields are grouped by the relative orbits of their respective source image pairs. Pairs of images acquired from the same relative orbit are hereafter referred to as *repeat-track pairs*, and those from different relative orbits *cross-track pairs*. For any given outlet glacier, certain combinations of orbit pairs may have anywhere from a few to >100 velocity fields. A reference flow field is constructed using the median U and V velocity values (in x and y EPSG:3413 Polar Stereographic North grid directions) from all (2017-2021) repeat-track velocity fields. Before processing, uncorrected velocity fields are filtered using a 3 × 3 median filter to reduce noise. For individual velocity fields, the expected displacement is calculated from these reference flow fields and the temporal baseline of the scene pair. The offset between the uncorrected displacement and the expected displacement is then calculated. Empirical orbit-pair offset fields are generated as the median offset for each orbit pair. Where these offet fields are constructed from fewer than five velocity fields, their quality is notably degraded. Hence, where a particular orbital pair has fewer the five observations, an empirical offset field is not constructed and the velocity fields are not processed further.

We note that over the course of the study period, ongoing glacier surface elevation change will continue to change Δh and hence the orthorectification error will not be constant. However, our method of estimating orthorectification error across the entire 2017-2021 study period implicitly assumes a constant Δh. We could improve this assumption by assessing offsets on shorter timescales – such as annually – but shorter timescales (smaller sample sizes) result in a notable reduction in the number of cross-track pairs available for correction (i.e. satisfying our threshold of 5 available velocity fields), and a lower quality of offset fields even for cross-track pairs where sufficient data is available. However, given the large initial Δh values, surface elevation change over the study period likely has a negligible impact on the orthorectification error. Using surface elevation change values from Smith *et al.* (2020) as approximations, maximum surface elevation change rates at our study sites range between -0.3 m a^{-1} (Store Glacier) to -3.6 m a^{-1} (Kangerlussuaq Glacier). Maximum estimated $|\Delta h|$ values (Section 3.1), which roughly correlate with these values, range between ~160 and ~400 m. Over the four-year study period, this results in a potential time-dependent error in Δh of between 0.3 and 2.2% at our study glaciers. Applying these uncertainties to typical maximum offsets (directly correlated with Δh) of between 40 and 60 metres shows that, even using worst-case assumptions, offset vector errors range between ±0.2 and ±1.3 m, values which are subsumed by other error terms (e.g. miscorrelation and coregistration).

2.1.3 Velocity correction and filtering

Once orbit-pair offsets have been constructed, uncorrected velocity fields are converted to absolute displacement, corrected using the appropriate orbit-pair displacement offset field, and converted back to velocity. Displacements are corrected only over ice as defined in the GIMP mask, as our correction scheme is designed to correct for surface elevation change which we assume is largely negligible outside of the ice sheet boundaries. Due to changing ice boundaries at marine-terminating locations, areas within the GIMP ocean mask are not filtered or removed, but ice velocities beyond the extent of the GIMP ice mask should not be considered reliable.

To remove erroneous velocity measurements, areas within the GIMP ice mask are filtered where flow directions are >20˚ offset from the reference flow field. If, after filtering, no data remains (<1% of the ice area has valid velocity measurements) the field is discarded and no output data is generated.

2.1.4 Error assessment

A first-order estimate of error is taken as the root mean square error (RMSE) of the absolute velocity of the bedrock area (as defined by the GIMP bedrock mask). RMSEs tended to be low, with the median RMSE consistently beneath <0.5 m d^{-1} for the study glaciers discussed in this paper (fig. S1). Additionally, the mean and standard deviation of the *U* and *V* velocity fields of the bedrock area are also recorded, in order to assess systematic error within individual flow fields due to e.g. poor co-registration. Where the mean of the *U* or *V* velocity is greater than one standard deviation away from zero, the field is considered to have a systematic error and is not included for presentation in this study (section 2.2).

2.2 Data presentation

2.2.1 Sampling

To present time series' of glacier surface velocity, we sample four sites of increasing distance from the calving front at each of our sample glaciers (Figure 1; Table S1). We sample across a 1 × 1 km area, calculating the median velocity across this sample region. We filter out data points where < 70% of the sample region contains data, or where the RMSE error estimate is > 5 m d^{-1}. We further filter fields where the temporal baseline is only two days, where errors were significantly greater than any other baselines (Figure S3).

2.2.2 Gaussian process regression

The output from our velocity correction process produces a dense time-series of varying error estimates. Hence, we use Gaussian process (GP) regression (Rasmussen and Williams, 2006, Hugonnet *et al.* 2021) to estimate a continuous uncertainty-bounded time-series velocity from our sampled time-series observations. Two particular properties of GP regression make it useful for the current application: (i) the Bayesian nature of the method accommodates the incomplete velocity record, producing a smooth, nonlinear, interpolated output; and (ii) the probabilistic model can incorporate uncertainty estimates and provides an empirical confidence interval to predictions.

We performed GP regression using the `GaussianProcessRegressor()` implementation in the Python `scikit-learn` library. We model ice velocity as the sum of two kernels (covariance functions) representing seasonal and short-term variability respectively. Our implementation follows Rasmussen and Williams (2006), ignoring long-term variability over the five-year period we assess. Seasonal variability is incorporated as the product of an exponential sine squared kernel (`ExpSineSquared()`) and a radial basis function kernel (`RBF()`). The exponential sine squared kernel has a fixed periodicity of one year and the length-scale as a free parameter. The radial basis function kernel has a free length-scale. We implement short-medium term variability as a rational quadratic kernel (`RationalQuadratic()`), with free length-scale and alpha (controlling the diffusivity of the length-scale) parameters. Finally, we incorporate velocity error estimates (section 2.1.3) into modelling directly via the `alpha` parameter of the `GaussianProcessRegressor()`.

2.3 Supplementary data

As orthorectification errors scale with Δh, we compare generated offset maps to DEM difference as a proxy for Δh. The first DEM is the *Viewfinder Panorama 3" DEM for Greenland (Version 1)* (de Ferranti, 2014; hereafter the 'Viewfinder DEM'). Kääb *et al.* (2016) inferred that the Viewfinder DEMs likely shared source data with the PlanetDEM for a test site in Northern Norway. The second DEM is the *GIMP DEM from GeoEye and WorldView Imagery (Version 1)* (Howat *et al.* 2014; 2017; hereafter the 'GIMP DEM'), which was produced from imagery between 2009 – 2015.

We compare our generated velocity fields to two other public datasets. To compare with medium-resolution SAR velocity fields, we make use of MEaSUREs *Greenland 6 and 12 day Ice Sheet Velocity Mosaics from SAR (Version 1)* velocity fields from speckle-tracked Sentinel-1 data (Joughin *et al.* 2018; Joughin, 2021a). These were downloaded from the NSIDC data portal. To compare with optical velocity fields derived from Landsat 8 imagery, we make use of the ITS_LIVE dataset (Gardner *et al.* 2018, 2021). These were downloaded using the ITS_LIVE API, and, to match our Sentinel-2 data thresholds, filtered to velocity fields where (i) the maximum interval between data pairs was 30 days; and (ii) at least 1% of the data contained valid pixels. For both Sentinel-1 and Landsat 8 fields, we extracted time-series data in the same way as for Sentinel-2 data: i.e. filtering to data that covers at least 70% of the 1 × 1 km sample region, and < 5 m d^{-1} error.

3 Results

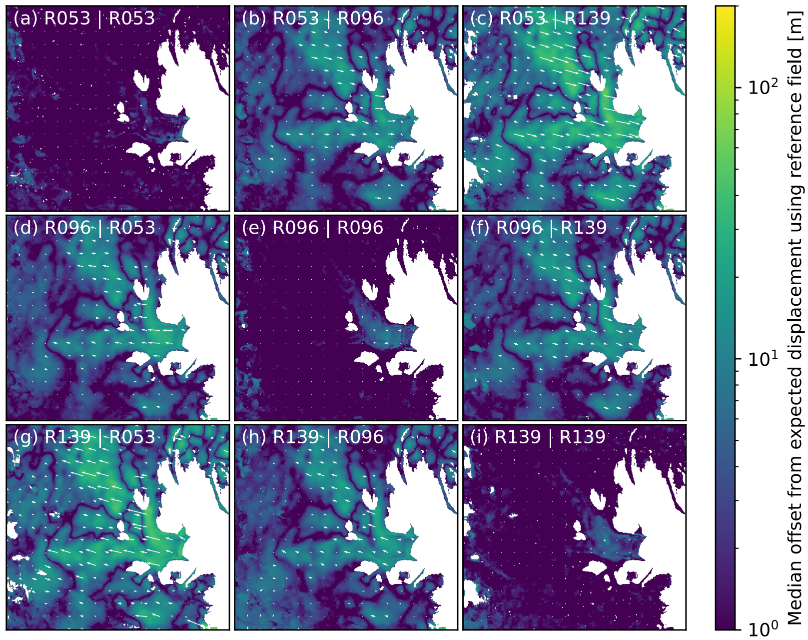

Figure 3: Median offset between measured displacement and expected displacement (from reference velocity field) for different orbital pairs at Helheim Glacier. Magnitude is shown in colour; vectors are shown as white arrows. The orbital geometry of the constituent relative orbits is visualised in Figure S2.

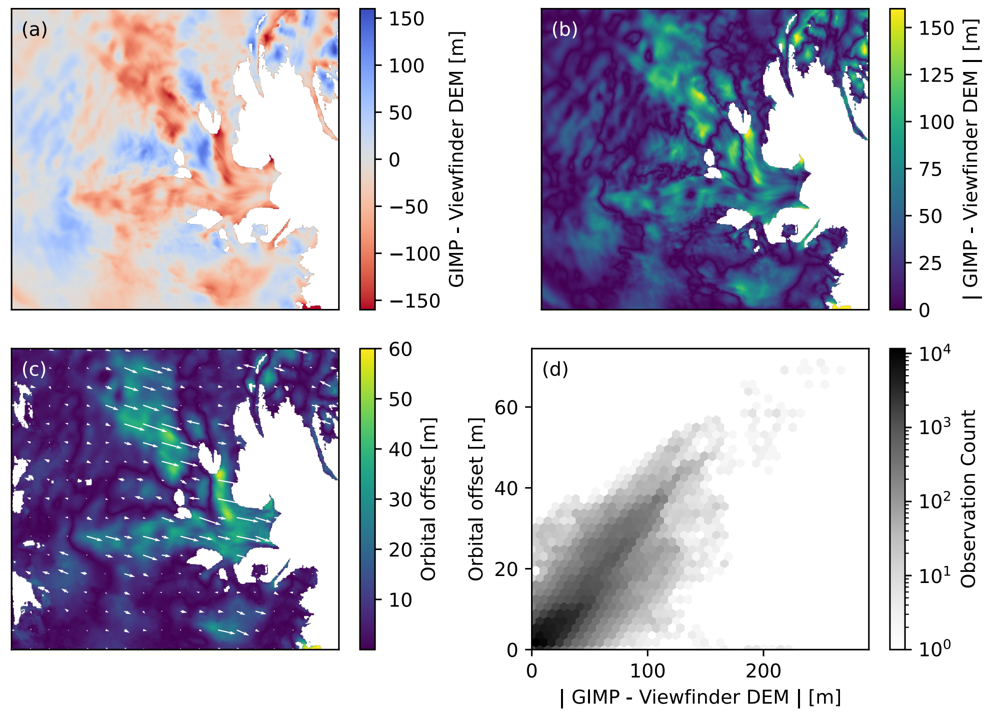

Figure 4: Comparison between DEM vertical differences and velocity offset for Helheim Glacier. (a) DEM difference between the Viewfinder DEM and the GIMP DEM. Areas where the GIMP DEM has a lower elevation value are in red, and a higher elevation value in blue. (b) Absolute vertical difference between the Viewfinder DEM and GIMP DEM. (c) Median offset between the measured displacement and expected displacement for velocity fields taken from relative orbits 053 and 139 (identical to fig. 3c). (d) Hexbin density plot comparing the absolute vertical differences between the Viewfind DEM and GIMP DEM with the empirically determined orbital displacement.

3.1 Orbital offset generation

Empirically determined systematic offsets can reach values on the order of tens of metres at outlet glaciers around Greenland. Comparing offset fields constructed for various orbit pairs at Helheim Glacier (fig. 3) shows that the behaviour of empirical offset fields are consistent with theory. Offset fields for repeat orbit pairs (fig. 3a, e, i) have negligible offset, whilst the offset is largest for the cross-track pairs from relative orbits 053 and 139 (fig 3c, g), which have the greatest distance between their respective orbital paths (fig. S2). Empirical offset vectors are, to within reasonable error, uniformly parallel and occur along the epipolar line of the satellite pair viewing geometry (fig. 3 cf. S2), consistent with the hypothesis that these systematic errors occur due orthorectification error in the off-nadir direction.

If the measured offsets are due to orthorectification error, they should directly correlate to the *Δh* between the times of DEM and image acquisition. We can thus validate that our offsets are meaningful by comparing them to the DEM difference between the Viewfinder DEM, which we use as a proxy for the PlanetDEM (section 2.3), and version 2 of the GIMP DEM, which acts as our reference elevation, or at least a closer approximation to the surface elevation at the time of image acquisition (Fig. 4).

The spatial pattern of |*Δh*| (fig. 4b) matches the magnitude of the orbital offset (fig. 4c), with a strong, positive correlation ($R^2 = 0.67$, $p < 0.01$) between the two values (fig. 4d). The direction of the offset (white vectors in fig. 4c) is also predicted by the direction of *Δh* (fig. 4a): where *Δh* values are negative, offset error occurs in the ESE direction (towards relative orbit 053), whereas positive *Δh* values occur where offset errors are in the WNW direction (towards relative orbit 139).

The efficacy of the orbital correction fields is visualised for an example case at Jakobshavn Isbræ (fig. 5), which shows the relative ability of corrected and uncorrected glacier velocity fields to properly capture the magnitude and extent of a summer slowdown occurring in late July 2019 (see also figs. 6 and 7 for time series of this event). In the uncorrected velocity fields (fig. 5a and b), orthorectification error in the cross-track velocity fields (relative orbits 025 / 111 and 068 / 025 respectively) introduce an apparent difference in the final velocity field in excess of 10 m d^{-1} at the calving front and 2 m d^{-1} even tens of kilometres inland, a rate of change that is unphysical. After correction (figs. 5d and e), this change is reduced to ~2-3 m d^{-1} at the front and negligible amounts inland (fig. 5f), an observation that is in line with contemporaneous Sentinel-1 observations (fig. 5g-i).

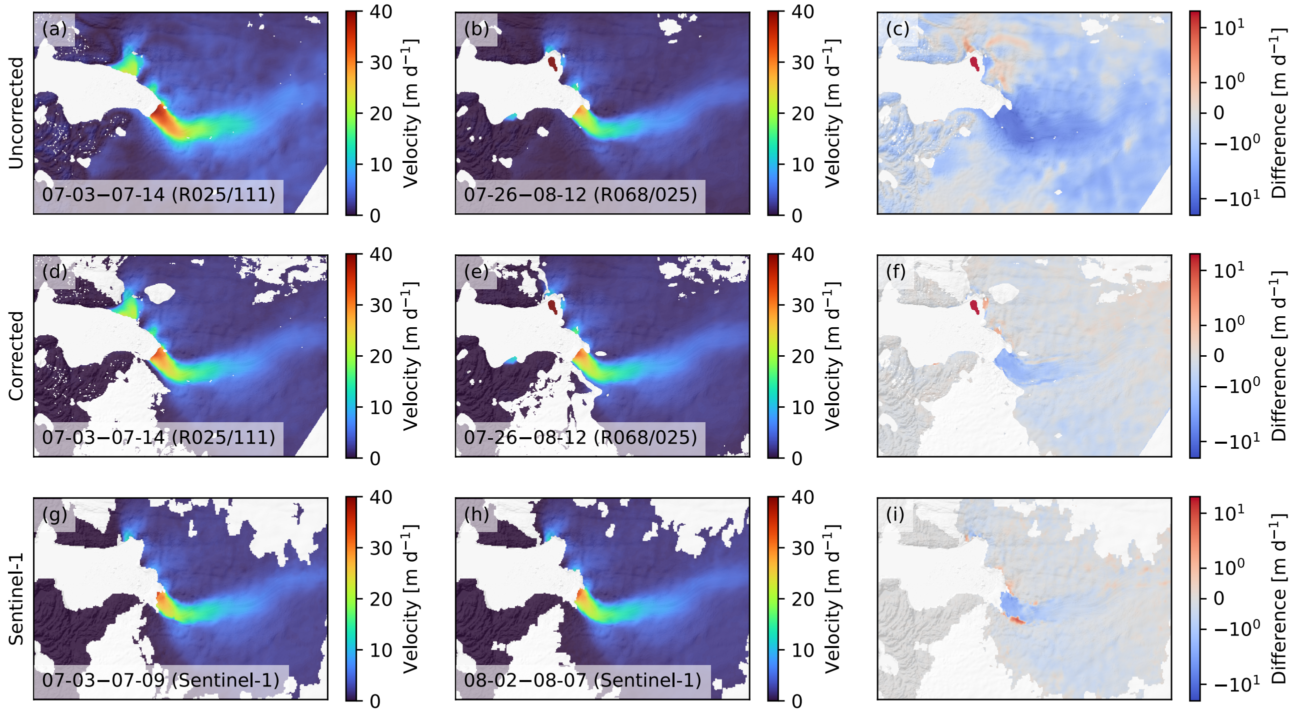

Figure 5: Velocity difference maps at Jakobshavn over a period of summer slowdown in 2019 (see figs. 6 and 7). (a) Velocity field pre-slowdown, from uncorrected Sentinel-2 feature tracking. (b) Velocity field post-slowdown, from uncorrected Sentinel-2 feature tracking. (c) Difference between pre- and post-slowdown from uncorrected Sentinel-2 feature tracking. (d-f) Same as (a-c), but for corrected Sentinel-2 feature tracking. (g-i) Same as (a-c), but for contemporaneous Sentinel-1 speckle tracking. Note log scale used on difference maps. Backgrounds are GIMP DEM hillshades (Howat, 2017).

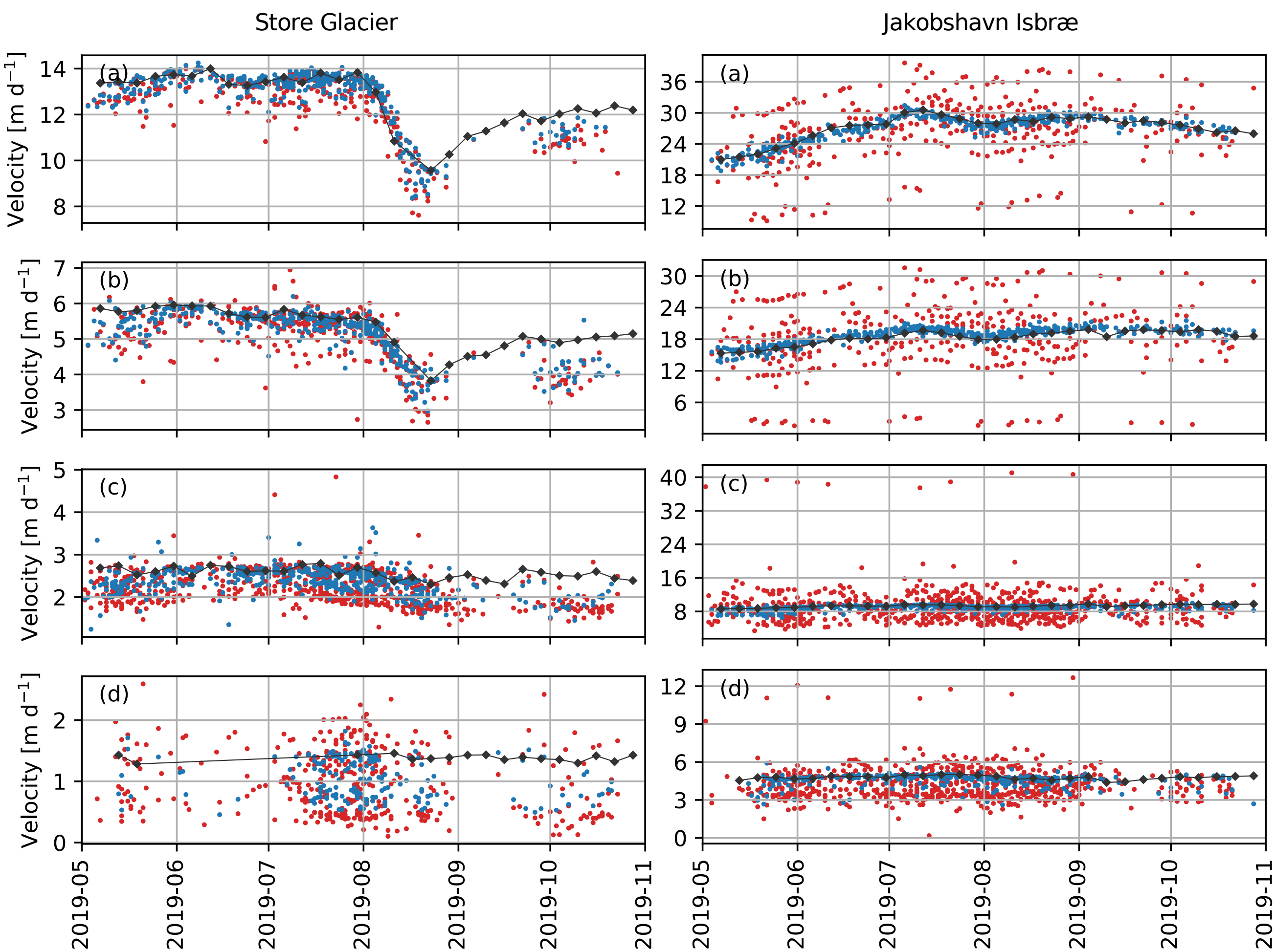

Figure 6: Corrected (blue) and uncorrected (red) velocity time-series for four sample glaciers in 2019. Top to bottom are points a-d in fig. 1. Black line and diamonds mark the MEaSUREs Sentinel-1 velocity dataset for comparison.

3.2 Ice velocity time series

For the sampled 1 × 1 km sectors of our four study glaciers, we compare our corrected Sentinel-2 data against other data sources. Comparing our corrected velocity data to the uncorrected dataset shows the improvement that our empirical correction has (fig. 6). Uncorrected data shows a characteristic error distribution, with points offset from the Sentinel-1 -derived velocities based on their constituent orbit pairs and their temporal baseline. Short-baseline pairs are increasingly offset from the reference dataset value, even where orbit pairs are identical (fig. S3, S4). This is because orthorectification error is an absolute displacement offset, and as such becomes a higher relative error component of short-baseline velocities. In all cases, the corrected velocity data converges near the reference Sentinel-1 time-series. The correction is greatest for observations at Jakobshavn Isbræ and Kangerlussuaq Glacier, where Δh values are the largest.

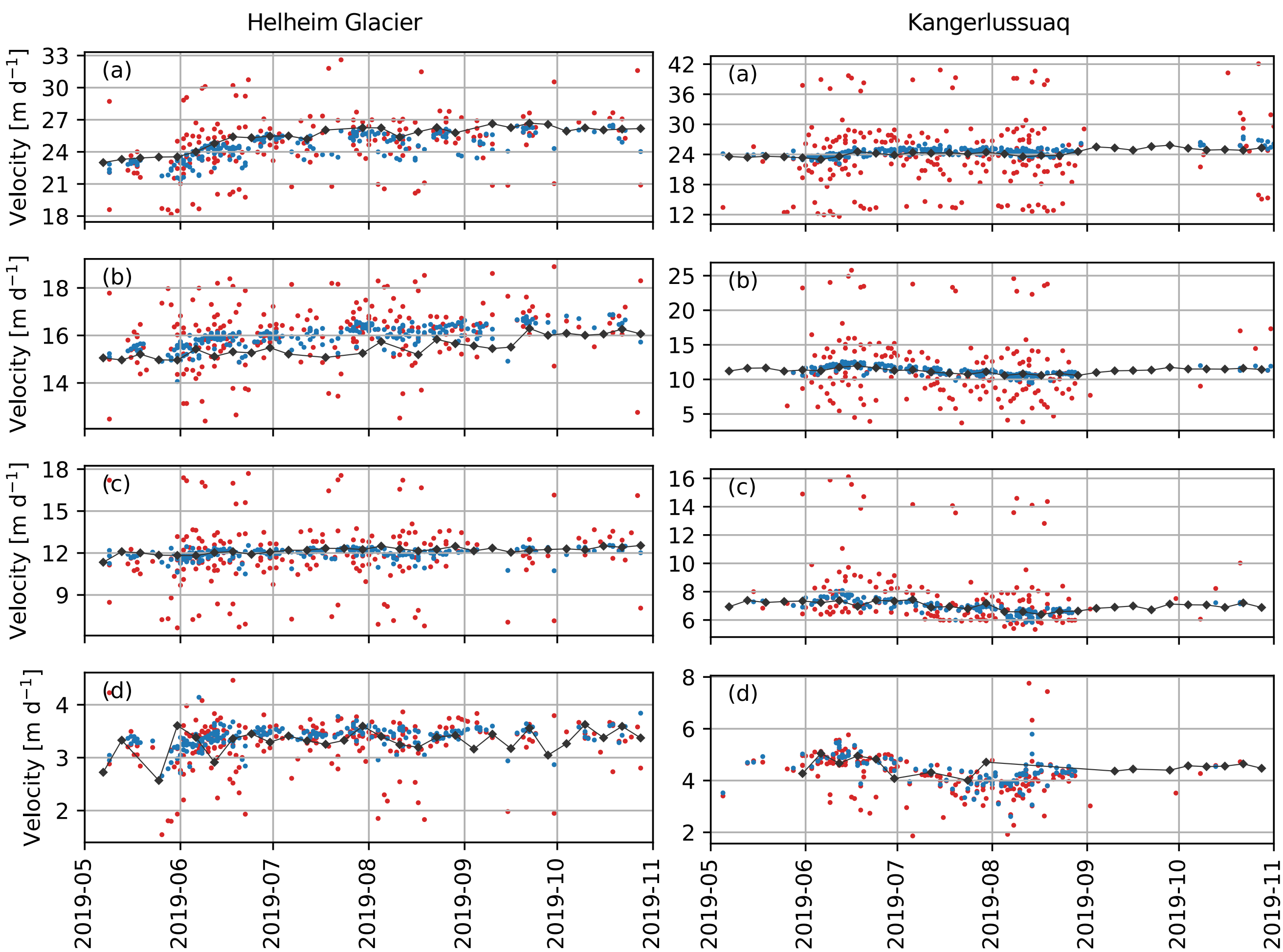

Figure 6 (cont).

The high density of data points results in a noisy dataset compared to the SAR-derived data. The use of GP regression highlights an effective way of interpolating a continuous time-series from this dataset (blue line in fig. 7), whilst accounting for sample-specific error values and time-variable data densities. Across all sites, the median difference between the Sentinel-1 record and the GP fit is 0.08 m d^{-1}; 68% (95%) of values lie within 0.41 (0.73) m d^{-1} of one another; and 90.5% of values lie within their errors (Table S2). These differences, on the order of decimetres, align with our error estimates from off-ice displacement of between 0.3 and 0.5 m d^{-1} on average (fig. S1). The lowest level of agreement between the two datasets occurs at Helheim Glacier point b, where only 52.3% of Sentinel-1 velocity values lie within error of the Sentinel-2 GP fit, and the Sentinel-1 record is, on average, 0.47 m d^{-1} slower than that of the GP fit.

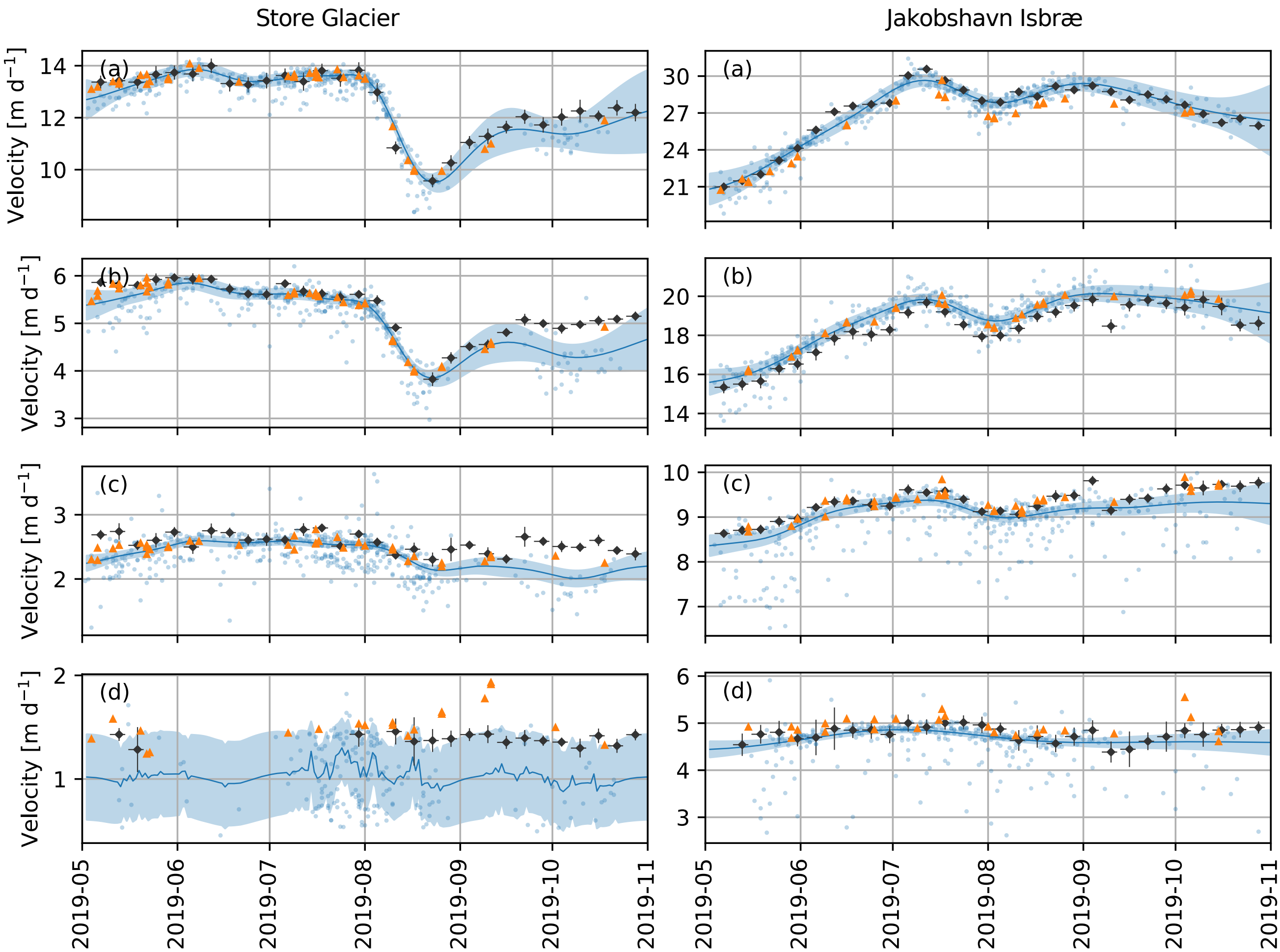

Figure 7: Corrected Sentinel-2 (blue dots), ITS_LIVE Landsat-8 (orange triangles), and MEaSUREs Sentinel-1 (black diamonds) velocity time-series for four sample glaciers in 2019. Top to bottom are points a-d in fig. 1. Blue line marks the output of the Gaussian process regression, with the blue shading marking the 2-sigma uncertainty bound.

This high level of agreement also occurs in comparison with the ITS_LIVE dataset derived from optical Landsat 8 imagery (orange triangles in fig. 7). Across all sites, the median difference between the ITS_LIVE record and the GP fit is 0.12 m d^{-1}; 68% (95%) of values lie within 0.29 (0.63) m d^{-1} of one another; and 87.0% of values lie within error of each other (Table S3). The increased density of the Sentinel-2 record relative to the Landsat 8 record allows for a higher temporal precision in identifying rapid drainage events. For instance, the summer slowdown at sites a and b at Store Glacier are not as not well captured in the Landsat 8 or Sentinel-1 records, where a sparse record means that the timing of the slowdown often appears to take more than a month, whilst the dense Sentinel-2 dataset and the GP fit shows the slowdown to occur on a timescale of ~2 weeks (fig. S5). The Sentinel-2 dataset is also of sufficient precision to be able to assess changing strain rates across the same slowdown (fig. S6). However, the Landsat 8 record performs better at lower velocity values (sub-4 m d^{-1} – e.g. site d at Store

Glacier), likely because any errors in the empirically-derived displacement field have greater relative influence at low flow velocities.

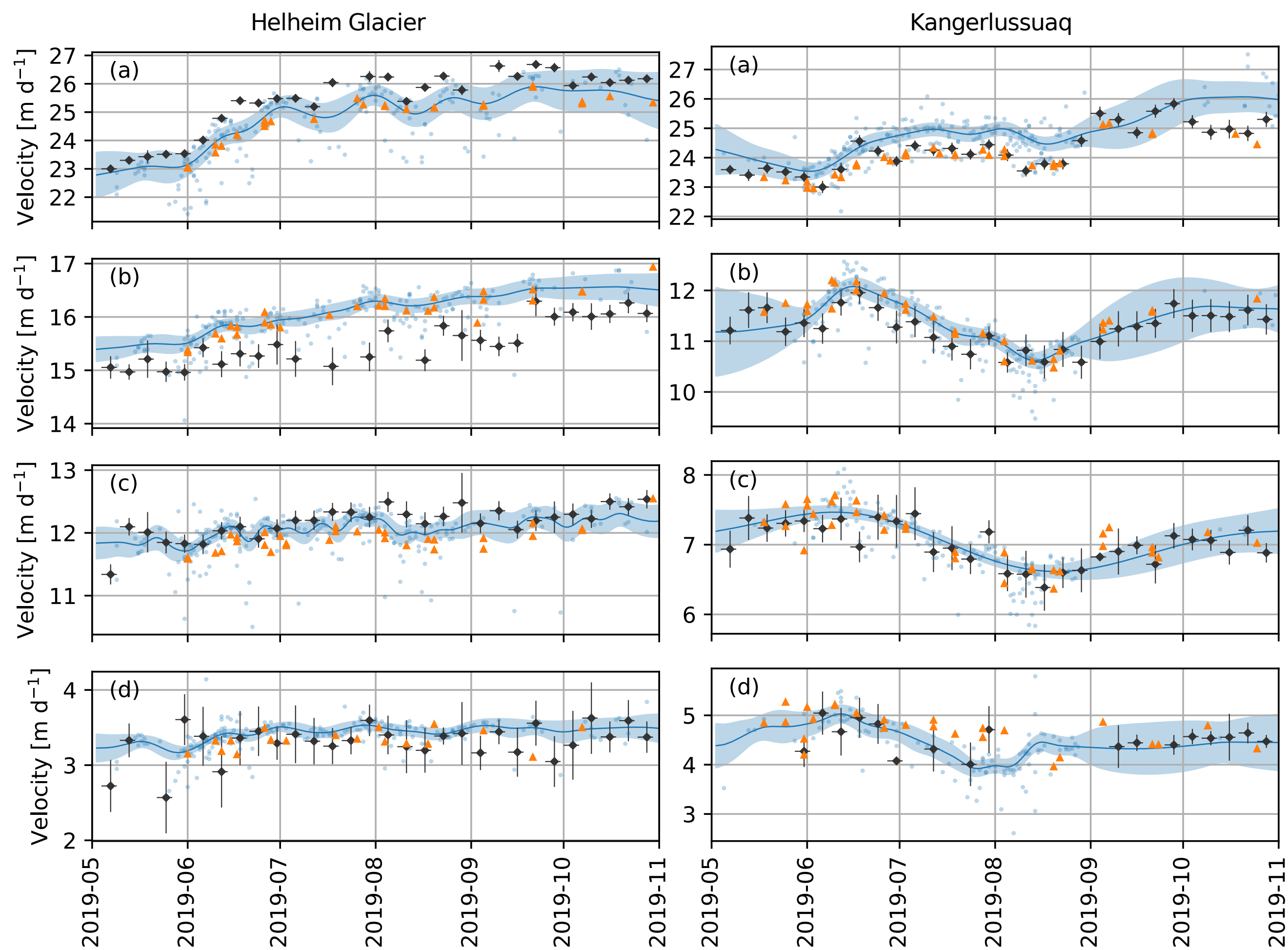

Figure 7 (cont).

4 Discussion

4.1 Data quality

In the absence of empirical correction, orthorectification errors in glacier velocity can result in errors in excess of 10 m d^{-1} in short-baseline velocity fields at Greenlandic outlet glaciers (figs. 5, 6). As orthorectification error is linearly proportional to Δh between the orthorectification DEM and the true glacier surface (fig. 4), errors are greatest at glaciers where ice surface elevation has changed the most. This is exemplified by Store Glacier, which is notable for its long-term stability induced by a sill at the glacier terminus (e.g. Morlighem *et al.* 2016), and has relatively low orthorectification errors prior to correction (fig. 6).

The empirical corrections described in this paper demonstrably reduce the influence of systematic orthorectification error on Sentinel-2-derived glacier velocity fields, reducing the variance from the order of metres to decimetres (figs. 6, 7). This allows for the incorporation of cross-track pairs into the glacier velocity time series, increasing the temporal density of observations, in particular for pairs with short (<10 day) baselines (e.g. fig. S4b). The temporal density of observations means that short term variations in speed, such as the slowdown at the terminus of Store Glacier, can be well resolved (fig. S5). The noise of the dataset highlights the utility of effective filtering and time series modelling, such as GP regression (fig. 7), in extracting a continuous velocity estimate – following other studies that have made use of, for example, Kalman filtering (King *et al.* 2018) to similar effect – and provide the advantage of a greater accuracy through the synthesis of multiple estimates. GP fits over the four glaciers assessed here differ from Sentinel-1 estimated by, on average, only 0.08 m d^{-1}, and 90.5% of estimates lie within error (table S2). Unlike monthly averaging, the use of GP regression retains the ability to capture short-term variation at ~weekly timescales and provides a time-varying confidence interval based on training data uncertainty and coverage.

The data display high agreement with comparative datasets sourced from Sentinel-1 and Landsat 8 (fig. 7). Although Sentinel-1 has been presented as a 'true' reference velocity in this study, it is of note that when only two of the three records presented agree with each other, which two will vary: contrast, for instance, Helheim site b (Sentinel-2 and Landsat 8 agree); Jakobshavn site a (Sentinel-2 and Sentinel-1 agree); and Kangerlussuaq site a (Sentinel-1 and Landsat 8 agree). All three methods are orthorectified using different DEMs whose ages, accuracies, and resultant Δh values are variable. As such, it is likely that systematic biases are present in all three datasets and are spatially variable on a km-scale. However, the characteristics of the cross-track Sentinel-2 dataset will make it advantageous over Landsat 8-derived datasets in scenarios where a high density of observations is required, as the increased observation frequency will allow for a greater chance of successful image pairs over critical dynamic periods, such as the melt season acceleration and late summer slowdown. In contrast, individual Landsat 8 velocity fields appear to be more precise, and may be preferable when the accuracy of individual velocity fields are necessary (e.g. comparing early and late season velocities). Sentinel-2 also provides a temporal advantage over Sentinel-1 datasets, which are limited to a fixed 6-day repeat cycle, and may also provide a valuable alternative for applying to small and steep glaciers, such as in high mountain regions, where the use of synthetic aperture radar for velocity extraction can be challenging (e.g. Paul *et al.* 2021).

4.2 DEM origin

The differencing of orbital pair offsets from a reference flow field has been shown to be an effective method of reconstructing systematic orthorectification errors when the underlying DEM is not available. Furthermore, we show that orthorectification errors over Helheim Glacier are consistent with Sentinel-2 data being orthorectified using data sources present in the Viewfinder Panorama 3" DEM (de Ferranti, 2014). This finding agrees with that of Kääb *et al.* (2016), who found similar agreement between a Δh estimated from the Viewfinder DEM and Sentinel-2 offsets for a test site in Northern Norway. Further testing is required to establish if the Viewfinder DEM and PlanetDEM continue to share high-latitude data sources, but if they do then the freely available Viewfinder DEM opens new avenues for deriving high-latitude Sentinel-2 orthorectification error from analytical methods (Ressl & Pfeifer, 2018). To do so would require not only the Viewfinder DEM but a 'true' glacier surface elevation DEM contemporaneous to image acquisition, which is difficult to acquire in rapidly-changing sectors of the cryosphere. Here, we highlight the utility of time-evolving DEMs and the potential of e.g. ArcticDEM strips (Porter *et al.* 2018) and monthly/annual composites thereof to address this challenge in the future.

4.3 Future datasets

The empirical offsets derived in this study are only applicable up to the 23rd August 2021. On this date (or the 30th March 2021 for coverage of Europe and Africa), L1C/L2A Sentinel-2 products switched to an improved geometric refinement with two major changes: (i) co-registration of scenes to a Global Reference Image (GRI); (ii) the use of a new DEM, the Copernicus DEM at 90 m resolution ("GLO-90"), for topographic correction. This new DEM is based on 2011 – 2015 radar satellite data acquired during the TanDEM-X mission, meaning there will be a discontinuity in Δh between the pre- and post-August 2021 data.

There is no currently announced reprocessing of the Sentinel-2 archive with the new geometric refinements (cf., for example, Landsat Collection 2), so this discontinuity between pre- and post-2021 data will remain into the future. As such, the method outlined here remains applicable for the correction of Sentinel-2-derived glacier velocity across the wider cryosphere for Sentinel-2 L1C and L2A data between 2015 – 2021. The more recent dataset will reduce the value of Δh and as such the orthorectification error in the new datasets - although as the DEM used for orthorectification consists of pre-2015 data, some orthorectification error will remain. As such, offset fields such as those calculated in this study will need to be calculated separately for the pre- and post-2021 data. Additionally, the lower initial Δh value will likely invalidate the assumption we make in this study that further elevation change will have a negligible impact on our assumption of stable geometry. However, GLO-90 is available for public download (European Space Agency, 2021), allowing for correction via analytical methods. In combination with the opportunities we highlight in Section 4.2, this provides framework for the production of a continuous Sentinel-2 velocity dataset from 2016-present with cross-track velocities corrected analytically according to their underlying DEM.

5 Conclusion

By taking advantage of the complete Sentinel-2 record between 2017 – 2021, we demonstrate an empirical method of correcting for systematic orthorectification error in glacier velocity fields derived from cross-track image pairs in large-scale datasets. The method is simple and computationally undemanding, whilst allowing for spatially continuous glacier velocity fields to be constructed without limitations introduced by satellite and flow geometry. This method is used to produce a complete dataset of glacier velocity for four key marine-terminating glaciers across the Greenland Ice Sheet. Comparison with alternative datasets highlights the key advantages of the high temporal frequency of Sentinel-2 imagery, but this density of data also necessitates the use of statistical techniques to account for noise and uncertainty in the dataset. Furthermore, we use this method to identify a likely publicly available data source for the DEM data used to correct Sentinel-2 scenes over Greenland between 2015 – 2021, which will be useful for future studies concerned with orthorectification error in the region. The transition to a new geometric refinement method for Sentinel-2 scenes beyond August 2021 will require new correction offsets to be generated, but the transition to a publicly available DEM (and the identification of 2015 – 2021 data sources in this study) provides the opportunity for this to occur through analytical rather than empirical methods.

Data availability. The velocity datasets described here will be able to be accessed through the NSIDC upon final publication. The NSIDC also hosts access to the supplementary data used in this work: ITS_LIVE data (doi:10.5067/IMR9D3PEI28U), MEaSUREs Sentinel-1 velocity data (doi:10.5067/6JKYGMOZQFYJ), GIMP DEM (doi:10.5067/H0KUYVF53Q8M), and the GIMP mask (doi:10.5067/B8X58MQBFUPA). Version 1 of the 3” Viewfinder Panorama DEM for Greenland can be found at http://viewfinderpanoramas.org/GL-ReadMe.html. Sentinel-2 scenes were accessed via the AWS Registry of Open Data (https://registry.opendata.aws/sentinel-2). Velocity processing was performed using SETSM, available at https://github.com/setsmdeveloper/SETSM.

Author contributions. TRC: conceptualisation, methodology, investigation, formal analysis, investigation, writing – original draft, visualisation. IMH: conceptualisation, methodology, writing – review & editing, supervision, project administration, funding acquisition. BNY: investigation, resources, data curation, writing – review and editing. MJN: investigation, methodology, software, writing – review and editing.

Competing interests. The authors declare they have no conflict of interest.

Acknowledgements. This project was supported by grants 80NSSC18M0078 and 80NSSC18K1027 from the National Aeronautics and Space Administration programme. The velocity fields contain modified Copernicus Sentinel data (2017–2021), processed by ESA. We are grateful to the scientific editor, Etienne Berthier, as well as reviewers Jeremie Mouginot and Bas Altena for the constructive remarks on the manuscript.

References

Altena, B., Haga, O. N., Nuth, C., and Kääb, A.: Monitoring sub-weekly evolution of surface velocity and elevation for a high-latitude surging glacier using sentinel-2, XLII-2/W13, 1723–1727, https://doi.org/10.5194/isprs-archives-XLII-2-W13-1723-2019, 2019.

Altena, B. and Kääb, A.: Elevation Change and Improved Velocity Retrieval Using Orthorectified Optical Satellite Data from Different Orbits, Remote Sens, 9, 300, https://doi.org/10.3390/rs9030300, 2017.

Dehecq, A., Gourmelen, N., Gardner, A. S., Brun, F., Goldberg, D., Nienow, P. W., Berthier, E., Vincent, C., Wagnon, P., and Trouvé, E.: Twenty-first century glacier slowdown driven by mass loss in High Mountain Asia, Nature Geosci, 12, 22–27, https://doi.org/10.1038/s41561-018-0271-9, 2019.

European Space Agency: Copernicus GLO-90 Digital Surface Model, https://doi.org/10.5069/G9028PQB, 2021.

de Ferranti, J.: Viewfinder Panorama 3" DEM, http://www.viewfinderpanoramas.org/dem3.html, 2014.

Friedl, P., Seehaus, T., and Braun, M.: Global time series and temporal mosaics of glacier surface velocities derived from Sentinel-1 data, Earth Syst Sci Data, 13, 4653–4675, https://doi.org/10.5194/essd-13-4653-2021, 2021.

Gardner, A. S., Moholdt, G., Scambos, T., Fahnstock, M., Ligtenberg, S., van den Broeke, M., and Nilsson, J.: Increased West Antarctic and unchanged East Antarctic ice discharge over the last 7 years, Cryosphere, 12, 521–547, https://doi.org/10.5194/tc-12-521-2018, 2018.

Howat, I.: MEaSURES Greenland Ice Velocity: Selected Glacier Site Velocity Maps from Optical Imagery, Version 3, https://doi.org/10.5067/RRFY5IW94X5W, 2020.

Howat, I., Negrete, A., and Smith, B.: MEaSUREs Greenland Ice Mapping Project (GIMP) Digital Elevation Model from GeoEye and WorldView Imagery, Version 1, https://doi.org/10.5067/H0KUYVF53Q8M, 2017.

Howat, I. M., Negrete, A., and Smith, B. E.: The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets, Cryosphere, 8, 1509–1518, https://doi.org/10.5194/tc-8-1509-2014, 2014.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., and Kääb, A.: Accelerated global glacier mass loss in the early twenty-first century, 592, 726–731, https://doi.org/10.1038/s41586-021-03436-z, 2021.

Joughin, I.: MEaSUREs Greenland 6 and 12 day Ice Sheet Velocity Mosaics from SAR, Version 1, https://doi.org/10.5067/6JKYGMOZQFYJ, 2021a.

Joughin, I.: MEaSUREs Greenland Ice Velocity Annual Mosaics from SAR and Landsat, Version 3, https://doi.org/10.5067/C2GFA20CXUI4, 2021b.

Joughin, I., Smith, B. E., and Howat, I.: Greenland Ice Mapping Project: ice flow velocity variation at sub-monthly to decadal timescales, Cryosphere, 12, 2211–2227, https://doi.org/10.5194/tc-12-2211-2018, 2018.

Kääb, A., Winsvold, S. H., Altena, B., Nuth, C., Nagler, T., and Wuite, J.: Glacier Remote Sensing Using Sentinel-2. Part I: Radiometric and Geometric Performance, and Application to Ice Velocity, Remote Sens, 8, 598, https://doi.org/10.3390/rs8070598, 2016.

King, M. D., Howat, I. M., Jeong, S., Noh, M. J., Wouters, B., Noël, B., and van den Broeke, M. R.: Seasonal to decadal variability in ice discharge from the Greenland Ice Sheet, Cryosphere, 12, 3813, 2018.

King, M. D., Howat, I. M., Candela, S. G., Noh, M. J., Jeong, S., Noël, B. P. Y., van den Broeke, M. R., Wouters, B., and Negrete, A.: Dynamic ice loss from the Greenland Ice Sheet driven by sustained glacier retreat, Commun Earth Environ, 1, 1–7, https://doi.org/10.1038/s43247-020-0001-2, 2020.

Mankoff, K. D., Colgan, W., Solgaard, A., Karlsson, N. B., Ahlstrøm, A. P., van As, D., Box, J. E., Khan, S. A., Kjeldsen, K. K., Mouginot, J., and Fausto, R. S.: Greenland Ice Sheet solid ice discharge from 1986 through 2017, Earth Syst Sci Data, 11, 769–786, https://doi.org/10.5194/essd-11-769-2019, 2019.

Moon, T., Joughin, I., Smith, B., van den Broeke, M. R., van de Berg, W. J., Noël, B., and Usher, M.: Distinct patterns of seasonal Greenland glacier velocity, Geophys Res Lett, 41, 7209–7216, https://doi.org/10.1002/2014GL061836, 2014.

Morlighem, M., Bondzio, J., Seroussi, H., Rignot, E., Larour, E., Humbert, A., and Rebuffi, S.: Modeling of Store Gletscher's calving dynamics, West Greenland, in response to ocean thermal forcing, Geophys Res Lett, 43, 2659–2666, https://doi.org/10.1002/2016GL067695, 2016.

Nagy, T., Andreassen, L. M., Duller, R. A., and Gonzalez, P. J.: SenDiT: The Sentinel-2 Displacement Toolbox with Application to Glacier Surface Velocities, Remote Sens, 11, 1151, https://doi.org/10.3390/rs11101151, 2019.

Noh, M.-J. and Howat, I. M.: Applications of High-Resolution, Cross-Track, Pushbroom Satellite Images With the SETSM Algorithm, IEEE J Sel Top Appl, 12, 3885–3899, https://doi.org/10.1109/JSTARS.2019.2938146, 2019.

Poinar, K. and Andrews, L. C.: Challenges in predicting Greenland supraglacial lake drainages at the regional scale, Cryosphere, 15, 1455–1483, https://doi.org/10.5194/tc-15-1455-2021, 2021.

Porter, C., Morin, P., Howat, I., Noh, M.-J., Bates, B., Peterman, K., Keesey, S., Schlenk, M., Gardiner, J., Tomko, K., Willis, M., Kelleher, C., Cloutier, M., Husby, E., Foga, S., Nakamura, H., Platson, M., Wethington, M., Williamson, C., Bauer, G., Enos, J., Arnold, G., Kramer, W., Becker, P., Doshi, A., D'Souza, C., Cummens, P., Laurier, F., and Bojesen, M.: ArcticDEM, 10.7910/DVN/OHHUKH, 2018.

Ressl, C. and Pfeifer, N.: Evaluation of the elevation model influence on the orthorectification of Sentinel-2 satellite images over Austria, Eur J Remote Sens, 51, 693–709, https://doi.org/10.1080/22797254.2018.1478676, 2018.

Rosenau, R., Scheinert, M., and Dietrich, R.: A processing system to monitor Greenland outlet glacier velocity variations at decadal and seasonal time scales utilizing the Landsat imagery, Remote Sensing of Environment, 169, 1–19, https://doi.org/10.1016/j.rse.2015.07.012, 2015.

Smith, B., Fricker, H. A., Gardner, A. S., Medley, B., Nilsson, J., Paolo, F. S., Holschuh, N., Adusumilli, S., Brunt, K., Csatho, B., Harbeck, K., Markus, T., Neumann, T., Siegfried, M. R., and Zwally, H. J.: Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes, 368, 1239–1242, https://doi.org/10.1126/science.aaz5845, 2020.

Solgaard, A., Kusk, A., Merryman Boncori, J. P., Dall, J., Mankoff, K. D., Ahlstrøm, A. P., Andersen, S. B., Citterio, M., Karlsson, N. B., Kjeldsen, K. K., Korsgaard, N. J., Larsen, S. H., and Fausto, R. S.: Greenland ice velocity maps from the PROMICE project, Earth Syst Sci Data, 13, 3491–3512, https://doi.org/10.5194/essd-13-3491-2021, 2021.

Vijay, S., King, M. D., Howat, I. M., Solgaard, A. M., Khan, S. A., and Noël, B.: Greenland ice-sheet wide glacier classification based on two distinct seasonal ice velocity behaviors, J Glaciol, 1–8, https://doi.org/10.1017/jog.2021.89, 2021.