Cross-Correlation Stacking for Robust Offset Tracking Using SAR Image Time-Series

Shiyi Li, *Graduate Student Member, IEEE*, Silvan Leinss, *Senior Member, IEEE*, and Irena Hajnsek, *Fellow, IEEE*

***Abstract*—Offset tracking is widely applied for measuring ground surface displacements from remote sensing data. Displacements are determined by the offset where two image templates match best. The match can be evaluated with normalized cross correlation (NCC), in which the height and location of the NCC peak represent the matching quality and the corresponding offset. Attaining robust offset estimations requires an unambiguous tracking of the peak in the NCC noise floor. To improve offset estimations, we propose a cross-correlation stacking method that can significantly suppress the noise floor of NCC. Instead of deriving offsets from each pairwise NCC, we stack a series of consecutive pairwise NCCs and determine the offset after averaging the NCC stack. Thereby, tracking benefits from the redundant information in multiple NCCs and is more robust to noise. We assessed the method by measuring the flow velocity of the Great Aletsch Glacier in Switzerland using image time series collected by the synthetic aperture radar satellites TanDEM-X and Sentinel-1 A. Using relatively small templates of 48×48 pixels combined with a stack of seven pairwise NCCs of TanDEM-X images, we obtain velocity fields whose spatial coverage are almost equivalent to the coverage of velocity fields obtained with templates of 96×96 pixels applied on a single image pair. Similar improvements in spatial coverage are observed for Sentinel-1 A. The results demonstrate that the stacking method can greatly enhance both the spatial resolution and the coverage of the obtained velocity fields.**

***Index Terms*—Cross correlation, glacier velocity, offset tracking, stacking, synthetic aperture radar (SAR).**

I. Introduction

MEASURING ground surface displacement using either optical or synthetic aperture radar (SAR) images is of interests to many environment-related studies, and two mostly used methods for this purpose are offset tracking and differential interferometric SAR. The focus of this article is offset tracking, which has been successfully applied to many studies such as detecting landslides [1]–[3], estimating coseismic slips [4], [5], and measuring glacier surface velocities [6]–[15].

Offset tracking is a template matching method which can be robustly applied to either optical or SAR images [6], [16]. It is often referred as feature tracking when applied to optical images [17]. For SAR images, it is also called coherence tracking when using SAR single-look complex images, or intensity tracking when only using real-valued intensity images [6]. The concept of offset tracking is to estimate displacement offsets through measuring the similarity between two images taken over the same area and at different times. In practice, image pairs are first partitioned into templates, and then the agreement between two templates are calculated based on certain similarity measures, such as least square difference [18], maximum likelihood [19], and normalized cross correlation (NCC) [16], [20]. The templates are shifted against each other until the best agreement is found, for which the shift defines the estimated offset.

So far, plenty of efforts have been made in order to improve offset tracking. The first family of improvement are mainly focused on the image preprocessing steps. For instance, Leprince *et al.* [17] proposed a software named COSI-CORR to achieve automatic and precise orthorectification, coregistration, and subpixel correlation for optical images; Lange *et al.* [21] suggested applying spatial high-pass filters to SAR imagery before cross correlation to focus the offset tracking on small surface features; Debella-Gilo and Kääb[22] developed a technique to locally optimize template sizes for image partition.

Techniques for the postprocessing of offset fields have also been proposed in previous studies, which mainly exploit redundant offset measurements to attain robust offset fields (e.g., smoothing, detecting outliers, median filters). Ahn and Howat [23] proposed a multiple-image multiple-chip algorithm, in which multiple NCCs are generated with template pairs under different configurations (i.e., template sizes, convolution filters, etc.), and a population of offset vectors are obtained from every individual pairwise NCC so that a final offset vector can be voted from the population. Similarly, Stumpf *et al.* [3] proposed a multiple pairwise image correlation technique, which generates multiple offsets for each pixel location using pairwise NCC and then summarizes the stack of offsets using developed indicators.

The previous attempts have greatly pushed forward the performance of offset tracking for displacement measurement. These works, focusing either on preprocessing of images or on postprocessing of offset fields, derive all information from a pairwise NCC and use the apparent peak of the NCC field to estimate offsets [16]. Unfortunately, for a pairwise NCC with low signal-to-noise ratio (SNR), the apparent peak could be found at a false location when the NCC is dominated by noise.

Manuscript received November 18, 2020; revised February 5, 2021 and March 25, 2021; accepted April 2, 2021. Date of publication April 9, 2021; date of current version May 26, 2021. This work was supported by ETH Zurich. *(Corresponding authors: Shiyi Li; Silvan Leinss.)*

Shiyi Li and Silvan Leinss are with the Institute of Environmental Engineering, ETH Zürich, 8093 Zürich, Switzerland (e-mail: shiyi.li@ifu.baug.ethz.ch; leinss@ifu.baug.ethz.ch).

Irena Hajnsek is with the Institute of Environmental Engineering, ETH Zürich, 8093 Zürich, Switzerland, and also with the German Aerospace Center (DLR), Microwaves and Radar Institute, 82234 Wessling, Germany (e-mail: irena.hajnsek@dlr.de).

Digital Object Identifier 10.1109/JSTARS.2021.3072240

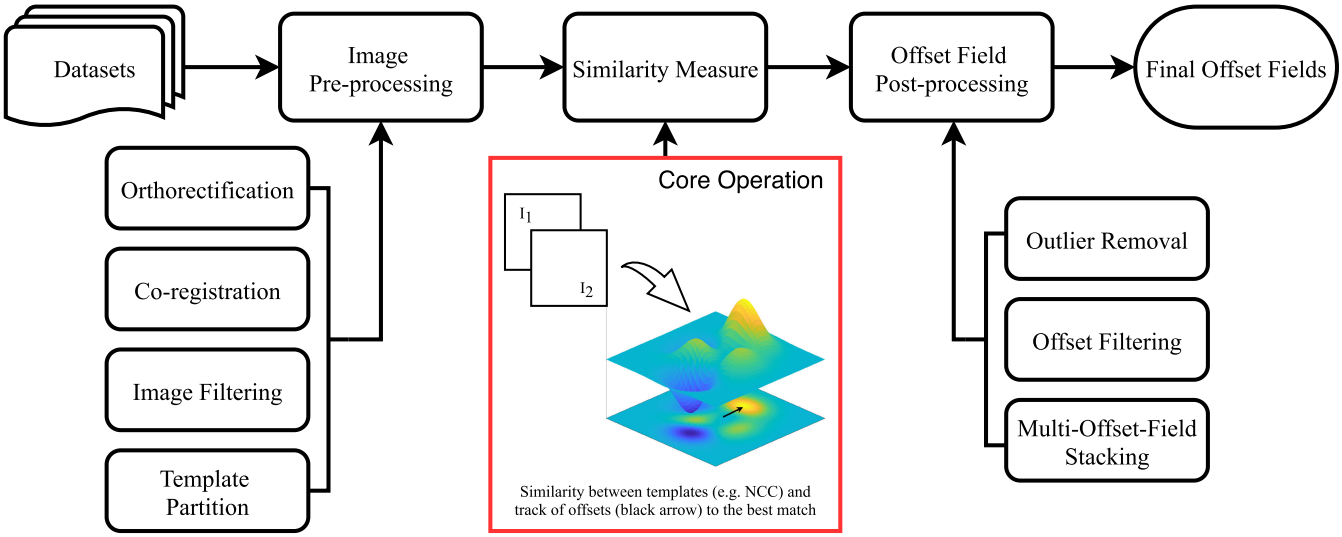


Fig. 1. Common procedure for offset tracking. The core operation of offset tracking is highlighted with the red box.

Even though redundant offsets estimated from multiple pairwise NCCs can be used to reduce false detection (e.g., [3]), directly separating the time-invariant NCC component from noise has almost never been explored for SAR offset tracking.

In this article, we propose a method to improve the performance of SAR incoherent offset tracking by first creating a stack of pairwise NCCs from an image time series and then averaging the NCC stack for offset estimation. Thereby, we make the NCC more robust for tracking by exploiting the entire information of the NCC to reduce the noise floor and to enhance the SNR. The concept has been adopted for particle image velocimetry [24], [25], for medical image tracking [26], [27], and recently for ice fall tracking with optical imagery [28]. In these works, different terms were used for the concept, including mean cross correlation, ensemble tracking, ensemble correlation, or ensemble matching. We prefer to refer to the method as "cross-correlation stacking" (or specifically "NCC stacking"), because these words can best represent the core operation of the method and correspond to a widely used terminology for SAR image time-series analysis [29]–[32]. So far, NCC stacking has not been adopted for SAR images, where we expect significant improvement by reduction of speckle noise.

In this article, we described the NCC stacking method for SAR imagery, and assessed its effectiveness by measuring the flow velocities of the Great Aletsch Glacier in Switzerland using SAR image series collected by the satellites TanDEM-X [33] and Sentinel-1 [34].

II. Methodology

A. NCC Stacking

1) NCC Decomposition: Classical offset tracking algorithms commonly use NCC to evaluate the similarity between a pair of image templates with a workflow similar to Fig. 1. The NCC can be calculated in either the spatial domain or the frequency domain. In the frequency domain the NCC between the two templates I_1 and I_2, extracted from two SAR intensity images taken over the same area and at different times, can be conveniently calculated as

$$\hat{\gamma} = \frac{\hat{I}_1^* \cdot \hat{I}_2}{\sqrt{E\{|\hat{I}_1|^2\} \cdot E\{|\hat{I}_2|^2\}}} \tag{1}$$

with $\hat{I}_i = \mathcal{F}(I_i)$ the Fourier transform of the template I_i, * the complex conjugation, and $E\{|\hat{I}_1|^2\}$ and $E\{|\hat{I}_2|^2\}$ the total energy of the two templates. The obtained spectrum $\hat{\gamma}$ is transformed back to obtain the NCC in the spatial domain

$$\gamma = \mathcal{F}^{-1}(\hat{\gamma})$$

where the offset is estimated by tracking the location of the peak in γ.

Robust tracking relies on unambiguous identification of the NCC peak, which requires the template pair to contain abundant trackable features, such as rigidly shifted geometric structures. On the other hand, noise that hampers the successful offset estimation is added to the NCC field by temporally uncorrelated contents, such as incoherent speckle, nonrigidly translated features (e.g., collapsing crevasses, distorted landscape), and temporally changing surface properties (e.g., snow cover, water content, or vegetation change).

From the perspective of offset tracking, an image template, and also its Fourier transform, can be considered as being composed of signal $\hat{s}$ and noise $\hat{n}$

$$\hat{I} = \hat{s} + \hat{n}.$$

The signal $\hat{s}$ corresponds to the correlated content (i.e., $\hat{s}_1 = \hat{s}_2$) that generates the signal of the NCC, whereas the noise $\hat{n}$ represents the uncorrelated content. Then, we can rewrite (1) as

$$\hat{\gamma} = \frac{\hat{s}_1^* \hat{s}_2 + \hat{s}_1^* \hat{n}_2 + \hat{n}_1^* \hat{s}_2 + \hat{n}_1^* \hat{n}_2}{\sqrt{E\{|\hat{I}_1|^2\} \cdot E\{|\hat{I}_2|^2\}}} \tag{2}$$

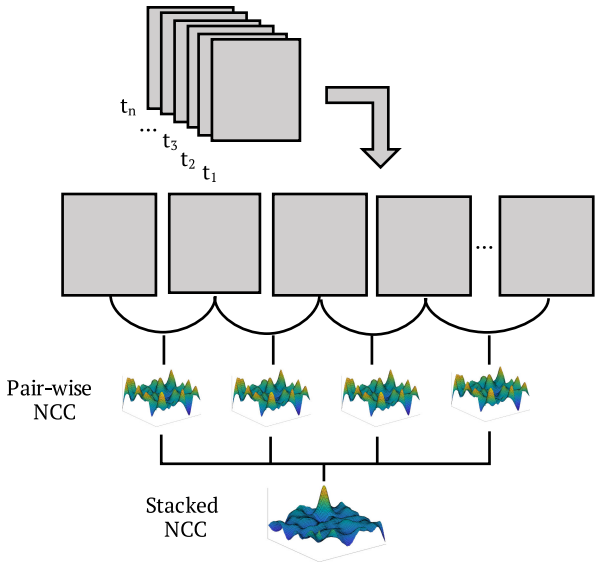


Fig. 2. Conceptual workflow of the NCC stacking method. With the acquired image time-series, an NCC stack is obtained by calculating pair-wise NCCs consecutively, and then the stack is averaged to track the peak. The shown stacking workflow can be directly inserted into the core operation in Fig. 1(red box) to replace the pairwise NCC.

which decomposes $\hat{\gamma}$ into four elementary cross correlations, including a shifted autocorrelation $\hat{s}_1^* \hat{s}_2$, and three cross correlations between signals and noise ($\hat{s}_1^* \hat{n}_2$, $\hat{n}_1^* \hat{s}_2$, and $\hat{n}_1^* \hat{n}_2$). Since the noise in one template is independent from either the noise or the signal in the other template, we can consider the latter three components as noise contributing to $\hat{\gamma}$. Therefore, in order to attain reliable offset estimations, the Fourier transform of $\hat{s}_1^* \hat{s}_2$ must present a dominant peak in γ that greatly surpasses the superposition of all noise components. The dominance of this peak can be characterized by the SNR of the NCC, defined as

$$\text{SNR} = 10 \log_{10} \left(\frac{c_p^2}{\overline{c_f^2}} \right) \tag{3}$$

with c_p the NCC peak height and $\overline{c_f^2}$ the average of the squared ambient NCC field γ. To ensure reliable tracking, the SNR should be sufficiently high. Otherwise the tracking delivers false results because the peak that represents the true offset can not reveal itself from the noise floor (i.e., the SNR is too low).

2) NCC Stacking Using Image Time-Series: To make offset tracking more robust against noise, we can create a stack of pairwise NCCs using image time-series and then average the NCC stack to suppress the noise floor. An overview of our general workflow is presented in Fig. 2.

Suppose a short time-series of N coregistered image is collected with a repeat interval of t. With integer multiples a of the repeat interval t, we can calculate a stack of $N - a$ pairwise NCCs from (1) using all template pairs with a time interval of $T = at$. In this stack, each pairwise NCC captures a snapshot of the displacement represented by the template pair. Assuming constant surface velocities during acquisition of the time-series, we can expect that all pairwise NCCs record identical offsets. Hence, it can be presumed that the peaks generated by $\hat{s}_1^* \hat{s}_2$ in all pairwise NCCs γ are located at more or less the same position, whereas the ambient noise field, generated by $\hat{s}_1^* \hat{n}_2 + \hat{n}_1^* \hat{s}_2 + \hat{n}_1^* \hat{n}_2$, averages out in the stack.

This principle can be easily extended to multiple image time-series, such as image series acquired in multiple different years, in different spectral or polarimetric channels, or even from different imaging sensors. For that, suppose M image time-series are collected and each time-series i consists of N_i ($i = 1 \ldots M$) images, we can then create a total stack that consists of M substacks of pairwise NCCs, each of which consists of $N_i - a$ pairwise NCCs. Assuming identical surface velocities for all individual time-series, we can average the total stack of pairwise NCCs in spatial domain to get the averaged NCC in which the offset is tracked

$$\overline{\gamma} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N_i - a} \gamma_{i,j}}{\sum_{i=1}^{M} (N_i - a)}. \tag{4}$$

Particularly, seasonal velocity variations, or the so-called "phase-averaged velocity fields" [25], can be resolved with (4) when substacks for each phase of a year (e.g., months) are averaged over multiple years. Equation (4) also permits to track step-like displacement (e.g., coseismic slips) using a stack of correlation pairs in which each pair is composed of pre- and postevent images.

B. Offset Tracking With Subpixel Precision

In practice, precise tracking requires subpixel precision, which is achieved by first fitting the NCC peak with a continuous function (e.g., polynomial or Gaussian function), and then finding the location of the maximum of the fitted function [35]. In this article, we used a Gaussian function $g(x, y)$ in the form of (5) to fit the NCC peak

$$g(x, y) = A e^{(-P(x-x_0)^2 + Q(x-x_0)(y-y_0) + R(y-y_0)^2)} + b \tag{5}$$

where

$$P = \frac{\cos^2 \theta}{2\sigma_x^2} + \frac{\sin^2 \theta}{2\sigma_y^2}$$

$$Q = \frac{-\sin 2\theta}{2\sigma_x^2} + \frac{\sin 2\theta}{2\sigma_y^2}$$

$$R = \frac{\sin^2 \theta}{2\sigma_x^2} + \frac{\cos^2 \theta}{2\sigma_y^2}.$$

The Gaussian function $g(x, y)$ is parameterized by its center location (x_0, y_0), standard deviation (σ_x, σ_y), the maximum value A, the rotation parameter θ, and the vertical shift b. In order to fit the Gaussian function, we first extracted a window Ω of a few pixels size (see Section III-B) centered at the NCC peak, and then up-sampled the extracted window by a factor of ten using bilinear B-Spline interpolation. Within the up-sampled window Ω', we solve a nonlinear least-squares problem to minimize the

cost function

$$F = \frac{1}{2} \sum_{(x,y)\in\Omega'} (\mathrm{NCC}(x,y) - g(x,y))^2 \tag{6}$$

with x, y coordinates of the up-sampled window Ω'. To ensure convergence, the window size for extraction and initial value of parameters must be carefully selected. In this article, we used the following values for initialization:

1) (x_0, y_0): the center location of the up-sampled window;
2) (σ_x, σ_y): 1/4 of the width and height of the up-sampled window respectively;
3) A: the NCC peak height;
4) θ: zero;
5) b: mean value of the ambient field.

For successful fits, the peak position $\vec{p}_{\mathrm{fit}} = (x_{0,\mathrm{fit}}, y_{0,\mathrm{fit}})$ is used. If minimization fails, the initial values $\vec{p} = (x_0, y_0)$ are kept. Once the peak location is determined, the offset vector $\vec{D} = (D_x, D_y)$ can be estimated by measuring the shift from the NCC center $\vec{c} = (x_c, y_c)$ to the peak with $\vec{D} = \vec{p} - \vec{c}$, and the obtained offset vector field can be converted to a velocity field by

$$\vec{v} = \frac{\vec{D}}{T}. \tag{7}$$

C. Evaluating the Obtained Velocity Field

To quantify the performance gain between the stacked and the pairwise NCC, we calculated the following three parameters from the obtained velocity fields. As these parameters vary strongly for different image pairs, we evaluated for each stack size all possible combinations of correlation pairs and obtained mean, minimum, and maximum values for the spatial coverage and the velocity residuals.

1) Spatial Coverage: In offset tracking applications, the obtained velocity field is often postprocessed by removing outliers and unreliable estimations. In this article, we used two thresholds for this purpose, one is based on the maximum velocity magnitude and the other is based on the minimum acceptable SNR (values are defined in Section III-B). Thresholding the velocity field leads to voids in the velocity maps, and this allows us to evaluate the robustness of different methods by measuring the spatial coverage R_{cov} of the velocity map as

$$R_{\mathrm{cov}} = \frac{A_{\mathrm{map}}}{A_{\mathrm{total}}} \tag{8}$$

with A_{map} the area of the map excluding voids and A_{total} the total area of the study region (black glacier outline in Fig. 3).

2) Velocity Residuals: The accuracy of offset tracking can be evaluated by examining velocity residuals over static ground, assuming that velocities of such region should be equal to zero. For that, we define the residual ratio within selected static regions as

$$R_{\mathrm{res}} = \frac{A_{\mathrm{residual}}}{A_{\mathrm{static}}} \tag{9}$$

with A_{residual} the area covered by residual velocities within the static ground and A_{static} the total area of the static ground (red dashed rectangles in Fig. 3). Particularly, we define residual velocity as the velocity vector whose corresponding offsets larger than one pixel in either $x-$ or $y-$ directions.

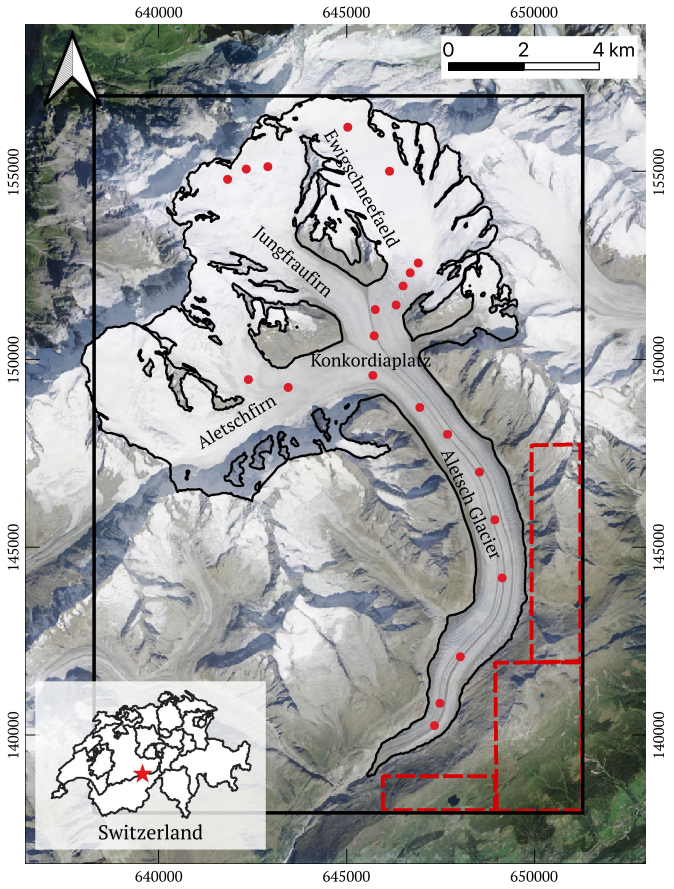


Fig. 3. Overview of the Great Aletsch Glacier. The glacier outline is delineated in black according to [36]. The area within the black glacier outline was used to calculate the spatial coverage of successful velocity estimates. Locations of *in situ* velocity measurements by GPS are marked with red dots. Glacier-free regions used for residual velocity evaluation are indicated by red dashed rectangles. The processed area of the SAR images is shown by the black rectangle. The approximate location of the Great Aletsch Glacier within Switzerland is indicated by the star in the inset. The base image is taken from SWISSIMAGE 25, 2017 2019 swisstopo (JD100042).

3) SNR Gain: Reliable tracking can be characterized by the SNR value. To evaluate the improvement of confidence level on the obtained velocity field, we can calculate the SNR gain by

$$\mathrm{SNR}_{\mathrm{gain}} = \mathrm{SNR}_{\mathrm{stack}} - \mathrm{SNR}_{\mathrm{pair}}. \tag{10}$$

With the SNR gain, it is possible to examine the change of SNR at every template location in the image scene. Specifically, $\mathrm{SNR}_{\mathrm{gain}} > 0$ indicates increased confidence level, whereas $\mathrm{SNR}_{\mathrm{gain}} < 0$ indicates decreased confidence level.

III. Study Site and Data

A. Study Site

To evaluate NCC stacking, we derive the surface velocity of the Great Aletsch Glacier located in central Switzerland. The glacier represents one of the largest alpine glacier systems

TABLE I
ACQUISITION FOR THIS ARTICLE

	TanDEM-X	Sentinel-1A	
Acquisition Dates	2017-01-10	2017-01-03	2018-01-10
	2017-01-21	2017-01-15	2018-01-22
	2017-02-01	2017-02-08	2018-02-03
	2017-02-12	2017-02-20	2018-02-15
	2017-02-23	2017-02-27	2018-02-27
	2017-03-06	2017-03-04	2018-03-11
	2017-03-17	2017-03-16	2018-03-23
	2017-03-28	2017-03-28	

in the European Alps (see Fig. 3). The Great Aletsch glacier consists of three main tributaries, including the Aletschfirn, the Jungfraufirn, and the Ewigschneefaeld, and these tributaries merge at Konkordiaplatz and flow into the main stream of the Great Aletsch Glacier.

The Great Aletsch glacier has been used as a test site for many studies to explore methods for glacier surface velocity estimations. For instance, Prats *et al.* [37] used airborne interferometric repeat pass SAR data in L- and P-band to measure the surface velocity field around Konkordiaplatz; Erten *et al.* [38] proposed a method based on polarimetric similarity measure for velocity tracking and tested it here; Schubert *et al.* [39] compared a wavelet-based and correlation-based image matching methods for glacier velocity retrieval using repeat TerraSAR-X stripmap and spotlight images acquired in August 2009 and obtained velocity estimates over the strongly crevassed area (ice fall, tongue) of Aletsch Glacier.

An almost complete velocity map has been generated by [40] based on pairwise cross correlation using a time series of about 130 TanDEM-X acquisitions. They also provided the *in situ* GPS velocity measurement at 22 locations on the glacier (see red dots in Fig. 3). In our study, we used these GPS measurement to validate our velocity results. For that we calculated the root-mean-square error (RMSE) as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_i^{\mathrm{track}} - v_i^{\mathrm{meas}})^2} \tag{11}$$

with v_i^{track} one component (x, y) of the tracked velocity, v_i^{meas} the corresponding measured velocity, and n the total number of velocity points that is kept after thresholding.

B. SAR Data, Preprocessing, and Thresholding

To test the proposed stacking method, we processed two distinct SAR datasets. The first dataset was collected by TanDEM-X, which is a high-resolution interferometric X-band SAR mission launched at June 21, 2010 [33]. Our dataset is a subset of the dataset used in [40] and consists of eight dual-pol stripmap images acquired between 2017-1-10 and 2017-3-28 with equal repeat interval of 11 days (see Table I). All acquisitions were made from orbit 154 (descending) at an incident angle of 32°, resulting in a ground range resolution of 6.6×2.2 m (az x rg) [41]. The second dataset is collected from Sentinel-1 A, which is a C-band SAR mission launched by European Space Agency in April, 2014 [34]. This dataset includes two image series acquired from the same orbit (66, descending). One consists of eight images acquired between 2017-1-3 and 2017-3-28, and the other consists of seven images acquired between 2018-1-22 and 2018-3-23 (see Table I). All images were acquired in Interferometric Wide Swath mode with dual polarization (VV and VH) and at identical incident angle of 41°, resulting in a resolution of 22×4.7 m (az x rg) [42]. Acquisitions were made with equal repeat interval of 12 days within each time series.

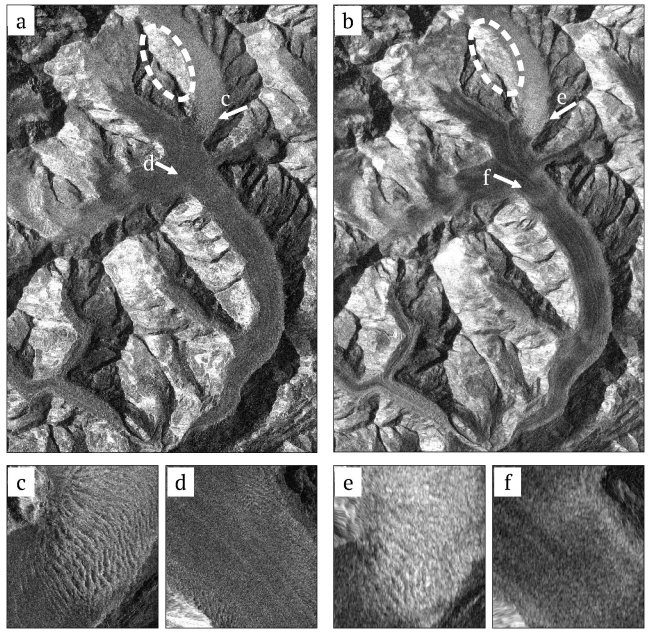


Fig. 4. Example orthorectified SAR images taken from (a) TanDEM-X and (b) Sentinel-1 A. No smoothing filter was applied for speckle removal to preserve the spatial resolution. The zoom-in figures (c)– (f) show the effect of different resolution. Very bright mountain slopes are affected by layover (e.g., the east-facing slope of Ewigschneefaeld as indicated by the dashed oval).

For each sensor, the images were coregistered to a common reference scene followed by orthorectification using the SwissAlti3D elevation model obtained from the Federal Office of Topography (swisstopo). As all data were acquired in two polarizations, we averaged the intensity of the two polarizations to reduce SAR speckle. For Sentinel-1 A we weighted the VH polarization by the ratio of mean intensities, $\langle \mathrm{VV} \rangle / \langle \mathrm{VH} \rangle$; for TanDEM-X, we did not apply any weighting because the VV and HH polarization have very similar backscatter intensities. We did not apply oversampling of the amplitude before intensity calculation [43] as both dataset are already oversampled compared to their native radar pixel spacing. No multilooking was applied to neither of the datasets, though the two interpolation steps of coregistration and orthorectification can be considered as multilooking to some grade. After the preprocessing, the pixel spacing of TanDEM-X imagery is 2×2 m, and that of Sentinel-1 A is 5×5 m. All images were converted to log-scale before cross correlation. Examples of SAR images from the two datasets are presented in Fig. 4. The figure illustrates how different image resolution impacts visible surface feature details [see Fig. 4(c) and (d)].

Before cross correlation, we applied high-pass filters to both datasets to better focus on small scale features [21]. The high-pass filter is implemented using a Gaussian kernel whose size is 51×51 pixels and standard deviation is 17×17 pixels. To study the influence of template sizes on the behavior of the stacked NCC, we used a large template size of 96×96 pixels and a small template size of 48×48 pixels for image partition. For tracking the NCC peak with subpixel accuracy, extraction windows of 11×11 pixels and 7×7 pixels were used for TanDEM-X and Sentinel-1 A images, respectively. The chosen window sizes ensured that the curve fitting procedure successfully converged at 99% of the templates for the two datasets.

For TanDEM-X data, image pairs were grouped with time interval of 11 days ($a = 1$) to calculate pairwise NCCs, and thus, we obtained in total seven pairwise NCCs from the acquisitions listed in Table I. For Sentinel-1 A, due to its larger pixel sizes compared to TanDEM-X, the time interval used to form image pairs was set to be 24 days ($a = 2$) to accommodate larger offsets. Therefore, we obtained in total 11 pairwise NCCs for Sentinel-1 A imagery, including six from the time-series of 2017 and five from the time-series of 2018.

For both datasets, considering the maximum velocity measured in previous studies [37], [38], [40], the threshold of maximum velocity magnitude was set to $v_{\max} = 1\ \mathrm{m \cdot d^{-1}}$. Higher velocities were considered as outliers and were removed. Threshold on minimum SNR for unreliable estimation removal was set as $\mathrm{SNR_{min}} = 10$ dB. These thresholds were not optimized for glacier velocity filtering, as the major focus of this article is to compare the performance of different algorithms rather than obtain optimum velocity fields for glacier dynamic studies. The threshold to identify residual velocities in the glacier-free regions (see Fig. 3) was set to be $0.2\ \mathrm{m \cdot d^{-1}}$, corresponding to an offset of about one pixel, for both datasets.

IV. Results

A. Velocity Fields: PairWise Versus Stacked Cross Correlation

Velocity fields obtained from TanDEM-X imagery are presented in Fig. 5 for pairwise and stacked NCCs with large (96×96 pixels, 2 m pixel spacing) and small (48×48 pixels) templates. Black voids in the figure are caused by removing outlier and unreliable estimates. The two velocity fields of pairwise NCC [see Fig. 5(a) and (c)] are produced with images acquired at 2017-1-21 and 2017-2-1. This image pair represent the *best* coverage of reliable estimates (92% and 74%, respectively) among the seven other velocity fields generated with pairwise NCC. The averaged velocity maps of all seven pairwise velocity fields (see Fig. 12 in the appendix) show less noise in the glacier free area, but the coverage over the glacier is greatly reduced than the selected ones. Either picking the best one or averaging all velocity maps are user-defined postprocessing steps, and thus, it is difficult to determine, which one represents the fairest comparison between NCC stacking and pairwise NCC velocity estimation. We consider our choice of picking the best pairwise map for comparison to be of disadvantage for NCC stacking which, nevertheless, shows a better spatial coverage.

In Fig. 5(a)–(d), velocity magnitude fluctuations along the glacier central line are consistent in all results, with the highest velocities at the ice fall between Ewigschneefaeld and Konkordiaplatz. The orientation of velocities [see Fig. 5(e)–(h)] is a good indicator to verify the consistency of velocity fields and our results show generally consistent orientation patterns that conform with glacier geometries.

The velocity fields generated by NCC stacking [see Fig. 5(b) and (d)] show much less noise than results generated by the single pairwise NCC [see Fig. 5(a) and (c)] for both template sizes. Focusing on areas indicated by red arrows in Fig. 5(b) and (f) and comparing them to Fig. 5(a) and (e), the stacked NCCs produce much smoother fields for both velocity magnitude and orientation than a pairwise NCC. For small templates and pairwise NCC [see Fig. 5(c) and (g)], areas of arrows are mostly covered by widely spreading voids, suggesting that it is difficult to obtain accountable velocity measurements over these areas.

Velocity fields obtained from Sentinel-1 A imagery are shown in Fig. 6 for large and small templates (96×96 and 48×48 pixels at 5 m pixel spacing). Results of pairwise NCCs shown in Fig. 6(a), (c) and (e), (g) are produced from the image pair 2017-1-15 and 2017-2-08, which represent again the *best* coverage (86% and 53%, respectively) among the 11 results of pairwise NCCs. The average of the 11 pairwise velocity maps (see Fig. 12 in the appendix) shows a comparable coverage for large templates (96 px) and a slightly better coverage for small templates (48 px). However, the additionally covered area shows many velocity artifacts. Hence, the single pairwise NCC with the best coverage was chosen for further study. Compared to Fig. 5, the Sentinel-1 A results show much lower spatial resolution and also less spatial coverage. Nevertheless, results obtained by NCC stacking show reduced noise compared to pairwise NCCs, and velocity fields at specific locations [see red arrows in Fig. 6(b) and (f)] are smoother for stacked NCCs than for pairwise NCCs.

B. Spatial Coverage of Velocity Fields

For the high elevations in the accumulation area of all glacier tributaries, velocity estimation from TanDEM-X failed using pairwise NCCs [see Fig. 5(a)]. In contrast, the uncovered region shrink to only the upper part of Ewigschneefaeld when applying NCC stacking [see Fig. 5(b)]. When decreasing template sizes, the spatial coverage decreases for pairwise and stacked NCC [see Fig. 5(c) and (d)], but the pairwise NCC shows a more serious decline of spatial coverage than the stacked NCC. Remarkably, the velocity map of stacked NCCs with small templates [see Fig. 5(d) and (h)] reaches a level of coverage that is almost as good as the coverage of pairwise NCCs and big templates [see Fig. 5(a) and (e)]. Noteworthy, the small templates doubled the spatial resolution compared to the big ones.

For Sentinel-1 A, when applying pairwise NCCs and large templates [see Fig. 6(a)], velocities can only be properly estimated in areas showing large crevasses including the ice fall, the glacier tongue, as well as the upper part of the Aletschfirn and the Jungfraufirn. However, with NCC stacking [see Fig. 6(b)], the whole body of the Great Aletsch Glacier, except for the upper part of Ewigschneefaeld, is successfully tracked. For pairwise

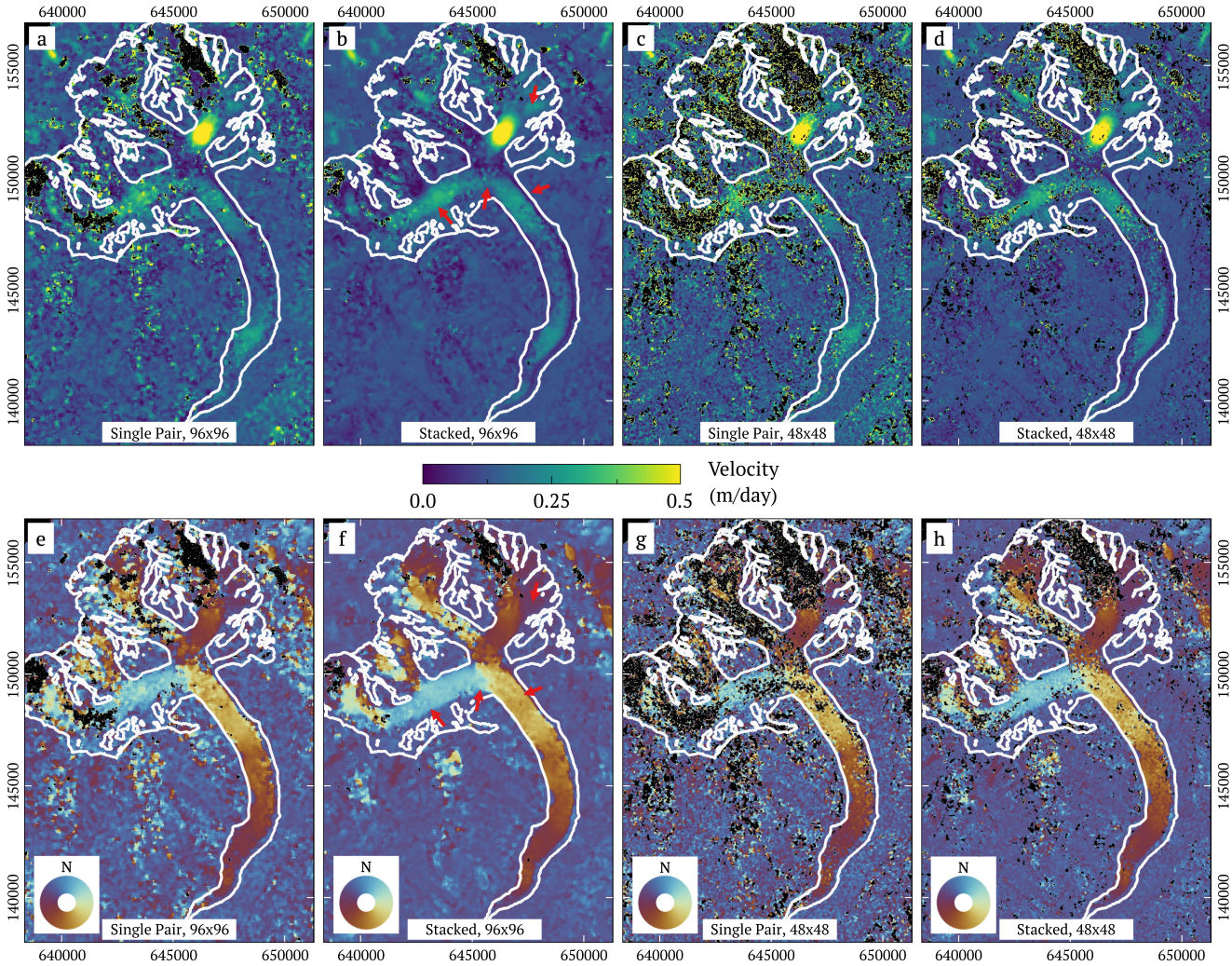


Fig. 5. Magnitude (top) and direction maps (bottom) of velocity fields produced using TanDEM-X SAR images. The left four figures, (a), (e) and (b), (f), are generated with large template (96 × 96 pixels), the right four figures, (c), (d) and (g), (h), with small templates (48 × 48 pixels). (a), (c) and (e), (g) are produced with a single pairwise NCC for which the *best* coverage among all pairwise results was obtained (2017-1-21 versus 2017-2-1); (b), (d) and (f), (h) are generated by stacking seven pairwise NCCs calculated from acquisitions between 2017-1-10 and 2017-3-28 (see Table I). Black voids represent removed outlier and results with an SNR below 10 dB. The glacier outline is delineated in white. Axes are labeled with Swiss Coordinates (CH1903/LV03) of unit meters.

NCCs and small templates [see Fig. 6(c)], the successfully tracked area is further reduced comparing to Fig. 6(a). Only the ice fall can still be tracked due to well visible crevasses [see Fig. 4(c)]. With stacked NCC the area of reliable estimates is significantly increased at the glacier tongue, the top of the Aletschfirn, and the Jungfraufirn [see Fig. 6(d)].

Fig. 7 shows the systematic analysis of the influence of the stack size on the spatial coverage of reliable velocity estimates. In general, the coverage ratio increases when increasing stack sizes for both template sizes and for both datasets. For pairwise NCC applied on TanDEM-X data (light and dark red bars in Fig. 7), the average coverage ratio is 85% for templates of 96 × 96 pixels and 68% for templates of 48 × 48 pixels. They are increased to 97% and 88% with a stack size of seven. For Sentinel-1 A (see light and dark blue bars in Fig. 7), the average coverage ratio is increased from 83% to 96% for templates of 96 × 96 pixels, and from 49% to 77% for templates of 48 × 48 pixels.

For both datasets, the biggest relative coverage improvement are observed when employing stack size of three. At this point, the best coverage ratio of TanDEM-X results increased by 4% to 96% for large templates and by 11% to 83% for small templates. With Sentinel-1 A imagery, the best coverage ratio increased by 7% to 93% for large templates, and by 17% to 70% for small templates. Further increasing the stack size keeps improving the coverage ratio but it gradually saturates at a certain level. For large templates, almost fully coverage for both datasets is obtained, whereas for small templates the coverage seems to saturate at about 89% for TanDEM-X and 79% for Sentinel-1 A. This implies that for the given dataset the stacking NCC has mostly exploited the information content in the image series that contributes to the successful tracking. In the remaining uncovered areas, mostly located in the snow covered accumulation zone, no trackable features seem to be present. Higher coverage could be expected for image pairs where a certain level of coherence is maintained so that also

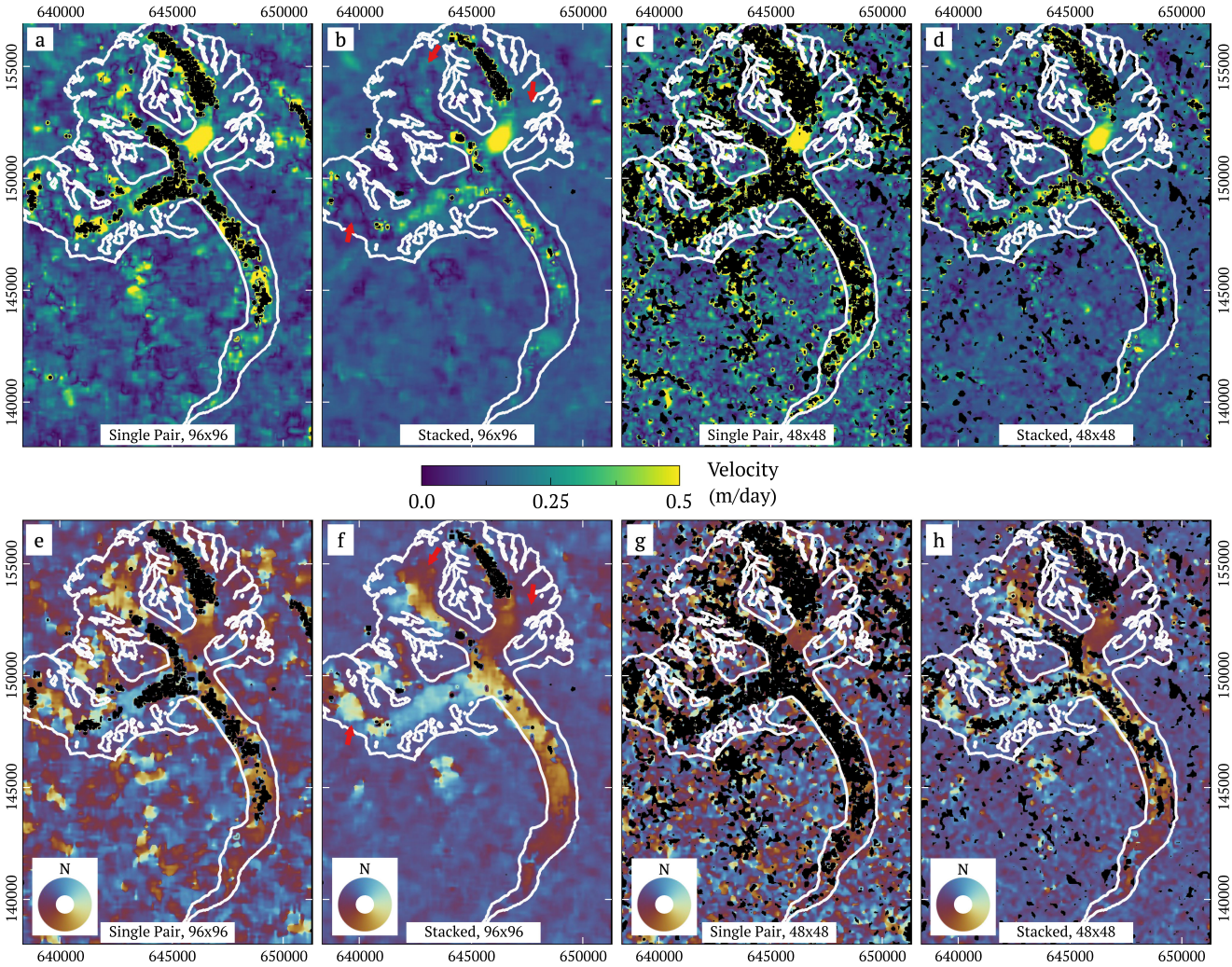


Fig. 6. Magnitude (a)–(d) and direction map (e)–(f) of velocity fields produced using Sentinel-1 A images. Panels are arranged the same way as Fig. 5. The results of pairwise NCC [(a), (e) and (c), (g)] are produced using the image pair from which the *best* spatial coverage of 11 NCC pairs was obtained (2017-1-15 and 2017-2-08). Results of stacked NCC [(b), (f) and (d), (h)] are produced by stacking 11 pairwise NCC calculated from winter acquisitions between 2017-1-3 and 2018-3-23 (see Table I). Black voids in velocity fields are caused by removing outlier and unreliable estimates. The glacier outline is delineated in white.

speckle pattern can be tracked in the feature-free snow-covered areas.

Figs. 5 and 6 show that the coverage of successful velocity estimates using NCC stacking and small templates was very similar to the coverage of the single NCC pair estimates where, however, doubled template size was used. This is confirmed by the statistics shown in Fig. 7. Specifically, the difference of coverage ratio between NCC stacking with small templates and single pairwise NCC with large templates is only 4% (88% versus 92%) for the TanDEM-X results and 9% (77% versus 86%) for the Sentinel-1 A results.

C. Residual Velocities of Velocity Fields

In the glacier free areas of Figs. 5 and 6, the pairwise NCC results show stronger noise and more residuals than the results of NCC stacking. The systematic analysis of these velocity residuals in Fig. 8 shows that the residual ratio, (9), decreases with the increase of stack size for both datasets and both template sizes.

Similar to the spatial coverage ratio, the biggest relative decrease of residual ratio is observed when stacking three pairwise NCCs. For TanDEM-X and at stack size of three, the average residual ratio (red in Fig. 8) decreased from 9% to less than 1% with large templates and from 23% to 5% with small templates. With Sentinel-1 A, it decreased from 18% to 3% with large templates, and from 37% to 13% with small templates (see blue in Fig. 8). For large templates, the minimum residual ratios already reduce to almost zero with stack size of three for both TanDEM-X and Sentinel-1 A. Very big ranges between the maximum and minimum residual ratios are observed for pairwise NCCs. However, with NCC stacking, the ranges quickly narrow to much smaller extents for both datasets.

D. SNR Gain

Fig. 9 illustrates the SNR gain of stacked NCC compared to the pairwise NCC with the best coverage ratio for both datasets. For most areas, the SNR increases regardless of template sizes

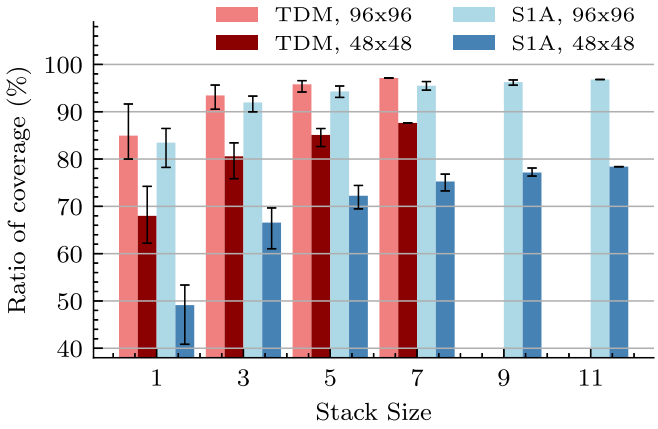


Fig. 7. Change of spatial coverage when using different stack sizes (i.e., number of pairwise NCC in the stack). A single pairwise NCC corresponds to a stack size of one. In each group, all possible combinations of pairwise NCC for stacking are evaluated. The bar height indicates the mean value of the group, and the black error bars indicate the maximum and minimum value in each group.

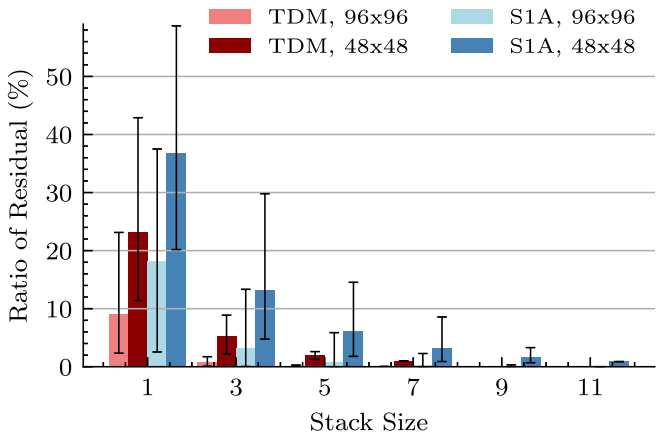


Fig. 8. Same as Fig. 7 but for the change of residual ratios with difference stack sizes.

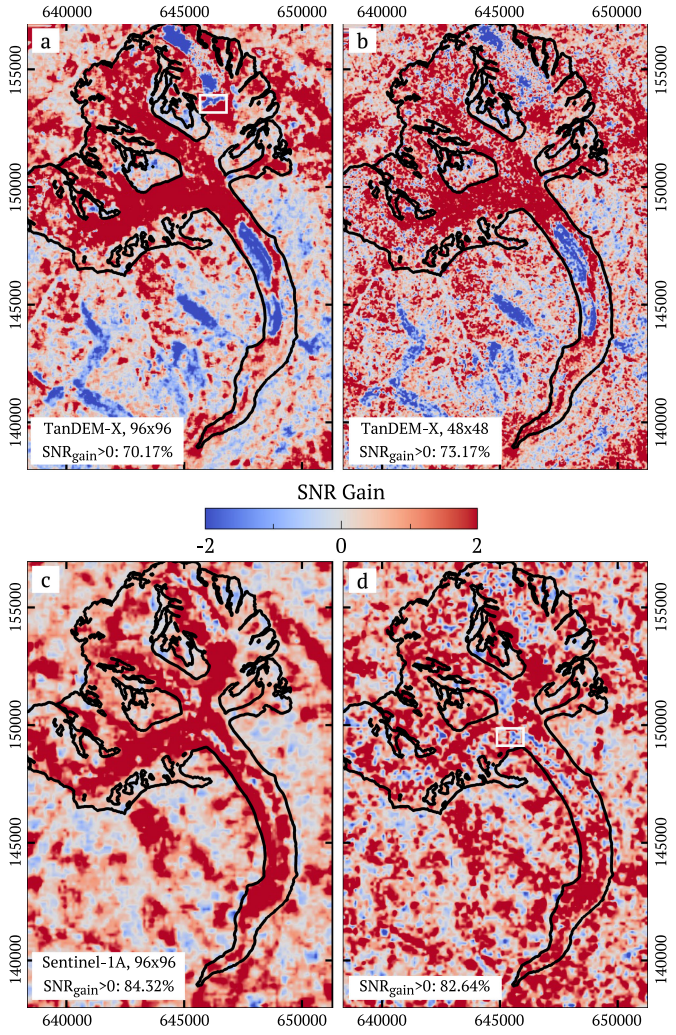


Fig. 9. SNR gain of stacked compared to pairwise NCC. (a) TanDEM-X with 96 × 96 templates. (b) TanDEM-X with 48 × 48 templates. (c) Sentinel-1 A with 96 × 96 templates. (d) Sentinel-1 A with 48 × 48 templates.

and sensor. For TanDEM-X data and both template sizes, large areas of strong SNR gain (in red) are found at the Aletschfirn and the Jungfraufirn. In these areas velocity fields are noisy and contain many gaps for the pairwise results shown in Figs. 5 and 6 whereas the stacking results show increased spatial coverage and much smoother velocity orientations. Areas showing an SNR decrease (in blue) seem to be located in areas where the velocity maps do not show much improvement either because the area is well traceable already with a single pair or because the area is not traceable, neither with pairwise nor with stacked NCC. For the well traceable areas (e.g., center of the Aletsch Glacier), the SNR dramatically decreases, because the pairwise NCC with best coverage already represents the highest SNR in the NCC stack, and thus averaging the stack brings the SNR down.

To gain insight into the reason of SNR change, we chose two representative templates with an SNR increase and decrease, selected from the red and blue area within the white boxes in Fig. 9(a) and (d). For these templates, the pairwise and stacked NCC fields are presented in Fig. 10. In the upper row of the figure, the SNR decrease for both TanDEM-X and Sentinel-1 A can be attributed to the drop of the averaged NCC peak, which is caused by the variation of peak strength within the NCC stack. Moreover, for Sentinel-1 A, the bright diagonal pattern in the stacked NCC indicates that the major energy of pairwise NCC are evenly distributed along the diagonal, and image features within templates very likely have periodic patterns. Although not shown, but comprehensible from lower row of Fig. 10, the SNR reduction in very difficult areas can also be caused by an apparent peak picked from the noise, which can show a larger SNR in pairwise NCC compared to the SNR of the true peak identified in the stacked NCC. In this sense, the SNR improvement in the lower row of Fig. 10 for both TanDEM-X and Sentinel-1 A shows the successful suppression of ambient noise with stacking.

E. Validation Against In Situ *Measurements*

Velocities obtained by the different methods are validated by comparing velocity magnitudes and components along the

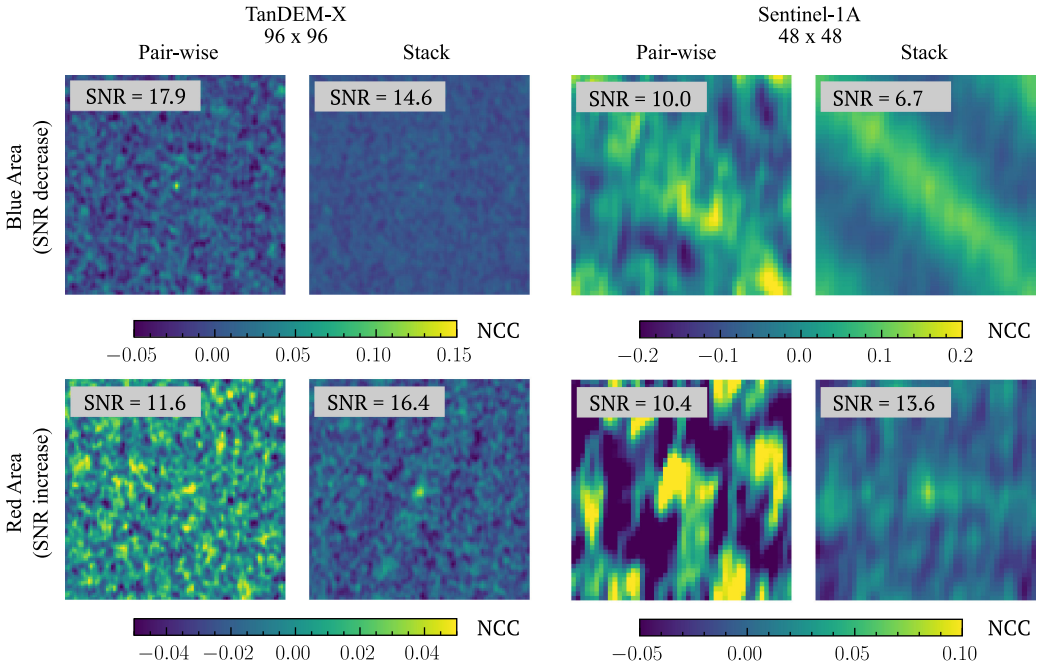


Fig. 10. Example NCC fields extracted from blue and red areas within the black box of Fig. 9(a) and (d). The pairwise NCC shown here represent the highest SNR in all pairwise NCCs.

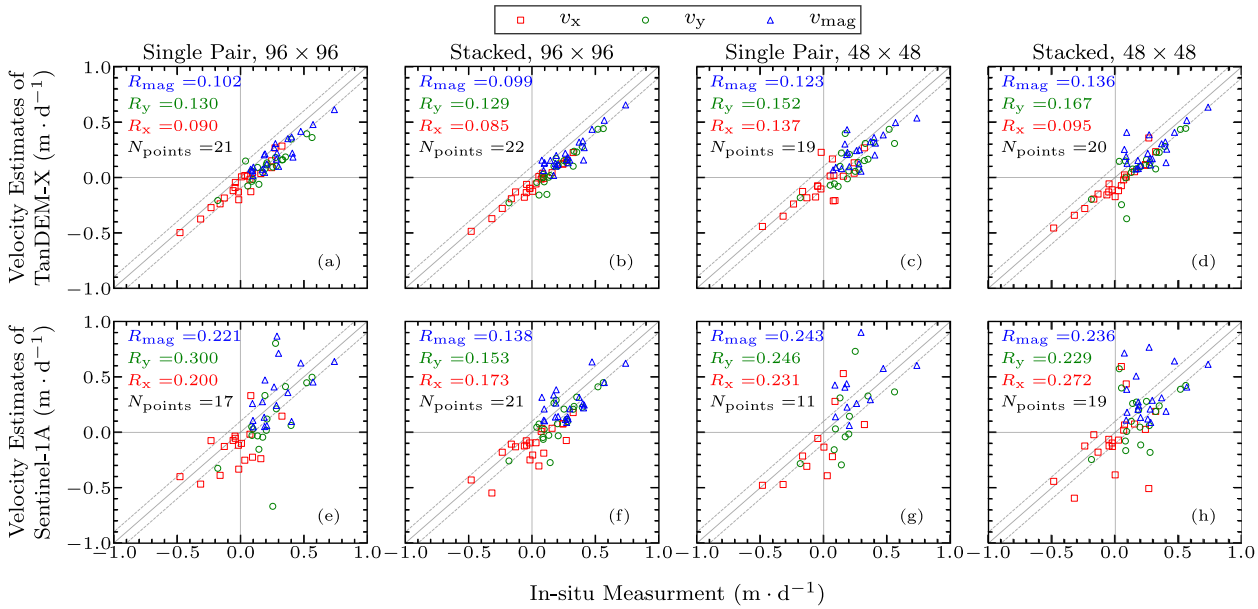


Fig. 11. Validation of estimated velocities with respect to *in situ* measurements for TanDEM-X (upper row) and for Sentinel-1 A (lower row). The four columns shows result of a single pairwise NCC and large templates (96 × 96 pixels) (a), (e), stacked NCC and large templates (b), (f), a single pairwise NCC and small templates (48 × 48 pixels) (c), (g), as well as stacked NCC and small templates (d), (h). The diagonal solid grey line indicates a 1:1 match, and the two dashed grey lines indicate a mismatch of $\pm 0.1 \text{m} \cdot \text{d}^{-1}$. RMSE values for velocity magnitude and components are reported in each panel with corresponding color. Due to the incomplete velocity coverage, not all of the 22 measurement points are used for validation. The number of points (N_{points}) used for RMSE calculation are reported in associated panels.

easting (v_x) and northing (v_y) with the *in situ* measurement. Results are presented in Fig. 11. Positive v_x indicates east pointing velocity vector, and positive v_y indicates south pointing velocity vector. For both TanDEM-X and Sentinel-1 A, the RMSE value for stacked NCCs are in general smaller than that for pairwise NCCs.

For TanDEM-X, most velocities estimated are somewhat smaller than the GPS data, in both $x-$ (easting) and $y-$ (northing) directions, shown by the right-wards biased point cloud in Fig. 11(a)–(d). Velocities estimated with Sentinel-1 A imagery have higher uncertainty than that with TanDEM-X, as shown by the more randomly distributed point cloud. However, more

TABLE II
RMSE VALUES FOR DIFFERENT GROUP OF VALIDATION POINTS IN THE RESULTS OF SENTINEL-1 A

	Points added to the result of stacking	Points kept in both results
96×96 pixels	$R_{\text{mag}} = 0.196$ $R_{\text{y}} = 0.078$ $R_{\text{x}} = 0.278$	$R_{\text{mag}} = 0.120$ $R_{\text{y}} = 0.166$ $R_{\text{x}} = 0.138$
48×48 pixels	$R_{\text{mag}} = 0.312$ $R_{\text{y}} = 0.310$ $R_{\text{x}} = 0.350$	$R_{\text{mag}} = 0.159$ $R_{\text{y}} = 0.144$ $R_{\text{x}} = 0.196$

than half of the points are biased to the right side of the 1:1 line, showing that results of Sentinel-1 A are also somewhat smaller than the *in situ* measurement.

For both sensors, points of high velocity magnitude v_{mag} show less scatter from the 1:1 line than points of small velocity magnitude, indicating that fast moving surfaces are estimated with higher confidence. For TanDEM-X, RMSE values in northing (R_{y}) are in general higher than that in easting (R_{x}). In contrast, although R_{y} for Sentinel-1 A are also higher than R_{x} with pairwise NCCs, they turn to be smaller than R_{x} with stacked NCCs. The RMSE of Sentinel-1 results are in general higher than the RMSE of TanDEM-X results. This is also illustrated by the increased outliers contained in Fig. 11(e)–(h) comparing to Fig. 11(a)–(d). This is partly determined by the different image resolution of the two sensors. Moreover, as the stacked NCC greatly extended the coverage of the velocity map (see Fig. 6), the number of points being used for validation are increased for the stacking results [see Fig. 11(f) and (h)]. As shown in Table II, the added points are estimated with larger uncertainties comparing to the points that are kept in both methods. Hence, the outliers in Sentinel-1 results are increased.

V. DISCUSSION

A. Spatial Coverage

The spatial coverage of reliable offset estimates is a crucial parameter for the obtained offset field. As seen from the comparison between results of large and small templates in Figs. 5 and 6, it is clear that using large templates can effectively increase reliable estimates. This is because large templates can accommodate more image features than small ones, and thus, can enhance the autocorrelation peak height and the SNR of the NCC. For offset tracking using SAR images, larger templates are required, compared to optical images, to compensate for the noise introduced by temporally uncorrelated (incoherent) SAR speckles.

However, using large templates inevitably depreciates the spatial resolution of the obtained offset field, and, thus, is not of advantage for studying small-scaled problems (e.g., small glaciers or landslides). In addition, large templates are also not suitable for studying nonhomogeneous velocity fields such as shear zones, where strong velocity gradient can cause nonuniform shifts of image contents within one template and, thus, can lead to dispersed peaks in the NCC.

When large templates are not applicable, our results suggest that NCC stacking can provide a powerful solution for robust tracking at high spatial resolution. As shown in Fig. 7, the spatial coverage of NCC stacking with templates of 48×48 pixels is almost equivalent to that of pairwise NCC with templates of 96×96 pixels for both sensors. This suggests that NCC stacking permits to use smaller templates to get velocity fields of equivalent coverage as using big templates for pairwise NCC.

B. Estimation Accuracy

In this article, we first evaluated the accuracy of velocity estimation using residuals over static ground. Comparing to the validation against the *in situ* measurement, this is an indirect method of evaluating the estimation accuracy, as it is difficult to precisely quantify to what degree the accuracy over glacier-free area actually represents that over glacier covered regions. However, many studies have used this method due to the lack of *in situ* measurement [13]. In Fig. 8, the decrease of the residual ratio with increasing stack size shows that NCC stacking effectively reduced the uncertainty of velocity estimation. In addition, when using pairwise NCC, the range between the maximum and minimum residual ratios is quite big, suggesting that different imaging conditions have a strong influence on the estimation accuracy. With stacking, the ranges quickly narrow down, showing that temporally stacking pairwise NCCs can average out the change of image conditions and, thus, make it more accurate for velocity estimation than just a single pairwise NCC.

Using the *in situ* measurement, we also directly quantified the error of estimation at selected locations on the glacier. In general, estimated velocities show quite small RMSE values with respect to the *in situ* measurement, but the satellite-based velocities are, in average, slightly below the *in situ* data. Velocity time series indicate that the seasonal velocity increase starts around early May [40]. Considering that *in situ* velocities were measured between 2019-4-30 and 2019-5-06, it is reasonable that the *in situ* measurement appear slightly higher than the winter velocities obtained by offset tracking.

The confidence level of offset tracking depends on surface structures but also on the sensor resolution. We observed a higher confidence for fast moving regions where strong and well visible features are induced by fast glacier flow (e.g., big crevasses). Differences of confidence levels also occur between velocity components along the northing and easting, which is likely caused by the different resolution along the azimuth (mostly along the northing) and the slant-range (mostly along the easting) direction of the used SAR images. At the Great Aletsch Glacier, the image scene is rotated by about 10° in average between northing-easting and azimuth-range due to orthorectification. The resolution (az $\times$ rg) is about 6.6×2.2 m for TanDEM-X [41] and about 22×4.7 m for Sentinel-1A [42]. Therefore, RMSE along the northing R_{y} is in general higher than RMSE along the easting R_{x}. Despite the lower resolution along the northing, we obtain comparable uncertainties for v_{x} and v_{y}.

C. Stack Size

The results show that NCC stacking is able to sufficiently improve the tracking performance by increasing both spatial coverage and tracking accuracy. Although the performance keeps improving with the increased stack sizes, we found the biggest relative improvement of the performance for a stack size of three. This indicates that the most prominent relative performance gain is not necessarily achieved using many pairwise NCCs for stacking. Therefore, when the size of NCC stack is limited by the amount of available data (i.e., too little images available), stacking over a limited number of pairwise NCC can still provide considerable performance enhancement. A small NCC stack is also of advantage when high temporal resolution is desired. For instance, when studying temporal velocity variations (e.g., glacier surge, seasonality, or fast landslides), large stacks would either average out transient velocities or, in worst case, could lead to strong blurring of the NCC peak such that it cannot be tracked anymore. Using relatively small NCC stack can better compromise the need of improving spatial coverage without sacrificing too much temporal resolution.

For example, studying the seasonal fluctuations of a glacier's flow velocity requires high temporal resolution. With the increasingly available SAR images of short revisit time, it is possible to collect many images within a few months during which the seasonal velocity variations are limited. Hence, we can compose an image series within one season to make velocity variations between seasons traceable. Another example is to apply the stacking method to fast-flowing glaciers. Fast-flowing glaciers often behave complex dynamics with strong acceleration and deceleration NCC, and thus, the NCC peaks are shifted in the NCC series. In this case, we can adjust the stacking size such that the velocity fluctuation within one stack is minimized.

D. TanDEM-X Versus Sentinel-1 A

The comparison of Figs. 5 and 6 shows that velocity fields obtained from TanDEM-X imagery have better spatial coverage than that of Sentinel-1 A, for the same template size. We think that the coverage difference could originate from the different resolutions of the sensors. High resolution means that the image template could contain more distinguishable features than templates of the same pixel size but generated from low resolution data. A similar effect has been observed for optical imagery, where low resolution sensors often underestimate the velocities observed by high resolution sensors [44]. For example, the two meter resolution of TanDEM-X can better resolve features like crevasses, which are often only a few meters large. For Sentinel-1, it is more likely that multiple crevasses are contained in one resolution cell and, thus, add up to speckle patterns by interference of the radiation scattered at different locations of the crevasse. The transformation of features to speckle degrades the correlation process.

The high resolution of TanDEM-X obviously improved the accuracy of offset estimation. In Fig. 11, RMSE values for Sentinel-1 A are in general twice as large as for TanDEM-X, which is comparable to the difference of resolutions of the two sensors. Moreover, different pixel sizes also affect the choice of time intervals between image pairs. In the experiment, due to the large pixel sizes, time intervals between Sentinel-1 A image pairs were set to be $T = 24$days, which are about twice as long as the repeating time of TanDEM-X. Taking long time interval ensures that image features shift far enough for detection. If the time interval were not sufficiently large, image features would not move out of one resolution cell for low resolution sensors such as Sentinel-1. In this case, the NCC would not be able to track the displacement. However, larger time intervals also increases temporal decorrelation between the image pair [40], and thus, degrades the SNR of the NCC.

Although TanDEM-X has the advantage of having high resolution, the data coverage of TanDEM-X is not as applaudable as Sentinel-1. Sentinel-1 has very good global coverage and constant revisiting time of 12 days (six days for Sentinel-1 A/B together), making it especially preferred for time-series collection. We found that the Great Aletsch Glacier is almost at the limit for Sentinel-1 A offset tracking with pairwise NCC, but the stacked NCC shows promising results. The NCC stacking method could show further improvement when combining very short sequences of Sentinel-1 A/B, e.g., average 6-day SAR image pairs of Sentinel-1 A/B before calculating cross correlation over significantly larger time intervals.

VI. Conclusion

In this article, we presented a cross-correlation stacking method to improve offset tracking. Instead of estimating displacement offsets from a single pairwise NCC, we first calculate a stack of multiple pairwise NCCs from image time-series, and then average the NCC stack before tracking the NCC peak. With the assumption that the surface object moves with constant velocity during the image series acquisition, we can effectively suppress the noise floor and meanwhile maintain the NCC peak height.

The proposed stacking method is assessed by measuring the flow velocity of the Great Aletsch Glacier using both TanDEM-X and Sentinel-1 A images. The result shows that the SNR of the NCC are greatly improved by stacking, leading to extended coverage of velocity field and more precise velocity estimates. Remarkably, the coverage of NCC stacking with small templates is equivalent to the coverage of pairwise NCC with big templates. This makes NCC stacking the preferred method when the template size is limited by a study area with small glaciers. Assessing the performance improvement with respect to different stack sizes shows that prominent performance gain does not necessarily rely on a large number of pairwise NCCs in the stack, which suggests that the stacking method can also be applied to small NCC stacks when only short image time-series is available or high temporal resolution is required.

In this article, we evaluated NCC stacking with equally spaced time-series data of a single sensor. However, as long as the spatial offset can be assumed to be the same for all NCC pairs, various data sources can be used as input, e.g., different spectral or polarization channels, data from different sensors (e.g., SAR and optical), data from the same period of the year or different image combinations of before and after seismic events. Resampling of the cross-correlation function before stacking could even allow

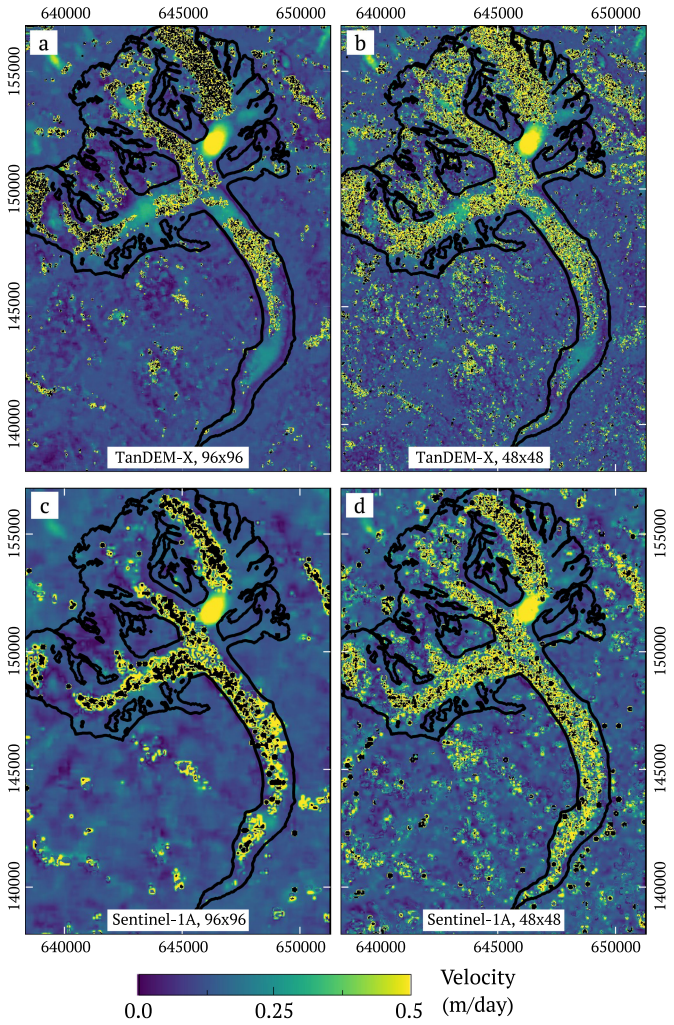


Fig. 12 Averaged velocity magnitude maps produced using TanDEM-X [(a), (b)] and Sentinel-1 A images [(c), (d)] with large (96 × 96 pixels) and small (48 × 48 pixels). Acquisitions are listed in Table I). Black voids represent removed outlier and results with an SNR below 10 dB. The spatial coverage ratio of each panel are listed in Table III. The glacier outline is delineated in black. Axes are labeled with Swiss Coordinates (CH1903/LV03) of unit meters.

TABLE III
SPATIAL COVERAGE OF THE AVERAGED VELOCITY MAPS

Template Size	96 × 96 pixels	48 × 48 pixels
TanDEM-X	88.77%	89.91%
Sentinel-1A	89.7%	91.8%

for NCC stacking over image series with unequal time intervals With the increasingly growing archives of remote sensing data, cross-correlation stacking provides a promising method to benefit from the acquired time series for robust offset tracking.

APPENDIX

Using the pairwise NCC, we obtained series of velocity maps for the collected TanDEM-X and Sentinel-1 A images. After applying the same thresholds with $v_{\max} = 1$ m·d^{-1} and $\mathrm{SNR}_{\min} = 10$dB, the velocity map series are averaged to produce the averaged velocity maps (see Fig. 12). The spatial coverage ratio of the averaged velocity maps are presented in Table III. For TanDEM-X results, the coverage ratio of the averaged velocity maps [see Fig. 12(a) and (b)] are lower than the best coverage as presented in Fig. 5. For Sentinel-1 A, although the coverage ratio of the averaged velocity maps [see Fig. 12(c) and (d)] are somewhat higher than the selected map with the most spatial coverage (see Fig. 6), the additionally covered area in the averaged velocity maps reveals very noisy velocity data.

ACKNOWLEDGMENT

The authors would like to thank the two anonymous reviewers for their valuable comments on this article. They also like to thank A. Kääb for his valuable suggestion on including Sentinel-1 A data in this article, and also like to thank O. Frey for fruitful discussion of the method and interpretation of the results. The TanDEM-X data used in this article was provided by the German Space Agency DLR via proposal XTI_GLAC6780.

REFERENCES

[1] A. Singleton, Z. Li, T. Hoey, and J.-P. Muller, "Evaluating sub-pixel offset techniques as an alternative to D-InSAR for monitoring episodic landslide movements in vegetated terrain," *Remote Sens. Environ.*, vol. 147, pp. 133–144, 2014.

[2] X. Shi, L. Zhang, T. Balz, and M. Liao, "Landslide deformation monitoring using point-like target offset tracking with multi-mode high-resolution TerraSAR-X data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 105, pp. 128–140, 2015.

[3] A. Stumpf, J.-P. Malet, and C. Delacourt, "Correlation of satellite image time-series for the detection and monitoring of slow-moving landslides," *Remote Sens. Environ.*, vol. 189, pp. 40–55, 2017.

[4] E. J. Fielding, P. R. Lundgren, T. Taymaz, S. Yolsal-Çevikbilen, and S. E. Owen, "Fault-slip source models for the 2011 m 7.1 van earthquake in Turkey from SAR interferometry, pixel offset tracking, GPS, and seismic waveform analysis," *Seismological Res. Lett.*, vol. 84, no. 4, pp. 579–593, 2013.

[5] Y. Himematsu and M. Furuya, "Fault source model for the 2016 kumamoto earthquake sequence based on ALOS-2/PALSAR-2 PIXEL-OFFSET DATA: Evidence for dynamic slip partitioning," *Earth, Planets Space*, vol. 68, no. 1, pp. 1–10, 2016.

[6] T. Strozzi, A. Luckman, T. Murray, U. Wegmuller, and C. L. Werner, "Glacier motion estimation using SAR offset-tracking procedures," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2384–2391, Nov. 2002.

[7] I. Joughin, "Ice-sheet velocity mapping: A combined interferometric and speckle-tracking approach," *Ann. Glaciol.*, vol. 34, pp. 195–201, 2002.

[8] A. Luckman, D. Quincey, and S. Bevan, "The potential of satellite radar interferometry and feature tracking for monitoring flow rates of himalayan glaciers," *Remote Sens. Environ.*, vol. 111, no. 2/3, pp. 172–181, 2007.

[9] D. Scherler, S. Leprince, and M. R. Strecker, "Glacier-surface velocities in alpine terrain from optical satellite imagery-accuracy improvement and quality assessment," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3806–3819, 2008.

[10] A. Dehecq, N. Gourmelen, and E. Trouvé, "Deriving large-scale glacier velocities from a complete satellite archive: Application to the Pamir-Karakoram-himalaya," *Remote Sens. Environ.*, vol. 162, pp. 55–66, 2015.

[11] T. Schellenberger, T. Dunse, A. Kääb, J. Kohler, and C. Reijmer, "Surface speed and frontal ablation of kronebreen and kongsbreen, nw svalbard, from SAR offset tracking," *Cryosphere*, vol. 9, no. 6, pp. 2339–2355, 2015.

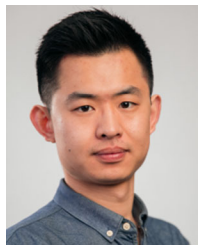
[12] F. Paul *et al.*, "The glaciers climate change initiative: Methods for creating glacier area, elevation change and velocity products," *Remote Sens. Environ.*, vol. 162, pp. 408–426, 2015.

[13] V. Round, S. Leinss, M. Huss, C. Haemmig, and I. Hajnsek, "Surge dynamics and lake outbursts of Kyagar glacier, Karakoram," *Cryosphere*, vol. 11, no. 2, pp. 723–739, 2017.

[14] T. Li, Y. Liu, T. Li, F. Hui, Z. Chen, and X. Cheng, "Antarctic surface ice velocity retrieval from modis-based mosaic of antarctica (MOA)," *Remote Sens.*, vol. 10, no. 7, 2018, Art no. 1045.

[15] J. P. Merryman Boncori *et al.*, "Intercomparison and validation of SAR-based ice velocity measurement techniques within the greenland ice sheet CCI project," *Remote Sens.*, vol. 10, no. 6, p. 929, 2018. [Online]. Available: https://www.mdpi.com/about/announcements/784

[16] T. Heid and A. Kääb, "Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery," *Remote Sens. Environ.*, vol. 118, pp. 339–355, 2012.
[17] S. Leprince, S. Barbot, F. Ayoub, and J.-P. Avouac, "Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground deformation measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1529–1558, Jun. 2007.
[18] M. Debella-Gilo and A. Kääb, "Measurement of surface displacement and deformation of mass movements using least squares matching of repeat high resolution satellite and aerial images," *Remote Sens.*, vol. 4, no. 1, pp. 43–67, 2012.
[19] E. Erten, A. Reigber, O. Hellwich, and P. Prats, "Glacier velocity monitoring by maximum likelihood texture tracking," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 394–405, Feb. 2009.
[20] T. A. Scambos, M. J. Dutkiewicz, J. C. Wilson, and R. A. Bindschadler, "Application of image cross-correlation to the measurement of glacier velocity using satellite image data," *Remote Sens. Environ.*, vol. 42, no. 3, pp. 177–186, 1992.
[21] R. d. Lange, A. Luckman, and T. Murray, "Improvement of satellite radar feature tracking for ice velocity derivation by spatial frequency filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2309–2318, Jul. 2007.
[22] M. Debella-Gilo and A. Kääb, "Locally adaptive template sizes for matching repeat images of earth surface mass movements," *ISPRS J. Photogrammetry Remote Sens.*, vol. 69, pp. 10–28, 2012.
[23] Y. Ahn and I. M. Howat, "Efficient automated glacier surface velocity measurement from repeat images using multi-image/multichip and null exclusion feature tracking," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2838–2846, Aug. 2011.
[24] R. D. Keane and R. J. Adrian, "Theory of cross-correlation analysis of PIV images," *Appl. Sci. Res.*, vol. 49, no. 3, pp. 191–215, 1992.
[25] C. D. Meinhart, S. T. Wereley, and J. G. Santiago, "A PIV algorithm for estimating time-averaged velocity fields," *J. Fluids Eng.*, vol. 122, no. 2, pp. 285–289, 2000.
[26] L. Bohs, B. J. Geiman, K. Nightingale, C. Choi, B. Friemel, and G. Trahey, "Ensemble tracking: A new method for 2D vector velocity measurement," in *Proc. IEEE Ultrasonics Symp. Proc. Int. Symp.*, 1995, vol. 2, pp. 1485–1488.
[27] L. N. Bohs, B. J. Geiman, M. E. Anderson, S. M. Breit, and G. E. Trahey, "Ensemble tracking for 2d vector velocity measurement: Experimental and initial clinical results," *IEEE Trans. Ultrasonics, Ferroelect., Freq. Control*, vol. 45, no. 4, pp. 912–924, Jul. 1998.
[28] B. Altena and A. Kääb, "Ensemble matching of repeat satellite images applied to measure fast-changing ice flow, verified with mountain climber trajectories on Khumbu icefall, mount everest," *J. Glaciol.*, vol. 66, no. 260, pp. 905–915, 2020.
[29] N. Yague-Martinez, F. De Zan, and P. Prats-Iraola, "Coregistration of interferometric stacks of Sentinel-1 tops data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1002–1006, Jul. 2017.
[30] H. Ansari, F. De Zan, and R. Bamler, "Sequential estimator: Toward efficient INSAR time series analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5637–5652, Oct. 2017.
[31] F. De Zan, "Coherent shift estimation for stacks of SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 6, pp. 1095–1099, Nov. 2011.
[32] F. De Zan and P. Lopez-Dekker, "SAR image stacking for the exploitation of long-term coherent targets," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 502–506, May 2011.
[33] G. Krieger *et al.*, "TANDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3317–3341, Nov. 2007.
[34] R. Torres *et al.*, "GMES Sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9–24, 2012.
[35] M. Debella-Gilo and A. Kääb, "Sub-pixel precision image matching for measuring surface displacements on mass movements using normalized cross-correlation," *Remote Sens. Environ.*, vol. 115, no. 1, pp. 130–142, Jan. 2011.
[36] F. S. Paul *et al.*, *GLIMS Glacier Database*. Boulder, CO, USA: National Snow and Ice Data Center, 2019.
[37] P. Prats, R. Scheiber, A. Reigber, C. Andres, and R. Horn, "Estimation of the surface velocity field of the aletsch glacier using multibaseline airborne SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 419–430, Feb. 2009.
[38] E. Erten, "Glacier velocity estimation by means of a polarimetric similarity measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3319–3327, Jun. 2013.
[39] A. Schubert, A. Faes, A. Kääb, and E. Meier, "Glacier surface velocity estimation using repeat TerraSAR-X images: Wavelet- vs. correlation-based image matching," *ISPRS J. Photogrammetry Remote Sens.*, vol. 82, pp. 49–62, Aug. 2013, Art no. 00039.
[40] S. Leinss and P. Bernhard, "TanDEM-X: Deriving InSAR height changes and velocity dynamics of great aletsch glacier," *IEEE J. Sel. Topics Remote Sens.*, to be published, doi: 10.1109/JSTARS.2021.3078084.
[41] T. Fritz *et al.*, "TerraSAR-X - ground segment - basic product specification document," DLR, Tech. Rep. tx-gs-dd-3302; 2013.
[42] M. Bourbigot, H. Johnsen, R. Piantanida, G. Hajduch, J. Poullaouec, and G. Hajduch, "Sentinel-1 product definition," ESA, Tech. Rep. S1-RS-MDA-52-7440, 2016.
[43] C. Werner, U. Wegmuller, T. Strozzi, and A. Wiesmann, "Precision estimation of local offsets between pairs of SAR SLCS and detected SAR images," in *Proc. Int. Geosci. Remote Sens. Symp.*, 2005, vol. 7, Art no. 4803.
[44] R. Millan *et al.*, "Mapping surface flow velocity of glaciers at regional scale using a multiple sensors approach," *Remote Sens.*, vol. 11, no. 21, 2019, Art. no. 2498.



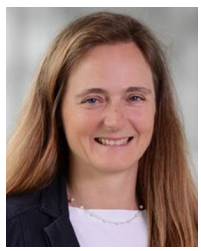
Shiyi Li (Graduate Student Member, IEEE) received the B.S. degree in geology from Nanjing University, Nanjing, China, in 2014, and the M.S. degree in applied geophysics from Delft University of Technology, Delft, The Netherlands, from ETH Zürich, Zürich, Switzerland, and from RWTH Aachen University, Aachen, Germany, in 2018. He is currently working toward the Ph.D. degree in environmental engineering with ETH Zürich (focusing on monitoring glacier dynamics with synthetic aperture radar).

He was a Research Assistant with the Geothermal Energy and Geofluids group and with the Center for Law and Economics, ETH Zürich in 2018.

Silvan Leinss (Senior Member, IEEE) received the Dipl. degree in laser physics from the University of Constance, Konstanz, Germany, in 2008 and the Ph.D. degree in radar remote sensing from ETH Zurich, Zurich, Switzerland, in 2016.

In 2009, he was working in the field of coherent light scattering at the University of Constance and in 2010, in the field of light scattering statistics in Bose-Einstein-Condensates at ETH Zurich. Since 2021, he has been with the Laboratoire d'Informatique, Systémes, Traitement de l'Information et de la Connaissance (LISTIC), Polytech Annecy-Chambéry, Université de Savoie Mont Blanc, Chambéry, France.

Dr. Leinss was the recipient of the Helmut-Rott-Prize for his excellent contributions for remote sensing of the cryosphere.



Irena Hajnsek (Fellow, IEEE) received the Dipl. Geography degree (Hons.) from the Free University of Berlin, Berlin, Germany, in 1996 and the Dr. rer. nat. degree (Hons.) from the Friedrich Schiller University of Jena, Jena, Germany, in 2001.

Since November 2009, she has bene a Professor of Earth Observation with the Swiss Federal Institute of Technology (ETH), Zürich Institute of Environmental Engineering, Zurich, Switzerland, and the Head of the Polarimetric SAR Interferometry Research Group, German Aerospace Center Microwaves and Radar Institute, Wessling, Germany. Since 2010, she has been the Science Coordinator of the German satellite mission TanDEM-X and proposed satellite mission Tandem-L. Her main research interests include electromagnetic propagation and scattering theory, radar polarimetry, SAR and interferometric SAR data processing techniques, and environmental parameter modeling and estimation.

Dr. Hajnsek was the Technical Program Co-Chair of the IEEE IGARSS 2012 in Munich, Germany, and 2019 in Yokohama, Japan. Since 2013, she has been a member of the IEEE GRSS AdCom, and from 2016 to 2020, she was the Vice President of the IEEE GRSS Technical Committees.