Science of Remote Sensing 6 (2022) 100070

Contents lists available at ScienceDirect

Science of Remote Sensing

journal homepage: www.sciencedirect.com/journal/science-of-remote-sensing

Improved surface displacement estimation through stacking cross-correlation spectra from multi-channel imagery

Bas Altena [a,*], Silvan Leinss [b,c,1]

[a] Institute for Marine and Atmospheric Research, Utrecht University, Utrecht, the Netherlands
[b] Institute of Environmental Engineering, ETH Zürich, Zurich, Switzerland
[c] LISTIC, Polytech Annecy-Chambéry, Université Savoie Mont Blanc, Annecy, France

ARTICLE INFO

Keywords:
Time-resolved PIV
Ensemble matching
Correlation stacking
Imaging geodesy
Swell movement
Surface displacement
Multi-spectral image matching
RapidEye

ABSTRACT

Studying sporadic and complex geophysical surface flows, like earthquakes or sea surface circulation, are challenging cases. If a satellite is able to image an event, it becomes essential to pull out as much information as possible. In this contribution we demonstrate a method to increase the coverage and signal-to-noise ratio for displacement estimation, making such surface flow estimates more complete. We leverage upon the redundant offset information acquired by multi-channel push-broom imagery. The individual cross-correlation spectra (cross power spectral density; Fourier transform of the cross-correlation function) of different spectral bands are averaged in the frequency domain before sub-pixel offset-estimation by phase-plane fitting. The method is demonstrated near Kaiko¯ura, where in 2016 a surface rupture occurred. RapidEye data from two different dates were used to reconstruct the displacement. In addition, the circulation along the coast is estimated from data from a single date where multiple spectral bands were acquired within seconds which made stacking of cross-correlation spectra possible. The demonstrated methodology is applied to data from the already decommissioned RapidEye constellation, but can be adopted to other pushbroom systems, such as the Landsat legacy or Sentinel-2.

1. Introduction

Many of the processes occurring at, on, or below the Earth's surface are sporadic. For example, coastal circulation is shaped by forces varying at different temporal and spatial scales. Other sporadic surface displacements are a result from an often catastrophic release after a build-up of forces, such as earthquakes. Fortunately, imagery from optical satellites are recording large stretches of the Earth surface, and by doing so, have a good chance of recording such sporadic phenomena or complex surface dynamics.

Pairs of images separated by a time interval can be used to extract surface displacements. For that, a small region of a satellite scene is used to find the most similar region in a second scene. It is common practice to use a single spectral band with the highest spatial resolution for this analysis. However, the quality of the resulting displacement estimate is heavily influenced by the similarity of the compared scene. Repetition of the recording to enhance the quality is in many cases not feasible. Thus it is of the utmost importance to extract the best displacement estimate out of an imaging pair.

In this study we demonstrate the improvement achievable by stacking cross-correlation spectra. This is possible because of the multitude of bands that are recorded near-simultaneously by multi-spectral instruments on board of optical satellites, as the geometric configuration between band recordings is highly similar. Neither is the method limited to multi-spectral data, nor to optical imagery at all, in essence any data with different channels can be used, be it different spectral bands or polarization. The method is of interest for many disciplines, though here we limit ourselves to a demonstration of the examples resulting from displacements extracted from an earthquake and sea surface flow. Multi-spectral imagery from the RapidEye constellation are used over the Kaiko¯ura region in New Zealand, where a large co-seismic displacement occurred in November 2016.

2. Background

A typical workflow for offset tracking would first pre-process the

* Corresponding author.
E-mail address: b.altena@uu.nl (B. Altena).
[1] Now at: GAMMA Remote Sensing and Consulting AG, Gümlingen, Switzerland.

https://doi.org/10.1016/j.srs.2022.100070
Received 3 June 2022; Received in revised form 3 October 2022; Accepted 4 October 2022
Available online 18 October 2022

satellite imagery, or specifically select a spectral band. For example, in glaciology one aims to extract the ice surface flow and not the drifting surface snow cover or the moving cast shadows. Thus in this discipline, it is customary to reduce such effects by high-pass filters (Ahn and Howat, 2011), band-pass filters (Leprince et al., 2007), or principle component analysis (Stearns and Hamilton, 2005). Sometimes a lower resolution band is used, as these might have better gain settings or less instrument noise, as seems to be the case for Landsat 7 (Heid, 2011). In the domain of landslides and earthquakes, similar pre-processing procedures are used.

There can also be technical or physical considerations to process a specific spectral band. For example, Stumpf et al. (2018) uses the red band of Sentinel-2 as this is the reference band for interband co-registration. Similarly, the spectral properties might also be of interest; Redpath et al. (2013) uses the blue band, as this gave the best results on debris-covered glaciers for the specific lithology in their study region. Also, near-infrared imagery is popular for polar studies, as in this part of the spectrum snow shows a large variability of the reflectivity, so the radiometric spread of the signal is higher. Necsoiu et al. (2009) proposes the selection of the near-infrared band in dune migration studies as this band has a large spectral bandwidth and limited atmospheric absorption. Differences in the displacement results do not seem to be observable for solutions from different Sentinel-2 bands (Lacroix et al., 2018), indicating an accurate sensor calibration.

Lastly, there are alternative (post-)processing strategies to enhance the quality of displacement fields. Filtering can be done either through pixel-per-pixel group-statistics of displacement estimates from different band pairs (Klinger et al., 2018), or adjacent pixels (Paul et al., 2015). Since displacement magnitudes have long tailed distributions, robust statistics and constraining methods like e.g. temporal displacement networks are needed to create clear products (Altena et al., 2019; Charrier et al., 2021).

In this study the term redundancy refers to multiple observations of (almost) the same offset with different spectral bands, or time intervals. It is this redundant displacement information present in the ensemble of bands in multi-spectral imagery that is exploited, in a similar but improved way to Altena and Kääb (2020) and Li et al. (2021). In these studies temporal ensembles (or stacks) of cross-correlations from optical and SAR imagery were used. Different to post-processing of sets of velocity estimates, these methods - known as ensemble matching or cross-correlation stacking - average sets of cross-correlation functions before detection of the displacement peak, as suggested by Meinhart et al. (2000) for time resolved PIV. These average correlation methods usually reduce the temporal resolution when sequentially acquired cross-correlation functions are averaged. Therefore, in this contribution we apply the same principle but average cross-correlation spectra from almost simultaneously acquired multi-spectral bands which preserves the temporal resolution. In contrast to Meinhart et al. (2000), Altena and Kääb (2020), Li et al. (2021), we implement the stacking one step earlier in the frequency domain of an offset tracking pipeline. By doing the stacking in the frequency domain, a phase-plane can be estimated. A phase-plane can be used for sub-pixel displacement estimation instead of sub-pixel estimation from the cross-correlation peak in the spatial domain which is sensitive to integer-pixel-locking (Foroosh et al., 2002) and, in addition, the computational load is reduced compared to earlier ensemble methods. Independent on implementation in the spatial or frequency domain, stacking enhances the signal-to-noise ratio of the cross-correlation function, hence making detection of the displacement peak in the cross-correlation more robust and successful. As a consequence, the completeness of a displacement field is increased. More specifically, it enhances the visibility of the cross-correlation peak in the spatial domain or, equivalently in the frequency domain, it enhances the visibility of the phase plane and extends the plane to higher frequencies, making it possible to use smaller template sizes than what would normally be achievable for single-pair matching. The method limits the post-processing efforts (and options) as redundancy is already exploited during correlation rather than during post-processing.

3. Methodology

To extract displacements from imagery, the principles of pattern matching are used. In this procedure a small image subset ($\mathbf{I}_1$) at a time instance and spectral range is taken from a satellite scene at a regular grid interval. At the same map location a similar or larger subset of a second image ($\mathbf{I}_2$) is extracted. Here 1 and 2 denote the template and search chip, respectively. The displacement is then estimated through a similarity metric. This procedure is repeated for every point of the grid, resulting in a field of potential planimetric displacements with associated similarity scores.

Matching of imagery can be done in both the spatial and the frequency domain. Here we implement the stacking in the frequency domain, and here we follow and extend the procedure as presented in Leprince et al. (2007), as this procedure is currently present within the popular COSI-Corr[2] package. A schematic of a matching setup is shown in Fig. 1.

3.1. General image matching

Here and throughout this work the boldcase capitals denote a matrix. It starts with an image subset ($\mathbf{I}$) of all bands b of a multi-spectral image. Each band of the subset is transformed ($\mathcal{F}_2$, i.e. a 2D Fourier transform) into a representation based on frequencies (ω) along its two spatial axis (x, y), resulting in a complex valued spectrum ($\mathbf{S}$). A multi-spectral image with acquisition number p and band identification a, is transformed as,

$$\mathbf{S}_{t=p}^{b=a}(\omega_x, \omega_y) = \mathcal{F}_2(\mathbf{I}_{t=p}^{b=a}(x, y)). \tag{1}$$

This spectrum is then weighted with a raised cosine filter ($\mathbf{V}$ (Stone et al., 2001; Leprince et al., 2007)), in order to suppress high frequency components in the spectrum and reduce the influence of the template borders. Here the first template has a stronger cut-off frequency putting emphasis on the center of the image template ($\mathbf{V}_1(\beta_1)$, $\beta_1 = .35$), while the search template is less suppressed ($\mathbf{V}_2(\beta_2)$, $\beta_2 = .50$), as suggested by Leprince et al. (2007).

The unadjusted complex spectrum can be used, but additional modifications can also be applied to the image spectrum at this stage. One approach is to normalize both image spectra ($\widehat{\mathbf{S}} = \mathbf{S}/|\mathbf{S}|$), which is known as phase correlation (Kuglin and Hines, 1975; Foroosh et al., 2002) or phase-only correlation (Horner and Gianino, 1984). Another approach is to normalize only the first spectrum, and keep the second spectrum unadjusted, this methodology is named symmetric phase-only filtering (Wernet, 2005). If the amplitude of the second spectrum is suppressed even more ($\widehat{\mathbf{S}}_2 = \mathbf{S}_2/|\mathbf{S}_2|^2$) then this procedure is known as amplitude compensated correlation (Mu et al., 1988). The next step is to calculate the (weighted) cross-correlation spectrum ($\mathbf{Q}$),

$$\mathbf{Q}_{p,q}^{a,b}(\omega_x, \omega_y) = \left(\mathbf{V}_1 \odot \mathbf{S}_{t=p}^{b=a}(\omega_x, \omega_y)\right) \odot \left(\mathbf{V}_2 \odot \mathbf{S}_{t=q}^{b=b}(\omega_x, \omega_y)\right)^{\star}. \tag{2}$$

Here the * denotes the complex conjugate and $\odot$ an element-wise multiplication, or hadamard product. If no normalization is applied to the cross-spectrum, this is the cross power density as obtained after Fourier-transformation of the normal cross-correlation function.

3.2. Cross-correlation spectra stacking

When the same time interval is present between pushbroom recordings, ensemble methods can be applied. If cross-correlation stacking is implemented (Li et al., 2021), the cross-correlation spectrum is

[2] available at: https://tecto.gps.caltech.edu/slip_history/spot_coseis/.

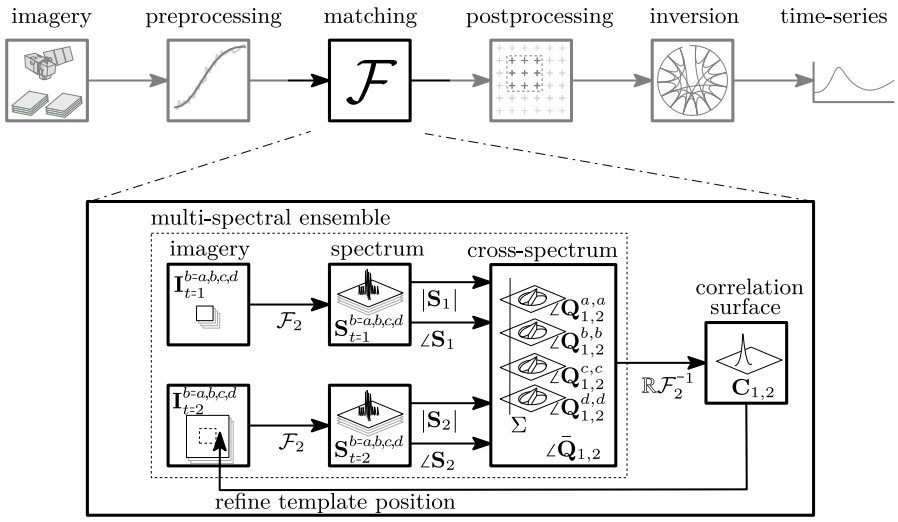


Fig. 1. Schematic showing at the top the building blocks of a geodetic imaging pipeline. The zoom-in shows the matching back-bone implemented in this study.

converted back to the spatial domain via an inverse Fourier transform. Then the different correlation functions ($\mathbf{C}(x,y) = \Re[\mathcal{F}_2^{-1}(\mathbf{Q})]$) are stacked and averaged. However, here we apply the stacking one step earlier, within the frequency domain. For that, different cross-correlation spectra ($\mathbf{Q}_{p,q}^{a,b}$) are calculated and can be summed to get a mean cross-correlation spectrum (here denoted by an overbar),

$$\overline{\mathbf{Q}}(\omega_x,\omega_y) = \Sigma_i^{N-1}\frac{1}{(N-1)}\mathbf{Q}_{p,p}^{i,i+1}(\omega_x,\omega_y). \tag{3}$$

Here i is the channel or band identifier and N is the total number of used pair combinations. This formulation is for a configuration when sequentially acquired band combinations from the same overpass p. The bands i and $i+1$ are a few seconds apart, and these are also known as interband pairs. A similar formulation can be given for multispectral imagery with p and q having a longer timespan, for example when two multispectral images from different acquisitions p, q where pairs of equal bands are used,

$$\overline{\mathbf{Q}}(\omega_x,\omega_y) = \Sigma_i^{N}\frac{\mathbf{Q}_{p,q}^{i,i}(\omega_x,\omega_y)}{N}. \tag{4}$$

In this study cross-correlation spectra from different spectral bands are combined (but spectral bands can also be interchanged by different polarizations acquired by SAR sensors), hereafter called interdate pairs. Fig. 2 illustrates that when an increased number of pairs are combined, a wider area of the phase plane ($\angle\overline{\mathbf{Q}}$) surpasses the noise level.

Equations (3) and (4) average the correlation spectra sequentially, which is computationally efficient (as each spectrum needs to be computed only once). Though this mean spectrum is composed of individual cross-spectra that carry information of the cross-power spectral amplitudes (see Equation (2)). The spread of these amplitudes for individual frequencies in a cross-correlation spectrum can range several orders of magnitude. Consequently amplitude information is transferred into this weighting of Equations (3) and (4), through element wise multiplication. Thus when phase correlation is used weighting is equal for all frequencies and bands. Emphasis on large amplitudes is enhanced for the case of ordinary cross-correlation processing, or something intermediate with amplitude compensated correlation. These different processing approaches are studied later on.

3.3. Phase plane fitting

After the cross-power spectral mean is calculated, the phase angle ($\angle\overline{\mathbf{Q}}$) can then be used to directly estimate the displacement. This is possible through several methods, see Tong et al. (2019) for an extensive description, but for the results presented here the procedures as in Leprince et al. (2007) are adopted. Briefly, this is done in a two-step estimation procedure, where the first displacement estimate is done in the spatial domain ($\mathbf{C}(x,y) = \Re[\mathcal{F}_2^{-1}[\angle\mathbf{Q}(\omega_x,\omega_y)]]$) to estimate the displacement with pixel-accuracy. Hereafter, the second template is shifted towards the location of the highest peak and a new template is sampled. With this second cross-power spectrum the phase angle is used to estimate the sub-pixel displacement (Δ_x, Δ_y) (Stone et al., 2001). This is done by minimizing the phase angle difference ($\Delta\varphi(\omega_x, \omega_y)$).

$$\Delta\varphi(\omega_x,\omega_y) = \mathbf{W}\odot\|\mathbf{Q}(\omega_x,\omega_y) - e^{j\left(\omega_x\Delta_x+\omega_x\Delta_y\right)}\|^2 \tag{5}$$

The minimization is done through a gradient descent method, with an iteratively adjusted weighting matrix ($\mathbf{W}$) that is dependent on the goodness of fit of the phase angle difference (Leprince et al., 2007). It

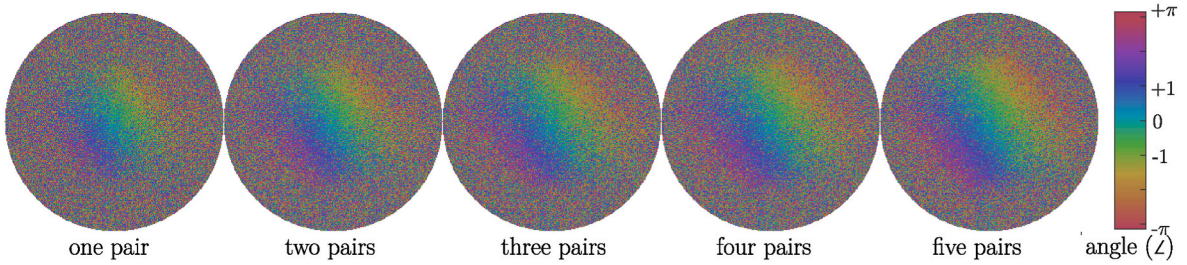


Fig. 2. Phase angle ($\angle$) of cross-correlation spectra of a subset of RapidEye imagery for an increasing number of N of stacking pairs ($\overline{\mathbf{Q}}$). Averaging an increasing number of band combinations (Eq. (4)) enhances the visibility of the phase plane in the center of the spectrum. The cross-power spectra are disks due to the raised cosine filtering that cuts-off high-frequencies.

starts with a binary selection on the cross-spectrum, where all log-transformed frequencies ($\log_{10}\|\mathbf{Q}(\omega_x, \omega_y)\|$) above the mean are used. This weighting matrix is then adjusted at every iteration through,

$$\mathbf{W}_{i+1}(\omega_x, \omega_y) = \mathbf{W}_i(\omega_x, \omega_y)\left(1 - \frac{\Delta\varphi_i(\omega_x, \omega_y)}{4}\right)^n \tag{6}$$

where n is typically set to 6. After convergence, the signal to noise ratio (SNR) can be computed based upon the element-wise sum over the final weighting matrix (Leprince et al., 2007),

$$\mathrm{SNR} = 1 - \frac{\Sigma_{\omega_x}\Sigma_{\omega_y}\Delta\varphi}{4\Sigma_{\omega_x}\Sigma_{\omega_y}\mathbf{W}} \tag{7}$$

Such a measure describes the angular noise in relation to an ideal phase plane (SNR = 1), since the domain for computation of this measure is governed by the weighting matrix but this influence is not represented within this measure. Another proxy, that is introduced here, is the amount of support. That is the spread or amount of phase angle information that is used for the estimation, and is calculated through

$$\mathrm{Support} = \Sigma_{\omega_x}\Sigma_{\omega_y}\mathbf{W}/N_{\mathbf{W}} \tag{8}$$

where $N_{\mathbf{W}}$ denotes the amount of elements in the weighting matrix. Hence, this measure ranges from zero (when no angles are used) to one (when all angles are taken into account).

In this study we analyzed both combinations of equation (3) and 4 with the aim of demonstrating the improved extraction of surface displacements from multi-spectral images. The interband cross-correlation spectra stacking (Eq. (3)), is highlighted on the application of sea surface circulation. While the stacking of cross-correlation spectra for interdate pairs (Eq. (4)) is demonstrated on co-seismic displacements.

4. Study region & data

To evaluate the method of cross-correlation stacking correlation stacking for multispectral bands, we use two RapidEye acquisitions of Kaiko¯ura. Kaiko¯ura is situated at the northeastern coast of the southern island of New Zealand (Fig. 4). On the November 13, 2016 an earthquake (M_w 7.8) with an associated 180 km long surface rupture occurred along the Kekerengu fault, with displacements up to 12 m. The earthquake caused surface displacements, landslides, coastal up-lift, landslips and a tsunami. We use this displacements to asses the performance of stacking of multispectral interdate correlations. Documentation of the planimetric displacements through optical remote sensing data has been done by several satellite system (Kääb et al., 2017; Hollingsworth et al., 2017; Howell et al., 2020). Details and analysis on the tectonic significance can be found therein, as interpretation of the kinematics is not the scope of this study. Due to the sequential imaging of different spectral band with RapidEye, we also use this data to demonstrate spectral stacking of interband-cross correlations. This makes it possible to map the wave speed of the coastal upper surface.

4.1. Repeat pass RapidEye satellite imagery

The imagery used to map the displacement of the Kaiko¯ura earthquake are acquired on the 31st of October and the November 17, 2016 by the RapidEye system. This constellation is composed of five identical mini-satellites (Tyc et al., 2005). They are equally seperated within the same orbital plane, resulting in a repeat period of 5.5 days. In 2008 the constellation became operational until the start of 2020, when they were decommissioned. Each RapidEye satellite has a multi-spectral imaging system onboard, that is composed of five parallel CCD arrays mapping different spectral bands. The filters in front of each array are centered around the blue (440–510 nm), green (520–590 nm), red (630–685 nm), red edge (690–730 nm) and near infrared (760–850 nm) wavelengths. The nadir ground sample distance is 6.5 m, but the orthorectified products are sampled onto a product with a 5 m resolution, see also Fig. 3 for the geometric sensing configuration.

4.2. Interband recordings

A general overview for configurations of surface flow extraction at different time-scales from a suite of optical satellite configurations is given by Kääb and Leprince (2014). Here we use data from the RapidEye satellite constellation from the November 17, 2016, which has shown to be capable of generating interband displacements (Krauß et al., 2013) (in that case traffic flow estimation). The pushbroom configuration has three bands in the forward looking position, and two looking slightly backwards with an incidence angle difference of two degrees (see Fig. 3). Consequently, the maximum time separation achievable with the RapidEye system is 3 s ($\Delta t_{red-blue}$) when the first (red) and last (blue) band is used. However, the interval between the CCD arrays in forward or backward position can also be used, resulting in a shorter time interval($\Delta t_{blue-green}$ = 0.4s), but providing three more combinations to stack.

The short consecutive time delay from these interband recordings between adjacent sensor arrays can be used (3 combinations) to estimate the swell along the coast of Kaiko¯ura through different methodologies, though all are based upon the same principle: through the analysis of pair-wise cross-correlation spectrum (Yurovskaya et al., 2019). Here we use the stacking of interband cross-correlation spectra as outlined above and the implementation is similar to the approach presented by De Michele et al. (De Michele et al., 2012), which is based upon Leprince et al. (2007), but only uses a single band pair.

5. Results

Both the land and ocean displacements are estimated with a template size of 16 pixels, corresponding to a template of 80 by 80 m. The displacements are estimated without overlap, so a movement is estimates on a grid with an 80 m interval. This is half the template edge length used in the frequency domain matching by Leprince et al. (2007) who

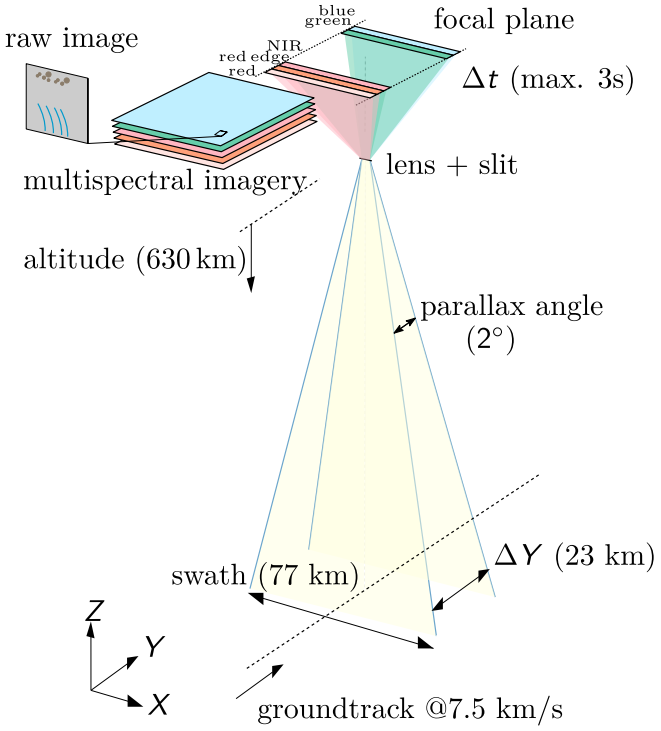


Fig. 3. Schematic of the geometry of the multispectral recording of the RapidEye satellite system. Configuration is not scaled uniformly.

found a heuristic lower limit of 32 pixels. This constrain on the template edge length is a result from the noise at high frequencies in the phase plane (Fig. 2), in combination with a limited resolution for the phase angle estimation as phase angle further away from the center of the phase plane can make the displacement estimation more precise. However, due to the stacking presented here, the SNR of the phase plane is significantly enhanced at these frequencies making it feasible to use smaller template sizes.

5.1. Near-shore currents from interband pairs

To the best of our knowledge, observations are lacking of sea surface swell over the study region. However a general modelled geostrophic (sub-)surface currents from GlobCurrent at the time of overpass is available (plotted in Fig. 4 on the left) and can give context to our results. The vector field shows the d'Urville current flowing through Cook Strait that is situated between northern and southern island. Similarly in the bottom a part of the Southland current travels upwards, and an eddy forms just east of the study region. New Zealand is part of a submerged continent, and therefore surrounded by a shallow shelf sea, consequently wind and tides are a key driver of the wave regime (Stevens et al., 2019). Thus the shown geostrophic flow (Fig. 4, left) that appears to be perpendicular to the observed swell direction (Fig. 4, middle) will not describe this region very well, and might be less informative for our study region.

The swell displacements from the RapidEye data along the near shore of Kaiko¯ura are shown in Fig. 4 (middle). Velocities up to 3 $m \cdot s^{-1}$ are estimated (and displayed as tints), while the directions are shown as streamlines. The swell angle arriving at the coast can be clearly seen, as well as, the short deflection at Kaiko¯ura and Mangamaunu. Another striking feature is the bending line present in the southern part of the study region. This swing is caused by a large ship generating wakes, the western part after this line has a swell angle similar to the direction of the Kelvin wake. Suggesting both wind and boat generated swell velocities are sensed. When looked at the correlation score (Fig. 4, right), the phase-stacking of multiple pairs has a significant advantage over single pair processing. Especially close to the coast where the swash zone is situated the matching score is significantly increased, show casting the advantage of phase-stacking.

5.2. Co-seismic displacement from interdate pairs

The displacement results over land are shown in Fig. 5, here the main rupture over the Kekerengu Fault is clearly observable in the north-south displacement. Only the eastern part of the rupture is visible, as a cloud was situated in the imagery and resulted in a large patch with low correlations scores. A clear distinction between the single band pair and the stacked solution can be observed. When zooming into the Fig. 5 (left) a far more crisper appearance and more details is being present, which is also manifested in the significantly higher correlation scores (middle). The colour scheme in Fig. 5 is the same as Fig. 4 but in general the correlation scores are lower. This might stem from appearance change due to different orbits or incidence angles in combination with temporal decorrelation within the 17 days between the acquisitions. Nevertheless, the coverage is greatly improved and when looked closer to the displacement results, the stacked results only seem to have trouble with river beds and streams.

A quantitative assessment of different processing procedures is done here through statistics. The cumulative histograms of the estimated matching support (Eq. (8)) are shown in Fig. 6a. A clear improvement can be seen between the use of one or multiple band pairs. From the implementations that use multiple band pairs, phase correlation is underperforming, and is not assessed any further hereafter. While the normal procedure of cross-correlation seems to have the highest absolute scores in the higher regions of its distribution, but the amplitude compensated correlation is able to push low scoring matches slightly towards a higher score (the cross-over point is around 0.24). Hence stacking with standard cross-correlation is able to improve already clear matches, while amplitude compensated correlation is able to shove some matches above the noise level. This can also be seen in Fig. 6b, where the cumulative histograms of the local vector coherence (Burgess et al., 2013) are plotted. Since the estimated displacements are a combination of surface displacement and orthorectification offsets (Fig. 10), it is difficult to compare against one another. Here the choice for local vector coherence is made since both displacement phenomena are slowly changing in the direct neighborhood in respect to the resolution

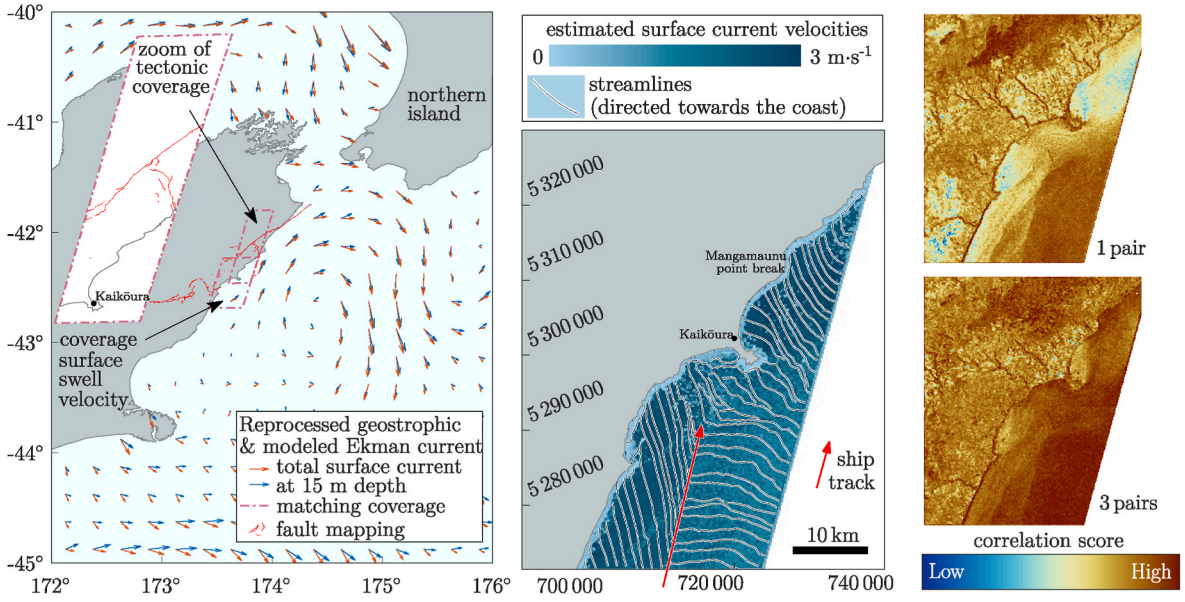


Fig. 4. Left) Subset of two different levels of geostrophic currents around the time of acquisition on the 17th of November, based upon Copernicus GlobCurrent data. The tectonic fault mapping around Kaikoura as also displayed in Figs. 5 and 7, are shown in red. From this red footprint only the lower part covers the ocean, hence the extend of the ocean results are therefore cropped and showed alongside. Middle) Streamlines of the estimated swell from on the November 17, 2016. Right) Correlation scores (Eq. (7)) of the displacement estimate for single pair and phase-stacking of three pairs, using a template size of 16 by 16 pixels. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

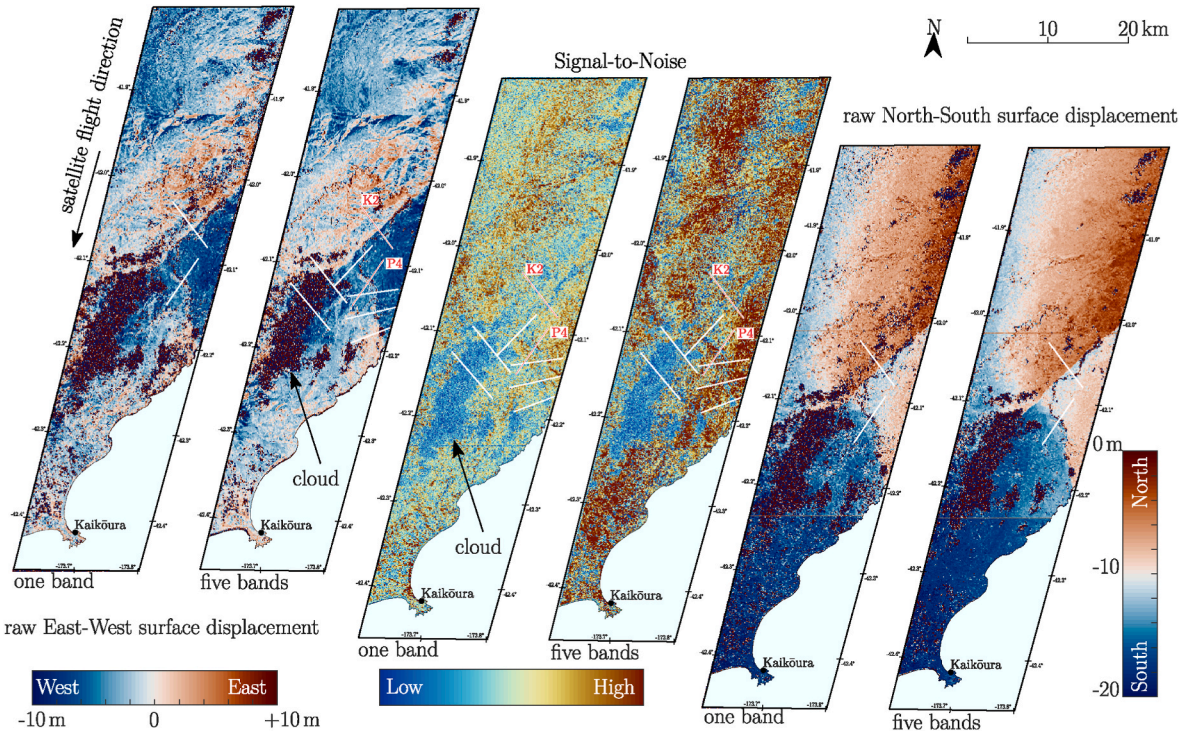
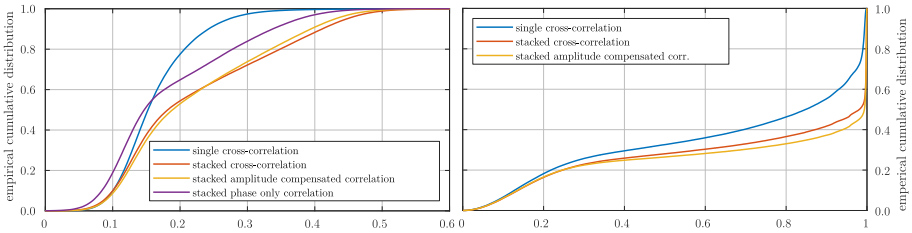


Fig. 5. Estimated co-seismic displacement for single band pair matching and cross-power spectra stacking of five band pairs, using a template size of 32 by 32 pixels. The footprint location is illustrated in Fig. 4.



(a) Matching support (Equation 8) (b) Local vector coherence

Fig. 6. Cumulative histogram of matching metrics over the land displacement estimates, for different processing approaches. Matching support describes the amount of frequencies in the spectra used for the displacement estimation, while local vector coherence is a metric related to smoothness in the direct spatial surrounding of the displacement estimate. The spatial distribution of these metrics are shown in Fig. 7.

of the displacement product, thus the local vector coherence can be an appropriate proxy to assess the performance of different matching and stacking methodologies.

The spatial distribution of both proxies are shown in Fig. 7. A clear improvement can be seen between single or multiple band processing. The pattern and amplitude of the support metric is similar to the signal-to-noise (see Fig. 5). While these metrics do describe a different property, the similar distribution might stem from less repetitive and an unique image structure. If a match is precise it also has good support. The support for the amplitude compensated results has a shorter long tail, while the cross-correlation results seems to have more scattered and slightly higher scores. However, the amplitude compensated results might be preferred, which becomes more clear in the vector coherence plots on the righthand side.

5.3. Displacement verification

The relative displacement estimates are compared against another independent optical remote sensing data-set, in order to assess the measurement precision. For this we use the displacement estimates from Zinke et al. (2019). These displacements are based upon oblique very high resolution WorldView imagery at 50 cm resolution. The method presented in Zinke et al. (2019) estimated planar displacement and elevation change simultaneously, but here only the former is used. Since, our RapidEye displacement estimates in the across-track direction are dominantly corrupted by orthorectification offsets with local variations, only the along-track displacements are analyzed. In the along-track direction variation is still present, but these seem to be more subtile, more on this in the discussion. We show results in Fig. 8 over similar transects as presented in Klinger et al. (2018).

The WorldView data generates a displacement field at 8 m resolution, this high resolution creates detail that is clearly visible in the results shown in Fig. 8. While the RapidEye data is less detailed due to its lower resolution, the presence of noise or outliers seems to be very limited or absent, while the WorldView data has been cleaned up by a zonal median filtering. When looked at the linear regression of the fault displacement, there is a good agreement. The absolute scale is different, but this is due to absolute coregistration errors. However, the along-track displacement seems to deviate further away from the fault. This might be caused by the systematic shear from CCD-array mis-

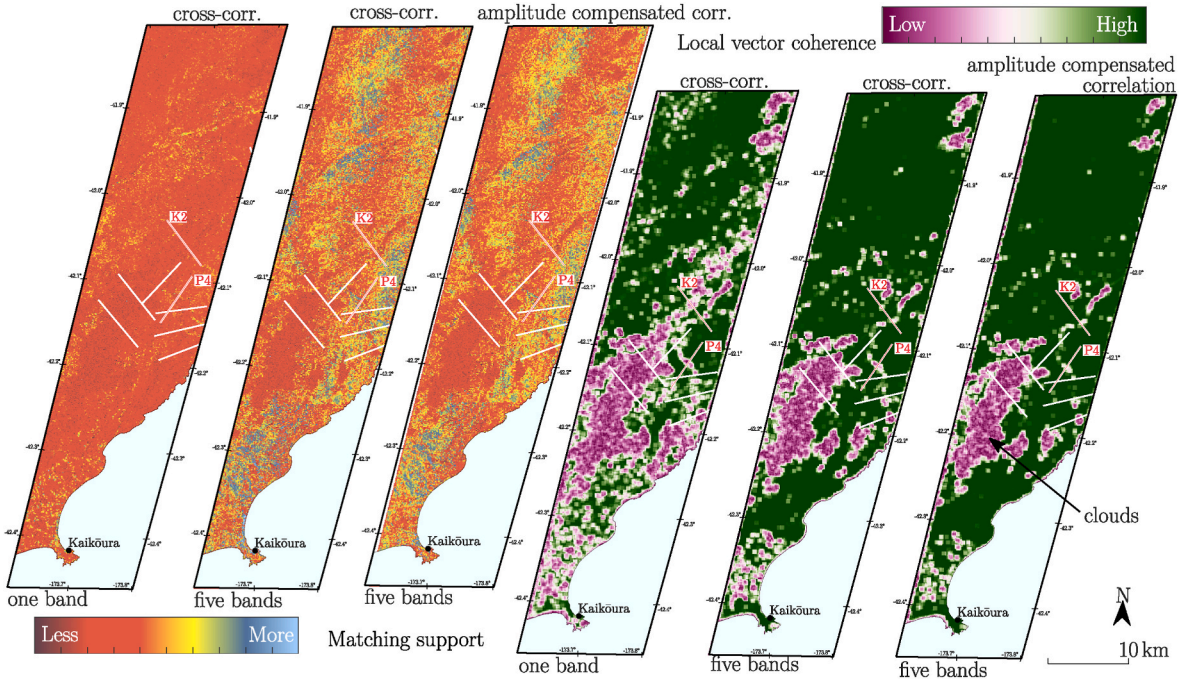
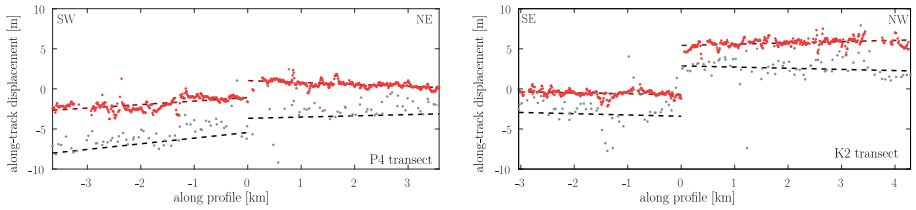


Fig. 7. Estimated support and local vector coherence for phase correlation with single band pair matching, phase correlation stacking of five band pairs and amplitude compensated matching. The footprint location is illustrated in Fig. 4, while their histograms are shown in Fig. 6.



(a) P4 transect (b) K2 transect

Fig. 8. Along-track displacement over two transects, their location is shown in Fig. 10. The dots are individual displacement estimates. Red dots correspond from 50 cm WorldView satellite data, while the grey dots come from 5 m RapidEye imagery. The dashed lines are linear regression at either side of the fault. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

registration as discussed previously, or jitter.

Verification of the swell velocities with the surface currents is not done here. This would mean downscaling of the geostrophic flows and upspinning of, for example, a SWAN model (Simulating WAve Nearshore). Since, this study is based upon the same methodology as the tectonic displacement, the assessment might be transferable. In any case, the example is shown to present that more information can be drawn out of a single multispectral RapidEye image when multi-spectral cross-correlation stacking is applied compared to a single pair-wise cross-correlation.

5.4. Computational performance of image matching

In order to assess the performance of stacking in the frequency domain in relation to other ensemble matching approaches, a small test is conducted. This analysis is based upon a machine with 8 cores at 1.8

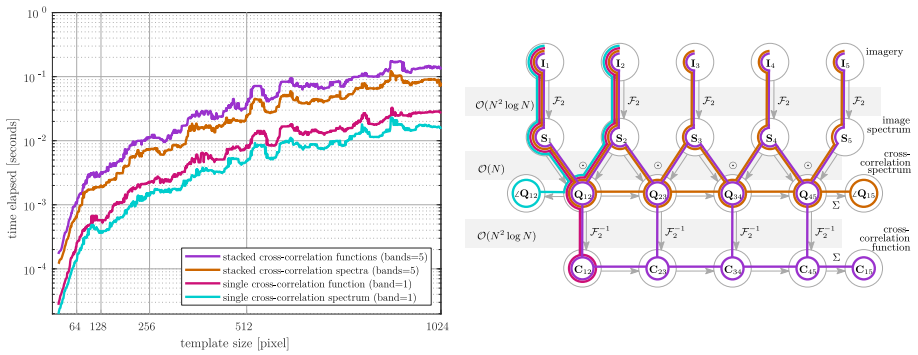


Fig. 9. At the left side is a performance graph of processing times for variable window sizes, through different processing approaches. On the right is a processing diagram, where the processing steps of the different approaches are colored the same way as shown in the graph. For both, single pairs as well as five cross-correlation pairs, stacking of cross-correlation spectra is about a factor 2 faster than stacking cross-correlation functions (left) as the final (inverse) Fourier-transform is not required.

GHz with 64 bit and 12 GB RAM memory. The timing of displacement calculation is recorded for different window sizes. A visualization which computation are done extra or can be neglected is schematically illustrated in a flow diagram within Fig. 9. The run time experiment, shows the stacking procedure scales with a constant factor depending on the amount of image pairs. An improvement of a factor two is achieved when the (inverse) Fourier transform operation is not needed, as is the case for the stacking cross-correlation spectra instead of cross-correlation functions. Furthermore, the runtime for the calculation of cross-correlation function for one band, does align to the same runtime for a five band cross-correlation spectra with a power of two smaller template size. Our results show stacking of cross-correlation spectra is able to enhance the matching coverage, hence it can provide a displacement field at a higher resolution, at the same computational time as a normal processing pipeline.

6. Discussion

In this contribution we have introduced the stacking of the cross-correlation in the frequency domain. This has the advantage of estimating displacement through the phase plane ($\angle\mathbf{Q}$) compared to transforming the phase plane back to the spatial domain and do a sub-pixel localization of the correlation surface ($\mathbf{C}$). A first reason to do so, is because the discrete version of the convolution theorem are taylored towards an integer displacement (Foroosh et al., 2002). A second reason is rooted in the versatility of frequency methods, as translational and affine components are decoupled in the phase and magnitude of the spectrum (Krüger and Calway, 1998). This will ease matching when an affine model is more suited, which is an interesting direction to explore. Hence these reasons make us prefer to do the stacking and displacement estimation within the frequency domain.

6.1. Potential aspects influencing matching performance

Many approaches that stack cross-power spectra seem to outperform single pair processing. A reason why this is the case, might be due to the different spectral or polarized reflections of the surface. Surface reflections are different between different bands, altering the image structure. However, also the geometric sensing configuration might play a role, since bi-directional reflection functions (BDRF) also alter the appearance. Both effects can cause a different image configuration from band to band, and through that put emphasis on other frequency components. This gradually enhance the visibility of the phase plane with every pair, as more frequencies emerge above the noise level (e.g. Fig. 2).

When amplitude compensated correlation is used, a more complete solution is given (e.g. Fig. 6). Explanation for this can be found in the amplitudes of the Fourier transform. When phase correlation is used, the weighting of the mean over all frequency components is equal since the cross-spectrum is normalized. However, if energy is present at a certain frequency and direction, it is an essential part of the signal, thus amplitude dependent weighting in the stacking process is beneficial. However, the amplitudes can be of sufficient magnitude ranging several orders of magnitude, especially if these are multiplied, as is the case for cross-correlation ($\mathbf{S}_1 \odot \mathbf{S}_2$). Amplitude compensated correlation ($\widehat{\mathbf{S}}_1 \odot \mathbf{S}_2$) seems to be a middle ground, where the weighting of frequency components is not amplified. For locations where the signal of the phase plane is close to the noise level, it is able to extract the common orientation. While stacking of cross-correlation spectra mostly seem to enhance a signal that is already moderately observable at a single pair calculation.

6.2. Geometric performance of RapidEye

Each spectral band of a RapidEye satellite is composed of a single CCD array. This is different from systems such as Sentinel-2 or Landsat 8, where several ramp errors within an acquisition can be present (Kääb et al., 2016). However, CCD-array mis-registration might still occur (Ayoub et al., 2008), and this seems to be most present in the N-S displacements. The differences are more pronounced when the displacements are oriented towards an along- and across-track coordinate system ($\parallel$ & $\perp$, respectively), as in Fig. 10. The along-track displacement show the tectonic displacement, but most pronounced is a consistent shear that can be attributed to the sensor, see also Stumpf et al. (2018).

More local undulations, due to uncompensated attitude distortions, might still be present. Instruments on the ZiYua-3 satellite have a somewhat similar sensor specifics as RapidEye and for this satellite the residual jitter can be in the order of 0.3 pixels (Tong et al., 2015). Interband matching can be used to estimate and compensate these effects (Ye et al., 2019), but we do not see these sub-pixel oscillations in our single pair matching results. They might thus be small, though these fluctuations occur at different attitudes and when these spectra are stacked, smearing occurs. Jitter patterns have been identified for Sentinel-2 (Kääb et al., 2016), and it can thus be expected that stacking of such multi-spectral data will influence the precision for such data, as jitter dilutes the phase signal. Though, this will be less the case for other imaging configurations that acquire simultaneously through a dichroic filter, frame camera or a polarized synthetic aperture radar (SAR) imaging system.

In the across-track direction the patterns appear to have a local character. Since the acquisitions are sensed from two different orbits, it has terrain correlated displacements in across-track direction (Altena and Kääb, 2017). When looked at the terrain, these patterns correspond to the across-slope as well as the parallel slope direction. Hence, it is a combination of errors stemming from the absolute location of satellite, and orthorectification offsets due to the earthquake induced terrain movement.

The different vantage points from the forward and backward looking sensor arrays result in a small parallax in along-track direction (Delon and Rougé, 2007). If for our results, the ramp error present in the along-track displacement ($d_{\parallel}$) is removed via a planar fit, then the residual ($\widetilde{d}_{\parallel}$) shows correspondence with the topographic signal, as is also shown in Fig. 11. If this is set out against the aspect of the terrain, against

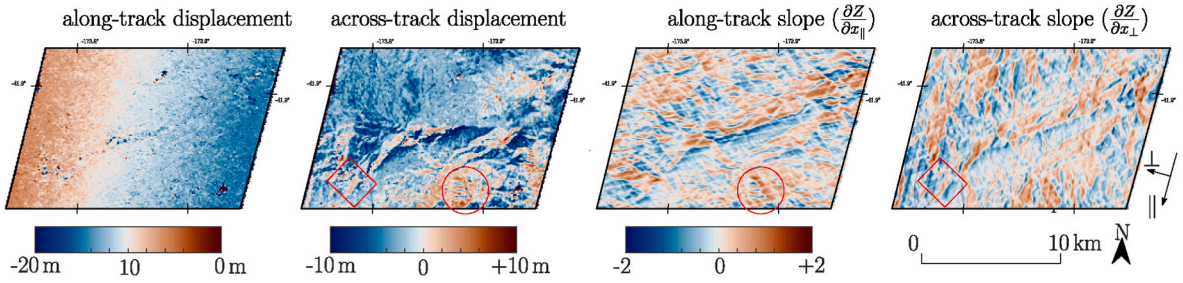


Fig. 10. Zoom in of the displacements estimates as shown in Fig. 5, but now turned towards a satellite centered coordinate frame. Similarly, the gradients of the terrain over the same region are plotted.

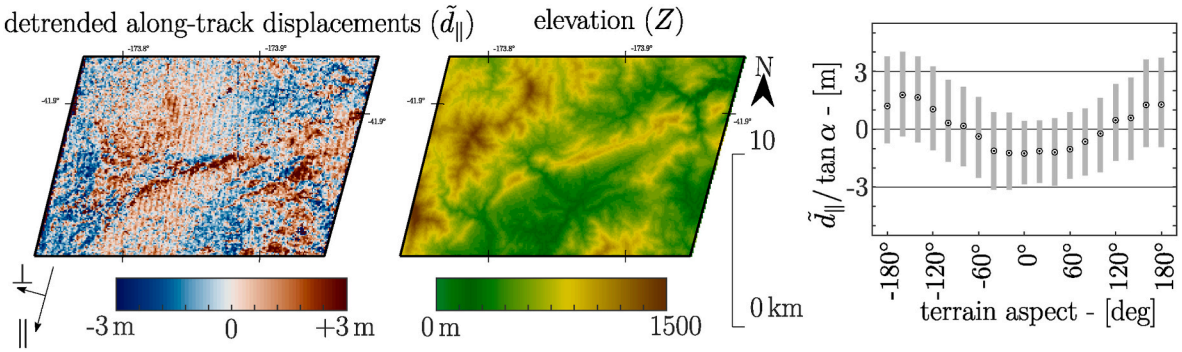


Fig. 11. Detrended along-track displacements estimates as shown in Fig. 10. The subfigure in the middle shows the elevation, while the rightmost subfigure illustrates the aspect dependency of these along-track displacements over this sub-region.

the along-track displacement normalized by terrain slope, through a similar procedure as Nuth and Kääb (2011), a sinusiodal relation is present. This hints towards a translational bias still present in the data, while stereographic aspects are less visible in our results.

Literature on the analysis of the geometric performance of RapidEye is limited, and if present, studies seem to focus on the absolute geolocation of the imagery (Chander et al., 2013; Behling et al., 2014; Scheffler et al., 2017). Co-registration is implemented by a constant offset, excluding the influence of orthorectification and jitter effects. This is remarkable, since the RapidEye constellation has changing orbits and thus acquisitions from different view points. Fortunately, the native 6.5 m resolution Level-1 data is also available, so such errors can be accounted for. Though this is not done here, as the main innovation of this work is the demonstration of a method. Instead of orthorectified imagery, as used here, it is preferred to use raw imagery and proper ground control. Though ideal, such an approach might not always be possible when raw imagery is unavailable (e.g.: Sentinel-2, Landsat9). Consequently, orthorectification effects are present in our results, but do not influence the performance of the method.

7. Conclusion

In this contribution we demonstrate an improved extension of image matching to include multiple spectral bands. Enhancement of the displacement signal is achieved by taking the mean of several band-wise correlation spectra. This improves both the initial displacement, as well as, the sub-pixel localization, creating a more complete displacement field. The computational load is low, as both displacement estimates are implemented in the frequency domain, and are not transformed back to the spatial domain. Neither is there a need to carefully select a specific band to match the dominant spectral signature in the scenes, as information from any channel is taken into account. Its computational load is limited and reduces or diminishes the importance of post-processing routines for outlier detection.

This methodology of correlation-stacking, or ensemble matching, has previously been demonstrated on SAR (Li et al., 2021) and optical imagery (Altena and Kääb, 2020), though this was based upon data from several sequential time intervals. While for this technique the multi-spectral nature of the optical sensors makes single timespan processing possible when the spectral bands are acquired at subsequent times.

Our methodology of cross-correlation spectra stacking is demonstrated on RapidEye data over the Kaikō ura region, New Zealand. Surface ruptures from the 2016 earthquake where clearly mapped and in agreement with very-high resolution WorldView results. In addition, other artifacts related to sensor geometry and processing become more apparent than single pair processing. We observe our methodology works best when amplitude compensated cross-correlation spectra are stacked. When cross-correlation spectra stacking is applied, smaller templates can be used and post-processing or other types of filtering are less needed. Furthermore, the mapping capabilities for swell movements were demonstrated, mapping a complete velocity field of several kilometers along the coast and outwards into the sea using interband cross-correlation spectra stacking.

Funding

B.A. is funded through a Nederlands Wetenschappelijk Onderzoek (NWO) project (ALWGO.2018.044). S.L. is funded through the French ministry of Europe and foreign affairs, within the framework of the 'Make Our Planet Great Again' program for postdoctoral researchers.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

We thank three anonymous reviewers and handling editors Shunlin Liang and Yong Wang for providing insightful comments that helped improve this article. We also would like to thank Robert Zinke for providing the high resolution WorldView co-seismic displacement estimates, and Planet for providing the RapidEye data via their Ambassadors Program.

References

Ahn, Y., Howat, I., 2011. Efficient automated glacier surface velocity measurement from repeat images using multi-image/multichip and null exclusion feature tracking. IEEE Trans. Geosci. Rem. Sens. 49 (8), 2838–2846. https://doi.org/10.1109/TGRS.2011.2114891.

Altena, B., Kääb, A., 2017. Elevation change and improved velocity retrieval using orthorectified optical satellite data from different orbits. Rem. Sens. 9 (3), 300. https://doi.org/10.3390/rs9030300.

Altena, B., Kääb, A., 2020. Ensemble matching: measuring ice flow from repeat satellite images over fast-changing glacier parts. Demonstration for Khumbu icefall, Mount Everest. J. Glaciol. 66 (260), 905–915. https://doi.org/10.1017/jog.2020.66.

Altena, B., Scambos, T., Fahnestock, M., Kääb, A., 2019. Extracting recent short-term glacier velocity evolution over southern Alaska and the Yukon from a large collection of Landsat data. Cryosphere 13, 795–814. https://doi.org/10.5194/tc-13-795-2019.

Ayoub, F., Leprince, S., Binet, R., Lewis, K., Aharonson, O., Avouac, J.-P., 2008. Influence of camera distortions on satellite image registration and change detection applications. IGARSS 2008. IEEE International. In: Geoscience and Remote Sensing Symposium, 2. IEEE, II–1072, 2008.

Behling, R., Roessner, S., Segl, K., Kleinschmit, B., Kaufmann, H., 2014. Robust automated image co-registration of optical multi-sensor time series data: database

generation for multi-temporal landslide detection. Rem. Sens. 6 (3), 2572–2600. https://doi.org/10.3390/rs6032572.
Burgess, E., Forster, R., Larsen, C., 2013. Flow velocities of Alaskan glaciers. Nat. Commun. 4 (1), 2146. https://doi.org/10.1038/ncomms3146.
Chander, G., Haque, M., Sampath, A., Brunn, A., Trosset, G., Hoffmann, D., Roloff, S., Thiele, M., Anderson, C., 2013. Radiometric and geometric assessment of data from the RapidEye constellation of satellites. Int. J. Rem. Sens. 34 (16), 5905–5925. https://doi.org/10.1080/01431161.2013.798877.
Charrier, L., Yan, Y., Koeniguer, E., Leinss, S., Trouvé, E., 2021. Extraction of velocity time series with an optimal temporal sampling from displacement observation networks. IEEE Trans. Geosci. Rem. Sens. 60 (4302810), 1–10. https://doi.org/10.1109/TGRS.2021.3128289.
De Michele, M., Leprince, S., Thiébot, J., Raucoules, D., Binet, R., 2012. Direct measurement of ocean waves velocity field from a single SPOT-5 dataset. Rem. Sens. Environ. 119, 266–271.
Delon, J., Rougé, B., 2007. Small baseline stereovision. J. Math. Imag. Vis. 28 (3), 209–223. https://doi.org/10.1007/s10851-007-0001-1.
Foroosh, H., Zerubia, J., Berthod, M., 2002. Extension of phase correlation to subpixel registration. IEEE Trans. Geosci. Rem. Sens. 11 (3), 188–200. https://doi.org/10.1109/83.988953.
Heid, T., 2011. Deriving Glacier Surface Velocities from Repeat Optical Images. Ph.D. thesis, University of Oslo. URL urn.nb.no/URN:NBN:no-30427.
Hollingsworth, J., Ye, L., Avouac, J.-P., 2017. Dynamically triggered slip on a splay fault in the Mw 7.8, 2016 Kaikoura (New Zealand) earthquake. Geophys. Res. Lett. 44 (8), 3517–3525. https://doi.org/10.1002/2016GL072228.
Horner, J.L., Gianino, P.D., 1984. Phase-only matched filtering. Appl. Opt. 23 (6), 812–816.
Howell, A., Nissen, E., Stahl, T., Clark, K., Kearse, J., Van Dissen, R., Villamor, P., Langridge, R., Jones, K., 2020. Three-Dimensional surface displacements during the 2016 Mw 7.8 Kaikoura earthquake (New Zealand) from photogrammetry-derived point clouds. J. Geophys. Res. Solid Earth 125 (1), e2019JB018739.
Kääb, A., Leprince, S., 2014. Motion detection using near-simultaneous satellite acquisitions. Rem. Sens. Environ. 154, 164–179. https://doi.org/10.1016/j.rse.2014.08.015.
Kääb, A., Winsvold, S., Altena, B., Nuth, C., Nagler, T., Wuite, J., 2016. Glacier remote sensing using sentinel-2. Part I: radiometric and geometric performance, and application to ice velocity. Rem. Sens. 8 (7), 2072–4292. https://doi.org/10.3390/rs8070598.
Kääb, A., Altena, B., Mascaro, J., 2017. Coseismic displacements of the 14 November 2016 Mw7.8 Kaikoura, New Zealand, earthquake using the Planet optical cubesat constellation. Nat. Hazards Earth Syst. Sci. 17, 627–639.
Klinger, Y., Okubo, K., Vallage, A., Champenois, J., Delorme, A., Rougier, E., Lei, Z., Knight, E., Munjiza, A., Satriano, C., Baize, S., Langridge, R., Bhat, H., 2018. Earthquake damage patterns resolve complex rupture processes. Geophys. Res. Lett. 45 (19), 10–279.
Krauß, T., Stätter, R., Philipp, R., Bräuninger, S., 2013. Traffic flow estimation from single satellite images. ISPRS Archives 40 (WG I/4), 241–246.
Krüger, S., Calway, A., 1998. Image registration using multiresolution frequency domain correlation. In: British Machine Vision Conference, 32, pp. 1–10. https://doi.org/10.5244/C.12.32.
Kuglin, C., Hines, D., 1975. The phase correlation image alignment method. In: Proceedings of the IEEE International Conference on Cybernetics and Society, pp. 163–165.
Lacroix, P., Bièvre, G., Pathier, E., Kniess, U., Jongmans, D., 2018. Use of Sentinel-2 images for the detection of precursory motions before landslide failures. Rem. Sens. Environ. 215, 507–516. https://doi.org/10.1016/j.rse.2018.03.042.
Leprince, S., Barbot, S., Ayoub, F., Avouac, J.-P., 2007. Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground deformation measurements. IEEE Transactions on Geosci. Rem. Sens. 45 (6), 1529–1558. https://doi.org/10.1109/TGRS.2006.888937.
Li, S., Leinss, S., Hajnsek, I., 2021. Cross-correlation stacking for robust offset tracking using SAR image time-series. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 14, 4765–4778. https://doi.org/10.1109/JSTARS.2021.3072240.
Meinhart, C., Wereley, S., Santiago, J., 2000. A PIV algorithm for estimating time-averaged velocity fields. J. Fluid Eng. 122 (2), 285–289. https://doi.org/10.1115/1.483256.
Mu, G.-G., Wang, X.-M., Wang, Z.-Q., 1988. Amplitude-compensated matched filtering. Appl. Opt. 27 (16), 3461–3463.
Necsoiu, M., Leprince, S., Hooper, D., Dinwiddie, C., McGinnis, R., Walter, G., 2009. Monitoring migration rates of an active subarctic dune field using optical imagery. Rem. Sens. Environ. 113 (11), 2441–2447. https://doi.org/10.1016/j.rse.2009.07.004.
Nuth, C., Kääb, A., 2011. Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. Cryosphere 5 (1), 271–290. https://doi.org/10.5194/tc-5-271-2011.
Paul, F., Bolch, T., Kääb, A., Nagler, T., Nuth, C., Scharrer, K., Shepherd, A., Strozzi, T., Ticconi, F., Bhambri, R., et al., 2015. The glaciers climate change initiative: methods for creating glacier area, elevation change and velocity products. Rem. Sens. Environ. 162, 408–426.
Redpath, T., Sirguey, P., Fitzsimons, S., Kääb, A., 2013. Accuracy assessment for mapping glacier flow velocity and detecting flow dynamics from ASTER satellite imagery: Tasman Glacier, New Zealand. Rem. Sens. Environ. 133, 90–101. https://doi.org/10.1016/j.rse.2013.02.008.
Scheffler, D., Hollstein, A., Diedrich, H., Segl, K., Hostert, P., 2017. AROSICS: an automated and robust open-source image co-registration software for multi-sensor satellite data. Rem. Sens. 9 (7), 676. https://doi.org/10.3390/rs9070676.
Stearns, L., Hamilton, G., 2005. A new velocity map for Byrd Glacier, East Antarctica, from sequential ASTER satellite imagery. Ann. Glaciol. 41 (1), 71–76.
Stevens, C., O'Callaghan, J., Chiswell, S., Hadfield, M., 2019. Physical oceanography of New Zealand/Aotearoa shelf seas–a review. N. Z. J. Mar. Freshw. Res. 1–40. https://doi.org/10.1080/00288330.2019.1588746.
Stone, H., Orchard, M., Chang, E.-C., Martucci, S., 2001. A fast direct Fourier-based algorithm for subpixel registration of images. IEEE Trans. Geosci. Rem. Sens. 39 (10), 2235–2243. https://doi.org/10.1109/36.957286.
Stumpf, A., Michéa, D., Malet, J.-P., 2018. Improved co-registration of Sentinel-2 and Landsat8 imagery for Earth surface motion measurements. Rem. Sens. 10 (2), 160. https://doi.org/10.3390/rs10020160.
Tong, X., Li, L., Liu, S., Xu, Y., Ye, Z., Jin, Y., Wang, F., Xie, H., 2015. Detection and estimation of ZY-3 three-line array image distortions caused by attitude oscillation. ISPRS J. Photogrammetry Remote Sens. 101, 291–309. https://doi.org/10.1016/j.isprsjprs.2015.01.003.
Tong, X., Ye, Z., Xu, Y., Gao, S., Xie, H., Du, Q., Liu, S., Xu, X., Liu, S., Luan, K., Stilla, U., 2019. Image registration with fourier-based image correlation: a comprehensive review of developments and applications. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 12 (10), 4062–4081. https://doi.org/10.1109/JSTARS.2019.2937690.
Tyc, G., Tulip, J., Schulten, D., Krischke, M., Oxfort, M., 2005. The RapidEye mission design. Acta Astronaut. 56 (1–2), 213–219. https://doi.org/10.1016/j.actaastro.2004.09.029.
Wernet, M., 2005. Symmetric phase only filtering: a new paradigm for DPIV data processing. Meas. Sci. Technol. 16 (3), 601.
Ye, Z., Xu, Y., Tong, X., Zheng, S., Zhang, H., Xie, H., Stilla, U., 2019. Estimation and analysis of along-track attitude jitter of ZiYuan-3 satellite based on relative residuals of tri-band multispectral imagery. ISPRS J. Photogrammetry Remote Sens. 158, 188–200. https://doi.org/10.1016/j.isprsjprs.2019.10.012.
Yurovskaya, M., Kudryavtsev, V., Chapron, B., Collard, F., 2019. Ocean surface current retrieval from space: the Sentinel-2 multispectral capabilities. Rem. Sens. Environ. 234, 111468.
Zinke, R., Hollingsworth, J., Dolan, J., Van Dissen, R., 2019. Three-dimensional surface deformation in the 2016 Mw 7.8 Kaikoura, New Zealand, earthquake from optical image correlation: implications for strain localization and long-term evolution of the Pacific-Australian plate boundary. G-cubed 20 (3), 1609–1628. https://doi.org/10.1029/2018GC007951.