

COPERNICUS BIG DATA AND GOOGLE EARTH ENGINE FOR GLACIER SURFACE VELOCITY FIELD MONITORING: FEASIBILITY DEMONSTRATION ON SAN RAFAEL AND SAN QUINTIN GLACIERS

M.Di Tullio, F. Nocchi, A.Camplani, N. Emanuelli, A. Nascetti, M. Crespi

Geodesy and Geomatics Division, University of Rome La Sapienza, Italy (andrea.nascetti, mattia.crespi)@uniroma1.it

WG III/9

KEY WORDS: Big Data, Google Earth Engine, Sentinel-1, Glacier velocity field, SAR off-set tracking, San Rafael Glacier, San Quintin Glacier

ABSTRACT:

The glaciers are a natural global resource and one of the principal climate change indicator at global and local scale, being influenced by temperature and snow precipitation changes. Among the parameters used for glacier monitoring, the surface velocity is a key element, since it is connected to glaciers changes (mass balance, hydro balance, glaciers stability, landscape erosion). The leading idea of this work is to continuously retrieve glaciers surface velocity using free ESA Sentinel-1 SAR imagery and exploiting the potentialities of the Google Earth Engine (GEE) platform. GEE has been recently released by Google as a platform for petabyte-scale scientific analysis and visualization of geospatial datasets. The algorithm of SAR off-set tracking developed at the Geodesy and Geomatics Division of the University of Rome La Sapienza has been integrated in a cloud based platform that automatically processes large stacks of Sentinel-1 data to retrieve glacier surface velocity field time series. We processed about 600 Sentinel-1 image pairs to obtain a continuous time series of velocity field measurements over 3 years from January 2015 to January 2018 for two wide glaciers located in the Northern Patagonian Ice Field (NPIF), the San Rafael and the San Quintin glaciers. Several results related to these relevant glaciers also validated with respect already available and renown software (i.e. ESA SNAP, CIAS) and with respect optical sensor measurements (i.e. LANDSAT8), highlight the potential of the Big Data analysis to automatically monitor glacier surface velocity fields at global scale, exploiting the synergy between GEE and Sentinel-1 imagery.

1. INTRODUCTION

In the last century, because of global warming the glaciers masses with few exception have decreased at unprecedented rates. This phenomenon may be a risk for more than a hundred million people worldwide living within a meter of mean sea level. Furthermore, when salt water intrudes into freshwater aquifers, it threatens sources of drinking water and makes raising crops problematic. Today the goal is to understand the glacier-climate interactions to best predict for local, regional, and global implications of glaciers changes and mountain ice sheets.

In particular, the glacier surface velocity represents one of most critical parameters for the glacier monitoring, being connected to glaciers changes (mass balance, hydro balance, glaciers stability, landscape erosion) (Nascetti et al., 2016). Remote sensing is highly suited for monitoring in inaccessible areas such as high mountains and cold regions, where mass movement processes such as glacier flow, rockglacier creep and rockslides (Debellagilo and Käab, 2012).

The use of satellite synthetic aperture radar (SAR) imagery has been shown to be a feasible method to measure glacier surface velocities, with significant advantages over other traditional methods such as ground measurements, by which only point information at low observation rate. (Jiang et al., 2012). In general the estimation of glacier velocity fields is possible with repeat optical satellite imagery and SAR data using speckle and/or feature tracking methods. They are usually called image matching in the optical domain and offset-tracking in the microwave domain (Paul et al., 2015). For these methods the main constraints are the revisit time of the images, since the displacements must be bigger than accuracy of the method, and variations of the surface, which

has to be limited in order to not impair the image matching. In the case of SAR imagery, high accuracy is possible up to the potential of differential SAR interferometry to map displacements at centimeter resolution, resulting in spectacular new results in glaciology (Strozzi et al., 2002). However interferometric techniques shows some drawbacks, they need short baselines (days) and are less suitable for operational and global spatial/temporal scale application.

We have to consider that the last decade has seen the increasing development of glacier remote sensing monitoring techniques supported by the latest higher spatial temporal resolution SAR sensor carried by the last generations of EO satellites. This ever-increasingly volumes of data acquired by satellites, as SAR ESA Sentinel-1, supply a solid basis for more accurate long-time analysis of glaciers dynamic.

2. SENTINEL-1 AND GOOGLE EARTH ENGINE

With European Copernicus program of Earth Observation, European Space Agency has made possible to exploit a huge amount of SAR and optical data thanks to the launch of the new Sentinel constellation since 2014. The main issue is to handle and exploit the huge amount of these EO data, since it is not technically feasible or financially affordable to transfer the data and use traditional local processing methods (e.g. image file download). To address these challenge, new generation of IT processing platforms are becoming available. Unlike other high-performance computing systems, the GEE platform, makes it easy to access high-performance computing resources for processing very large geospatial datasets (Gorelick et al., 2017). Considering the high potential of EO Big Data, since 2015

Google included the whole archive of Copernicus Sentinel-1 and Sentinel-2 data in GEE, which consists of a multi-petabyte analysis-ready data catalog co-located with a high-performance, intrinsically parallel computation service (Gorelick et al., 2017). The GEE archives provides Sentinel-1 Ground Range Detected (GRD) scenes ready-to-use, preprocessed by Sentinel-1 Toolbox to generate a calibrated, ortho-corrected product and weekly updated. The short revisit time offered by the Sentinel-1 constellation (up to 6 days), good spatial resolution (10 m) and the high-performance computing resources of GEE, makes it possible to determine glacier surface velocity field on large area and with suitable spatial and time resolution.

The aim of this work is to test the methodology implemented into the cloud based infrastructure to obtain glacier velocity field estimation, with automated processing of large stack of Sentinel-1 GRD imagery, highlighting the potentials and weaknesses of this approach. In order to develop this infrastructure it was chosen the virtual environmental provided by Google Cloud Platform (GCP), because highly connected to GEE database, and capable of providing greater versatility.

3. METHOD

The algorithm of SAR Intensity Off-set Tracking was designed and developed at the Geodesy and Geomatic Division of DICEA, University of Rome "La Sapienza".

The implemented software, entirely written in Python, uses the GEE APIs to interface and access to the huge GEE database of satellite imagery, and it has been developed within the Google Cloud Platform (GCP) and Google Cloud Storage (GCS) in order to minimize costs and data transfer times. It is based on GDAL, an open source library to handle raster and vector geospatial data and OpenCV (Bradski and Kaehler, 2008), an open source computer vision library.

The algorithm consists of the following steps:

- identification of the area covered by each glacier through the GLIMS Glacier Database (Raup et al., 2007) and Randolph Glacier Inventory (RGI) (Pfeffer et al., 2014)
- selection on GEE of the Sentinel-1 imagery for the period of interest covering the glacier, imagery transfer on GCS and coregistration refinement at sub-pixel accuracy
- estimation of the glacier surface velocity with the Intensity Offset-Tracking method and global Least Squares Estimation (LSE) to exploit the benefits of data redundancy (outliers detection, surface velocity field regularization)

3.1 Glacier area identification and sampling

The GLIMS Glacier Database has been designed to store geospatial information about glaciers (Raup et al., 2007). This database provides the glacier outlines in standard shape-file format in WGS84 reference frame. Starting from the glacier outlines, the area of the glacier is covered with a regular grid (200 x 200 meters posting) obtaining a set of 2D sampled points.

3.2 Images stack acquisition and coregistration refinement

GEE APIs permit user to make requests to Earth Engine server and to transfer very quickly data into GCP bucket. We used images stacks with a temporal resolution of 12 to 48 days in order to

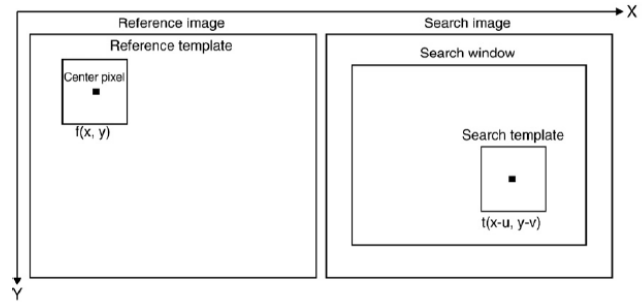


Figure 1. Normalized Cross Correlation

guarantee enough speckle correlation between images and to detect the glacier motion considering the Sentinel-1 images ground resolution of 10 m. This step is quite critical for the huge dataset size, depending upon the GEE API for geographical filtering used to manage the GEE database. We performed an images selection based on the intersection between the upper left and lower right corners of RGI shape-files and all Sentinel-1 images available for the period of interest on GEE database.

Moreover several tests have shown that a significant residual misalignment at sub-pixel level can still remain between images already coregistrated by ESA software SNAP ToolBox. This residual misalignment involves errors in the displacements estimation, specially if these displacements and the temporal baseline are quite small. The situation can be reduced to a typical image alignment problem, where there are different algorithms for the motion parameters estimation. For this purpose we adopted the recently proposed Enhanced Correlation Coefficient (ECC) (Evangelidis and Psarakis, 2008) introduced in OpenCV 3 as similarity measure. This algorithm involves a reduced computational complexity compared to other ones, allowing to perform more accurate alignments with a faster convergence.

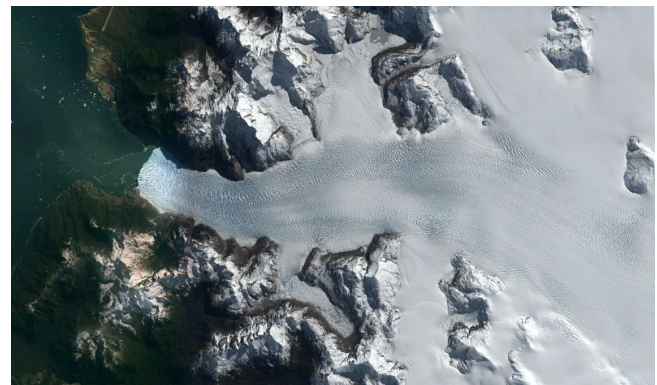


Figure 2. San Rafael Glacier, Northern Patagonian Ice Field (NPIF)

3.3 Surface velocity fields computation and global Least Square Estimation

The glacier displacements between two coregistrated images were estimated using the Intensity Offset-Tracking method, which rely on the image matching, a technique for finding areas of an image that match to a template (patch). In the study on glacier velocities, Normalized Cross Correlation (NCC) is a matching method that is often used, due to its simplicity (Heid and Käab, 2012). Each of sampled points on the regular grid has become the patch centre on which template matching has been implemented. The NCC is calculated by convolving a

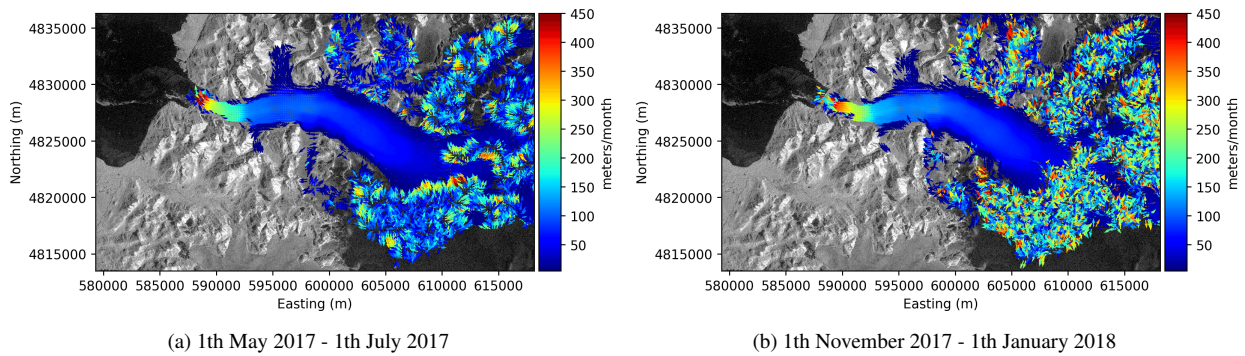


Figure 3. San Rafael Glacier surface velocity field: seasonal variations

portion of the reference image (the reference template) with the corresponding larger subset (the search window) in the search image (Ravanelli et al., 2017). The size of search window depends on the maximum expected displacement; moreover, in order to improve the matching, it was determined the optimal size of 95 pixel for the reference template.

Downstream phases of the matching process, we achieved a redundant set of velocity measures for the same glacier point, so that the problem was solved using the least squares method. In particular, it was implemented a robust linear model, in order to mitigate the effect of outliers, obtaining a more robust velocities estimates.

4. RESULTS

In order to assess the feasibility of a possible global application of the methodology on long time intervals, an important glacial region has been investigated, the Northern Patagonian Ice Field (NPIF). Here the results are discussed of surface velocity analysis on a time interval of 3 years over two huge glaciers: San Quintin Glacier and San Rafael Glacier.

The NPIF is a trace of the Patagonian Ice Sheet, an extensive ice sheet that covered all of Chilean Patagonia and the westernmost parts of Argentine Patagonia during the Quaternary glaciations. Today, with its Glaciers largely in retreat and covering an area of 4.200 km², it is still the second largest continuous mass of ice outside of the polar regions; this ice field encompasses glaciers characterised by very different behaviours. Some details of the two investigated glaciers are hereafter reported:

- **San Rafael Glacier:** it is a unique tidewater glacier of the NPIF (Figure 2), located at the lowest latitude in the world, with an area of 760 km², it calves into the Laguna San Rafael and is contained within Laguna San Rafael National Park.
- **San Quintin Glacier:** is the largest glacier of the NPIF, with an area of 790 km² in 2001. The glacier extends 50 km from the ice divide in the center of the ice cap. It is comparable in size to San Rafael Glacier but terminates in a pro-glacial lake. According to previous investigations, this glacier reaches its maximum surface velocity about 29 km from the ice front, in the narrowest section of the terminal valley at equilibrium line altitude (ELA), then the surface velocity reduces below 350 m/yr in the terminus region (Pelto, 2016). Its terminus has been undergoing a steady retreat with decay of its piedmont lobe.

4.1 San Rafael Glacier

A phenomenon that can clearly evidenced is the seasonal variation of velocities. About that, for instance it is clear the difference between the winter (Figure 3a) and the summer (Figure 3b) seasons, in the latter there is an evident increase of velocity, since during the summer the temperatures are higher.

In general Figure 4 and Figure 3 show velocities of about 150 m/month in the upper zone and higher velocities of 400 m/month on the snout, according with the existent measurements by NASA's Spaceborne Imaging Radar-C (SIR-C) in 1994 and with recent measurements (Willis et al., 2012) using 7 day repeat Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data from year 2007.

It is difficult to measure the velocity of its terminus with other SAR sensors because their time repeat is too long to maintain signal coherence (e.g., ice moves 1km in 46 days).

We can see that matching algorithms do not directly provide perfect results, in this case several false matches can be present especially in the glacier accumulation area, where there is not a well defined movement. In these case Off-tracking method is not able to give robust results also because the glacier surface is uniformly covered by snow during all the season causing a too uniform speckle texture. Therefore errors and outliers cannot be avoided because of non-perfect image and ground conditions (Paul et al., 2015). These inconsistent results obtained from matching step, causes a significant slowdown in the remaining least squares processing phase taking account of the huge number of measurement points (16700 samples) considered for the San Rafael Glacier analysis.

4.2 San Quintin Glacier

The area of interest is very large, so that the 200 x 200 meters posting grid involves more than 20000 points. As a consequence of the huge extension, in the whole dataset downloaded by GEE, there are not many Sentinel-1 images that cover the entire area, which, on its own, causes a low data redundancy the effect of the low data redundancy is shown in the Figure 5, where it is evident that the surface velocity field has not been correctly reconstructed in some areas; moreover, there is clearly a higher variability in the velocity vectors. This will require in the future to improve the images selection by a mosaic procedure of all Sentinel-1 images potentially available on GEE database; and to test different GEE filter methods, in order to obtain more robust velocities estimation.

Nevertheless the results obtained for certain periods are in perfect accordance with other researchs using aerial photographs and aerial survey photographs as (ANIYA, 2017, Pelto, 2016). In fact

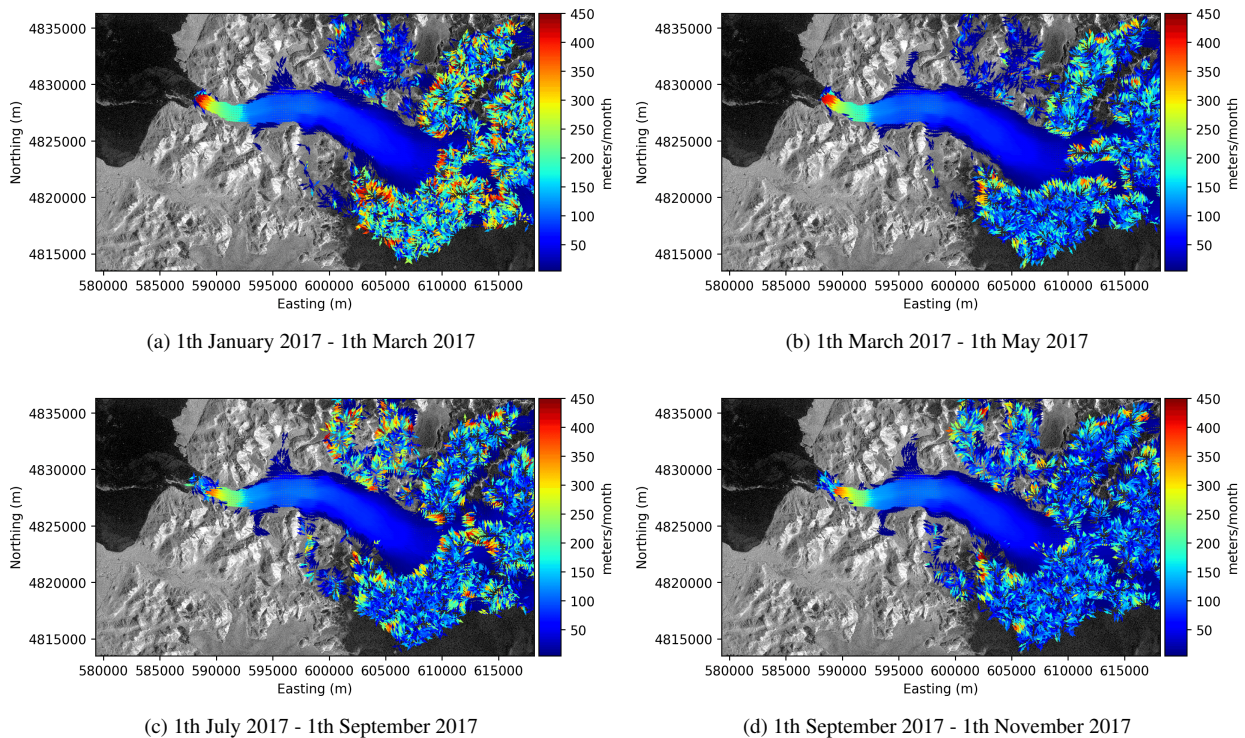


Figure 4. Temporal series of San Rafael Glacier surface velocity field

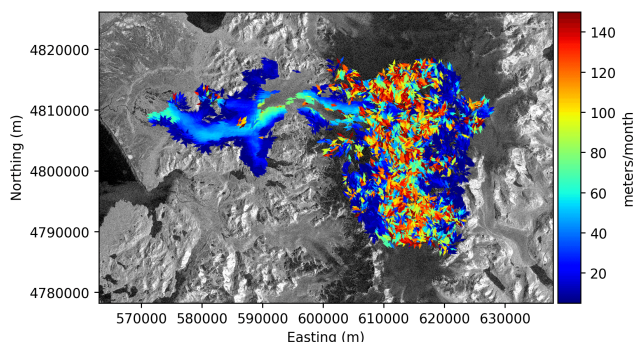


Figure 5. San Quintin Glacier surface velocity field

the **Figure6** shows that the velocities are high in the upper area and along the glacier central axis, while they were lower towards glacier terminus, moreover it is clearly visible that the glacier margins are essentially stationary compared with the center due to the ice friction against the surrounding rocks. The highest velocities reached about 95 m/month (1.14 km/yr), while close to the terminus they were 25-30 m/month (360 m/yr). We have to consider that this glacier is very rapid, in the case of too long temporal spam (revisiting time of consecutive images 48 days up 72) coherence can be insufficient to estimate reliable displacements. It should be noted also the different behaviour between San Quintin and San Rafael, in the latter case velocities increases from top to bottom. This behaviour is a consequence of the tide-water nature of San Rafael glacier; in fact, sea water has a very strong influence on glacier movements (San Quintin Glacier terminates in a pro-glacial lake).

5. CONCLUSIONS AND PROSPECTS

In this work it is described an innovative methodology, to routinely and continuously estimate glacier surface velocity fields

exploiting large stack of Sentinel-1 images in GRD format, and the first results are reported. The methodology was developed under an infrastructure based on Google Earth Engine and within Google Cloud Platform, which allow to leverage high speed and intensive processing capacity provided by Google. This allows us to download, manage and analyse freely and quickly enough a huge imagery datasets for long periods (here 3 years from 2015 to 2018 are considered).

The whole process is subdivided in three different steps: glacier area identification, on the basis of a standard global database (RGI); imagery selection for the period of interest from Sentinel-1 repository within GEE; surface velocity estimation on a regular grid (here with 200 x 200 meters posting) for the period of interest (here from 2015 to 2018) on a regular interval basis (here each 2 months). Such a continuous in space and time surface velocity field estimation allows to highlight both the velocity gradient across the glacier, and the seasonal variations of the velocities.

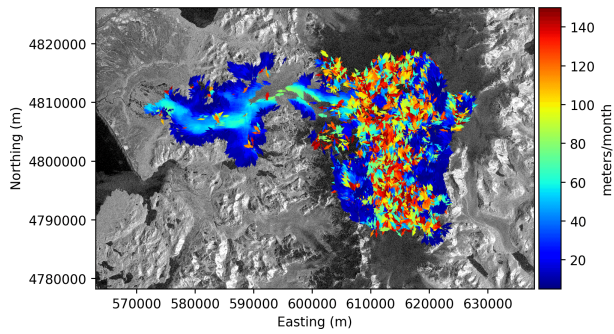
Overall the analysis of the results demonstrates that the designed algorithm and implemented software allow to obtain a robust and continuous estimation of glacier surface velocity fields guaranteed by the available huge Sentinel-1 dataset. Moreover, thanks to the high potentialities of GEE and GCP, this first application demonstrates the feasibility of a possible wider application, up to global scale.

On the other side, this first test points out some problems, which are now under consideration and have to be solved in the next future. The first one is the need of a filtering procedure downstream of the matching step, in order to remove outliers, that is non-physically based surface velocity estimates. This would allow to reduce time processing and to improve the velocity estimation through LSE. A second important question is the possible low redundancy of images in case of wide glaciers, as experienced during the San Quintin analysis; this will require to develop a more refined approach for images selection. A third relevant issue is the masking of foreshortening and layover areas, basing on a Digital Elevation Model and incidence angle meta-

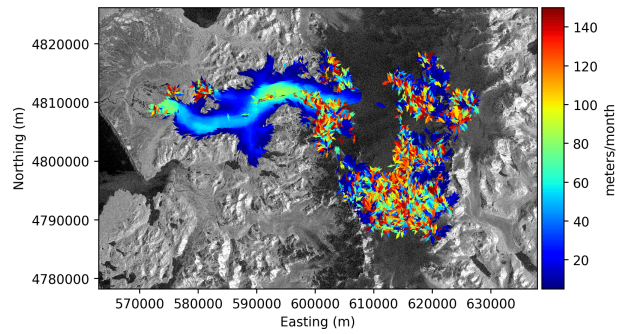
data and the parallelization of the different processes, in order to significantly decrease the computational time. Finally, other applications over much more different glaciers are needed to clearly assess the worldwide potential of the implemented procedure.

REFERENCES

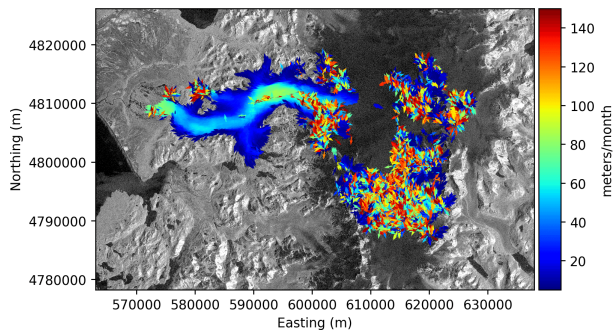
- ANIYA, M., 2017. Glacier variations of hielo patagónico norte, Chile, over 70 years from 1945 to 2015. *Bulletin of Glaciological Research* 35, pp. 19–38.
- Bradski, G. and Kaehler, A., 2008. *Learning OpenCV: Computer vision with the OpenCV library*. ” O’Reilly Media, Inc.”.
- Debella-Gilo, M. and Käab, A., 2012. Measurement of surface displacement and deformation of mass movements using least squares matching of repeat high resolution satellite and aerial images. *Remote sensing* 4(1), pp. 43–67.
- Evangelidis, G. D. and Psarakis, E. Z., 2008. Parametric image alignment using enhanced correlation coefficient maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(10), pp. 1858–1865.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R., 2017. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202, pp. 18 – 27. Big Remotely Sensed Data: tools, applications and experiences.
- Heid, T. and Käab, A., 2012. Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery. *Remote Sensing of Environment* 118, pp. 339–355.
- Jiang, Z.-l., Liu, S.-y., Peters, J., Lin, J., Long, S.-c., Han, Y.-s. and Wang, X., 2012. Analyzing yengisogat glacier surface velocities with alos palsar data feature tracking, karakoram, china. *Environmental Earth Sciences* 67(4), pp. 1033–1043.
- Nascetti, A., Nocchi, F., Camplani, A., Di Rico, C. and Crespi, M., 2016. Exploiting sentinel-1 amplitude data for glacier surface velocity field measurements: Feasibility demonstration on baltoro glacier. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 41, pp. 783.
- Paul, F., Bolch, T., Käab, 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. *Remote Sensing of Environment* 162, pp. 408–426.
- Pelto, M., 2016. *Recent Climate Change Impacts on Mountain Glaciers*. John Wiley & Sons.
- Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C. et al., 2014. The randolph glacier inventory: a globally complete inventory of glaciers. *Journal of Glaciology* 60(221), pp. 537–552.
- Raup, B., Racoviteanu, A., Khalsa, S. J. S., Helm, C., Armstrong, R. and Arnaud, Y., 2007. The glims geospatial glacier database: a new tool for studying glacier change. *Global and Planetary Change* 56(1-2), pp. 101–110.
- Ravanelli, R., Nascetti, A., Di Rita, M., Belloni, V., Mattei, D., Nisticò, N. and Crespi, M., 2017. A new digital image correlation software for displacements field measurement in structural applications. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 42, pp. 139.
- Strozzi, T., Luckman, A., Murray, T., Wegmuller, U. and Werner, C. L., 2002. Glacier motion estimation using sar offset-tracking procedures. *IEEE Transactions on Geoscience and Remote Sensing* 40(11), pp. 2384–2391.
- Willis, M. J., Melkonian, A. K., Pritchard, M. E. and Ramage, J. M., 2012. Ice loss rates at the northern patagonian icefield derived using a decade of satellite remote sensing. *Remote Sensing of Environment* 117, pp. 184–198.



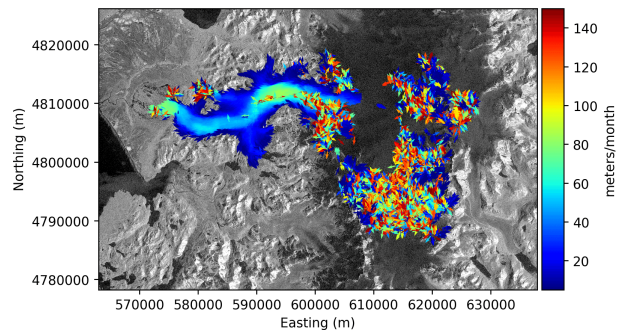
(a) 1st September 2015 - 1st November 2015



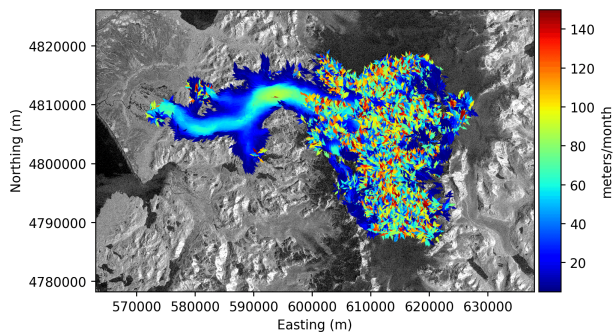
(b) 1st January 2016 - 1st March 2016



(c) 1st March 2016 - 1st May 2016



(d) 1st May 2016 - 1st July 2016



(e) 1st July 2016 - 1st September 2016

Figure 6. Temporal series of San Quintin Glacier surface velocity field