Time-lapse optical flow regularization for geophysical complex phenomena monitoring

Héla Hadhri, Flavien Vernier, Abdourrahmane Atto, Emmanuel Trouvé

Time-lapse optical flow regularization for geophysical complex phenomena monitoring

Hadhri Hela[1], Vernier Flavien[1], Atto Abdourrahmane M.[1], Trouvé Emmanuel[1]

Univ. Savoie Mont Blanc, LISTIC, F-74000 Annecy, France

Abstract

In this paper, we introduce a framework for the tracking of geophysical complex phenomena via time-lapse images. It includes the regularization of the derived surface motion maps time series. The proposed processing chain addresses five main challenges: undesired camera movement, missing frames, important photometric changes, weak/repetitive texture and model-free dense spatial transformations. In the proposed framework the motion maps time series are obtained via robust pre-processing steps and optical flow computing. The contribution consists of regularizing the resulting velocity and position time series to minimize a *temporal closure error* in a subsequent stage. This step serves to alleviate the limitations of existing methods in the context of geophysical monitoring. The temporal closure errors are formulated as linear mappings to inverse using signal priors and two formulations are defined along with illustrative cases. Related methods are discussed and extensive experimentation on simulated datasets is carried out to validate the approach and compare between the different proposed formulations and resolution schemes. Experimental results are presented on time series acquired by ground-based cameras used for the monitoring of Alpine glaciers. The algorithm is computationally efficient, even considering the quantity of processed and generated data, and is run in parallel on multiple cores for speed-up.

Keywords: Computer Vision, Time series, Temporal regularization, Remote/proximal sensing, Natural Outdoor Environment, Glaciers monitoring, Tracking

1. Introduction

Motion tracking and scene understanding from time-lapse images acquired by automatic ground-based cameras is a task that has been studied in many applications, including medical imaging tracking systems, video surveillance or geophysical monitoring. With a controlled environment scenario, this motion is characterized by the flow of pixels with constant intensity. This assumption of intensity conservation is typically unrealistic for outdoor geophysical monitoring data because illumination difference between images is relatively important, and more photometric invariance must be introduced. There are two global families of computation techniques to achieve this: (1) Differential techniques that compute velocity from spatio-temporal derivatives of pixel intensities while incorporating other constancy assumptions and regularity constraints [1][2]; (2) Region-based offset-tracking techniques, which compute the trajectory of

Email addresses: **hela.hadhri@univ-smb.fr** (Hadhri Hela), **flavien.vernier@univ-smb.fr** (Vernier Flavien), **abdourrahmane.atto@univ-smb.fr** (Atto Abdourrahmane M.), **emmanuel.trouve@univ-smb.fr** (Trouvé Emmanuel)

pixels via the identification of local displacements that provide optimal similarity of two image patches and which usually create robustness for intensity variation via normalization. In this paper, optical flow (OF) methods refer to the first type of methods while the second type is referred to as normalized cross-correlation (but other similarity measures exist) [3][4]. Moreover, geophysical monitoring data exhibit other particularities, such as geometric non-rigid deformations, repetitive patterns, low texture, occlusion, frame mis-registration and weather artifacts (such as snow, fog, water drops/sun beam reflection on the screen, clouds, shadows, etc.) [3]. The main challenge in these applications is to *automatically* estimate accurate long term motion time series from pairwise image registration despite the data contamination. To overcome the limitations of existing techniques in such conditions, we propose a data processing methodology which includes several steps: (1) pre-processing steps: to discard some of the wrong measurements; (2) parameterized motion models: to handle frame mis-registration due to camera motion; and (3) robust optical flow: to estimate apparent motion of the object of interest over time. To introduce further robustness, we propose a regularization step that takes the generated incomplete pixel-wise Velocity and Position Time Series (VTS and PTS) and forces temporal closure, which is the main contribution of the paper. This regularization is inspired by the Small Baseline Subset (SBAS) technique for Synthetic Aperture Radar (SAR) interferometry [5]. In this domain, time series of complex SAR images are acquired to observe ground surface deformations [6]. Redundant displacement fields are obtained by computing phase differences from image pairs selected according to their temporal and spatial baselines and the observed phenomena: urban and land subsidence [7][8][9], water level monitoring [10], volcanic deformation [11], slope movements [12] and so on. These initial observations suffer from different error sources such as noise, atmospheric artifacts and phase unwrapping errors. The SBAS approach consists of deriving a final displacement time series with reduced uncertainty by combining these observations in a linear system inverted by different techniques [13][14]. An overview of the parameters of the SBAS algorithm can be found in [11]. Similarly to the SBAS workflow, the proposed regularization is formulated as an inverse problem where a temporal closure error within an appropriate range is minimized to reduce uncertainty via redundancy.

2. Related methods

2.1. Data regularization

Major time series regularization techniques are linked and include modelling, query by content, anomaly correction, motif discovery, prediction, clustering, classification, dimensionality reduction, indexing and segmentation. A survey of these methods can be found in [15]. Hence, this regularization is often done by:

- rolling statistics [16]
- filter like methods such as a non-linear median filtering or extended Kalman filters [17][18]
- other stochastic approaches such as Multiple Imputation and Bootstrapping [19]
- model fitting methods; such as Self-Organizing Maps [20], parametric/non-parametric regression [21], AutoregRessive Moving Average (ARMA) models [22], etc.,
- supervised learning methods, such as RANdom SAmple Consensus (RANSAC), Support Vector Machine (SVM), etc.
- spectral methods such as principal component analysis/empirical orthogonal function decomposition, Discrete Wavelet Transform, etc. [23]

Geo-statisticians often use linear-weighted averaging methods, with data driven weighting functions, such as the interpolation/prediction kriging and collocation methods and deploy ground means, such as corner reflectors or GPS measurements to constrain the problem by fusing measurements from different sources. Recent works for the monitoring of geophysical phenomena via time-lapses include:

- adaptive Least Squares collocation to detect outliers in VTS and PTS during feature tracking [24],
- non-linear median filtering by averaging the values of Nearest Neighbors on surface change time series to correct random noise and fill missing observations [25].

Machine learning techniques have also been used to learn the expected behavior of time series in order to regularize observations [26][27][28][29] and it has been shown that fully Connected Networks are more suited for time series tasks as they allow for the input and output to have the same dimension and to process in a semantic way. A survey on Deep Learning methods for time series inquiry can be found in [30]. On the other hand, optical measurements are generally regularized via computer vision derived methods such as detection/recognition methods using interest points, features, HoGs (Histogram of oriented Gradients), etc. For example, the current best performing optical flow methods often use prior knowledge of the properties of the motion, like local smoothness assumptions [31], structure and motion adaptive assumptions [32] or the assumption that motion discontinuities are more likely at image edges [33] to ensure *spatial* coherence *over time*. In [34] the authors propose a method for optical flow adaptation to fluid monitoring using the continuity equation of 2D incompressible flows and passive scalar transport constraints. There are several methods to optimally balance data fidelity with regularization of the motion fields obtained by differential optical flow methods. Depending on the strength of the local image gradients, these methods yield a smooth trade-off between matching and interpolation, which can be very useful in the presence of noisy optical observations [2].

2.2. Inversion

Inversion consists of inferring the manifold of structural models that could yield some actual observations y through a given transformation A induced by a physical model. Theoretically, the goal of a reconstructive inversion is often to estimate a 2D or 3D field $x \in \mathbb{M}$: the Model space, from a set of observations or measurements $y \in \mathbb{D}$: the Data space, of the form $y = Ax + b$.
The measurement or model operator A can be linear or non-linear, full rank or rank-deficient and has a conditioning that can be quantified by means of the $L = 2$ condition number K defined as the ratio between the biggest and smallest singular values σ of A [35], such as:

$$K = \frac{\sigma_{\max}}{\sigma_{\min}},$$

which quantifies the sensitivity of the estimation to perturbations in the observations. b is the noise or measure uncertainty which hopefully will not be reconstructed by inversion. Inverting observations is now used extensively in many fields like remote sensing [36], deconvolution, optical tomography, 3D reconstruction [37][38] or geodesy [39]. A typical photogrammetric example would be recovering the descriptions of objects represented in image data x via optical measurements y corrupted by noise b by inverting the optical model A (back-projection). For this, a model misfit cost between y and Ax minimized. The used cost function is usually convex (unless a special penalizer is needed to create specific robustness) if the direct linear mapping is full-rank. Depending on the definition, sparseness and dimension of the direct problem, on the prior information on the estimated solution and on the choice of the cost function, direct

(spectral decomposition) or iterative (optimization via a solver) methods can be considered to stably recover x. In this paper, the physical model, that is encoded in A, and to be inverted, is the consistency of the displacement pattern over time: *the temporal closure constraint*. The inversion is thus done on the temporal axis of the computed OF maps and densely for every coordinate in order to obtain temporally regularized Optical Flow (OF) maps time series of the size of the images. The inversion is pixelwise and two formulations are proposed to encode the linear temporal closure model A.

2.3. Contribution

Despite the existence of discontinuity-preserving smoothness constraints to ensure spatial coherence of optical OF fields, one can notice the lack of long-term temporal constraints in recent models proposed for the Middlebury benchmark [40] due to computational requirements. Only a few methods make use of more than two frames and none of them incorporate temporal information in the underlying model. For off-line remote sensing of geophysical complex phenomena, we propose to revisit the Differential Synthetic Aperture Radar interferometry SBAS technique to minimize a temporal error from a Network of Optical Flow Fields (NOFF) forming the observations. The workflow to pre-process the raw data and compute relevant pairwise optical flow fields in object coordinates is presented followed by the paper's main contribution: the regularization of the pixelwise Velocity Time Series (VTS) and/or Position Time Series (PTS). To validate the whole pipeline under various conditions, 4 realistic datasets of simulated time lapses are constructed. To validate the temporal regularization step, different processing problems such as outliers, missing data and noise are further simulated on the observations and estimations are compared to ground truth solutions. Finally, results for the monitoring of alpine glaciers are presented.

3. Proposed pipeline

3.1. Overview

In a context such as stated before, the computed observations can be contaminated in different ways: bias can be introduced by the motion of the fixed camera; outliers by the photometric variance distorting the calculations in one of the layers of the automatic processing chain and/or the registration failing to capture the non-rigid spatial transformations; and gaps can be caused by targets being occluded and/or outliers due to weather artifacts being detected and removed. Figure 1 shows the layers of the proposed pipeline to process time-lapses in such a context: (A)Data is first acquired automatically via a programmable micro-controller; (B) a texture score is computed to dismiss images where the information is too corrupted and the remaining images are co-registered with regards to a master image to compensate for camera motion. (C) Optical flow (OF) fields $\vec{F} = (F_x, F_y)$ at every pixel are then computed between all the co-registered images within the chosen temporal closure range to form the Network of Optical Fields (NOFF) (D) The NOFF contain redundant information and makes it possible to minimize a temporal closure error to regularize the resulting VTS and PTS and interpolate values where observations are missing. The matrices in (C) and (D) are examples of observations in one pixel before and after regularization. The master image in figure 1 and all the illustrative examples of the paper being the first valid image of the time-lapse; I_0.

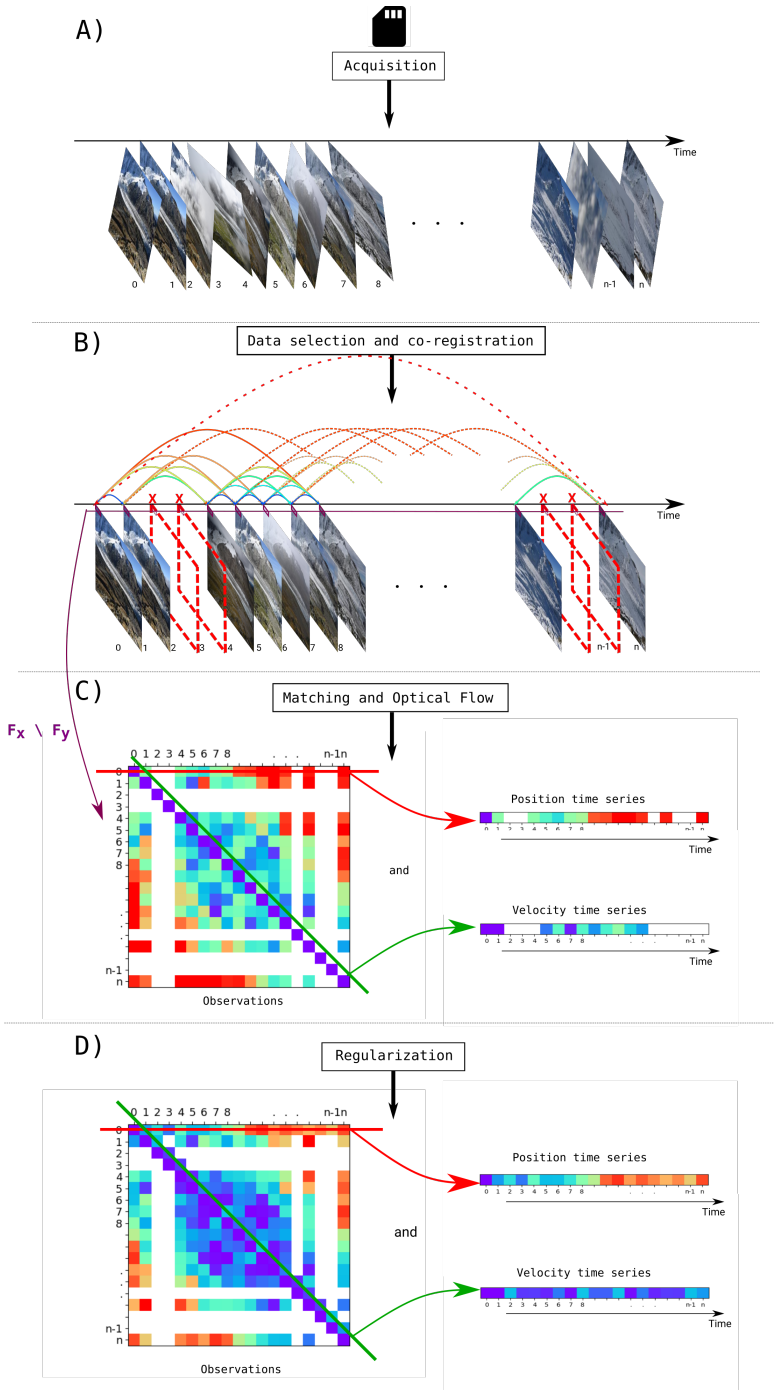


Figure 1: Proposed pipeline. (A) data acquisition. (B) data selection via texture detection. (C) matching and optical flow computation. (D) regularization of the obtained time series via the construction of a network of optical flow fields.

3.2. Data selection

After the acquisition of the time-lapse imagery, images are selected given their meta-data tags and according to the desired time step of the resulting time series. Textural information is quantified to dismiss corrupted images. This information is retrieved by the construction of a bank of multi-parameters Gabor filters designed manually to recognize the known textures (fringes, grains, tales, stripes, ice foliation, etc.) of the object of interest, if not occluded or highly deformed. This bank is constructed by varying some parameters (frequency, orientation and size) of the $2D$ Gabor function to create invariance to small differences between the images to be selected for further processing. A similar work can be found in [41]. For every image i, the sum of all the energies E of the obtained response maps is quantified. The energies of all the images in the time-lapse should theoretically follow a Gaussian distribution and the spurious images are detected by thresholding over median absolute deviation (MAD):

$$\tilde{E} = \text{median}(E_i), \tag{1}$$

$$\text{S}_{1i} = \text{median}(|E_i - \tilde{E}|) \tag{2}$$

S_1 is thus the texture score encoding the object of interest visibility in image I_i. This way, and as the known orientation of the gradients is also quantified by the designed filters, images that have a lack of textural information (foggy, occluded, damaged, cut off, etc.) are rejected and are no longer part of the pipeline. This classification gave a success rate of 96% for the rejection of data affected by adverse weather conditions and mis-acquisition.

3.3. Pairwise image processing

3.3.1. Co-registration

The stack of remaining images are co-registered by estimating a parametric geometric transformation regarding a chosen master image, the first relevant one for example. This step is done to compensate for the undesired geometric differences of the images in the time-lapse (see figure 1 (A)) and is thus done only on the static zone of the observed scene. This geometric difference is due to camera orientation and image plane instability due to extreme weather conditions. As the observed scene is distant, the images can be considered as the same planar surface in space, and a homography H can encode the translation, rotation, the scale and also the tilt (8 degrees of freedom) of the camera between two images I_i and I_j with 8 parameters $param$ (2 scaling, 2 rotation, 2 translation and 2 vanishing points parameters). The arguments $param$ defining the homography Φ are encoded in non-singular 3×3 matrix computed in terms of the Enhanced Correlation Coefficient criterion [42] on the static zones by iteratively minimizing the following normalized Euclidean distance:

$$E_{\text{ECC}_{(i,j)}} = \left\| \left(\frac{I_i(X)}{\|I_i(X)\|} - \frac{I_j(\Phi(X; param))}{\|I_j(\Phi(X; param))\|} \right) \right\|_2, \tag{3}$$

where $I_i(x)$ is the master image and $I_j(\Phi(X; param))$ is the image warped by inverse composition of pixels of coordinate X $(= (x_i, y_i, 1)^T$ (in homogeneous coordinates notation) with Φ:

$$\begin{pmatrix} x_j \\ y_j \\ 1 \end{pmatrix} = \Phi(X; param) = H^{-1}X = H^{-1} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \tag{4}$$

3.3.2. Matching and optical flow

Deep-Matching. For every pair of images in between which an Optical Flow (OF) field is computed, an initial matching is first quantified using a semi-dense matching algorithm called DeepMatching (DM) [43]. In DM, the correlation is quantified by means of Normalized Cross Correlation (NCC) of HOG descriptors of 4×4 pixel patches called atomic patches. Furthermore, this correlation is computed in a pyramidal scheme to, first, aggregate coarser correspondence by going up in the pyramid, and then, refine the correlation correspondences by optimizing the backtracking to the bottom of the pyramid and retrieve optimized matches of atomic patches. Besides the scale invariance insured by this multi-scale pyramidal scheme, the used operators for recursive construction of the correlation pyramid make further analysis weakly affected by non-rigidness and, more importantly, shows robustness against repetitive patterns. For speed-up, the parameters are tuned so the search area, the rotation and scale factors of a match are restricted to a range that can not be exceeded.

Optical flow computing. These dense matches are then used to initialize a further computation of a dense OF field $\vec{F}$ that minimizes a weighted energy functional based on two data terms E_{DB} and E_{DG}, from the brightness constancy assumption and the gradient constancy assumption respectively and a smoothness term E_{REG}. These assumptions respectively stipulate that the color and the edges between two successive frames are conserved, and the regularizer encourages the spatial smoothness of the flow field. For this, DeepFlow is used [44](DF). DF is similar to the well known Brox and Malik [45] algorithm but, in addition, takes a matching term E_M to further constrain the minimization, such as:

$$E_{\mathrm{OF}_{(\mathrm{i,j})}} = \int\int_{\Omega} (E_{\mathrm{DG}_{(i,j)}} + \alpha E_{\mathrm{DB}_{i,j}} + \sigma E_{\mathrm{reg}_{(i,j)}} + \beta E_{M_{(i,j)}}) \; dX, \tag{5}$$

Furthermore, DF have a weighting strategy that has also been shown to be robust to repetitive texture. A robust cost function $\psi(s^2) = \sqrt{S^2 + \epsilon^2}$ with $\epsilon = 0.001$ is used for optimization and for all terms. This makes the energy functional to minimize non-convex and non-linear. Let I_i and I_j be 2 consecutive frames defined on Ω with 3 RGB channels: $I_i, I_j : \Omega \to \mathbb{R}^3$. The estimated optical flow field is then $\vec{F} = (F_x, F_y) : \Omega \to \mathbb{R}^2$ and can be viewed as the change of structured light between the images, or in other terms, as the sub-pixelic displacement from an image I_i to an image I_j of the same scene at different time indexes t_i and t_j. This can be seen as the velocity of the observed scene projected onto a subspace orthogonal to the line-of-sight (image plane) if the set of dates are separated by unit intervals $\Delta t = t_{i+1} - t_i = 1$. OF fields are thus similar, in the physical meaning, to the phase observed in DinSAR images (which is projected onto the line of sight). The weights α, σ and β are tuned so to favor the gradient constancy assumption over the brightness constancy assumption. This makes the optical flow computation sturdier to photometric differences.

3.3.3. Scores and errors

To assess the pairwise OF maps computation steps, the following scores and error measures are computed:

- The mean texture score (from section 3.2): S_1.
- A correlation score per couple of images equal to the sum of all patch similarities along their DM back-tracking path: S_2.
- The inverse of the mean residual velocity amplitude $\frac{1}{\sum_{SZ}|F_{(i,j)}|}$ observed on the static zones: R_1.

3.3.4. Scaling and geo-referencing

For the result to be more reliable and to avoid parallax confusion on the computed Optical Flow maps $\vec{F}$, scaling, from sensor coordinates into object space metric coordinates, and geo-referencing is needed. For this purpose, and considering the pinhole camera geometric and optical equations, a ground sample distance map Δmm (in cm/pixel) can be computed from intrinsic camera parameters and the depth map Z with respect to the image against which the time-lapse has been co-registered:

$$\Delta mm = \frac{\mathbf{x_{mm}}}{f} Z, \tag{6}$$

where x_{mm} is the physical metric distance between pixels inside the hardware, or pixel pitch, and f is the lens focal length. This depth information Z was obtained via a shape from motion framework, whose seminal work can be found in [46], of stereo image pairs. This $3D$ photogrammetric workflow is stabilized by several in-situ control points that were measured on the static zones seen by the cameras. These in-situ measurements allowed us to constrain the search for the orientation of the cameras (extrinsic), to perform auto-calibration (intrinsic) and to geo-reference the obtained 3D model of the object of interest during the bundle-adjustment step. The depth maps from the camera centers are then retrieved by ray intersection from camera centers to the $3D$ points. This way, the components of $\vec{F}$ could be scaled to quasi-3D with known GPS location. Unfortunately, there are two main disadvantages to this technique. First, the flow of the scene that is projected onto the line of sight is neglected and, hence, the main direction of the sought after flow should be parallel to the image plane for an accurate augmentation to $3D$. Secondly, classic calibration algorithms (such as with calibration chessboards / corners) can introduce extra uncertainty in the case of geophysical monitoring due to ground constraints, important distances and the lack of geometrically stable features. Therefore, the in-situ measurements must be well scattered over the entire scene and measured with a high accuracy and precision for both the GPS and the corresponding pixel coordinates.

3.4. Regularization

3.4.1. Temporal closure errors

The cost functions for regularizing the obtained velocity and position time series (respectively VTS and PTS) are the norm of the temporal closure errors denoted b_{LF} and b_{CM} for the Leap-Frog and Common-Master formulations respectively. These temporal closure errors can be expressed as follows: Let N be the number of images and a the master image for the regularization step. Its first expression, the Leap-Frog formulation **LF**, is given by the forward and backward forms such that:

$$\begin{gathered}\forall\, a \in \{1..N\}, \forall \text{ ranges } n \in \{1..N-a\} \text{ and } \forall \mu \in \{1..a-1\}: \\ -F_{a,a+n} + \sum_{i=a}^{a+n-1} F_{i,i+1} = b_{\mathrm{LF}}, \\ F_{a,a-\mu} + \sum_{i=a-1}^{1} -F_{i,i+1} = b_{\mathrm{LF}}.\end{gathered} \tag{7}$$

The second expression is the Common-Master formulation **CM** and is given by the form such that:

$$\begin{gathered}\forall\, a \in \{1..N\}, \forall i \in \{a..N\}, \forall n \in \{1..N-i\}: \\ F_{i,i+n} + F_{a,i} - F_{a,i+n} = b_{\mathrm{CM}}.\end{gathered} \tag{8}$$

$F_{(i,j)}$ denotes either component F_x or F_y of the OF field between two images I_i and I_j with date indexes i and j and at a given pixel $\mathbf{p} = (x, y)^T$. These closed loop error measures should be equal to zero if the flow computation of the involved fields is flawless. In practice, temporal closure can be violated not only by the uncertainty introduced over the different layers of the pipeline but an abrupt event can also take place in the observed area inducing the flow computation to fail capturing the non-smooth large motion.

3.4.2. Temporal closure mappings

Given a master image a, to consider errors in Equations (7) and (8) for all temporal loop sizes n, we express them as the following pixelwise system of equations:

$$\begin{aligned} A_{\mathrm{LF}} \times F_{\mathrm{est}_{\mathrm{LF}}} - F_{\mathrm{obs}} &= b_{\mathrm{LF}}, \\ A_{\mathrm{CM}} \times F_{\mathrm{est}_{\mathrm{CM}}} - F_{\mathrm{obs}} &= b_{\mathrm{CM}}. \end{aligned} \tag{9}$$

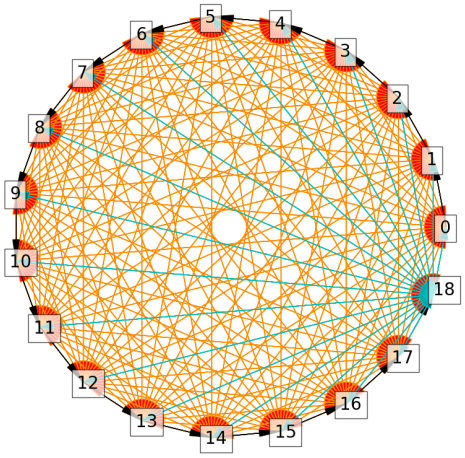
The processed observation vector $\vec{F}_{\mathrm{obs}}$ belongs to the Data space $\mathbb{D} = \mathbb{R}^{(N-m)(N-m-1)}$ and is the same for the CM and LF formulations. It consists of observed OF fields between all possible date index arrangements after the rejection of m unreliable images, or the NOFF values at a given pixel (all bi-directed orange arrows in the graphs in figure 2). The sought after estimates $\vec{F}_{\mathrm{est}_{\mathrm{LF}}}$ and $\vec{F}_{\mathrm{est}_{\mathrm{CM}}}$ should minimize a misfit error to the model spaces $\mathbb{M}_{\mathrm{LF}}, \mathbb{M}_{\mathrm{CM}} \subset \mathbb{R}^{N-1}$ respectively (black simple directed contour arrows of the graphs in figure 2). The model space $\mathbb{M}_{\mathrm{LF}}$ is the set of temporally coherent time series: $\vec{F}_{\mathrm{est}_{\mathrm{LF}}} = \{F_{i,i+1}, i \in (0, 1, ..., N-1)\}$ and $\mathbb{M}_{\mathrm{CM}}$ is the set of temporally coherent time series: $\vec{F}_{\mathrm{est}_{\mathrm{CM}}} = \{F_{a,i}, i \in (a+1, ..., a+N)\}$ relative to a Common Master image I_a. The output of inverting the Common Master (CM) formulation is thus, theoretically, the Riemann sum over the output estimated via the Leap-Frog (LF) inversion starting from the master index a such as:

$$\vec{F}_{\mathrm{est}_{\mathrm{CM}}}(t_k) = \sum_{i=a}^{k-1} \vec{F}_{\mathrm{est}_{\mathrm{LF}}}(t_i). \tag{10}$$

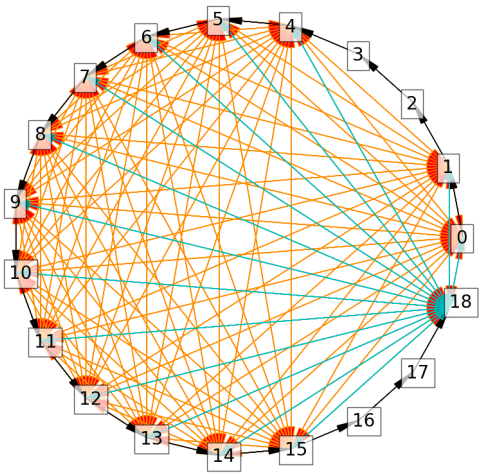
Both temporal closure error formulations in Equations (7) and (8) can thus be deducted from one another and whereas $\vec{F}_{\mathrm{est}_{\mathrm{LF}}}$ can be interpreted as the motion velocity between adjacent date indexes (at a given pixel), $\vec{F}_{\mathrm{est}_{\mathrm{CM}}}$ can be seen as the motion position relative to the Common Master date a. The output of the LF inversion is thus a VTS and the output of inverting the CM formulation is a PTS.

Illustration case. To illustrate the temporal closure mappings to inverse, let us consider a time-lapse of 18 images. Figure 2a shows the Network of Optical Flow Fields (NOFF) index graph, i.e. the graph of indexes in between which optical flow fields are computed (in orange) and the index of the estimations (in black). Figure 3 show the induced linear mappings and their corresponding Singular Values, respectively for the LF and CM formulations. We can see that both mappings are full-rank and have a low condition number (12.06 and 4.35 respectively for the LF and CM mappings) which suggests that the solution of the inversion process won't be too affected by noise in the observations.

Reducing the maximum range. Considering that the registration error increases with the temporal baseline and the important time span for geophysical monitoring, reducing the maximum range of the temporal closure loops allows to reduce mis-registration related errors. Figure 2b shows the obtained index graph and figure 4 the corresponding linear mappings. We can see that

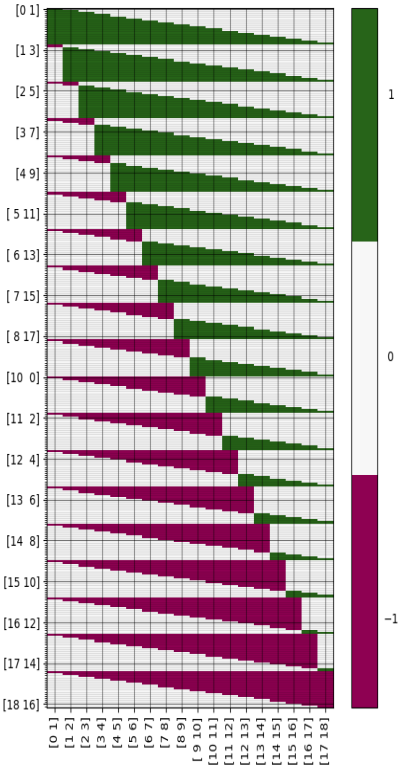


(a) Index graph for a full range problem: 18 vertices (size of F_{est}) and 342 directed edges (size of F_{obs}). In blue all observations including image I_{18}.

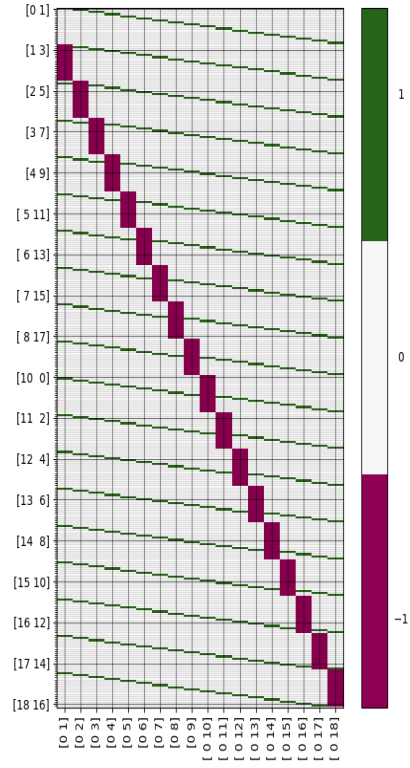


(b) Index graph for a problem with a reduced range: 18 vertices and 160 directed edges.

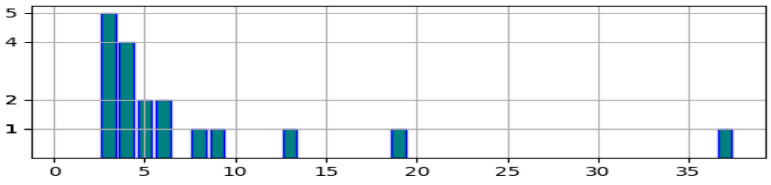
Figure 2: Polygonal index graph for a full-rank problem.



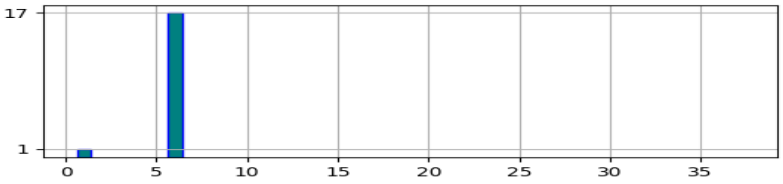
(a) matrix associated with the model in Equation (7): A_{LF} , size $= 342 \times 18$



(b) matrix associated with the model in Equation (8): A_{CM} , size $= 342 \times 18$



(c) Singular values (rounded) of the Leap-Frog linear mapping in figure 3a.

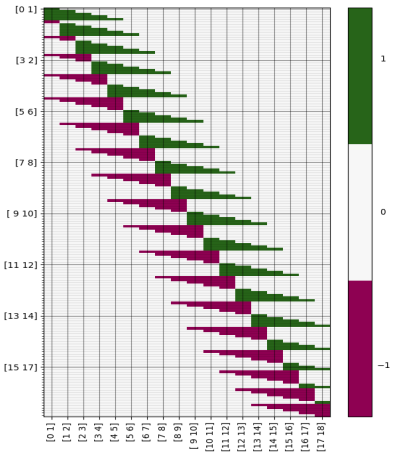


(d) Singular values (rounded) of the Common Master linear mapping in figure 4b.

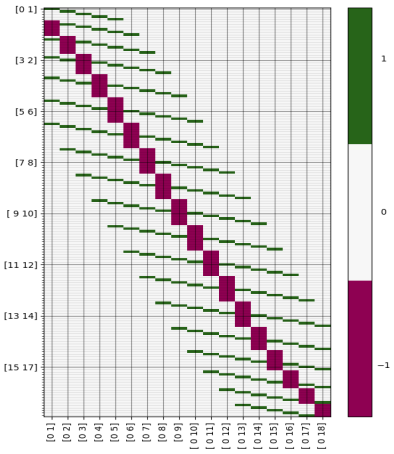
Figure 3: 3a: Respectively, mappings from the data space $\mathbb{D}$ to the model spaces $\mathbb{M}_{\mathrm{LF}}$ (left) and $\mathbb{M}_{\mathrm{LF}}$ (right) and thair histogram of singular values in figure 3c and 3d .The associated condition numbers are respectively 12.06 and 4.35. The rows and columns are arranged as in equations (11) and (12) but only 1 on 20 homologous indexes are displayed on the vertical axis.

this operation does not significantly affect the condition numbers of the mappings to inverse (5 and $7,79$ respectively for the Leap-Frog and Common Master mappings).

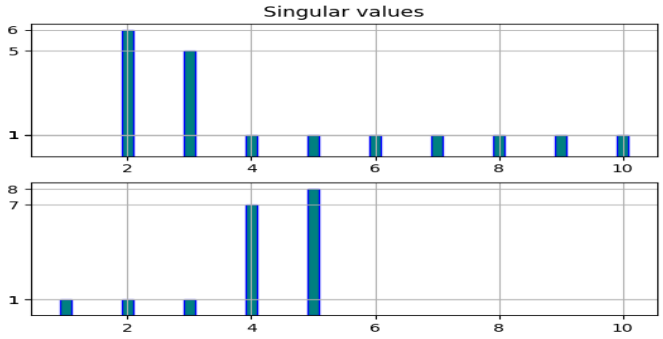
We can also see that the CM mapping presents smaller singular values for both the full and reduced range cases. As a practical matter, small singular values are usually a sign of ill-conditioning, in the one hand, and it can be difficult to distinguish zero and relatively small Singular Values in a large scale problem, on the other hand. For these reasons, we tend to say that, theoretically, inverting the LF formulation is less sensitive to errors in the observations.



(a) matrix associated with the model in Equation (7), size $= 160 \times 18$: A_{LF}



(b) matrix associated with the model in Equation (8), size $= 160 \times 18$: A_{CM}



(c) Singular values (rounded) of the linear mappings in figure 4a (top) and 4b (down).

Figure 4: 4a: Mapping from the data space $\mathbb{D}$ to the model space $\mathbb{M}_{\mathrm{LF}}$ and its histogram of singular values in figure 4c (top). 4b: Mapping from the data space $\mathbb{D}$ to the model space $\mathbb{M}_{\mathrm{CM}}$ and its histogram of singular values in figure 4c (bottom). The associated condition numbers are respectively 5.01 and 7.79.

Missing observations. Let us now suppose that images with date indexes 2, 3, 14, 16 and 17 have been discarded from pre-processing for lack of texture or high deformation 3.2. The index graph in figure 2b becomes as in figure 5 and the mappings become rank deficient (K is infinite). The obtained cost to minimize, respectively for the errors in Equations (7) and (8) and at a

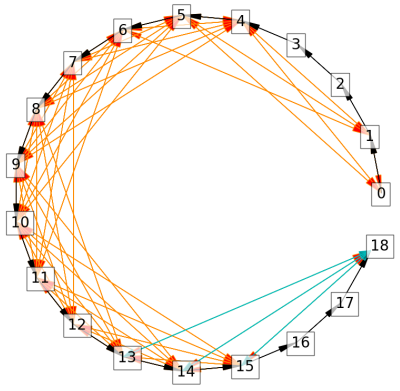


Figure 5: NOFF index graph for a rank-deficient problem: 18 vertices and 94 directed edges

given pixel p are:

$$\|b_{\mathrm{LF}}\|_L = \|\vec{F}_{\mathrm{obs}} - A_{\mathrm{LF}}\vec{F}_{\mathrm{est}}\|_L =$$

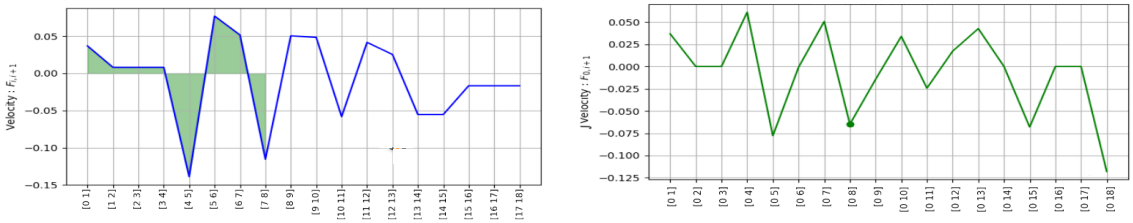
$$\left\| \begin{pmatrix} F_{\mathrm{obs}_{0,1}} \\ F_{\mathrm{obs}_{0,2}} \\ \vdots \\ F_{\mathrm{obs}_{0,5}} \\ F_{\mathrm{obs}_{1,0}} \\ F_{\mathrm{obs}_{1,2}} \\ \vdots \\ F_{\mathrm{obs}_{18,13}} \\ F_{\mathrm{obs}_{18,15}} \end{pmatrix} - \begin{pmatrix} 1&0&0&\dots&0&0&0&0&0 \\ 1&1&0&\dots&0&0&0&0&0 \\ & & & \vdots & & & & & \\ 1&1&1&\dots&0&0&0&0&0 \\ -1&0&0&\dots&0&0&0&0&0 \\ 0&1&0&\dots&0&0&0&0&0 \\ & & & \vdots & & & & & \\ 0&0&0&\dots&-1&-1&-1&-1&-1 \\ 0&0&0&\dots&0&0&-1&-1&-1 \end{pmatrix} \begin{pmatrix} F_{\mathrm{est}_{0,1}} \\ F_{\mathrm{est}_{1,2}} \\ F_{\mathrm{est}_{2,3}} \\ \vdots \\ F_{\mathrm{est}_{16,17}} \\ F_{\mathrm{est}_{17,18}} \end{pmatrix} \right\|_L, \tag{11}$$

$$\|b_{\mathrm{CM}}\|_L = \|\vec{F}_{\mathrm{obs}} - A_{\mathrm{CM}}\vec{F}_{\mathrm{est}}\|_L =$$

$$\left\| \begin{pmatrix} F_{\mathrm{obs}_{0,1}} \\ F_{\mathrm{obs}_{0,2}} \\ \vdots \\ F_{\mathrm{obs}_{0,5}} \\ F_{\mathrm{obs}_{1,0}} \\ F_{\mathrm{obs}_{1,2}} \\ \vdots \\ F_{\mathrm{obs}_{18,13}} \\ F_{\mathrm{obs}_{18,15}} \end{pmatrix} - \begin{pmatrix} 1&0&0&\dots&0&0&0&0&0 \\ 0&1&0&\dots&0&0&0&0&0 \\ & & & \vdots & & & & & \\ 0&0&0&\dots&0&0&0&0&0 \\ -1&0&0&\dots&0&0&0&0&0 \\ -1&1&0&\dots&0&0&0&0&0 \\ & & & \vdots & & & & & \\ 0&0&0&\dots&1&0&0&0&-1 \\ 0&0&0&\dots&0&0&1&0&-1 \end{pmatrix} \begin{pmatrix} F_{\mathrm{est}_{0,1}} \\ F_{\mathrm{est}_{0,2}} \\ F_{\mathrm{est}_{0,3}} \\ \vdots \\ F_{\mathrm{est}_{0,16}} \\ F_{\mathrm{est}_{0,17}} \\ F_{\mathrm{est}_{0,18}} \end{pmatrix} \right\|_L. \tag{12}$$

In this example, the Common Master image is the first valid one, i.e. I_0 and both linear mappings have rank 13. The dimension of the Nullity space of A equals the number of rejected image indexes leading to 5 linearly dependent columns out of 18. The condition-number is effectively infinite, the data space $\mathbb{D}$ has shrunk and further regularization or Model space $\mathbb{M}$ conditioning is needed to lead to a satisfying solution after inversion. To test the behaviour of

inverting both systems on the missing data indexes, a random observation vector was generated by assuming a gaussian process scaled by the temporal step. Figure 6 shows the simulated outputs. The values corresponding to missing observations are filled with a constant in $F_{\text{est}_{\text{LF}}}$ (VTS) and with zeros in $F_{\text{est}_{\text{CM}}}$ (PTS). To further illustrate how the solutions might be affected



(a) $F_{\text{est}_{\text{LF}}}$ (Velocity Time Series)

(b) $F_{\text{est}_{\text{CM}}}$ (Position Time Series)

Figure 6: Inverting a random observation vector: 6a: the output of inverting the LF temporal closure mapping is a VTS. 6b: the output of inverting the CM temporal closure mapping is a PTS. Images 2, 3, 14, 16 and 17 are discarded. The position of the tracked pixel in I_8 (green dot in figure 6a) is the green area in figure 6b.

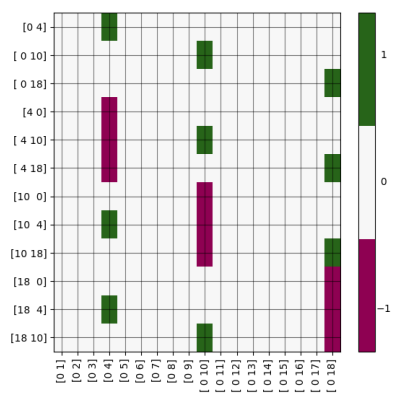
by missing observations, figure 7 illustrates the mappings when facing an extreme rejection case (that can occur in situations where the acquisition of the time-lapse is subject to extreme external conditions) and the obtained solutions $\vec{F}_{\text{est}}$ with manually simulated data that respects the temporal closure constraint but lacks several data.

3.4.3. Analytical aspect and optimization

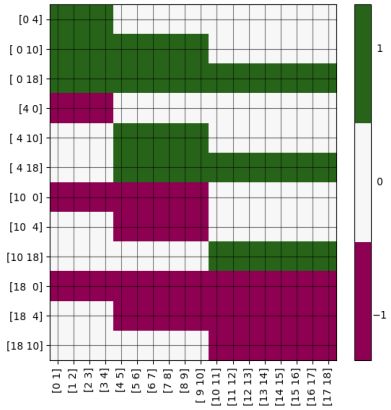
As stated before, the mappings from the model spaces $\mathbb{M}_{\text{LF}}$ and $\mathbb{M}_{\text{CM}}$ to the Data space $\mathbb{D}$ can be full rank, i.e. $\text{rank}(A) = N$, or rank deficient, depending on the number of missing images. Given that the rows of A are by definition redundant constraints, the row rank can never be full and hence the temporal closure linear mappings are never injective. This means that not every observation $\vec{F}_{\text{obs}} \in \mathbb{D}$ has a temporally coherent solution. Apart from specific situations where most of the data is rejected from preprocessing (see illustration case of figure 7), the problem is over-determined. Generally, for overdetermined problems, an optimal solution to the system of equations is sought by resolving the normal equation: $A^T A\vec{F}_{\text{est}} = A^T\vec{F}_{\text{obs}}$ either directly or by SVD decomposition. This is equivalent to solving the problem in Equation (13) where the cost function is the squared $L = 2$ norm $\Psi : x \to \|y - Ax\|_2 = \|y\|_2 - 2y^T Ax + \|Ax\|_2$. The linear mapping Ψ is twice differentiable and its Hessian is the positive semi-definite matrix $2A^T A$. The Least Squares minimization is thus convex and should lead to unique optimal VTS and PTS solutions, respectively, $\vec{F}_{\text{est}_{\text{LF}}}$ and $\vec{F}_{\text{est}_{\text{CM}}}$.

$$\underset{\vec{F}_{\text{est}}}{\arg\min}\|A \times \vec{F}_{\text{est}} - \vec{F}_{\text{obs}}\|_L, \quad \vec{F}_{\text{obs}} \in \mathbb{D} \tag{13}$$

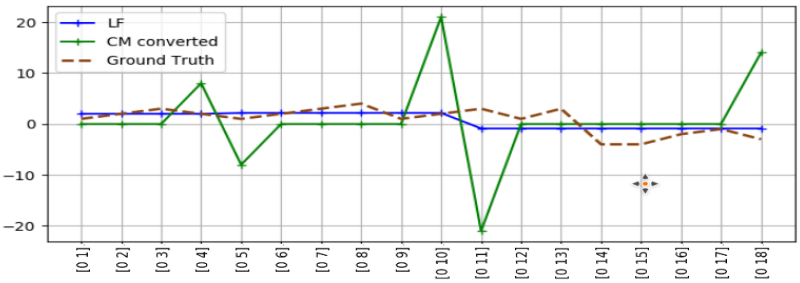
However, resolving the normal equation directly breaks down for the under-determined case (severe rejection cases) and the rank deficient case (missing observations). A truncated SVD decomposition is the most commonly used method for this scenario, and theoretically, gives the best unbiased Least Squares estimation. Nonetheless, even using a truncated SVD may lead to numerical difficulties as, for sparse and rank deficient problems, the problem is not strictly convex (presence of local minima) and this can lead to divergent solution when the observations are contaminated by noise and especially outliers. Most approaches for avoiding



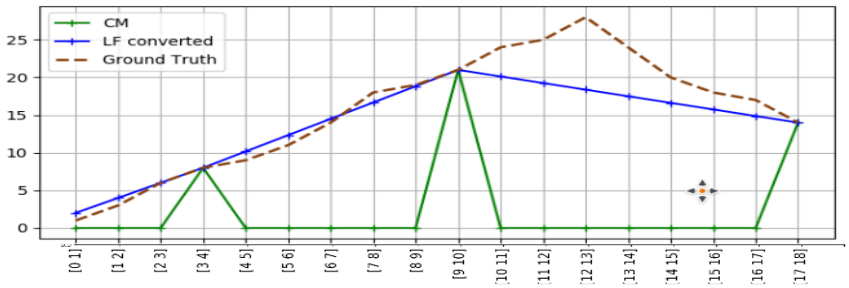
(a) matrix associated with the errors in Equation (7) with an important amount of missing data: A_{LF}



(b) matrix associated with the errors in Equation (8) with an important amount of missing data: A_{CM}



(c) Outputs, inpixels, of inverting matrices in Equations 7a and 7b: Velocity time series



(d) Outputs, in pixels, of inverting matrices in Equations 7a and 7b: Position time series

Figure 7: Illustration of a severe rank deficiency case: only images 0, 4, 10 and 18 are available over the 19 image forming the time lapse. The matrices of the LF (7a) and CM (7b) linear mappings became large (12×18) and of rank $r = 3$. 7c: velocity output of inverting a simulated observation vector $\boldsymbol{F}_{\mathrm{est}_{\mathrm{LF}}}$ and $\boldsymbol{F}_{\mathrm{est}_{\mathrm{CM}}}$ converted to a VTS. 7d: position output of inverting a simulated observation vector $\boldsymbol{F}_{\mathrm{est}_{\mathrm{CM}}}$ and $\boldsymbol{F}_{\mathrm{est}_{\mathrm{LF}}}$ converted to a PTS. .

these direct decomposition methods are based on dedicated incremental decomposition methods or on constrained/unconstrained linear and quadratic programming. Iteratively learning the inverse has the advantage of allowing to add equality/inequality constraints, to easily change the cost to be minimized to suit the prior knowledge on the solution and the errors, to add an initialization for the solution, and to optimize memory consumption by taking full advantage of the sparsity of the direct mappings. For minimizing a functional this way, one can extensively explore the model space, systematically (brute force), randomly (Monte Carlo), or use local properties of the cost function space to find a global minimum. There are many optimization algorithms and missing observations and rank-deficiency is shown to be handled differently according to the inner local properties of the used solver. The solvers that are relevant in this context were thus used in order to make the method tractable and will be cited and tested in Section 4.2 with simulated condition acquisition. More about direct and iterative methods for linear sparse problems can be found in [47][48][49].

Missing data and conditioning. As stated above, if there are no missing frames, the problem is full (column) rank. The convex resolution process theoretically leads to similar minimas with direct and iterative methods, after velocity/position conversion and for both LF and CM formulations. In case an image I_i has been rejected from preprocessing because of lack of information in the area of interest, the rows of A and $\vec{F}_{\text{obs}}$ involving the index i become obsolete and are deleted. For m deleted images over N, we have $2Nm - m(m+1)$ constraints less. The decay of the rank depends on m and on the contingency of the missing observations. In this case, for $r = rank(A)$ parameter to estimate the problem is still overdetermined (too many constraints involving the same unknowns) and for the $N-r$ values involving a rejected index, the problem is under-determined (redundant columns for LF system and zeroed columns for the CM system).

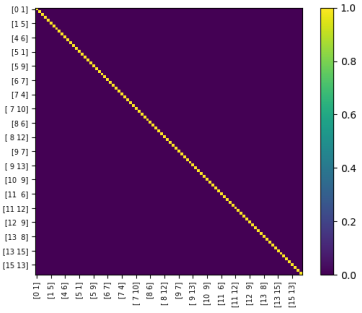
Reducing the maximum range and conditioning. It is important to note here that reducing the maximum range (as in figure 4) of the temporal closure loops can also have analytical benefits for the LF formulation. In fact, this operation prevents the matrix A_{LF} from having a dense row and, indeed, if a sparse matrix A have even one dense row, A^TA becomes full and iterative and direct methods using the normal equation are affected (much slower).

Error propagation simulation. Observational uncertainties (random errors) are assumed to be normally distributed with zero mean and with prior covariance Σ_{obs}. By "prior" we imply that they are not related to the inversion step but to the data and the prior layers of the pipeline. The uncertainties related to the observations propagate through the inversion process, leading to estimations $\vec{F}_{\text{est}_{\text{LF}}}$ and $\vec{F}_{\text{est}_{\text{CM}}}$ with covariances $\Sigma_{v_{\text{LF}}}$ (uncertainties of the LF estimated VTS) and $\Sigma_{p_{\text{CM}}}$ (uncertainties of the CM estimated PTS). The covariance of the estimations can be approximated by the rule of linear functions of random variables as in equation:

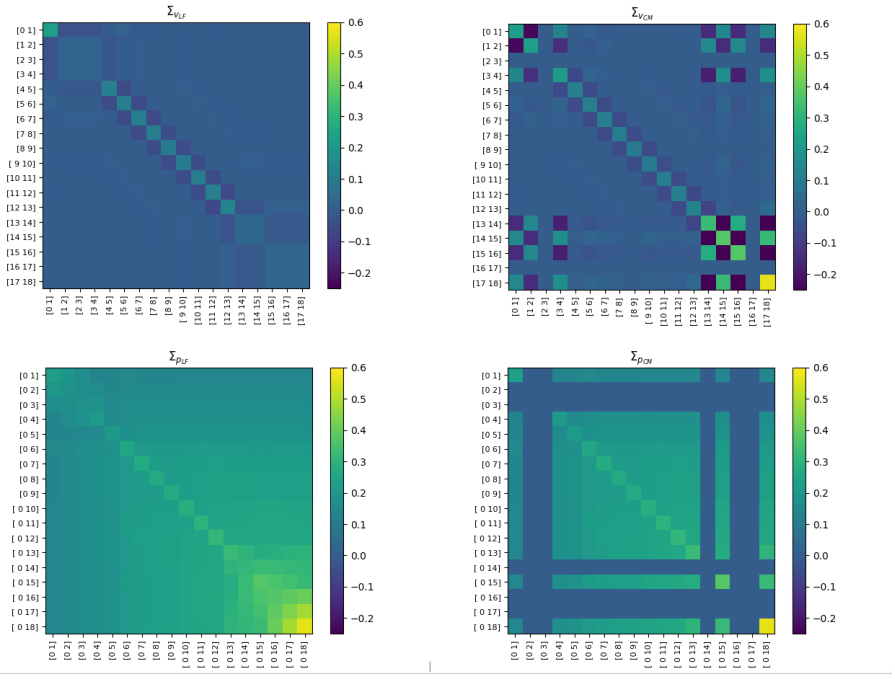
$$\Sigma_{\text{est}} = A^{\dagger}\ \Sigma_{\text{obs}}\ A^{\dagger^T}, \tag{14}$$

where † stands for the pseudo-inverse of A. As the mappings from VTS to PTS (and vice-versa) are linear, one can deduct $\Sigma_{p_{\text{LF}}}$ (uncertainties of the LF estimated PTS after conversion) from $\Sigma_{v_{\text{LF}}}$ and $\Sigma_{v_{\text{CM}}}$ (uncertainties of the CM estimated VTS after conversion) from $\Sigma_{p_{\text{CM}}}$. To ilustrate, let us consider statistically independent observational errors with variance 1, in which case $\Sigma_{\text{obs}} = \sigma^2_{\text{obs}}\mathbf{I}$. figure 8 shows the obtained covariance matrices of the estimates. The variances can be used to state data confidence bounds for the model parameters, e.g. $F_{\text{est}_{(\text{i,j})}} \pm 2\sigma_{F_{\text{est}_{(\text{i,j})}}}$ (but can be misleading in the case where the estimations are highly correlated). Figure 8 suggests

that VTS are less affected by uncertainty accumulation than PTS, on the one hand, and that the LF minimization is less prone to observational uncertainty propagation on the missing data indexes, on the other hand.



(a) Hypothetical uncertainties related to observations



(b) Uncertainties related to the estimated VTS (top) and PTS (bottom) through the LF (left) and CM (right) inversions.

Figure 8: Illustration of uncertainty propagation through inversion for the hypothetical data described in 3.4.2.

Weighting. In order to maximize the efficiency of the regularization and tackle heteroscedasticity, which is the fact that the observations are measured with different uncertainties, one can take into account the proper amount of influence over the observations by using the various residual error estimations and scores accumulated over the pipeline (Section 3.3). This is equivalent to preconditioning the matrix A or to minimizing the weighted norm $\|.\|_W$ of the temporal closure error b where W represents the weights of the inner-product defined by $< a, b >_W = b^* W a = b^T W a$ and thus must be a positive-definite bilinear *preferably* symmetric form. Practically, this can be implemented as follows:

$$\arg\min_{\vec{F}_{\text{est}}} \|W(A\vec{F}_{\text{est}} - \vec{F}\text{obs})\|_L. \tag{15}$$

The weighting matrix W is diagonal and the weights have to be estimated up to a proportionality factor. To focus accuracy, every observation $F_{\text{obs}_{(i,j)}}$ is given a weight $W_{(i,j)}$ corresponding to how well-registered the images I_i and I_j are (scores S_1 and S_2 in Section 3.3.3). To discount imprecision of the observations one would want to use an estimator that captures the information about how volatile $F_{\text{obs}_{i,j}}$ is (score R_1). In the Small Baseline Subset workflow, weighting is often done by multiplying by the precision matrix Σ_{obs}^{-1}.

Adding priors about the solutions. As the sought VTS and PTS are known to be smooth (quasi-stationary time series), a Tikhonov regularization term can also be incorporated in the cost function to avoid landing on "almost minimizers" that have a big norm. This term explicitly incorporates the regularity requirement (minimum variance) in the formulation of the problem. The damped system is such as:

$$\begin{aligned} \arg\min_{\vec{F}_{\text{est}}} \|W(A_{\text{LF}} \times \vec{F}_{\text{est}} - \vec{F}_{\text{obs}})\|_L + \lambda^2 \|\vec{F}_{\text{est}}\|_2 \\ \vec{F}_{\text{obs}} \quad \in \mathbb{D} \end{aligned} \tag{16}$$

In Equation (16), the first term measures the goodness of fit to the temporal closure constraints (in terms of the least L-norm cost), i.e. how well the estimations $\vec{F}_{\text{est}}$ predict the noisy observations $\vec{F}_{\text{obs}}$ and having important partial residuals $\|A_{\text{LF}_w} \times \vec{F}_{\text{est}} - \vec{F}_{\text{obs}_w}\|_L$ means that the solution does not solve the temporal closure problem. On the other hand, we should not make the residual smaller than the average size of the errors in $\vec{F}_{\text{obs}}$ as we do not want to fit the noise in the data. The second term $\lambda^2 \|\vec{F}_{\text{obs}_w}\|_2$ measures the regularity of the solution based on the knowledge that the naive estimations (without damping) can be dominated by high-frequency components with large amplitudes we want to avoid considering as solutions.

These assumptions and empirical experiments will determine the choice of λ that will best balance the problem, which is beyond the scope of this paper. The final resolution framework (choice of the solver for resolution and of the norm L) will be discussed in the result section (Section 4.1). The "regularization" term in this case (Tikhonov) is linear ($L = 2$ norm), but could have been any hand designed regularization priors, such as sparseness in the wavelet space and so on. Recent examples of cost function analysis and the Tikhonov regularization term for signal restoration by inversion can be found in [50] and [51].

4. Results on simulations

4.1. Simulated datasets

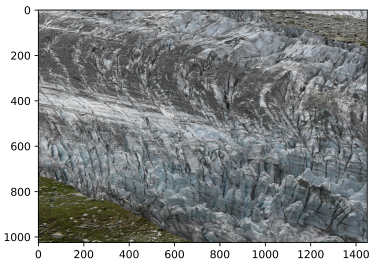
In order to assess the proposed pipeline in different controlled conditions, synthetic time-lapses are created by resampling a template image with different types of displacement fields.

This template image consists of a sub-image retrieved from real data. A simulated frameworks that consists of 4 synthetically generated time-lapses of a textured semi-rigid fluid (a glacier) with different parameters is obtained. These datasets have a full resolution (1100*1500 pixels covering a small sub-area of a glacier) and a simulated temporal range of $N = 6$. The different simulated displacement behaviors are:

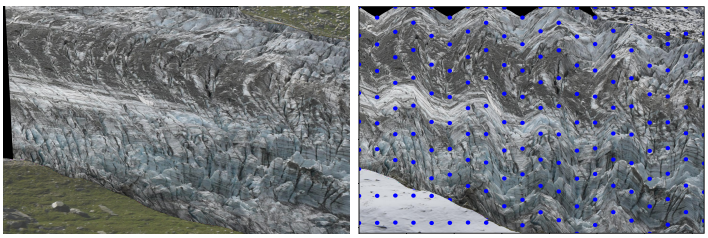
- Spatial rigidity: No spatial deformation (100% rigid) / Affine geometric transform **GA** (non-rigid, global)/ Piecewise-Affine geometric transform **GPA** (non-rigid, local).
- Temporal evolution of the PTS: Linear **LT** (VTS = constant) / Polynomial **PT** (non-linear PTS).

	Geometric transformation		Temporal Shift	
Dataset	**GA**	**PAG**	**LT**	**PT**
1	-	-	x	
2	x		x	
3	x			x
4		x		x

Table 1: Parameters of the generated datasets.



(a) Template image



(b) Warped image (affine) (c) Warped image (piecewise affine)

Figure 9: Example of piecewise affine transform on the dynamic zone. 9a: template image. 9b: affine transform warped image. 9c: piecewise affine transform warped image.

The template image I_0 (figure 9a) is retrieved from real data and is transformed to simulate images with indexes $i = (1, 2, 3, 4, 5)$. For dataset 1, the transformation is a sub-pixelic rigid motion (global translation) and the temporal evolution of the translation coefficient is linear (LT). For datasets 2 and 3, the transformation is a global affine transform (figure 9b) with greater rotation for dataset 3: for every image i in the simulated time-lapses 2 and 3, a rotation angle r_i (counter-clockwise direction), a shear factor sh_i and translation factors t_x and t_y are

defined with respect to image $i = 0$, which will set the six degrees of freedom of the affine transform. The temporal evolution of the translation coefficient is linear for dataset 2 (LT) and non-linear for dataset 3 (PT). The shear parameter of the affine transform increases along the vertical and temporal axis. This allows the analysis of the consequences of increasing non-rigid deformation on Normalized-Cross-Correlation (NCC) and Optical Flow (OF) computing as well as the ability of the regularization step to decrease the error on the computed VTS and PTS. These results are reported on figure 10 where the error related to the shear is approximated by summming over the temporal and horizontal axis, such as: $\sum_{x=0}^{1450}(\sum_{t=0}^{4}(\xi_t))$. ξ being the distance to the ground truth.
Dataset 4 is undergoing a piecewise-affine transform (figure 9c). Whereas affine transform exhibits several geometric properties that describe a non-rigid *global* motion as seen on the image plane, with a piecewise-affine transform, an image is modeled as a set of piecewise-convex areas (triangles), each of which is undergoing a different affine displacement. This geometric property grossly models the motion of a semi-rigid fluid subject to local compressions and dilatations. As for most remote sensing applications, only the object of interest is subject to motion, nearby zones remaining static. These static zones are retrieved from a real remote sensing dataset at different times to assess the invariance of the framework to environmental and photometric changes on motionless subimages.

4.2. Legend and evaluation strategy.

Several experiments are conducted to first evaluate the whole pipeline compared to a NCC approach with relatively clean data (Experiments A) and, later, to test the robustness of the inversion step to various acquisition scenarios and with the different proposed formulations (Experiments B, C and D):

- Experiments **A**: normal acquisition scenario
- Experiments **B**: noisy acquisition scenario
- Experiments **C**: missing observations
- Experiments **D**: systematic errors

Some other experiments and results will also be discussed.

The presented results are evaluated in terms of the temporal Root Mean Squared distance to the ground truth solution, denoted by the tilde, $\xi = \|\widetilde{F} - F_{\text{est}}\|_2$ averaged over a 500×1000 pixels central patch. For the LF formulation ξ refers to the error in the estimated Velocity Time Series (ξ_{VTS}) and for the CM formulation ξ refers to the error in the estimated Position Time Series (ξ_{PTS}). For every series of experiments, the iterative method is a weighted non-constrained damped least-squares minimization ($L = 2$) solved with the LSQR (Sparse Equations and Least Squares) algorithm [48][52]. This method proved to be the most appropriate as it is adapted to Sparse Linear systems (fewer vectors of working storage), to ill-conditioning and to both over and under determined problems. It is a conjugate-gradient type method that iteratively uses the Golub-Kahan bidiagonalization process to solve the normal equation. For the series of experiments presented in Subsections 4.5 and 4.6, other minimization routines that have shown an interesting behavior are presented (respectively; $L = 1$ norm for the cost function to deal with outliers and solvers that have shown a different strategy to deal with missing data in the observations).

*4.3. Experiments **A**: normal acquisition scenario*

In this series of experiments, the whole pipeline is evaluated and compared to Normalized Cross Correlation (NCC). Results with both the Leap-Frog and Common Master formulations

and with direct (truncated SVD) and iterative resolution are presented before and after regularization and with different weights (Non-weighted (NW), weighted with scores: $S = S_1 \times S_2$ (WS), R (WR) and mixed scores $S \times R$ (WSR) (see Section 3.3).

Dataset	OF	Leap Frog LF				
		Direct				Iterative
		NW	WS	WR	WSR	
1	0.01646	0.00171	0.00130	0.00115	0.00117	0.00289
2	0.08366	0.02005	0.16890	0.01590	0.01583	0.01457
3	0.10486	0.01737	0.01254	0.01097	0.01115	0.01036
4	0.04363	0.02834	0.02970	0.02970	0.02834	0.02115

(a) ξ_{VTS}

LF-converted to position			
NW	WS	WR	WSR
0.00339	0.00216	0.00135	0.00216
0.06105	0.03312	0.0315	0.02861
0.05623	0.02667	0.02759	0.01846
0.01346	0.01970	0.01750	0.02079

(b) ξ_{PTS}

Table 2: OF errors ξ_{VTS} (2a) and ξ_{PTS} (2b) before and after a weighted LF (WLF) regularization.

Dataset	OF	Common Master CM				
		Direct				Iterative
		NW	WS	WR	WSR	
1	0.05385	0.00339	0.00220	0.00135	0.00214	0.00527
2	0.38395	0.06115	0.03311	0.03149	0.02862	0.03450
3	0.42601	0.05623	0.02675	0.02763	0.02188	0.01850
4	0.09780	0.01349	0.01973	0.01748	0.02058	0.01345

(a) ξ_{PTS}

CM-converted to velocity			
NW	WS	WR	WSR
0.00171	0.00131	0.00116	0.00118
0.02008	0.01688	0.01585	0.01585
0.01737	0.01256	0.01099	0.01116
0.02837	0.02834	0.02968	0.02117

(b) ξ_{VTS}

Table 3: OF errors ξ_{PTS} (3a) and ξ_{VTS} (3b) before and after a weighted CM (WCM) regularization.

Dataset	NCC	Regularized NCC
1	0.11597	0.07570
2	0.63350	0.37497
3	2.43873	1.19262
4	1.71092	0.30109

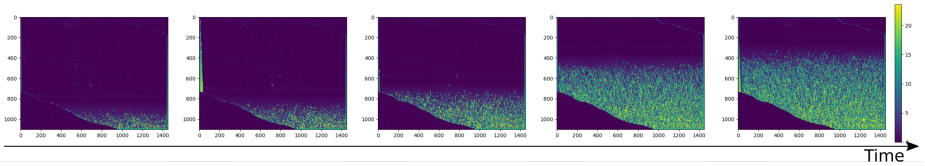
(a) ξ_{VTS}

Dataset	NCC	Regularized NCC
1	0.25562	0.19200
2	2.07360	0.43204
3	6.91797	3.83019
4	3.97034	0.51492

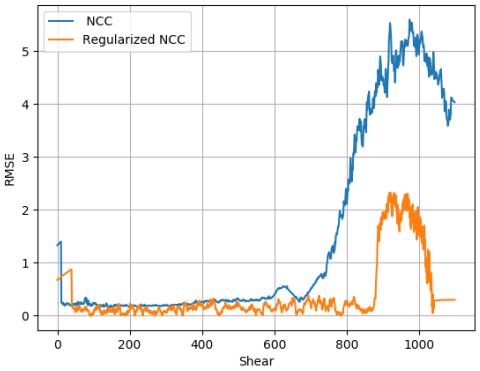
(b) ξ_{PTS}

Table 4: (NCC) errors ξ_{VTS} (4a) and ξ_{PTS} (4b) time series before and after (iterative - WLF) regularization.

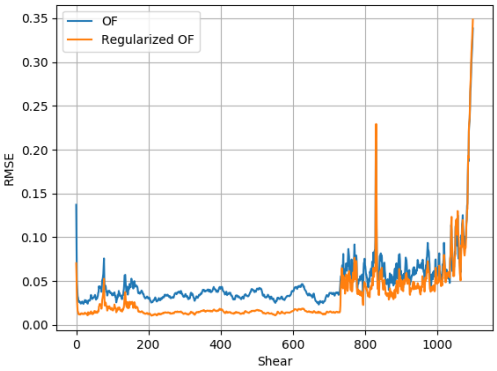
Conclusions A Results are reported in tables 2, 3 and 4. Unlike Optical Flow (OF), normalized Cross-Correlation (NCC), fails to capture non-rigid motion, especially on Dataset 3, which is characterized by a more distorting spatial transformation, and presents many outliers on Dataset 4, which is undergoing piecewise non-rigid spatial transformations (Table 4). Non-rigid deformation of the the processed images exhibited indeed abrupt changes in the correlation similarity measure. The error decreases by regularization for all datasets and weighting the minimization provides the solutions that give the smallest ξ_{VTS} and ξ_{PTS}. The mixed weighting with WSR gives better results most of the time (Tables 2, 3 and 4) and is therefore the one used for the iterative method and the rest of the experiments. The iterative damped Least-Squares resolution (LSQR) gives slightly better results than the Truncated SVD method and as the solver converges within 2 iterations, LSQR is a better choice for the task. The experiments with this acquisition scenario for minimizing the norm $L = 1$ of the temporal closure cost function shown results equivalent to those with $L = 2$ but with a slower convergence rate. The



(a) Non-regularized NCC spatial error maps with increasing shear along the y (directed to the bottom) and temporal axis.



(b) NCC RMSE with increasing shear



(c) OF RMSE with increasing shear

Figure 10: ξ_{VTS} in terms of the shear with NCC and OF for Dataset 2. The shear increases with the vertical and temporal axes resulting in an contamination of the observations.

inversion being pixelwise, the spatial transformation does not theoretically affect the amount of extra accuracy given by the temporal regularization, but we can see from figure 10 that when several observations are affected by registration errors (due to non-rigidity), inversion may fail to regularize the observations.

4.4. *Experiments* **B**: *noisy acquisition scenario*

In this series of experiments, an additive random white noise ω is injected spatially to *all* observations to simulate a measure uncertainty such as: $F_{\text{obs}_{(i,j)}} \leftarrow F_{\text{obs}_{(i,j)}} + \omega_{(i,j)}$. The $\omega_{(i,j)}$'s are independent, identically distributed and drawn from zero-mean normal distributions with different standard deviations σ_ω.

	WLF								
Dataset	$\sigma_\omega = 0.1$ pixel			$\sigma_\omega = 1$ pixel			$\sigma_\omega = 5$ pixel		
	OF $+\omega$	Direct	Iterative	OF $+\omega$	Direct	Iterative	OF $+\omega$	Direct	Iterative
4	0.13167	0.03364	0.03341	1.02832	0.53800	0.51000	5.02794	2.39173	1.38109

Table 5: OF erros ξ_{VTS} before and after a weighted LF regularization with different noise levels.

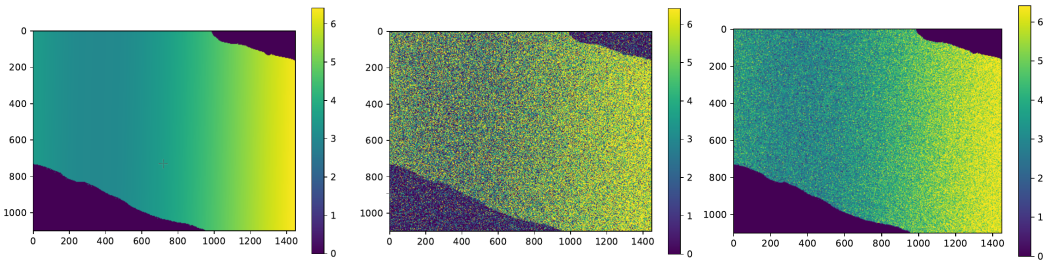
	WCM								
Dataset	$\sigma_\omega = 0.1$ pixel			$\sigma_\omega = 1$ pixel			$\sigma_\omega = 5$ pixel		
	OF $+\omega$	Direct	Iterative	OF $+\omega$	Direct	Iterative	OF $+\omega$	Direct	Iterative
4	0.20621	6.91810	0.03761	1.10082	7.27980	0.54335	5.09987	5.09320	2.26979

Table 6: OF errors ξ_{PTS} before and after a weighted CM regularization with different noise levels.

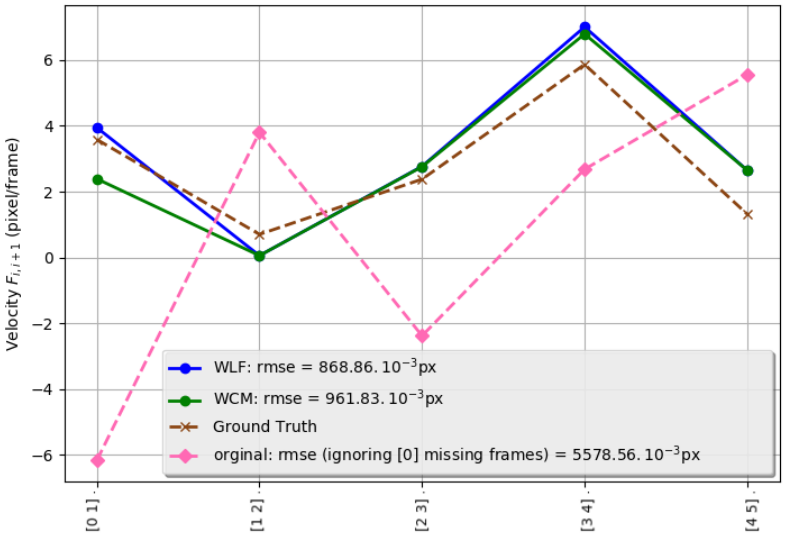
Conclusions B Results are reported in tables 5 and 6. The regularization process accurately denoises the obtained velocity fields from the added white noise and we can see in figure 11 that static zones have been cleared from residual displacement. Numerical instability associated with the SVD resolution can be seen in red in Table 6. The damped least-squares solution given by LSQR gives a "regularized" solution.

4.5. *Experiments* **C**: *missing observations*

In this series of experiments, images have been discarded from the input of the pipeline to imitate image rejection because of lack of texture (see Section 3.3). 1, 2 and 2 adjacent images are discarded over 6 to evaluate how the regularization process fills the gaps in the estimated VTS and PTS due to missing observations. The gap filling property is specific to the local approximations made by the used solver on the under-determined sub-systems induced by missing observations (local non-convexity). When solving with SVD or LSQR, gaps in the estimated VTS ($F_{\text{est}_{\text{LF}}}$) are filled with a constant value interpolated from neighboring data, and with zeros in the estimated PTS ($F_{\text{est}_{\text{CM}}}$) (see illustration case in figure 6). Other solvers were found to have other local properties for gap-filling. The COBYLA (Constrained Optimization BY Linear Approximation) algorithm [53][54][52] shows results similar to LSQR where no-data is missing and a gap filling strategy that gives the closest solutions to ground truth $\widetilde{F}$, at the cost of relatively slower execution time (approximately doubled). The iterative method in this subsection refers thus to the solution estimated by minimizing the damped least squares cost of the temporal closure errors with the COBYLA solver.



(a) Ground truth OF map (b) OF map with additional white noise (c) Regularized OF map



(d) OF VTS: LF and CM solutions at the central pixel of the simulated dataset 4.

Figure 11: Iterative LF inversion results after injection of a white noise of standard deviation $\sigma_\omega = 5$ pixel: Top: OF map between Images 0 and 1 of Dataset 4. Bottom: OF velocity pixel temporal tracking.

	WLF								
Dataset	1/6 missing image			2/6 missing images			2/6 adjacent missing images		
	OF	Direct	Iterative	OF	Direct	Iterative	OF	Direct	Iterative
4	0.03864+nod	0.53062	0.01667	0.04214+nod	0.79308	0.59328	0.03873+nod	1.66724	0.63470

Table 7: OF ξ_{VTS} before and after a weighted LF regularization with different configurations of missing observations. (nod = no data)

	WCM								
Dataset	1/6 missing image			2/6 missing images			2/6 adjacent missing images		
	OF	Direct	Iterative	OF	Direct	Iterative	OF	Direct	Iterative
4	0.12232+nod	4.22397	1.95198	0.09683+nod	6.52000	3.65185	0.13430+nod	6.50164	2.51473

Table 8: OF ξ_{PTS} before and after a weighted CM regularization with different configurations of missing observations. (nod = no data)

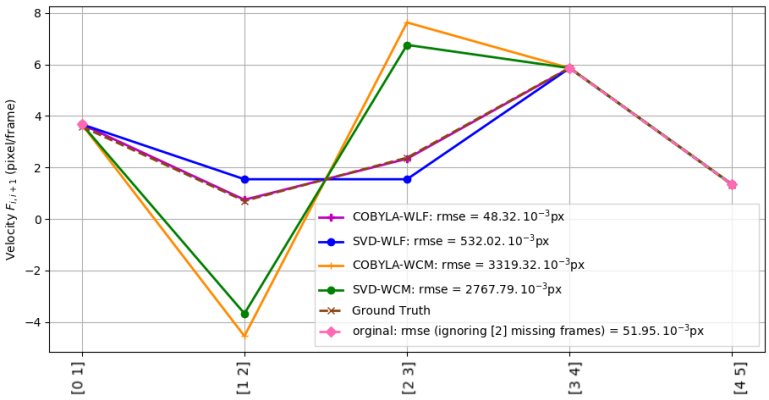


Figure 12: OF VTS: with SVD and COBYLA at the central pixel of the simulated dataset 4. Image 2 has been discarded creating a gap of 2 values on the original VTS. The WCM outputs (PTS) are converted to VTS.

Conclusions C Results are reported in tables 7 and 8. Empirical results show that the COBYLA method outperforms all the other solvers on all datasets and pixel samples through its local approximations. When data is missing, the Leap-Frog formulation of the temporal closure constraint gives better results than the Common-Master formulation (Figure 12) and the solutions from iterative methods are closer, in terms of the RMSE distance, to the ground truth velocity $\widetilde{F}$.

4.6. *Experiments* **D***: Systematic errors*

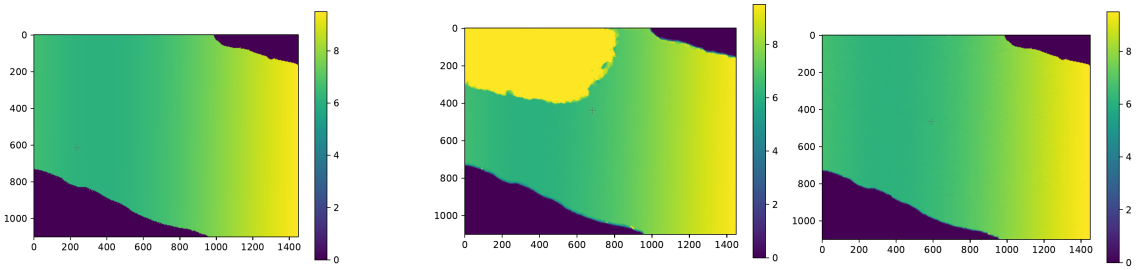
In this series of experiments, noise have been injected into (a part of) the processed velocity fields that include a common image I_i. The aim is to simulate bias induced by data, or systematic observational error, in the involved OF maps. This noise is drawn from a normal distribution with standard deviation $\sigma = 1$ and a mean value μ_0, which is the parameter we will vary in the experiments. In real data, this systematic observational error can be caused by a false detected event in the image I_i.

	WLF											
Dataset	$\mu_o = 5$ pix				$\mu_o = 10$ pix				$\mu_o = 20$ pix			
	OF + outliers	Direct	Iterative		OF + outliers	Direct	Iterative		OF + outliers	Direct	Iterative	
			$L=1$	$L=2$			$L=1$	$L=2$			$L=1$	$L=2$
4	2.25672	0.07824	0.02232	0.07824	4.49264	1.31144	0.02181	0.07824	8.96470	1.95816	0.02228	1.95662

Table 9: OF ξ_{VTS} before and after a weighted LF regularization with different outlier mean amplitude (associated with 1 image).

	WCM											
Dataset	$\mu_o = 5$ pix				$\mu_o = 10$ pix				$\mu_o = 20$ pix			
	OF + outliers	Direct	Iterative		OF + outliers	Direct	Iterative		OF + outliers	Direct	Iterative	
			$L=1$	$L=2$			$L=1$	$L=2$			$L=1$	$L=2$
4	2.25912	6.94773	0.01335	0.09298	4.49385	5.98537	0.01341	1.66244	8.96531	8.65959	0.01304	1.97657

Table 10: OF ξ_{PTS} before and after a weighted LF regularization with different outlier mean amplitude (associated with 1 image).



(a) Ground Truth OF map (b) OF map with additional outliers (c) Regularized OF map

Figure 13: Iterative LF inversion results after systematic error injection on the dynamic zone related to image 3. OF map between Images 0 and 3 of Dataset 4.

Conclusions D Results are reported in tables 9 and 10. The proposed weights rely strongly on the residual error in the static zones of the observed scene. In case the bias were injected in the dynamic zones (Figure 13), the weights might not "detect" the biased map and, as expected, using a cost function with $L = 1$ for the inversion is more robust against outliers (least absolute deviation). Using the extra regularization term (cost function in equation (16)) prevents the solution from diverging (Table 10).

4.7. Other experiments and conclusions

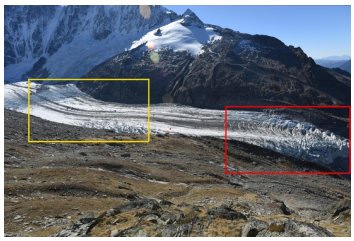
To summarize the results on the simulated datasets, the optimal resolution scheme depends on the nature of the suspected anomalies in the processed velocity and position time series. In the presence of outliers, such as when a high tolerance in the image rejection phase induced the presence of occluded zones in the scene, a norm $L = 1$ should be used to initialize a further regularization with a Least Squares cost function. Using the damped cost function in equation (16) has allowed us to avoid divergent solutions, which can happen in the presence of important noise, outliers or missing data. Extensive experiments showed that when combining two acquisition scenarios, the conclusions for both scenarios remain. Initializing with observations makes the convergence faster, in terms of the number of iterations. Introducing weights makes the optimization relatively slow because of the effect of floating point arithmetic on the numerical stability. As the weights must be determined up to a propto this day ity factor, one can consider scaling them up or rounding / truncating them.

5. Glacier monitoring

5.1. Motivation

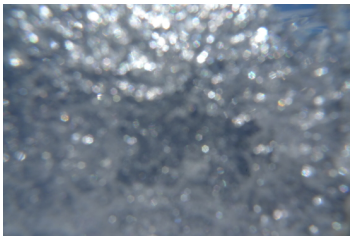
The motivation to apply our pipeline on alpine glacier monitoring is driven by the progressive awareness to climate changes as it is an indicator of the local effects of global warming. Glaciers are indeed thinning and melting after being subjected to a rise in the surrounding temperature and a spectacular retreat has been observed on most of the alpine glaciers in recent decades. The problem is that, to this day, it is still difficult to perform regular ground-based surveys of alpine glacier flow: only once or twice a year and only on a few spatial points due to their relatively unreachable location and dangerous geomorphology (crevasses). Remote studying of glaciers can thus lead us to a more complete understanding and modelling of global warming (along with meteorological data). On the other hand, studying the melting of glaciers is also

(a) Right camera on the Argentiere observation site



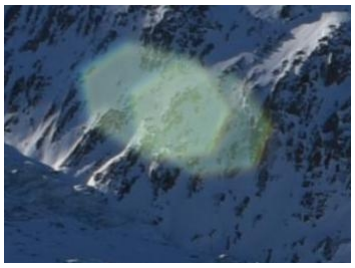
(b) Image example 1

(c) Image example 2



(d) Obstructed view

(e) Foggy view



(f) Sunbeam reflection

Figure 14: Argentiere site: data and artifact examples. In 14f a zoom is made to show a common artifact due to sunbeam reflections. This artifact does not have constant image coordinates.

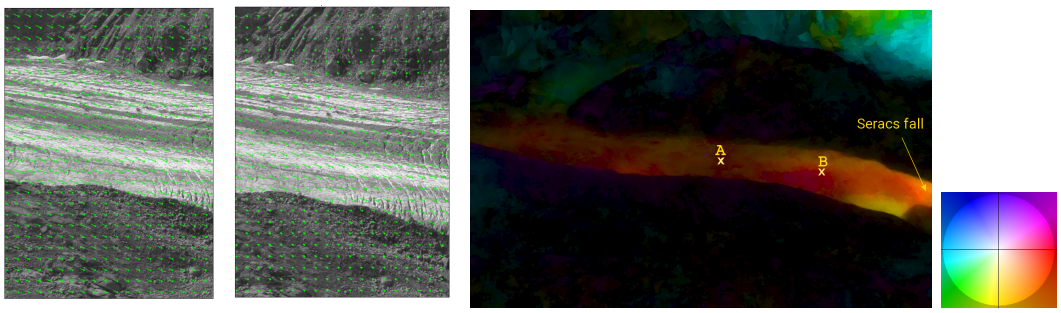
useful for economic and security reasons. Economic; because they also are water resources. Security; because the melting can cause hazardous events to the surrounding areas, such as destruction by seracs falls, or the so-called "glacier lake outburst flood" events, which occur when glacier lakes start to drain through a subglacial channel underneath the glacier causing flood in the downstream valleys. In this section we will show that daily surface measurements are possible via optical remote sensing if one can tackle the rather difficult radiometry. These surface measurements can later be used for a better understanding of the glacier's rheological [55] and morphological [56] parameters, and predict hazardous events [57].

5.2. Instrumentation and Set-up

The Argentiere glacier ($45°55$'N, $6°57$'E) is a temperate alpine glacier located in the Mont-Blanc massif in France. It is ~ 10km long with a surface of $\sim 12\text{km}^2$, and between altitudes 3400m (background) and 1600m (snout or Lognan seracs fall). Its surface flow has been studied numerous times by glaciologists [58] and remote sensors [59][60]. Given the ground constraints and the expected motion the choice of the acquisition system is made (camera sensor size, lens parameters range, disparity between the stereo views for depth maps computation, etc.). After the parametrization of the external (orientation, tilt and distances from the area of interest) and internal (consensus between focal length, depth of field and light dosage) camera parameters and finally, the instrumentation of the acquisition system (number of automatic images per day giving the object known velocity, electric power supply, etc.), the system gives us a stack of images taken automatically that will be inquired offline. To observe the largest area, without a "fish eye" effect and with the best resolution that we can embed in the installation, we chose a Nikon D810 with 20mm lens (f/2.8D) parametrized to have an infinite depth of field. It gives us 36.3-megapixel images with a high ground resolution, to observe as many details as possible,

but that makes code optimization needful. The expected inconveniences are indeed prominent in the raw data (see figure 14). The images also show a complex and repetitive texture pattern due to crevasses and cracked ice on the glacier zone. As stated before (section 3.3.4), the flow direction is mainly parallel to the image plane.The choice of the time step was made regarding the temporal closure range: The time separation between images within the temporal closure range must be long enough to increase the signal (the flow of the glacier) but short enough to preserve the features for tracking. A long time separation can indeed decorrelate the images to co-register and match for 3 main reasons: First, as stated before, photometric change, second; features on the glaciers may change due to melting, snowfall or changes caused by windblown snow. And last, good features for correlation, such as the crevasses and the ice foliation (mainly transverse), appear from time to time at the same places and with similar shapes. For example, crevasses open regularly upstream of a serac fall. For this reason, and as the Argentiere Glacier is a temperate glacier with a relatively high velocity, the time step between images in the time-lapse is 2 days (taken at 12AM to avoid drop shadows as much as possible) and the temporal closure range was chosen empirically to be within 10 images. This set-up makes the maximum time step between matched images (observations for the regularization) equal to 20 days. The time span of the presented results is from July 06 to December 06 2017. The computed depth map has an average uncertainty of ± 10 m compared to ground truth measurements.

5.3. Results on the Argentiere observation site

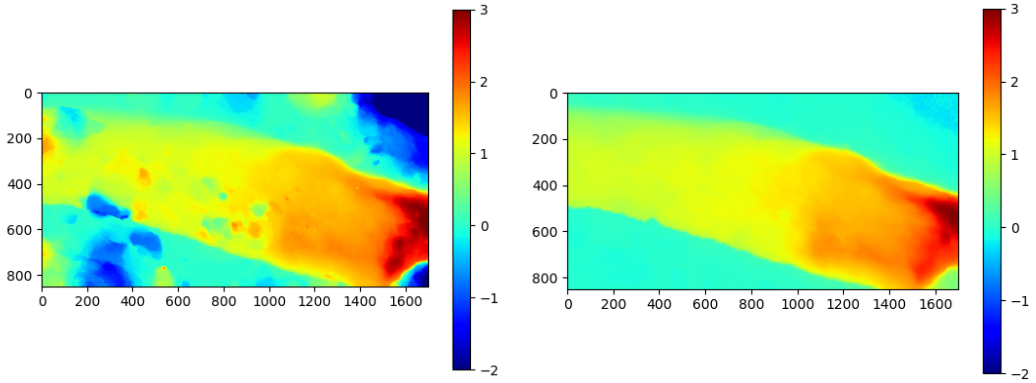


(a) Camera motion compensation (zoom)

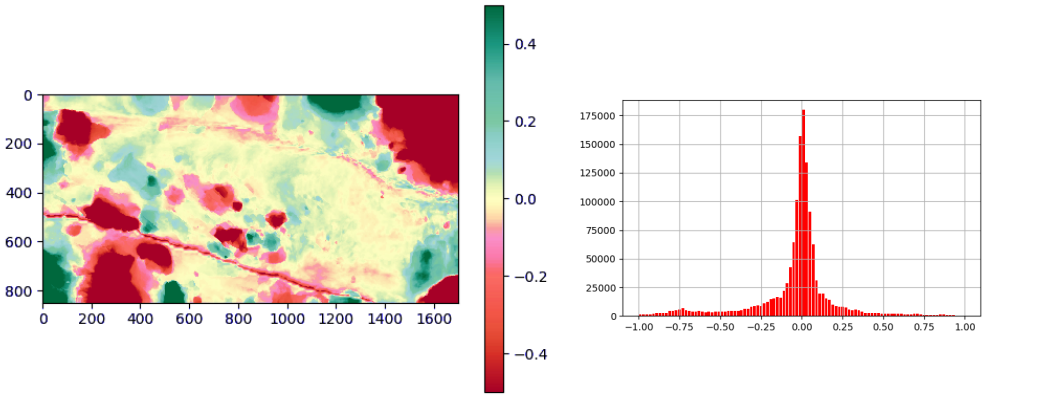
(b) Mean velocity map

Figure 15: Result on the Argentiere glacier. 15a: velocity map with and without camera motion compensation. This motion can be up to 20 pixels. 15b: mean velocity map in the HSV colorspace. Flow direction is encoded in the color (Hue) and flow amplitude is encoded in the Value.

The results obtained on the Argentiere site (Figure 14) are illustrated with figure 15b encoding the orientation and magnitude of the mean OF map in the HSV (Hue, Saturation, Value) colorspace. Images 27, 28, 37, 38, 57, 60, 61, 62, 63, 64, 72, 73 and 74 have been discarded from preprocessing mainly because of fog, clouds or droplets on the camera screen. These results were obtained by minimizing a damped Least-squares LF temporal closure error such as in Equation (16) using the LSQR solver. Data parallelism was performed to cope with the amount of pixels to process. A zoom on the results for the front and back zones of the glacier (respectively yellow and red rectangles in figure 14b are shown in figures 16 and 17. x and y components of the mean OF map are shown separately. We can clearly see the effect of the temporal regularization. On the static-zones, outliers caused by environmental changes, such as occlusion by clouds, have been removed. On the glacier zone, the mean OF maps shows a river-like flow

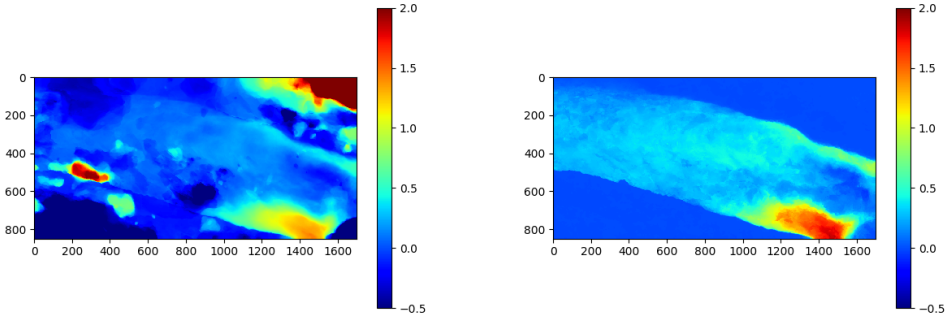


(a) Original (observed) mean OF map: x-component (b) Mean OF map after regularization: x-component



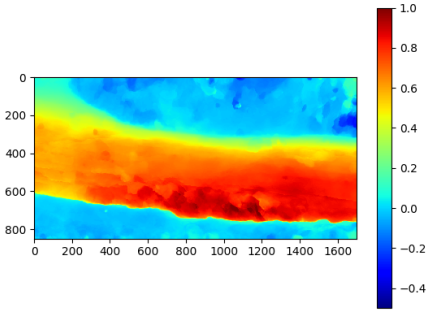
(c) Signed difference of maps in figure 16a and 16b: $F_{\text{obs}_{\text{mean}}} - F_{\text{est}_{\text{mean}}}$

(d) Histogram of values in the difference map in figure 16c

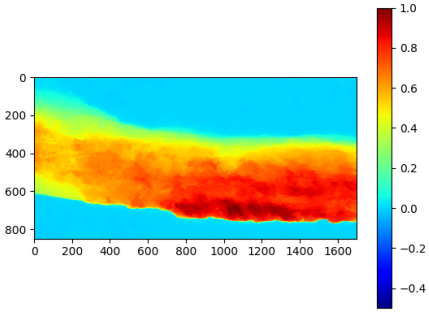


(e) Original (observed) mean OF map: y-component (f) Mean OF map after regularization: y-component

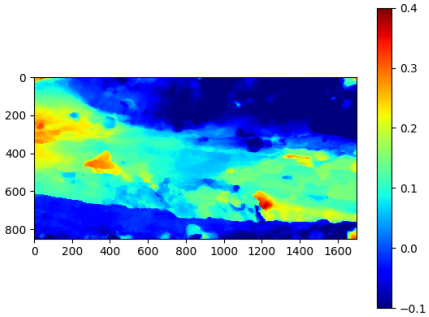
Figure 16: Results on the Argentiere glacier. Front zoom (in red on figure 14b).



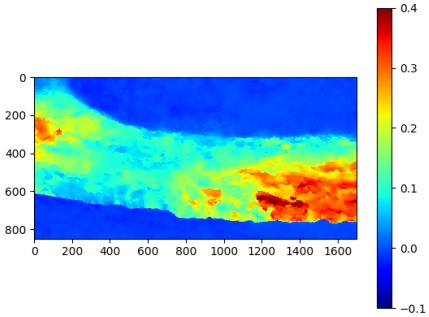
(a) Original (observed) mean OF map: x-component



(b) Mean OF map after regularization: x-component

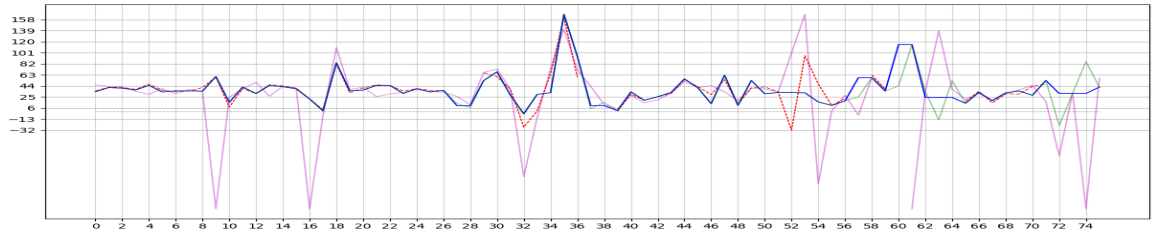


(c) Original (observed) mean OF map: y-component

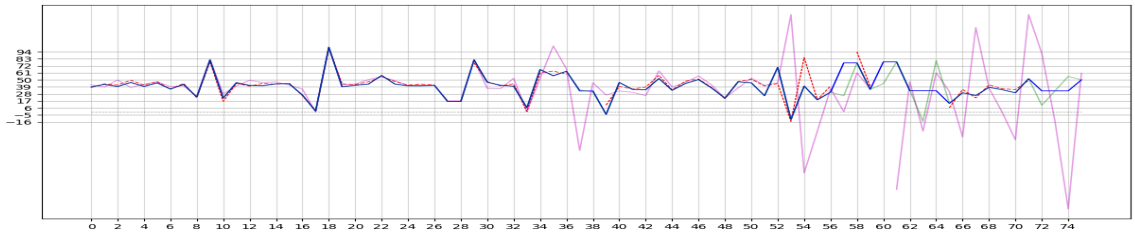


(d) Mean OF map after regularization: y-component

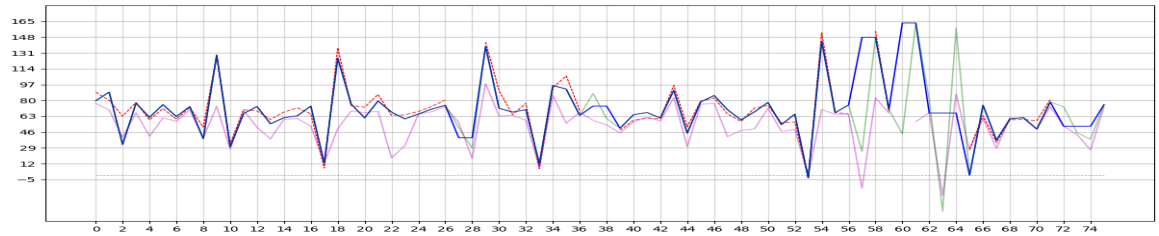
Figure 17: Results on the Argentiere glacier. Front zoom (in yellow in figure 14b). 17a and 17c: original mean OF maps. 17b and 17d: mean OF maps after regularization.



(a) $F_{\text{est}_{\text{LF}}}$ (VTS) for pixel A in figure 15b - converted to cm



(b) $F_{\text{est}_{\text{LF}}}$ (VTS) for pixel B in figure 15b - converted to cm



(c) $F_{\text{est}_{\text{LF}}}$ (VTS) for the seracs zone in figure 15b - converted to cm

Figure 18: Example of estimated VTS $\vec{F}_{\text{est}_{\text{LF}}}$ through the pipeline and on pixels showed by a yellow X in figure 15b. 18a: mean velocity has evolved from 37 cm/day (before VTS regularization) to 34 cm/day (after VTS regularization). 18b: mean velocity has evolved from 40 cm/day to 38 cm/day. 18c: mean velocity has evolved from 72 cm/day to 66 cm/day. x-component. **Legend.** Red: original VTS (OF), Pink: original VTS (NCC), blue: Regularized OF VTS (LSQR), green: regularized OF VTS (COBYLA).

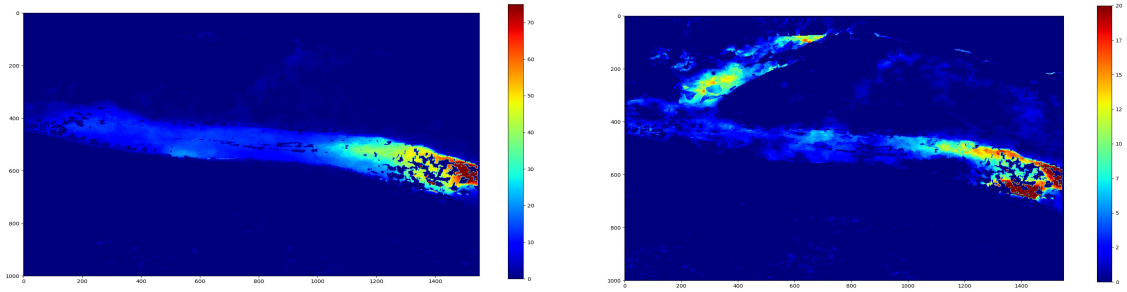


Figure 19: Example of regularized velocity map converted to cm. Between 06 and 07 June 2017. The gaps in the maps are due to missing depth data.

with more velocity at the center of the glacier caused by the friction on the rough edges within the bedrock. Figures 16f and 17d show OF maps more consistent with the longitudinal velocity and slope profiles presented in [61] on the same glacier. Examples of resulting time series for points A, B and the seracs fall zone in figure 15b are shown in figure 18.

	1996	2003	2007	2013	2017
Velocity ($cm.day^{-1}$) at altitude 2400m (cross-section 4 in [65][66])	24-41[62] 30[65][66]	23-50[63][56] 27[65][66]	18[65][66]	23(G)[64] 18(O)[64] 16[65] 15-30(SX) [64]	22 (us)
Mass balance averaged at altitude between 2380 – 2600m in $m.w.e.a^{-1}$ [67]	−3.05595	−4.9377857	−2.6223	−3.792	−4.993
cumulative precipitation in $mm.a^{-1}$ (MeteoFrance rainfall archives)	80	40	89	45	57

Table 11: Velocities of the Argentiere glacier from different sources. [62]: DinSAR between 10 & 11 March. [63]&[56]: SPOT5 between 23August & 18 Sept. [64]: multi-method, O for optical between 23 September & 7 October, SX for TerrasarX between 23 October & 3 November and G for GPS geocubes measurments (geocube 1018 in the paper) 16 to 21 September. [66][65]: GPS measurments, annual average. Us: optical photogrametry between June & December.

Spatially: we can see from figures 15b, 16, 17 and 19 that the amplitude of the surface velocity increases with the slope of the glacier and decrease with the altitude. This is in accordance with other velocity maps [56][62][63][64] on the glacier d'Argentiere and with a downslope flow of a fluid. This dynamic is valable for glaciers all over the world (example of Greenland glacier in [55]). Temporally: we see from figure 18 that the horizontal flow velocity time series don't have a noticeable trend at a given spatial point and during the inquired period of time (06 June to 06 December 2017). This can also be seen for the sliding velocity of the same glacier in [65] (figure 8.c) during the year 2013 along with further analysis of sudden accelerations peaks. The obtained results show a mean (spatial and temporal) horizontal velocity of 18 cm/day at the back of the glacier, as seen by our cameras, of 35 cm/day near the front of the glacier and up to 66 cm/day in the seracs fall zone. At an altitude of 2400 m (cross section 4 in [65] and [66]) the mean velocity is 22 cm/day. We reported published surface velocities over the Argentière Glacier on this cross-section since 1996 in table 11 and we can notice the temporal variability over time along with a general negative trend and local fluctuations. These temporal fluctuations in the surface flow velocity time series can also be seen in long mono-method time series such as in [66] (figure 4). In [65], the authors state that the negative trend is due to glacier thinning (weight loss) and local accelerations can be explained by a rise in temperature and episodes of precipitation. The inquired period (June - December) presents indeed higher temperature averages and several precipitation episodes (the rainy season began in August 2017). The semi-annual seasonality between dry/cool and hot/rainy seasons can be seen in [65] (figure 3). We can see that years 2003 and 2017 have similar mass balances [67] and velocity profiles [63] and are averaged over roughly the same period of the year, with a higher cumulative precipitation for year 2017. None of the referenced works over the Argentière Glacier gave spatially precise, temporally regularized time series of the surface velocity. These results can be used, for example, to estimate the glacier basal sliding or the ice thickness [56][67], which is of interest for potential hydro-power.

These experimental results highlight the potential of remote sensing via time-lapses, on the one hand, and the necessity to further regularize the automatically obtained results, on the other hand. As already highlighted section 3.3.4, numerous, precise and accurate in situ measurements are necessary to decrease the uncertainty related to the scaling to object coordinate and to geo-reference the observations our depth maps have an uncertainty of $\pm 10m$ which can be propagated

to the presented VTS.

6. Conclusion and future work

We have proposed a processing chain for accurate displacement time series generation from time-lapse images for complex geophysical phenomena monitoring. The pipeline includes robust pairwise OF maps computation and the formulation of a temporal closure error from a network of optical flow fields. Temporal closure mappings are formulated as an inverse problem where a weighted Least Squares solution is sought. Due to eventual ill-conditioning of the linear forward problem and the presence of noise in the observations, the inversion requires some extra "regularization" in order to generate physically plausible solutions. For this, a damping term was introduced to the cost function. Several formulations, weighting strategies, norms and solvers have been tested on simulated datasets to demonstrate the efficiency of the proposed approach under controlled and realistic conditions and to establish an automatic resolution framework given the a priori information on the observations. The proposed automatic method has been successfully applied to monitor the Argentiere glacier flow resulting in more complete and accurate OF maps time series all over the observed scene during a time span of 6 months. In future works, 2 extensions of the pipeline will be investigated; (1) replacing the manually constructed Gabor filter bank by learned convolutionnal filters, and (2) a scene flow [68] processing for deriving 3D motion fields from stereo regularized optical flow fields.

Acknowledgment

This work is supported by the ANR-15-CE23-0012 PHOENIX grant of the French *National Agency of Research*.

[1] E. Lannutti, M. G. Lenzano, C. Toth, L. Lenzano, and A. Rivera. Optical Flow Applied to Time-Lapse Image Series to Estimate Glacier Motion in the Southern Patagonia Ice Field. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 503–509, June 2016. doi: 10.5194/isprs-archives-XLI-B8-503-2016.

[2] C. Vogel, A. Bauder, and K. Schindler. Optical Flow for Glacier Motion Estimation. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 359–364, July 2012. doi: 10.5194/isprsannals-I-3-359-2012.

[3] J. Travelletti, C. Delacourt, P. Allemand, J.-P. Malet, J. Schmittbuhl, R. Toussaint, and M. Bastard. Correlation of multi-temporal ground-based optical images for landslide monitoring: Application, potential and limitations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 70:39 – 55, 2012. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2012.03.007. URL http://www.sciencedirect.com/science/article/pii/S0924271612000615.

[4] Adrian Schubert, Annina Faes, Andreas Kääb, and Erich Meier. Glacier surface velocity estimation using repeat terrasar-x images: Wavelet- vs. correlation-based image matching. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82:49 – 62, 2013. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2013.04.010. URL http://www.sciencedirect.com/science/article/pii/S0924271613001196.

[5] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti. A new algorithm for surface deformation monitoring based on small baseline differential sar interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11):2375–2383, Nov 2002. ISSN 0196-2892. doi: 10.1109/TGRS.2002.803792.

[6] Didier Massonnet and Kurt L Feigl. Radar interferometry and its application to changes in the earth's surface. *Reviews of geophysics*, 36(4):441–500, 1998.

[7] R. Lanari, P. Berardino, M. Bonano, F. Casu, C. De Luca, S. Elefante, A. Fusco, M. Manunta, M. Manzo, C. Ojha, A. Pepe, E. Sansosti, and I. Zinno. Sentinel-1 results: Sbas-dinsar processing chain developments and land subsidence analysis. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 2836–2839, July 2015. doi: 10.1109/IGARSS.2015.7326405.

[8] Jin-Woo Kim, Zhong Lu, Yuanyuan Jia, and C.K. Shum. Ground subsidence in tucson, arizona, monitored by time-series analysis using multi-sensor insar datasets from 1993 to 2011. *ISPRS Journal of Photogrammetry and Remote Sensing*, 107:126 – 141, 2015. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2015.03.013. URL http://www.sciencedirect.com/science/article/pii/S0924271615000933. Multitemporal remote sensing data analysis.

[9] Y. Yan, M. P. Doin, P. Lopez-Quiroz, F. Tupin, B. Fruneau, V. Pinel, and E. Trouve. Mexico city subsidence measured by insar time series: Joint analysis using ps and sbas approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4):1312–1326, Aug 2012. ISSN 1939-1404. doi: 10.1109/JSTARS.2012.2191146.

[10] S. H. Hong, S. Wdowinski, and S. W. Kim. Small temporal baseline subset (stbas): A new insar technique for multi-temporal monitoring wetland's water level changes. In *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages III – 550–III – 553, July 2008. doi: 10.1109/IGARSS.2008.4779406.

[11] Riccardo Lanari, Francesco Casu, Mariarosaria Manzo, Giovanni Zeni, Paolo Berardino, Michele Manunta, and Antonio Pepe. An overview of the small baseline subset algorithm: A dinsar technique for surface deformation analysis. In Detlef Wolf and José Fernández, editors, *Deformation and Gravity Change: Indicators of Isostasy, Tectonics, Volcanism, and Climate Change*, pages 637–661, Basel, 2007. Birkhäuser Basel. ISBN 978-3-7643-8417-3.

[12] Chloé Barboux, Tazio Strozzi, Reynald Delaloye, Urs Wegmüller, and Claude Collet. Mapping slope movements in alpine environments using terrasar-x interferometric methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 109:178 – 192, 2015. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2015.09.010. URL http://www.sciencedirect.com/science/article/pii/S0924271615002208.

[13] T. R. Lauknes, H. A. Zebker, and Y. Larsen. Insar deformation time series using an l_1 -norm small-baseline approach. *IEEE Transactions on Geoscience and Remote Sensing*, 49(1):536–546, Jan 2011. ISSN 0196-2892. doi: 10.1109/TGRS.2010.2051951.

[14] Kanika Goel and Nico Adam. An advanced algorithm for deformation estimation in non-urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73:100 – 110, 2012. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2012.06.001. URL http://www.sciencedirect.com/science/article/pii/S0924271612001013. Innovative Applications of SAR Interferometry from modern Satellite Sensors.

[15] Philippe Esling and Carlos Agon. Time-series data mining. *ACM Comput. Surv.*, 45(1):12:1–12:34, December 2012. ISSN 0360-0300. doi: 10.1145/2379776.2379788. URL http://doi.acm.org/10.1145/2379776.2379788.

[16] *Rolling Analysis of Time Series*, pages 313–360. Springer New York, New York, NY, 2006. ISBN 978-0-387-32348-0. doi: 10.1007/978-0-387-32348-0_9. URL https://doi.org/10.1007/978-0-387-32348-0_9.

[17] David M. Walker. *Kalman Filtering of Time Series Data, Chapter 6*, pages 137–157. Springer US, Boston, MA, 2002. ISBN 978-1-4615-0931-8. doi: 10.1007/978-1-4615-0931-8_7. URL https://doi.org/10.1007/978-1-4615-0931-8_7.

[18] Fernando Sedano, Pieter Kempeneers, and George Hurtt. A kalman filter-based method to generate continuous time series of medium-resolution ndvi images. *Remote Sensing*, 6(12):12381–12408, 2014. ISSN 2072-4292. doi: 10.3390/rs61212381. URL http://www.mdpi.com/2072-4292/6/12/12381.

[19] Bradley Efron. Missing data, imputation, and the bootstrap. *Journal of the American Statistical Association*, 89(426):463–475, 1994. ISSN 01621459. URL http://www.jstor.org/stable/2290846.

[20] Guilherme A. Barreto. *Time Series Prediction with the Self-Organizing Map: A Review*, pages 135–158. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-73954-8. doi: 10.1007/978-3-540-73954-8_6. URL https://doi.org/10.1007/978-3-540-73954-8_6.

[21] CHEN GEMAI, ABRAHAM BOVAS, and BENNETT GREG W. Parametric and non-parametric modelling of time series — an empirical study. *Environmetrics*, 8(1):63–74, 1997. doi: 10.1002/(SICI)1099-095X(199701)8:1<63::AID-ENV238>3.0.CO;2-B. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291099-095X%28199701%298%3A1%3C63%3A%3AAID-ENV238%3E3.0.CO%3B2-B.

[22] Ke Xu, Kar Wee Chia, and Adrian David Cheok. Real-time camera tracking for marker-less and unprepared augmented reality environments. *Image and Vision Computing*, 26(5):673 – 689, 2008. ISSN 0262-8856. doi: https://doi.org/10.1016/j.imavis.2007.08.015. URL http://www.sciencedirect.com/science/article/pii/S0262885607001266.

[23] Ghil M., Allen M. R., Dettinger M. D., Ide K., Kondrashov D., Mann M. E., Robertson A. W., Saunders A., Tian Y., Varadi F., and Yiou P. Advanced spectral methods for climatic time series. *Reviews of Geophysics*, 40(1):3–1–3–41. doi: 10.1029/2000RG000092. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000RG000092.

[24] Fabian Mathias Neyer. *Monitoring Rock Glaciers by Combining Photogrammetric and GNSS-Based Methods*. PhD thesis, ETH Zürich, 2016.

[25] Anette Eltner, Andreas Kaiser, Antonio Abellan, and Marcus Schindewolf. Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring. *Earth Surface Processes and Landforms*, 2017. doi: 10.1002/esp.4178.

[26] Ludmila I. Kuncheva and Juan J. Rodríguez. *An Experimental Study on Rotation Forest Ensembles*, pages 459–468. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-72523-7. doi: 10.1007/978-3-540-72523-7_46. URL https://doi.org/10.1007/978-3-540-72523-7_46.

[27] Durdu Oemer Faruk. A hybrid neural network and arima model for water quality time series prediction. *Eng. Appl. of AI*, 23:586–594, 06 2010.

[28] Sheng-Tun Li, Shu-Ching Kuo, Yi-Chung Cheng, and Chih-Chuan Chen. A vector forecasting model for fuzzy time series. *Applied Soft Computing*, 11:3125–3134, 04 2011.

[29] Marcus Schulz and Michael Matthies. Artificial neural networks for modeling time series of beach litter in the southern north sea. *Marine Environmental Research*, 98(Supplement C):14 – 20, 2014. ISSN 0141-1136. doi: https://doi.org/10.1016/j.marenvres.2014.03.014. URL http://www.sciencedirect.com/science/article/pii/S0141113614000634.

[30] John Cristian Borges Gamboa. Deep learning for time-series analysis. *CoRR*, abs/1701.01887, 2017. URL http://arxiv.org/abs/1701.01887.

[31] Berthold K.P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1):185 – 203, 1981. ISSN 0004-3702. doi: https://doi.org/10.1016/0004-3702(81)90024-2. URL http://www.sciencedirect.com/science/article/pii/0004370281900242.

[32] A. Wedel, D. Cremers, T. Pock, and H. Bischof. Structure- and motion-adaptive regularization for high accuracy optic flow. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1663–1668, Sept 2009. doi: 10.1109/ICCV.2009.5459375.

[33] Jérôme Revaud, Philippe Weinzaepfel, Zaïd Harchaoui, and Cordelia Schmid. Epicflow: Edge-preserving interpolation of correspondences for optical flow. *CoRR*, abs/1501.02565, 2015. URL http://arxiv.org/abs/1501.02565.

[34] Xu Chen, Pascal Zille, Liang Shao, and Thomas Corpetti. Optical flow for incompressible turbulence motion estimation. *Experiments in Fluids*, 56, 01 2015.

[35] Cucker Felipe Bürgisser, Peter. *Chapter 1: Normwise Condition of Linear Equation Solving & Chapter 5: Condition Numbers and Iterative Algorithms*. Springer International Publishing, 2014. doi: 10.1007/978-3-642-38896-5. URL https://www.springer.com/kr/book/9783642388958.

[36] Kui Zhang, Ruiqing Song, Hui Wang, Di Wu, and Hua Wang. Interferometric phase reconstruction using simplified coherence network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119:1 – 9, 2016. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2016.05.002. URL http://www.sciencedirect.com/science/article/pii/S0924271616300685.

[37] Bernard P. Wrobel. Least-squares methods for surface reconstruction from images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46(2):67 – 84, 1991. ISSN 0924-2716. doi: https://doi.org/10.1016/0924-2716(91)90017-P. URL http://www.sciencedirect.com/science/article/pii/092427169190017P.

[38] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[39] Albert Tarantola. A strategy for nonlinear elastic inversion of seismic reflection data. *GEOPHYSICS*, 51(10):1893–1903, 1986. doi: 10.1190/1.1442046. URL https://doi.org/10.1190/1.1442046.

[40] Simon Baker, Daniel Scharstein, J. P. Lewis, Stefan Roth, Michael J. Black, and Richard Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, Mar 2011. ISSN 1573-1405. doi: 10.1007/s11263-010-0390-2. URL https://doi.org/10.1007/s11263-010-0390-2.

[41] W. Li, K. Mao, H. Zhang, and T. Chai. Selection of gabor filters for improved texture feature extraction. In *2010 IEEE International Conference on Image Processing*, pages 361–364, Sept 2010. doi: 10.1109/ICIP.2010.5653278.

[42] Georgios Evangelidis and Emmanouil Psarakis. Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1858–1865, October 2008. URL https://hal.inria.fr/hal-00864385.

[43] Jerome Revaud, Philippe Weinzaepfel, Zaid Harchaoui, and Cordelia Schmid. DeepMatching: Hierarchical Deformable Dense Matching. *International Journal of Computer Vision*, 120(3):300–323, December 2016. doi: 10.1007/s11263-016-0908-3. URL https://hal.inria.fr/hal-01148432.

[44] Philippe Weinzaepfel, Jérôme Revaud, Zaid Harchaoui, and Cordelia Schmid. DeepFlow: Large displacement optical flow with deep matching. In *ICCV 2013 - IEEE International Conference on Computer Vision*, pages 1385–1392, Sydney, Australia, December 2013. IEEE. doi: 10.1109/ICCV.2013.175. URL https://hal.inria.fr/hal-00873592.

[45] T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):500–513, March 2011. ISSN 0162-8828. doi: 10.1109/TPAMI.2010.143.

[46] Carlo Tomasi and Takeo Kanade. Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9(2):137–154, Nov 1992. ISSN 1573-1405. doi: 10.1007/BF00129684. URL https://doi.org/10.1007/BF00129684.

[47] ÅKE BJURCK. Methods for sparse linear least squares problems. *Sparse Matrix Computations*, pages 177 – 199, 1976. doi: https://doi.org/10.1016/B978-0-12-141050-6.50015-5. URL https://www.sciencedirect.com/science/article/pii/B9780121410506500155.

[48] Christopher C. Paige and Michael A. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Softw.*, 8(1):43–71, March 1982. ISSN 0098-3500. doi: 10.1145/355984.355989. URL http://doi.acm.org/10.1145/355984.355989.

[49] J. A. Scales and A. Gersztenkorn. Robust methods in inverse theory. *Inverse Problems*, 4:1071–1091, October 1988. doi: 10.1088/0266-5611/4/4/010.

[50] C.J. Budd, M.A. Freitag, and N.K. Nichols. Regularization techniques for ill-posed inverse problems in data assimilation. *Computers and Fluids*, 46(1):168 – 173, 2011. ISSN 0045-7930. doi: https://doi.org/10.1016/j.compfluid.2010.10.002. URL http://www.sciencedirect.com/science/article/pii/S0045793010002707. 10th ICFD Conference Series on Numerical Methods for Fluid Dynamics (ICFD 2010).

[51] L. Mazzieri Gisela, D.Spies Ruben, and G. Temperini Karina. Regularization of inverse ill-posed problems with l2-bv penalizers and applications to signal restoration. In *Analysis, Modelling, Optimization, and Numerical Techniques*, pages 127–137, Cham, 2015. Springer International Publishing. ISBN 978-3-319-12583-1.

[52] Systems Optimization Laboratory, Huang Engineering Center, and USA Stanford, CA 94305-4121. Sol optimization library. http://web.stanford.edu/group/SOL/download.html.

[53] M J. D. Powell. A view of algorithms for optimization without derivatives. *Mathematics TODAY*, 43, 01 2007.

[54] M. J. D. Powell. *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*, pages 51–67. Springer Netherlands, Dordrecht, 1994. ISBN 978-94-015-8330-5. doi: 10.1007/978-94-015-8330-5_4. URL https://doi.org/10.1007/978-94-015-8330-5_4.

[55] Tavi Murray, Nick Selmes, Timothy D. James, Stuart Edwards, Ian Martin, Timothy O'Farrell, Robin Aspey, Ian Rutt, Meredith Nettles, and Tim Baugé. Dynamics of glacier calving at the ungrounded margin of helheim glacier, southeast greenland. *Journal of Geophysical Research: Earth Surface*, 120(6):964–982. doi: 10.1002/2015JF003531. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JF003531.

[56] Antoine Rabatel, Olivier Sanchez, Christian Vincent, and Delphine Six. Estimation of glacier thickness from surface mass balance and ice flow velocities: A case study on argentière glacier, france. *Frontiers in Earth Science*, 6:112, 2018. ISSN 2296-6463. URL https://www.frontiersin.org/article/10.3389/feart.2018.00112.

[57] E. Schwalbe, R. Koschitzki, and H.-G. Maas. Recognition of drainage tunnels during glacier lake outburst events from terrestrial image sequences. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B8:537–543, 2016. doi: 10.5194/isprs-archives-XLI-B8-537-2016. URL https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B8/537/2016/.

[58] C. Vincent, A. Soruco, D. Six, and E. Le Meur. Glacier thickening and decay analysis from 50 years of glaciological observations performed on glacier d'argentière, mont blanc area, france. *Annals of Glaciology*, 50(50):73–79, 2009. doi: 10.3189/172756409787769500.

[59] Fanny Ponton, Emmanuel Trouvé, Michel Gay, A. Walpersdorf, Renaud Fallourd, Jean-Marie Nicolas, Flavien Vernier, and Jean-Louis Mugnier. Observation of the argentière glacier flow variability from 2009 to 2011 by terrasar-x and gps displacement measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7:3274–3284, 2014.

[60] Lionel Benoit, Amaury Dehecq, Ha-Thai Pham, Flavien Vernier, Emmanuel Trouvé, Luc Moreau, Olivier Martin, Christian Thom, Marc Pierrot-Deseilligny, and Pierre Briole. Multi-method monitoring of Glacier d'Argentière dynamics. *Annals of Glaciology*, 56 (70):118–128, 2015. doi: 10.3189/2015AoG70A985. URL https://hal.archives-ouvertes.fr/hal-01154174.

[61] E. Trouve, G. Vasile, M. Gay, L. Bombrun, P. Grussenmeyer, T. Landes, J. M. Nicolas, P. Bolon, I. Petillot, A. Julea, L. Valet, J. Chanussot, and M. Koehl. Combining airborne photographs and spaceborne sar data to monitor temperate glaciers: Potentials and limits. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4):905–924, April 2007. ISSN 0196-2892. doi: 10.1109/TGRS.2006.890554.

[62] E. Trouve, G. Vasile, M. Gay, L. Bombrun, P. Grussenmeyer, T. Landes, J. M. Nicolas, P. Bolon, I. Petillot, A. Julea, L. Valet, J. Chanussot, and M. Koehl. Combining airborne photographs and spaceborne sar data to monitor temperate glaciers: Potentials and limits. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4):905–924, April 2007. ISSN 0196-2892. doi: 10.1109/TGRS.2006.890554.

[63] E. Berthier, H. Vadon, D. Baratoux, Y. Arnaud, C. Vincent, K.L. Feigl, F. Rémy, and B. Legrésy. Surface motion of mountain glaciers derived from satellite optical imagery. *Remote Sensing of Environment*, 95(1):14 – 28, 2005. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2004.11.005. URL http://www.sciencedirect.com/science/article/pii/S0034425704003463.

[64] Lionel Benoit, Amaury Dehecq, Ha-Thai Pham, Flavien Vernier, Emmanuel Trouvé, Luc Moreau, Olivier Martin, Christian Thom, Marc Pierrot-Deseilligny, and Pierre Briole. Multi-method monitoring of Glacier d'Argentière dynamics. *Annals of Glaciology*, 56 (70):118–128, 2015. doi: 10.3189/2015AoG70A985. URL https://hal.archives-ouvertes.fr/hal-01154174.

[65] CHRISTIAN VINCENT and LUC MOREAU. Sliding velocity fluctuations and subglacial hydrology over the last two decades on argentière glacier, mont blanc area. *Journal of Glaciology*, 62(235):805–815, 2016. doi: 10.1017/jog.2016.35.

[66] C. Vincent, A. Soruco, D. Six, and E. Le Meur. Glacier thickening and decay analysis from 50 years of glaciological observations performed on glacier d'argentière, mont blanc area, france. *Annals of Glaciology*, 50(50):73–79, 2009. doi: 10.3189/172756409787769500.

[67] C. Vincent, A. Soruco, M. F. Azam, R. Basantes-Serrano, M. Jackson, B. Kjøllmoen, E. Thibert, P. Wagnon, D. Six, A. Rabatel, A. Ramanathan, E. Berthier, D. Cusicanqui, P. Vincent, and A. Mandal. A nonlinear statistical model for extracting a climatic signal from glacier mass balance measurements. *Journal of Geophysical Research: Earth Surface*, 123(9):2228–2242. doi: 10.1029/2018JF004702. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JF004702.

[68] S. Vedula, P. Rander, R. Collins, and T. Kanade. Three-dimensional scene flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):475–480, March 2005. ISSN 0162-8828. doi: 10.1109/TPAMI.2005.63.