Spatio-temporal missing data reconstruction in satellite displacement measurement time series

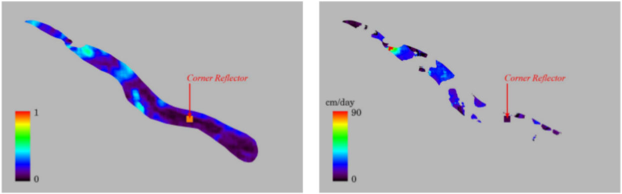
Alexandre Hippert-Ferrer[1 †], Yajing Yan[1], Philippe Bolon[1], Romain Millan[2]

[1] Laboratoire d'Informatique, Systèmes, Traitement de l'Information et la Connaissance (LISTIC), Annecy, France
[2] Institut des Géosciences de l'Environnement (IGE), Université Grenoble Alpes, CNRS, Grenoble, France
[†] Correspondence to: alexandre.hippert-ferrer@univ-smb.fr

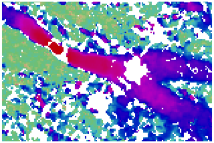
Thursday, May 7

Introduction

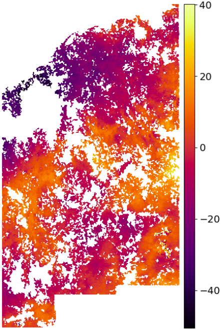
- Missing data is a frequent issue in displacement time series in both space and time.
- It can prevent the full understanding of the physical phenomena under observation.
- Causes : rapid surface changes, missing image, technical limitations.



Argentière glacier, offset tracking of TerraSAR-X in Summer 2010 (Fallourd et al., 2011)



Surface velocity over Fox Glacier, offset tracking of Sentinel-2 images in February 2018 (Millan et al., 2019)



Slow slip event (interferometry), Mexico (Maubant et al., 2020)

Motivation of the study

Handling missing data in displacement time series

- Classical approach : spatial **or** temporal interpolation
- Not exploited (yet) : spatio-temporal information

→ **Manage spatio-temporal missing data in time series** ←

Objective : propose a statistical gap-filling method addressing

1. Randomness and possible spatial, temporal and spatio-temporal correlation of
   - Noise
   - Missing data
2. Complex displacement behaviors (mixed frequencies)
3. Small time series

Extended EM-EOF

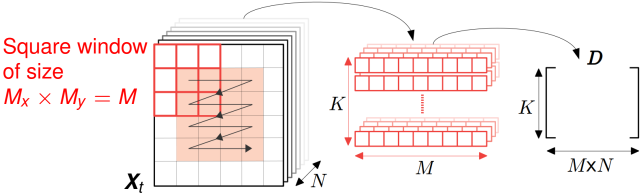
→ Extension of the EM-EOF method (Hippert et al., 2019, 2020) [3, 4]

Key features of the extended EM-EOF method :

- Low rank structure of the sample spatio-temporal covariance matrix.
- Displacement signal and noise decomposed in empirical orthogonal functions (EOFs).
- Reconstruction with an appropriate initialization of missing values.
- Expectation-Maximization (EM)-type algorithm for resolution.

Data representation

- Let $\boldsymbol{X}_t = \{x_{ij}(t)\}_{1 \le i \le P_x, 1 \le j \le P_y}$ be a spatial grid observed at time $t = 1, \ldots, N$.
- Some elements of $\boldsymbol{X}_t$ are missing.
- All $\boldsymbol{X}_t$ are stacked into a spatio-temporal data matrix $\boldsymbol{Y} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_N)$.



- Each $\boldsymbol{X}_t$ is augmented into a Hankel-block Hankel (HbH) matrix $\boldsymbol{D}_t$ of size $K \times M = K_x K_y \times M_x M_y$, with $K_x = (P_x - M_x + 1)$, $K_y = (P_y - M_y + 1)$.
- All $\boldsymbol{D}_t$ is stacked into a spatio-temporal matrix $\mathcal{D}$ of size $(K \times NM)$, that is $\mathcal{D} = (\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_N)$.

Covariance estimation and decomposition

- Sample spatio-temporal covariance is estimated :

$$\hat{\mathbf{C}} = \frac{1}{K}\boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{D}} \tag{1}$$

- The eigenvalue decomposition (EVD) of matrix $\hat{\mathbf{C}}$ yields to :

$$\hat{\mathbf{C}} \overset{\text{EVD}}{=} \sum_{i=1}^{NM} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T \tag{2}$$

Vectors $\boldsymbol{u}_i$ are the NM EOFs modes of matrix $\boldsymbol{\mathcal{D}}$. First modes capture the main spatio-temporal dynamical behavior of the signal, others represent perturbations.

- Reconstruction with an optimal number of EOF modes $R \ll NM$ is obtained as

$$\hat{\boldsymbol{\mathcal{D}}} = \boldsymbol{A}_R \boldsymbol{U}_R \tag{3}$$

$\boldsymbol{A}$ is the matrix of principal components, which are the projection of $\boldsymbol{\mathcal{D}}$ on each EEOF $\boldsymbol{u}$.

How do we find R?

Selection of the optimal number of EOF modes

1. Root-mean-square error (cross-RMSE) on cross-validated data $\mathcal{Y} \in \mathbf{Y}$:

$$\frac{1}{MN}||\hat{\mathcal{Y}}_k - \mathcal{Y}||_2 \tag{4}$$

- Requires no a priori information

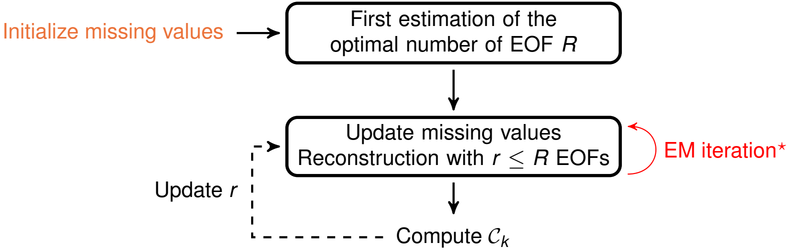
2. Confidence index associated with each eigenvalue of $\boldsymbol{\mathcal{D}}$:

$$\mathcal{C}_k = \frac{\max(\Gamma_k) - \Gamma_k}{\max(\Gamma_k) - \min(\Gamma_k)} \qquad k = 1, \ldots, NM \tag{5}$$

with $\Gamma_k = \log\left(\frac{\Delta\lambda_k}{\lambda_j - \lambda_k}\right)$.

- Investigation of eigenvalue degeneracy, which is linked to their uncertainties $\frac{\Delta\lambda_k}{\lambda_j - \lambda_k}$.
- Over-estimation of EOF modes is addressed by building metric $\mathcal{C}_k$.

General principle of the method



$\star$ For a fixed number of EOF modes, cross-RMSE is computed until it converges.

Determination of the spatial lag M for spatio-temporal covariance construction

- Trade-off between the amount of information extracted in the window (large M) and the number of repetitions of the window within each image (small M).
- Upper limit based on covariance estimation theory : $M < P/6$
- Lower limit :
  We use the spatial decorrelation decay τ defined as :

$$\tau = -\frac{\Delta P}{\log r} \tag{6}$$

r : lag-one auto-correlation
ΔP : spatial sampling rate, here 1 pixel.

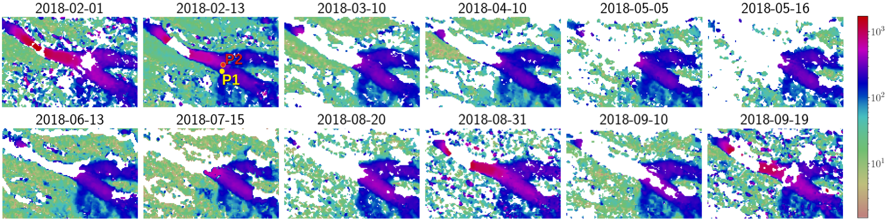
M can be approximated by $M \simeq P/\tau$ (Ghil et al., 2002) [1] which gives $M > P/20$ with $r < 0.95$.

Surface velocities on Fox Glacier, New Zealand

Period	Platform	Data type	Time series size	[min, max]% missing
02/2018-09/2018	Sentinel-2	Offset tracking	12	[10, 60]%

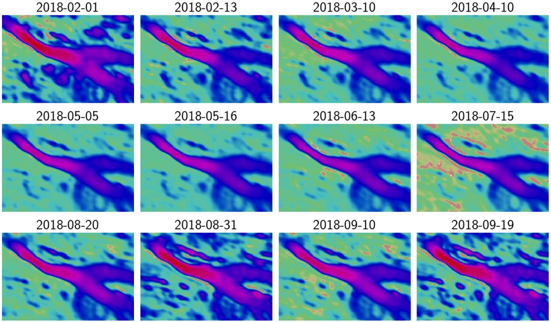
Time series description.

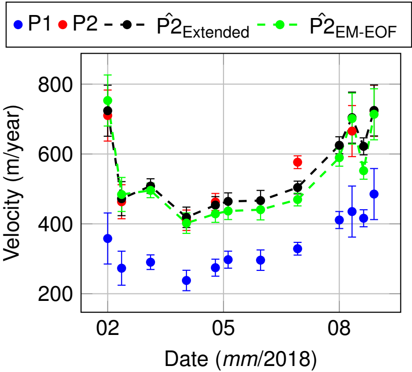
- Surface velocities computed from the study of (Millan et al., 2019) [5].



Surface velocities (m/year) on Fox Glacier. P1 and P2 locations are selected for temporal evolution analysis.

Reconstruction results

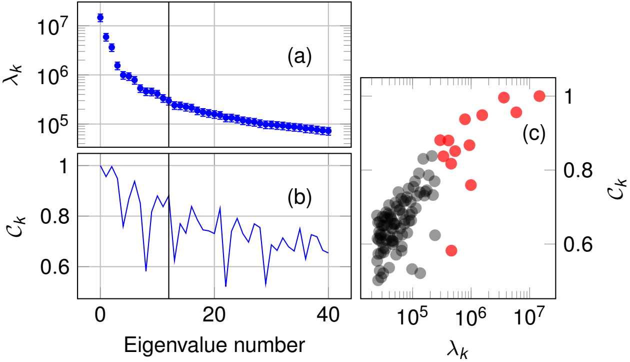




- 13 EOFs modes ; M=225 ; cross-validation data : 1% of observed values.
- Seasonal variation is retrieved, consistent values with the literature (4.5 m/day below the main ice fall in winter).
- Improved accuracy of ≃15m/year compared to the EM-EOF method.

Selection of the optimal number of EOF modes

- Optimal number of EOF modes (13) corresponds to a peak in $\mathcal{C}_k$ which coincides with a break in the eigenvalue spectrum.
- Eigenvalues multiplets are kept in the reconstructed data.



$$\mathcal{C}_k = \frac{\max(\Gamma_k) - \Gamma_k}{\max(\Gamma_k) - \min(\Gamma_k)}$$

$$\Gamma_k = \log\left(\frac{\Delta\lambda_k}{\lambda_j - \lambda_k}\right)$$

Conclusion

- Extension of the EM-EOF method to impute spatio-temporal missing values.
  - Can handle small time series with high incompleteness
  - Extraction of the displacement signal from heterogeneous perturbations (noise)
- Robust selection of the optimal number of EOF modes based on :
  - Iterative computation of the cross-validation error
  - Confidence metric based on eigenvalue uncertainties to address potential over-estimation due to eigenvalue degeneracy
- A range of spatial lag M is provided
- Limitations : potential edge effect due to spatial square window.

Perspective : Use a shaped window (adaptive spatial lag) instead of a square window.

Bibliography

Thank you for your attention.

[1] M. Ghil, M. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. Mann, A. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. Advanced spectral methods for climatic time series. **Review of Geophysics**, 40, 1 :1–41, 2002.

[2] N. Golyandina and K. Usevich. 2d-extension of singular spectrum analysis : Algorithm and elements of theory. **Matrix Methods : Theory, Algorithms and Applications**, pages 449–473, 2010.

[3] A. Hippert-Ferrer, Y. Yan, and P. Bolon. Gap-filling based on iterative EOF analysis of temporal covariance : application to InSAR displacement time series. In **IGARSS**, pages 262–265, 2019.

[4] A. Hippert-Ferrer, Y. Yan, and P. Bolon. EM-EOF : gap-filling in incomplete SAR displacement time series. **IEEE Trans. Geosci. Remote Sens.**, in review, 2020.

[5] R. Millan, J. Mouginot, A. Rabatel, S. Jeong, D. Cusicanqui, A. Derkacheva, and M. Chekki. Mapping surface flow velocity of glaciers at regional scale using a multiple sensors approach. **Remote Sensing**, 11(21), 2019.

This work has been supported by the Programme National de Télédétection Spatiale (PNTS, http ://www.insu.cnrs.fr/pnts), grant PNTS-2019-11, and by the SIRGA project.

Reconstruction averaging

Diagonal averaging, called *hankelization*, [2] is applied successively to each matrix $\boldsymbol{H}_{i,t}$ and to each matrix $\boldsymbol{D}_t$, so that we have the following averaging :

$$x_{ik}(t) = \frac{1}{\#\mathcal{A}_k} \sum_{(l,l')\in\mathcal{A}_k} x_{ll'}(t) \tag{7}$$

$$\boldsymbol{H}_{k,t} = \frac{1}{\#\mathcal{B}_k} \sum_{(l,l')\in\mathcal{B}_k} \boldsymbol{H}_{ll',t} \tag{8}$$

with $\mathcal{A}_k = \{(l,l') : 1 \leq l \leq K_y, 1 \leq l' \leq M_y, l + l' = k + 1\}$ and $\mathcal{B}_k = \{(l,l') : 1 \leq l \leq K_x, 1 \leq l' \leq M_x, l + l' = k + 1\}$.

Confidence index and effective sample size

North's et al. "rule of thumb" (North, 1982) to approximate the eigenvalue uncertainty :

$$\Delta\lambda_k \approx \sqrt{\frac{2}{L^*}}\lambda_k \qquad \Delta \boldsymbol{u}_k \approx \frac{\Delta\lambda_k}{\lambda_j - \lambda_k}\boldsymbol{u}_j \tag{9}$$

with $L^* = N^* M^*$.

- $N^* = N\left[1 + 2\sum_{k=1}^{N-1}(1 - \frac{k}{N})\rho(k)\right]^{-1}$ is the temporal ESS (Thiébaux, 1984)
- M^* is the spatial ESS within each spatial window of size M. We estimate it by :

$$M^* = M\left(1 + 2\sum_{k=1}^{M}(1 - \frac{k}{M})\nu(k)\right)^{-1} \tag{10}$$

Then $\Gamma_k = \log\left(\frac{\Delta\lambda_k}{\lambda_j - \lambda_k}\right)$ and $\mathcal{C}_k$ is computed as :

$$\mathcal{C}_k = \frac{\max(\Gamma_k) - \Gamma_k}{\max(\Gamma_k) - \min(\Gamma_k)} \qquad k = 1, \ldots, NM \tag{11}$$