Gap filling based on EOF analysis of temporal covariance of offset tracking displacement measurement time series

Alexandre Hippert-Ferrer\textsuperscript{1}, Yajing Yan\textsuperscript{1}, Philippe Bolon\textsuperscript{1}

\textsuperscript{1}Laboratoire d’Informatique, Systèmes, Traitement de l’Information et la Connaissance

13/09/2018
Contents

1. Context
2. Methodology
3. Synthetic simulations
4. Real data application
Regular availability of satellite images (e.g. Sentinel data)
Different displacement extraction techniques: (D-)InSAR, offset tracking...

Presence of missing data in displacement fields in space and time

Causes: image quality, missing image, possible limitations of displacement extraction techniques.
Motivation

- Inverse distance weighting (IDW), simple kriging (SK), regression problems
- Spatial interpolation only
- It is necessary to make use of temporal information

How can we manage the presence of spatio-temporal missing data in time series?

→ We propose a statistical method based on spatio-temporal characteristics of the displacement time series.
The EM-EOF method

- Based on Empirical Orthogonal Functions (EOFs) which allow a signal representation in terms of orthogonal functions which describe a temporal or spatial mode of variation (comes from PCA), see (Hannachi et al. 2007);
- Based on analyses of the temporal or spatio-temporal covariance of a time series of displacement measurement;
- Iterative processing with a first initialization of missing values by hand before EOF analysis.

→ Allows for a spatio-temporal interpolation
Workflow

- Initialization of missing values
- EOF analysis
- Reconstruction
- Reconstruction error
- Adjust number of EOFs
Data representation

Let $X(s, t)$ be a spatio-temporal field containing the values of $X$ at position $s$ and time $t$:

$$X = (x_1, x_2, \ldots x_n) = \begin{pmatrix} 
  x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\
  x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\
  x_{31} & x_{32} & x_{33} & \cdots & x_{3n} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  x_{p1} & x_{p2} & x_{p3} & \cdots & x_{pn} 
\end{pmatrix} \quad (1)$$

$(x_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$ is the value at position $s_i$ and time $t_j$ and may be missing.

We then initialize the missing values by a defined quantity.
Reconstruction using EOFs

- **Estimation of the temporal covariance**:

\[
C = \frac{1}{p-1} X^t X
\]

(2)

- Then we resolve:

\[
CS = SD
\]

(3)

→ $S$ contains the self-orthogonal eigenvectors $(s_i)_{0 \leq i \leq p}$ called EOFs.

- We reconstruct $X$ with $M$ number of EOFs:

\[
X = \sum_{i=1}^{r} a_i s_i^t \rightarrow \hat{X} = \sum_{i=1}^{M \ll r} a_i s_i^t
\]

(4)

with $a_i = Xs_i$ and $r \leq \text{min}(p, n)$. 

Key parameters

1. **Missing data**
   - % of missing values
   - Type: random, correlated

2. **Noise**
   - Signal-to-Noise Ratio (SNR)
   - Type: white Gaussian, correlated noise

3. **Missing data initialization**
   - Spatial mean
   - Spatial mean + noise
## Synthetic data

Different spatio-temporal fields are generated for the experience:

\[ f(r, t) = (1 - 2r)t + \sin\left(\frac{\pi}{2} t\right) \cos\left(\frac{\pi}{2} r\right) + 0.5 \cos\left(\frac{3\pi}{2} t\right) \cos(5\pi r) + 0.3 \sin\left(\frac{7\pi}{2} t\right) \sin(15\pi r) \]

<table>
<thead>
<tr>
<th>Order</th>
<th>1st order</th>
<th>2nd order</th>
<th>3rd order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="1st order" /></td>
<td><img src="image2.png" alt="2nd order" /></td>
<td><img src="image3.png" alt="3rd order" /></td>
</tr>
</tbody>
</table>
Some results on synthetic data

1 EOF - 15% gaps - SNR=3.18 - 27 iterations

1 EOF - 50% gaps - SNR=3.27 - 816 iterations
Some results on synthetic data

2 EOF - 30% gaps - SNR=2.24 - 203 iterations

2 EOF - 50% gaps - SNR=2.29 - 103 iterations
Some results on synthetic data: worst cases

1 EOF - 70% gaps - SNR=0.52 - 72 iterations

1 EOF - 70% gaps - SNR=0.52 - 326 iterations
Reconstruction error: RMSE vs. number of EOFs

- $\min\text{(RMSE)}$ is reached at $M \approx$ signal order
- $M$ greatly depends on missing data type and less on missing data initialization.
Reconstruction error

\[ \min \text{(RMSE)} \text{ for } \% \text{ of gaps vs. SNR with noise initialization} : \]

- **Best case**: high SNR (>3) and few missing values (<10%)
- The impact of SNR is more important than the % of missing values.
Time series of displacement fields

- 65 displacement fields of Mont-Blanc glaciers computed by *offset tracking* of Sentinel 1 SAR images
- Time period: September 2016 to December 2017
- Presence of missing data in space and time
Preliminary results

12 days-displacement [m] in radar LOS (5-30% gaps, 8 EOF, RMSE ≤10%)
Conclusion

1. EM-EOF method seems promising according to synthetic simulations:
   - The proposed method is robust against % of gaps (reconstruction still working for \( \geq 50\% \) gaps);
   - The reconstruction performs better with correlated gaps initialized with noise;
   - The reconstruction is more demanding in case of correlated noise;
   - The impact of the SNR is more significant than the quantity of gaps.

2. Satisfactory preliminary results over Argentiere glacier obtained;

3. Further investigation and validation by InSAR data in future work.
Questions ?

References :


Reconstruction error

Crossed validation reconstruction error:

\[
rmse = \frac{1}{N} \sqrt{\sum_{i=1}^{N} |x_r - x|^2}
\]  

(5)

where \(x_r\) and \(x\) are vectors containing \(N\) points randomly chosen in each image before reconstruction.
Singular Spectrum Analysis

Data representation:

\[ x_t = (x_t, x_{t+1}, \ldots, x_{t+M-1}) \]  \hspace{1cm} (6)

\[ \mathbf{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_M \\ x_2 & x_3 & \cdots & x_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-M+1} & x_{n-M+2} & \cdots & x_n \end{pmatrix} \]  \hspace{1cm} (7)

"Grand" covariance matrix:

\[ C = \frac{1}{n-M+1} \mathbf{x}^T \mathbf{x} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ C_{21} & C_{22} & \cdots & C_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1} & C_{p2} & \cdots & C_{pp} \end{pmatrix} \]  \hspace{1cm} (8)

where \((C_{ij})_{1 \leq i,j \leq p}\) is a delayed covariance matrix between points \(i\) and \(j\)

\[ C_{ij} = \frac{1}{n-M+1} \sum_{t=1}^{n-M+1} x_t^i x_t^j \]  \hspace{1cm} (9)
Singular Spectrum Analysis

Calcul des EOFs étendus $v_k$ :

\[ x = U\Theta V^t \]  \hspace{1cm} (10)

\[ x^T_t = \sum_{k=1}^{d} \theta_k u_k v_k \]  \hspace{1cm} (11)

Reconstruction du signal :

\[ RCs(t) = \begin{cases} 
\frac{1}{t} \sum_{j=1}^{t} \sum_{K} \theta_k v_k(t - j + 1)u_k, & 1 \leq t \leq M - 1 \\
\frac{1}{M} \sum_{j=1}^{M} \sum_{K} \theta_k v_k(t - j + 1)u_k, & M \leq t \leq n - M + 1 \\
\frac{1}{n-t+1} \sum_{j=t-n+M}^{M} \sum_{K} \theta_k v_k(t - j + 1)u_k, & n - M + 2 \leq t \leq n
\end{cases} \]
Space correlated noise

A partir d’une fonction d’auto-corrélation de type $c(r) = r^{-\beta}$ et d’une image de bruit blanc $b$ :

1. Calcul de la densité spectrale de puissance de $c$ : $\Gamma(c) = |\mathcal{F}\{c\}|$
2. Calcul de la TF de $b$ : $\mathcal{F}\{b\}$
3. Filtrage fréquentiel : $\mathcal{F}\{b\}\Gamma(c)$
4. Calcul de $\mathcal{F}^{-1}\{\mathcal{F}\{b\}\Gamma(c)\}$
Variance expliquée par les $M$ premiers modes :

$$\frac{100 \sum_{k=1}^{M} \lambda_k}{\sum_{k=1}^{r} \lambda_k} \%$$ (12)
min(rmse) vs. image number