A non-Gaussian decomposition of Total Water Storage (TWS), using Independent Component Analysis (ICA)

Ehsan Forootan and Jürgen Kusche
Astronomical Physical & Mathematical Geodesy, Bonn University
forootan@geod.uni-bonn.de

Geodätische Woche 2010, 07.10.2010
• Introduction & Motivation

• Method of EOF/PCA

• Rotation of PCA components towards a simplified structure

• Incorporate non-Gaussianity information in the frame of ICA

• Results & Discussion
1. Since 2002, GRACE has provided a valuable information about mass redistribution within the Earth system.

2. TWS anomalies represent integrated mass over a vertical column which is caused by different phenomena.

**Task**: to decompose the observed signals to its geophysically interpretable components.

Time series of the Total Water Storage (TWS) maps, derived from the process of ITG2010 solutions

\[ TWS = F(t, s) \begin{bmatrix} f_1, f_2, \ldots, f_m \end{bmatrix} \]

where:
- \( TWS \) represents the Total Water Storage
- \( F(t, s) \) is the function of time and solution number
- \( n \times m \) represents the number of solutions and number of grid points
PCA is the most widely used method which works based on eigenvalue decomposition (Lorenz, 1956).

- **PCA criterion**
  \[ \text{Max} \left( \sum_{i=1}^{n} e_{m} f_{i} \right)^{2} \]

- **Constrain**
  \[ e_{i} e_{j}^{T} = \delta_{ij} \]

- **Diagonalization of the auto-covariance matrix**
  \[ F(t, s)_{n \times m} = \sum_{k=1}^{m} \text{PCs}_{k}(t) \text{EOF}_{k}(s) \]

- **PCA de-correlates the dataset by decomposing it to the orthogonal components.**

- **Limitation**: Derived components are data modes and are not always physically interpretable.

- **Limitation**: Physical processes are not necessarily orthogonal. (Orthogonality problem)

- **PCA solution is optimized when the observed signals are Gaussian.**

- **Limitation**: Hydrological parameters associated to physical process models contain a significant level of non-Gaussianity. (Gaussianity problem)
Method of REOF

- **Rotated EOF**: is a solution for the PCA’s orthogonality constraint.
  - REOF technique is simply based on rotating the PCA’s components.
    - either rotating EOFs or PCs
    - orthogonal or oblique rotation
  - Rotation criteria are various, namely: VARIMAX, QUARTIMAX, QUARTIMIN,…

→ Varimax rotation, to derive simplified structure (Kaiser, 1958)
  - Choosing a subset of EOFs for rotation:
    
    Rotation of EOFs: \( U = E_k R \)

    Varimax criterion → \( f(U) = \text{Max}\left( \sum_{j=1}^{k} \left[ \sum_{i=1}^{p} u_{ij}^4 - \frac{1}{p} \left( \sum_{i=1}^{p} u_{ij}^2 \right)^2 \right] \right) \)

  - Varimax simplifies the structure in the data, (Richman, 1986).
• Non-Gaussian signals have non-zero high order statistical moments.

> Kurtosis: \( \frac{E(x^4)}{E(x^2)^2} - 3 \begin{cases} 
0 & \rightarrow \text{Gaussian signals} \\
< 0 & \rightarrow \text{Sub-Gaussian} \\
> 0 & \rightarrow \text{Super-Gaussian}
\end{cases} \)

> Around 60% of the TWS signals are non-Gaussian.

> This suggests to incorporate higher order moments within the decomposition procedure.

• Independence instead of uncorrelation:

> Independence implies uncorrelatedness but the reverse is not always true!

• For non-Gaussian signals, maximally independent signals are also likely approximately uncorrelated.

→ ICA algorithm: 1- perform the PCA to decorrelate, 2- to determine a suitable rotation to optimize an independence criterion.
Method of ICA

• Selected criterion: fourth order cumulant:  if \( \bar{x} = E(x) \)

\[
C(x_1, x_2, x_3, x_4) = E(\bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_4) - E(\bar{x}_1 \bar{x}_2)E(\bar{x}_3 \bar{x}_4)
\]
\[
- E(\bar{x}_1 \bar{x}_3)E(\bar{x}_2 \bar{x}_4) - E(\bar{x}_1 \bar{x}_4)E(\bar{x}_2 \bar{x}_3)
\]

• Similar to the JADE algorithm, (Cardoso, 1993)

Rotation of EOFs : \( U = E_k R \)

ICA criterion \( \rightarrow f(U) = \text{Max} \left( \sum_{j=1}^{k} C(u_j)^2 \right) \)

• Rotating EOFs \( \rightarrow \) Spatial ICA
• Temporal ICA \( \rightarrow \) Temporal ICA

• The ICA derived decomposition are more physically interpretable.
Simulation status

Signal:

Noise:

Jan 2003

July 2009

Temporal component 1

Temporal component 2
Simulation results
Results of decomposing ITG2010

PCA

Varimax REOF

07.10.2010
Results of decomposing ITG2010

Temporal ICA

Spatial ICA
Results & Discussion

• The EOF/PCA breaks the data into modes of variability, these modes are in nature data modes, and are not necessarily interpretable.

• The PCA's conventional extension (e.g. Varimax rotation) did not improve the mixing problem.

• The PCA method can be assumed as an initial step for the ICA method.
  ➢ This improves both the computational and interpretability of the decomposition procedure.

• Using the non-Gaussianinty information in the frame of ICA shows a better performance with compare to the ordinary PCA and Varimax to separate the GRACE's underlying signals even for such regions that exhibit different overlapping modes.
Thank you for your attention

- References:

