Data-driven Approaches for Ocean Remote Sensing:
from the Non-negative Decomposition of Operators
to the Reconstruction of Satellite-derived Sea Surface Dynamics

M. Lopez Radcenco, A. Pascual, L. Gomez-Navarro, P. Ailliot, A. Aissa-El-Bey, R. Fablet

SpatioTempMeteo Workshop
28-30 Nov. 2018. Rennes, France
CONTEXT AND MOTIVATION
**Data sources:** A great variety of heterogeneous sensors and models

**Key issue:** Exploit these data sources to improve our understanding of geophysical processes
Ocean Remote Sensing: What’s at stake?

Operational:
Monitoring and forecasting of ocean dynamics
► High socio-economical impact
► Management of ocean resources
► Environmental protection

Scientific:
Better understanding of ocean processes
► Model calibration and validation
► Estimation of geophysical quantities and parameters
► Improve model prediction/reconstruction

However:
► Strong reliance on physical models
► *Datasets: under-exploited potential?*

**FIGURE 1:** Remote sensing satellites
Problems of interest

Characterization and decomposition operators

\[ y_n = \sum_{k=1}^{K} \alpha_{nk} f_{\theta_k}(x_n) + \omega_n \]

Interpolation of altimetry fields from satellite data
1. NON-NEGATIVE DECOMPOSITION OF OPERATORS
   ► Models and Algorithms
   ► Applications
     - Forecasting of Dynamical Systems
     - Segmentation of Sea Surface Dynamics

2. INTERPOLATION OF SEA LEVEL ANOMALIES FROM SATELLITE DATA
   ► Problem formulation
   ► Observation System Simulation Experiment
   ► Results

3. VALORIZACION

4. CONCLUSION AND PERSPECTIVES
PART 1: NON-NEGATIVE DECOMPOSITION OF OPERATORS
Operator decomposition

Source separation:

\[ y_n = \sum_{k=1}^{K} \alpha_{nk} s_k + \omega_n \]

s. t. \( C(\alpha_{n1}, \ldots, \alpha_{nK}, s_1, \ldots, s_K) \)

Operator decomposition:

\[ y_n = \sum_{k=1}^{K} \alpha_{nk} f_{\theta_k} (x_n) + \omega_n \]

s. t. \( C(\alpha_{n1}, \ldots, \alpha_{nK}, f_{\theta_1}, \ldots, f_{\theta_K}) \)
Data-driven approaches for ocean remote sensing - López Radcenco et al.

**Decomposition of Geophysical Processes**

**Joint EOFs:**  
[Hotelling, 1933; Hannachi et al., 2007]

- Widely used in oceanography and atmospheric sciences  
- **Orthogonal** decomposition of cross-correlation  
- Find orthogonal modes that maximize explained variance

\[
C = \frac{1}{N} Y_1^T Y_2 = V \Lambda U^T = \sum_{r=1}^{R} \lambda_r v_r u_r^T
\]

**Dynamic Mode Decomposition:**  
[Schmid, 2010]

- Linearization + Eigendecomposition  
- **Choice of observation functions** \( g(x_n) \)  
- **Time invariance of operator** \( K \)

\[
x_{n+1} = f(x_n) \quad g(x_n) = Kg(x_{n-1})
\]

**FIGURE 2: DMD**
Decomposition of Geophysical Processes

Classically:
- Orthogonal decompositions (EOF, PCA)
- \( \ell_2 \)-norm penalization
- Strong hypotheses

But, no guarantees in terms of:
- Relevance
- Interpretability

Explore new decompositions

Blind source separation:
- New formulations
  - Non-negativity
  - Sparsity

\[
\mathbf{y}_n = \sum_{k=1}^{K} \alpha_{nk} \mathbf{s}_k + \mathbf{\omega}_n
\]
\[
s. \ t. \ C (\alpha_{n1}, \ldots, \alpha_{nK}, \mathbf{s}_1, \ldots, \mathbf{s}_K)
\]

\[
\mathbf{y}_n = \sum_{k=1}^{K} \alpha_{nk} f_{\theta_k} (\mathbf{x}_n) + \mathbf{\omega}_n
\]
\[
s. \ t. \ C (\alpha_{n1}, \ldots, \alpha_{nK}, f_{\theta_1}, \ldots, f_{\theta_K})
\]

Linear functions:
\( \mathbf{\beta}_k \)
MODELS AND ALGORITHMS
Relationship between observable variables $x_n$ and $y_n$

- Non-negative superposition of linear modes

\[ y_n = \sum_{k=1}^{K} \alpha_{nk} \beta_k x_n + \omega_n \]

Subject to
\[
\begin{align*}
\alpha_{nk} &\geq 0, & \forall k \in [1, K], \forall n \in [1, N] \\
\|\beta_k\|_F &= 1, & \forall k \in [1, K]
\end{align*}
\]

FIGURE 3: General model block diagram
Model calibration

\[ \forall n, \begin{cases} \left[ \hat{\alpha}_{nk}, \hat{\beta}_k \right] = \arg\min_{\alpha_{nk}, \beta_k} \sum_{m=1}^{N} W_{mn} \left( \left\| y_m - \sum_{k=1}^{K} \alpha_{nk} \beta_k x_m \right\|_F^2 \right) \\ \alpha_{nk} \geq 0, \quad \forall k \in [1, K] \\ \| \beta_k \|_F = 1, \quad \forall k \in [1, K] \end{cases} \]

Use multiple observation pairs for parameter estimation:

- Weights \( W_{mn} \) for considered observation pairs
  - Similarity to the current observation

- Non-linear
- Non-convex

Partial minimization over one set of parameters: \textbf{linear} and \textbf{convex}
**Alternating Least Squares (ALS) algorithm:** Alternate partial minimizations until convergence

- **\( \beta \) update**
  - Estimate \( \beta_k \) with fixed \( \alpha_{nk} \)
  - Normalization constraint

- **\( \alpha \) update**
  - Estimate \( \alpha_{nk} \) with fixed \( \beta_k \)
  - Non-negativity constraint

**Local linear operator decomposition:**

\[
\Theta_n \approx y_n \approx \Theta_n x_n
\]

\[
\Theta_n = \left( \sum_{m=1}^{N} W_n^m y_m x_m^T \right) \left( \sum_{m=1}^{N} W_n^m x_m x_m^T \right)^{-1}
\]

**Dictionary learning:**

- Wide variety of algorithms exist [Lee & Seung, 1999; Aharon et al, 2007; Hoyer, 2004]
- Constraints can be changed easily
Algorithm synthesis and comparison

Dictionary-based decomposition of local linear operators

- Best performance under favorable settings

ALS-direct/ALS-gradient:
- Less stable
- May prove useful under non-ideal settings

**FIGURE 4:** Distribution of nMSE for the estimation of model parameters
APPLICATIONS
Decomposition of upper ocean dynamics

\[ x_n = \mathcal{F}_T (\text{SST}) \quad y(s_i, t_i) = \sum_k \alpha_k(s_i, t_i) \beta_k x(s_i, t_i) \]
\[ y_n = \mathcal{F}_H (\text{SSH}) \]

Decomposition of forecasting operators

\[ s(t + \partial t) = A(s(t)) s(t) \quad \rightarrow \quad A(s(t)) = \sum_{k=1}^{K} \alpha_k(s(t)) \beta_k \]
Analog forecasting of Lorenz '96 dynamics
Case study: Lorenz ‘96 dynamical system

\[ \frac{\partial s_i}{\partial t} = (s_{i+1} - s_{i-2}) s_{i-1} - s_i + F \quad \forall i \in 1, L \]

- Developed to study predictability issues in weather forecasting
- Representative of chaotic geophysical dynamical systems (e.g. the atmosphere)

Forecasting:
- For a 40-variable Lorenz ‘96 time series \( s(t) \)

\[ s(t + \partial t) = A(s(t)) s(t) \]

**FIGURE 5:** Lorenz ‘96 model
Analog forecasting of Lorenz '96 dynamics

**FIGURE 6:** Analog forecasting

\[
\begin{aligned}
\mathbf{x}_n &= s(t) \\
\mathbf{y}_n &= s(t + \partial t)
\end{aligned}
\]

\[
\mathbf{y}_n = \Theta_n \mathbf{y}_n
\]
Analog forecasting of Lorenz ’96 dynamics

**FIGURE 7:** Non-negative decomposition of analog forecasting operators

\[ A(s(t)) = \sum_{k=1}^{K} \alpha_k(s(t)) \beta_k \]
Analog forecasting of Lorenz ’96 dynamics

FIGURE 8: RMSE vs number of analogs for classic analog forecasting and non-negative decomposition of the classic analog forecasting operator using $K=4$ modes.

(a) $N = 2 \times 10^5$
No noise

(b) $N = 2 \times 10^5$
$\sigma^2_{noise} = 0.1$

(a) $N = 2 \times 10^3$
No noise
Segmentation of SST/SSS Sea Surface Dynamics
Case study in the Alboran Sea:

- Daily WMOP synthetic anomaly images (2009-2012)
  - Sea surface temperature (SST)
  - Sea surface salinity (SSS)
- Cold water intake from the Atlantic through the Gibraltar Strait
- Strong seasonal patterns
- Inversion of SST/SSS correlation

**FIGURE 9:** SST/SSS fields on March 22nd, 2011
Segmentation of SST/SSS Sea Surface Dynamics

\[ SSS(t_n, p) = \Theta(t_n)vec(P_{SST}(t_n, p)), \forall p \]

\[ \Theta(t_n) = \sum_{k=1}^{K} \alpha_k(t_n)\beta_k \]

\textbf{Mode 1: inversion of correlation between SST/SSS fields}

\textbf{Mode 2: coherent SST/SSS relationship}

FIGURE 10: Modal SSS field predictions from the SST field on March 22\textsuperscript{nd}, 2011
FIGURE 11: Distribution of SST/SSS correlation coefficients
PART 2: INTERPOLATION OF SEA LEVEL ANOMALY FIELDS FROM SATELLITE-DERIVED REMOTE SENSING DATA
Interpolation of altimetry fields

Key issue in oceanography

**Main difficulty:** Irregular and partial sampling of the ocean surface
- Multiple data sources
  - Different spatio-temporal sampling strategies
- Missing data: up to 90%

**Current limitation:** Scales <100 km not accurately reconstructed
Beyond the 100 km limit

Model-based approaches:

*State-of-the-art: Optimal Interpolation* [Bretherton et al., 1976]
- Gaussianity
- Mean spatio-temporal covariance structures

*New approaches: Additional physically-motivated constraints*
- OI+Bathymetry [Escudier et al., 2013]
- Dynamic interpolation [Ubelmann et al., 2014]

New developments in instrumentation:

*SWOT* [Fu & Ferrari, 2008]

*SKIM* [Ardhuin et al., 2018]

We focus on exploring data-driven alternatives
Problem Formulation
Interpolation of altimetry fields

**State-space formulation**

\[
\begin{align*}
\mathbf{x}(t) &= \mathcal{M}(\mathbf{x}(t - \delta t)) + \epsilon(t) \\
\mathbf{y}(t) &= \mathcal{H}(\mathbf{x}(t), \Omega(t)) + \eta(t)
\end{align*}
\]

Dynamical model
Observation model

**Patch-based representation**

Interpolate each patch independently

**Multi-scale approach**

\[
\mathbf{x} = \mathbf{\bar{x}} + d\mathbf{x} + \zeta
\]

Scales >100 km
Resolved by OI
Analog data assimilation

Observation Successors Analogs Catalog

KNN search Weighted lin. Reg.

$x(t)$ $x^p(t + dt)$ $y(t)$ AnEnKF (Assimilation) $x^a(t + dt)$
Observing system simulation experiment
Satellite altimetry observations

**FIGURE 13:** SWOT satellite

**FIGURE 14:** Difference between AT and SWOT altimetry observations
Case study in the Western Mediterranean Sea, from 2010 to 2014

- Synthetic SLA (Sea Level Anomaly) fields simulated with the WMOP model
- Along-track observations: real satellite tracks (4 altimeters in 2014)
- Pseudo-SWOT observations: SWOT simulator (JPL-NASA)

Tested methods: Optimal Interpolation, DINEOF, Non-negative decomposition of linear interpolation operators, **Analog data assimilation**
**Pseudo-observations**: Observations accumulated on a $t_0 \pm D$ time window

**FIGURE 16**: Pseudo-observations obtained from along-track (AT) and SWOT data
Results
Results

**TABLE 1:** Root mean squared error (Correlation) for AnDA SLA and SLA gradient (rSLA) reconstruction from nadir along-track observations ($AT_D$) and wide-swath SWOT observations ($SWOT_D$). For each type of observations, both daily observations ($D=0$) and observations accumulated on a time window $t_0 \pm D$ with $D=5$ days are considered. Best result in **bold**.

<table>
<thead>
<tr>
<th>Setting</th>
<th>SLA</th>
<th>∇SLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$AT_0$</td>
<td>0.02395 (0.9186)</td>
<td>0.005507 (0.6989)</td>
</tr>
<tr>
<td>$AT_5$</td>
<td>0.01978 (0.9457)</td>
<td>0.004699 (0.7660)</td>
</tr>
<tr>
<td>$SWOT_0$</td>
<td><strong>0.01810 (0.9543)</strong></td>
<td><strong>0.004436 (0.7857)</strong></td>
</tr>
<tr>
<td>$SWOT_5$</td>
<td>0.01920 (0.9502)</td>
<td>0.004345 (0.7913)</td>
</tr>
<tr>
<td>OI</td>
<td>0.02927 (0.8451)</td>
<td>0.006655 (0.6052)</td>
</tr>
</tbody>
</table>
### TABLE 2: Root mean squared error (Correlation) for AnDA SLA and SLA gradient (rSLA) reconstruction from the fusion of nadir along-track observations ($AT_D$) and wide-swath SWOT observations ($SWOT_D$). For each type of observations, both daily observations ($D=0$) and observations accumulated on a time window $t_0 \pm D$ with $D=5$ days are considered. Best result in **bold**.

<table>
<thead>
<tr>
<th>Setting</th>
<th>SLA</th>
<th>∇SLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$AT_0 + SWOT_0$</td>
<td>0.01742 (0.9576)</td>
<td>0.004375 (0.7934)</td>
</tr>
<tr>
<td>$AT_5 + SWOT_5$</td>
<td>0.01876 (0.9523)</td>
<td>0.004318 (0.7952)</td>
</tr>
<tr>
<td>$AT_5 + SWOT_0$</td>
<td><strong>0.01687 (0.9607)</strong></td>
<td><strong>0.004286 (0.8051)</strong></td>
</tr>
<tr>
<td>OI</td>
<td>0.02927 (0.8451)</td>
<td>0.006655 (0.6052)</td>
</tr>
</tbody>
</table>

With respect to OI:

- **42% (14%) improvement** in terms of RMSE (correlation) for $SLA$
- **35% (33%) improvement** in terms of RMSE (correlation) for $∇SLA$
FIGURE 17: SLA fields interpolation results for Optimal Interpolation and for the AnDA assimilation of $AT_5$, $SWOT_0$ and $SWOT_0 + AT_5$ data. Ground-truth fields and observations included as reference.
FIGURE 18: ∇SLA fields interpolation results for Optimal Interpolation and for the AnDA assimilation of $AT_5$, $SWOT_0$ and $SWOT_0 + AT_5$ data. Ground-truth fields and observations included as reference.
VALORIZATION
5 International Conference papers
- M. Lopez-Radcenco, A. Aissa-El-Bey, P. Ailliot, and R. Fablet. Non-negative decomposition of geophysical dynamics. In *ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, Belgium, April 2017

2 National Conference papers

2 Journal papers
CONCLUSION AND PERSPECTIVES
Conclusions

Explore data-driven approaches for ocean remote sensing

**Non-negative decomposition of operators:**
- Relevant models
- Efficient and mathematically-sound algorithms
- Relevant applications in various scientific contexts
  - Segmentation of upper ocean dynamics from satellite data
  - Analog forecasting of dynamical systems

**Interpolation of SLA fields from satellite data:**
- Different sampling patterns: SWOT mission
- Data-driven fusion of AT and SWOT observations:
  - Clear performance gain from the fusion of AT and SWOT observations
Perspectives

Non-negative decomposition of operators:
► Further improve robustness of models
► Explore alternative constraints (sparsity)
► Explore non-linear model extensions
► Further study geophysical interpretation of model parameters
► Explore new applications (or extend previous ones)

Interpolation of SLA fields from satellite data:
► Filtering SWOT noise: key issue
  - Combine AnDA with current efforts to pre-process SWOT noise
► Complementary sources of altimetry data or alternative dynamical tracers (SST, SSS, etc.)
► Efficient exploitation of 2D information in SWOT:
  - Observation gradients
  - Finite size Lyapunov exponents
Thank you for your attention

That's all Folks!
It's QUESTION TIME!!