Décomposition Non-négative de Dynamiques Géophysiques

Manuel Lopez Radcenco, Abdeldjalil Aissa El Bey, Pierre Ailliot, Ronan Fablet

▶ **To cite this version:**

Manuel Lopez Radcenco, Abdeldjalil Aissa El Bey, Pierre Ailliot, Ronan Fablet. Décomposition Non-négative de Dynamiques Géophysiques. GRETSI 2017 : 26ème colloque du Groupement de Recherche en Traitement du Signal et des Images, Sep 2017, Juan-Les-Pins, France. hal-01611337

HAL Id: hal-01611337
https://hal.science/hal-01611337

Submitted on 4 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Décomposition Non-négative de Dynamiques Géophysiques

Manuel LÓPEZ-RADCENCO[1], Abdeldjalil AÏSSA-EL-BEY[1], Pierre AILLIOT[1,2], Ronan FABLET[1] *

[1] IMT Atlantique, UMR CNRS 6285 Lab-STICC, Université Bretagne Loire,
Technopôle Brest-Iroise CS83818, 29238 Brest Cedex 3, France

[2]Université de Brest, UMR CNRS 6205 LMBA,
6, Avenue Victor Le Gorgeu, B.P. 809, 29285 Brest Cedex, France

manuel.lopezradcenco@imt-atlantique.fr, abdeldjalil.aissaelbey@imt-atlantique.fr,
pierre.ailliot@univ-brest.fr, ronan.fablet@imt-atlantique.fr

Résumé – La décomposition des processus géophysiques en modes pertinents est un point clé pour les problèmes de caractérisation, prédiction et reconstruction dans le domaine des sciences de l'environnement. Par ailleurs, la séparation aveugle des contributions associées à différentes sources est un problème classique dans le domaine du traitement de signal et des images. Récemment, des progrès significatifs ont été obtenus avec l'introduction de formulations non-négatives et parcimonieuses. Dans ce travail, nous abordons la décomposition aveugle d'opérateurs linéaires ou de fonctions de transfert entre variables d'intérêt, en mettant l'accent sur un cadre non-négatif. Nous illustrons l'intérêt de ces décompositions pour l'analyse, la prédiction et la reconstruction de dynamiques géophysiques.

Abstract – The decomposition of geophysical processes into relevant modes is a key issue for characterization, forecasting and reconstruction problems in environmental sciences. Moreover, the blind separation of contributions associated with different sources is a classical problem in signal and image processing. Recently, significant advances have been reported with the introduction of non-negative and sparse formulations. In this work, we address the blind decomposition of linear operators or transfer functions between variables of interest, with an emphasis on a non-negative setting. We illustrate the relevance of these decompositions for the analysis, prediction and reconstruction of geophysical dynamics.

1 Introduction

La décomposition de processus géophysiques en modes pertinents et interprétables, notamment par des approches de décomposition orthogonale telles que l'analyse en composantes principales ou les fonctions orthogonales empiriques (PCA ou EOF, par ses sigles en anglais) [1], a été largement utilisée dans des applications d'analyse et de prédiction en océanographie et en météorologie. Outre les décompositions basées sur des EOFs, la séparation aveugle des contributions associées à différentes sources ou processus a été largement étudiée dans le domaine du traitement du signal et des images [2]. Dans ce contexte, des progrès significatifs ont été obtenus par l'introduction de formulations non-négatives [3] et parcimonieuses [4]. Ici, nous abordons la décomposition aveugle d'opérateurs ou de fonctions de transfert variant dans l'espace ou dans le temps. Les applications potentielles comprennent la décomposition de processus dynamiques ainsi que l'analyse des relations entre différentes variables d'intérêt (voir la Section 3 pour des illustrations). Dans un travail précédent [5], nous avons introduit une formulation non-négative qui permet la caractérisation des relations entre deux processus par la superposition de plusieurs relations linéaires. Nous généralisons cette approche et nous abordons ici une décomposition basée sur l'apprentissage de dictionnaires d'opérateurs linéaires locaux. Cet article est organisé comme suit. La Section 2 présente le modèle proposé et les algorithmes associés. Dans la Section 3, nous ilustrons des applications à l'analyse des dynamiques géophysiques. Enfin, nos conclusions et perspectives de travail sont présentées dans la Section 4.

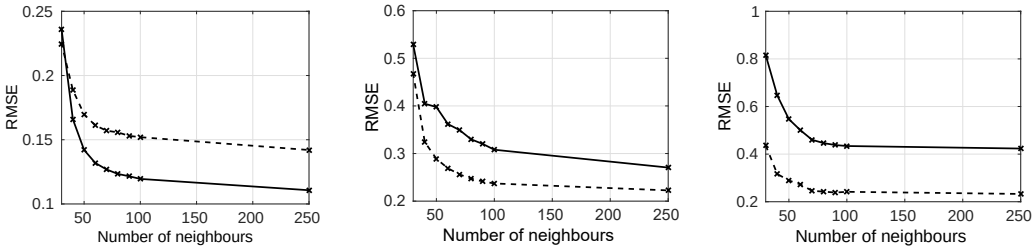
*Ce travail est partiellement financé par le projet SEACS Cominlabs, le projet OSTST MANATEE et la région Bretagne.

2 Formulation du modèle

Considérons un ensemble de données d'observations multivariées $\{\mathbf{x}, \mathbf{y}\}_n$, où $\mathbf{x}_n \in \mathbb{R}^J$, $\mathbf{y}_n \in \mathbb{R}^I$ désignent le $n^{ème}$ couple d'observations. En supposant qu'un opérateur linéaire lie $\mathbf{x}_n$ à $\mathbf{y}_n$, notre objectif est de décomposer cet opérateur linéaire comme la superposition non-négative de plusieurs relations linéaires. Suivant notre travail précédent [5], nous considérons le modèle suivant pour la relation entre les variables $\mathbf{x}_n$ et $\mathbf{y}_n$:

$$\begin{aligned} \mathbf{y}_n &= \sum_{k=1}^{K} \alpha_{nk} \boldsymbol{\beta}_k \mathbf{x}_n + \boldsymbol{\omega}_n \\ \text{s.c. } &\begin{cases} \alpha_{nk} \geq 0, & \forall\, k \in [\![1, K]\!],\, \forall n \in [\![1, N]\!] \\ ||\boldsymbol{\beta}_k||_F = 1, & \forall\, k \in [\![1, K]\!] \end{cases} \end{aligned} \tag{1}$$

où $\alpha_{nk} \in \mathbb{R}^+$ sont des coefficients de mélange non-négatifs qui modélisent la contribution de chaque mode linéaire à la



(a) $N = 2 \times 10^5$. Sans bruit (b) $N = 2 \times 10^5$. $\sigma^2_{noise} = 0.1$ (c) $N = 2 \times 10^3$. Sans bruit

FIGURE 1 – Performance de prédiction des dynamiques Lorenz '96 par des opérateurs analogues linéaires : nous représentons l'erreur quadratique moyenne de prédiction normalisée (RMSE) en fonction du nombre d'analogues utilisés pour estimer l'opérateur localement linéaire de prédiction de la dynamique de Lorenz '96, pour les opérateurs localement linéaires classiques (*ligne complète*) et pour la décomposition non-négative proposée pour des opérateurs localement linéaires avec $K = 4$ modes (*ligne pointillée*). Nous considérons trois scénarios expérimentaux : un scénario sans bruit avec une base d'apprentissage de 2×10^5 exemplaires, un scénario avec une base d'apprentissage bruitée de 2×10^5 exemplaires, un scénario sans bruit avec une base d'apprentissage de 2×10^3 exemplaires.

reconstruction de $\mathbf{y}_n$ à partir de $\mathbf{x}_n$, $\boldsymbol{\beta}_k \in \mathbb{R}^{I \times J}$ est la matrice de régression associée au mode k, $||\cdot||_F$ est la norme de Frobenius et $\boldsymbol{\omega}_n \in \mathbb{R}^I$ est un processus de bruit, généralement un bruit Gaussien. N et K désignent, respectivement, le nombre total d'observations et de modes. Dans [5], nous nous concentrons sur les cas d'étude où I, le nombre de dimensions de la variable $\mathbf{y}_n$, est assez grand pour garantir l'identifiabilité de la décomposition ($I > 1$). Nous étendons ici le modèle (1) sans contrainte spécifique sur I et proposons une approche de décomposition d'opérateurs linéaires par apprentissage de dictionnaires. Désignons par $\boldsymbol{\Theta}_n$ une estimation de l'opérateur linéaire reliant $\mathbf{x}_n$ et $\mathbf{y}_n$ (différentes stratégies d'estimation de $\boldsymbol{\Theta}_n$ sont présentées en Section 3). Le modèle (1) revient à une décomposition aveugle des opérateurs linéaires $\boldsymbol{\Theta}_n$ sous des contraintes de non-négativité :

$$\begin{cases} \left[\hat{\alpha}_{nk}, \hat{\boldsymbol{\beta}}_k\right] = \underset{\alpha_{nk}, \boldsymbol{\beta}_k}{\operatorname{argmin}} & \sum_{n=1}^{N} \left(\left|\left| \boldsymbol{\Theta}_n - \sum_{k=1}^{K} \alpha_{nk} \boldsymbol{\beta}_k \right|\right|_F^2 \right) \\ \alpha_{nk} \geq 0, & \forall n \in [\![1, N]\!], \forall k \in [\![1, K]\!] \\ ||\boldsymbol{\beta}_k||_F = 1, & \forall k \in [\![1, K]\!] \end{cases} \quad (2)$$

En considérant l'ensemble $\{\boldsymbol{\Theta}\}_n$ de tous les N opérateurs linéaires locaux et en les vectorisant, le problème de minimisation sous contraintes (2) peut être réécrit comme un problème de décomposition par apprentissage de dictionnaires :

$$\begin{cases} \left[\hat{\boldsymbol{\Gamma}}, \hat{\boldsymbol{D}}\right] = \underset{\boldsymbol{\Gamma}, \boldsymbol{D}}{\operatorname{argmin}} & ||\boldsymbol{\Phi} - \boldsymbol{D}\boldsymbol{\Gamma}||_F^2 \\ \boldsymbol{\Gamma}_{kn} \geq 0, & \forall k \in [\![1, K]\!], \forall n \in [\![1, N]\!] \\ ||\boldsymbol{D}_{:k}||_2 = 1, & \forall k \in [\![1, K]\!] \end{cases} \quad (3)$$

où $\boldsymbol{\Phi} \in \mathbb{R}^{IJ \times N}$ est la concaténation des opérateurs vectorisés $\boldsymbol{\theta}_n^v = vec(\boldsymbol{\Theta}_n)$ (c'est-à-dire $\boldsymbol{\Phi} = [\boldsymbol{\theta}_1^v | ... | \boldsymbol{\theta}_N^v]$), les lignes de la matrice $\boldsymbol{\Gamma} \in \mathbb{R}^{K \times N}$ contiennent les coefficients de mélange α_{nk} pour chaque mode k et les colonnes de la matrice $\boldsymbol{D} \in \mathbb{R}^{IJ \times K}$ (notées $\boldsymbol{D}_{:k}$) contiennent des versions vectorisées des matrices de régression linéaire modales $\boldsymbol{\beta}_k$.

Par rapport à [5], la formulation basée sur l'apprentissage de dictionnaires a une plus grande souplesse et adaptabilité. En particulier, les contraintes ou a priori formulés peuvent être modifiées aisément, par exemple par le choix d'une approche d'apprentissage de dictionnaires particulière (e.g. NMF [3], K-SVD [6], etc). Ici, pour un ensemble d'opérateurs donné $\{\boldsymbol{\Theta}\}_n$, nous résolvons la minimisation (3) en itérant deux étapes jusqu'à convergence, à savoir l'estimation au sens des moindres carrés de la matrice de dictionnaire $\boldsymbol{D}$ avec une contrainte de normalisation $||\boldsymbol{D}_{:k}||_2 = 1, \forall\, k \in [\![1, K]\!]$, et l'estimation des coefficients de mélange $\boldsymbol{\Gamma}$ à l'aide d'un opérateur proximal "hard-thresholding" ($\max(\boldsymbol{\Gamma}, 0)$) [7] pour imposer la non-négativité.

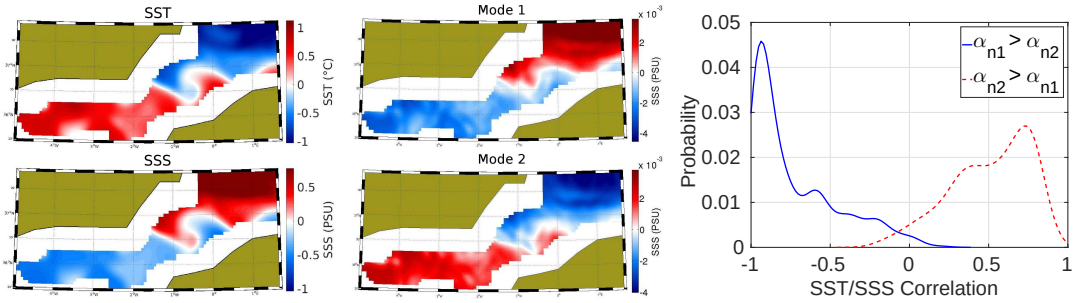
A partir d'un dictionnaire d'opérateurs $\{\hat{\boldsymbol{\beta}}\}_k$ estimés (matrice $\hat{\boldsymbol{D}}$ dans la formulation (3)), nous pouvons appliquer la décomposition proposée à tout nouvel ensemble de données d'observation $\{\mathbf{x}^*, \mathbf{y}^*\}_n$ pour estimer les coefficients de mélange $\hat{\alpha}_{nk}$ associés (matrice $\hat{\boldsymbol{\Gamma}}$ dans la formulation (3)). Deux approches peuvent être envisagées. De même que dans l'étape d'apprentissage du dictionnaire, nous pouvons d'abord estimer les opérateurs linéaires $\{\boldsymbol{\Theta}^*\}_n$ (voir la Section 3 pour plus de détails sur l'estimation pour chaque cas d'étude) et estimer ensuite les coefficients de mélange comme la projection de ces opérateurs sur le dictionnaire. Une deuxième approche revient à l'estimation directe des coefficients de mélange à partir du dictionnaire et des observations selon un critère de moindres carrés dérivé du modèle (1). Ce schéma ne nécessite pas l'estimation préalable des opérateurs linéaires $\{\boldsymbol{\Theta}^*\}_n$.

3 Applications

3.1 Décomposition d'opérateurs dynamiques

Nous illustrons d'abord l'intérêt de l'approche proposée pour la prédiction de systèmes dynamiques. Nous considérons comme cas d'étude le modèle de Lorenz '96, un modèle représentatif de dynamiques géophysiques chaotiques largement utilisé dans la communauté de sciences atmosphériques et océanographiques. Nous laissons le lecteur se référer à [8] pour une présentation détaillée du modèle de Lorenz '96.

Nous créons des jeux de données d'apprentissage et test à partir de séries temporelles 40-dimensionnelles indépendantes du modèle de Lorenz '96. Suivant [9], nous considérons un modèle analogue localement linéaire. Il s'agit d'ajuster localement des régressions linéaires multivariées $\Theta_n \in \mathbb{R}^{I \times J}$: $\mathbf{y}_n =$



(a) Images de SST et SSS (b) Prédictions modales de SSS (c) Distribution des corrélations SST-SSS

FIGURE 2 – Décomposition non-négative des relations température-salinité dans la Mer d'Alboran : (a) images de la température de surface (SST) et de la salinité de surface (SSS) de l'océan au 22 Mars, 2011, (b) Prédictions des images de SSS à partir des images de SST pour chaque mode de la décomposition à $K = 2$ modes considérée, (c) Distribution de la corrélation SST-SSS lorsque le mode 1 (resp. le mode 2) domine, c'est-à-dire $\alpha_{n1} > \alpha_{n2}$ (resp. $\alpha_{n2} > \alpha_{n1}$).

$\Theta_n \mathbf{x}_n$, où, pour une série temporelle donnée S, $\mathbf{x}_n = S(t)$ et $\mathbf{y}_n = S(t + \partial t)$. Les opérateurs linéaires locaux Θ_n sont estimés pour chaque paire d'observations n sous l'hypothèse que les paires d'observations les plus proches (selon la distance euclidienne entre des observations $\mathbf{x}_n$) dans le jeux de données d'apprentissage, appelées analogues, partagent le même opérateur Θ_n. Numériquement, l'estimation de Θ_n recourt à une estimation par moindres carrées pondérés à partir de l'ensemble de données formé par les analogues de la paire $(\mathbf{x}_n, \mathbf{y}_n)$, les pondérations prenant en compte des similitudes relatives.

Nous illustrons l'application de la décomposition non-négative proposée pour $K = 4$ modes. En considérant un dictionnaire $\{\hat{\boldsymbol{\beta}}\}_k$ estimé à partir du jeux de données d'apprentissage, nous calculons les opérateurs de prédiction du jeux de données de test selon le modèle (2). Il repose sur l'estimation des coefficients de mélange dans le modèle (2). Plus précisément, pour la paire d'observations actuelles $(\mathbf{x}_n^*, \mathbf{y}_n^*)$ dans l'ensemble de données de test, l'opérateur de prédiction par analogues associé implique la projection non-négative de l'opérateur localement linéaire Θ_n^* sur l'espace engendré par les modes estimés $\{\hat{\boldsymbol{\beta}}\}_k$. La prédiction au prochain pas de temps se ramène simplement à l'application de l'opérateur linéaire projeté à l'état actuel. La figure 1 illustre les performances de prédiction de l'approche proposée. Nous rapportons l'erreur quadratique moyenne (RMSE) de prédiction normalisée en fonction du nombre d'analogues utilisés pour la prédiction localement linéaire par analogues sans décomposition (ligne complète) et en utilisant une décomposition non-négative (ligne pointillée). Trois scénarios sont simulés : i) un scénario sans bruit avec un grand catalogue ($N = 2 \times 10^5$ exemplaires) (Fig. 1a), ii) un scénario avec un grand catalogue bruité (variance du bruit $\sigma^2_{noise} = 0.1$) (Fig. 1b), iii) un scénario sans bruit avec un petit catalogue ($N = 2 \times 10^3$ exemplaires) (Fig. 1c). La décomposition non-négative des opérateurs locaux réduit clairement l'erreur de prédiction lorsque les analogues sont échantillonnés à partir d'un catalogue bruité ou réduit. Cette décomposition peut être considérée comme une projection de l'opérateur d'origine dans un espace de dimension inférieure, ce qui rend plus robuste l'identification du modèle pour les petits ensembles de données et améliore la robustesse au bruit.

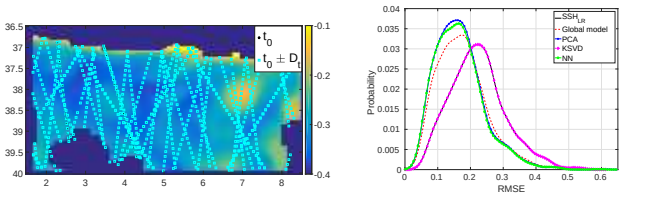
3.2 Décomposition des dynamiques de l'océan

Nous présentons une deuxième application du modèle proposé pour l'analyse des relations spatio-temporelles entre différentes variables géophysiques. Nous considérons ici les relations entre la température de la surface de l'océan (SST, pour son sigle en anglais) et la salinité de la surface de l'océan (SSS, pour son sigle en anglais) dans la Mer d'Alboran ($35°N$-$38°N$, $0°W$-$5°W$) (Figure 2a). Cette région présente de patterns saisonniers marqués avec des inversions de la corrélation entre les champs de SST et SSS, associées à l'apport d'eau froide de l'Atlantique à travers le détroit de Gibraltar.

Pour les expériences rapportées, nous exploitons des simulations opérationnelles dans la Mer d'Alboran. Nous analysons les relations entre des images journalières de SST et SSS (Figure 2a) en utilisant un modèle convolutionnel pour des patches d'image de taille 3×3, ce qui consiste à considérer que les valeurs de SSS pour un jour donné peuvent être approximées par une régression linéaire Θ_n sur les valeurs du patch de SST correspondant (de taille 3×3). Les opérateurs linéaires estimés à chaque jour sont ensuite décomposés sur $K = 2$ modes. La figure 2b présente les champs de SSS prédits par chaque mode. Le premier mode comprend clairement une inversion de la SST, tandis que le deuxième mode est associé à une relation SST-SSS sans inversion. Ceci est illustré plus clairement dans la figure 2c par les distributions des coefficients de corrélation des images SST et SSS suivant que le premier ou le deuxième mode domine. La bimodalité clairement observée illustre la capacité de la décomposition proposée à séparer avec précision les relations SST-SSS inversées et cohérentes, fournissant ainsi un outil efficace pour l'analyse des processus sous-jacents.

3.3 Super-résolution d'images SSH

Nous illustrons une troisième application du modèle proposé pour la super-résolution d'images haute résolution de la hauteur de la surface de l'océan (SSH, pour son sigle en anglais) à partir de données multi-sources. Le modèle proposé pour la super-résolution des champs altimétriques consiste à supposer que l'image SSH haute-résolution ($\mathbf{y}_n$ dans (1)) peut être re-



(a) Échantillonnage des images SSH

(b) Distribution de l'erreur de reconstruction de SSH

FIGURE 3 – Super-résolution des images SSH : (a) Illustration de l'échantillonnage satellitaire irrégulier des champs SSH : positions échantillonnées (carrés en cyan) dans une fenêtre de ± 10 jours autour du 20 avril, 2012. (b) Distribution de l'erreur de reconstruction (RMSE) pour des images journalières de SSH à haute résolution, pour un modèle convolutionnel global et pour des décompositions localement adaptées du modèle global utilisant PCA [1], KSVD [6] et une décomposition non-négative (NN) avec $K = 10$ classes. La distribution du RMSE pour des images SSH journalières à basse résolution est donnée comme référence (notée SSH_{LR}).

construite par un modèle convolutionnel local appliqué à une image SSH basse-résolution et une image SST haute-résolution ($\mathbf{x}_n$ dans (1)) (pour une présentation détaillée de ce modèle, nous référons le lecteur à [10]). Ce modèle peut être vu comme une extension multimodale des modèles de super-résolution d'images [11]. L'une des spécificités intéressantes réside dans la capacité à calibrer des opérateurs locaux de convolution à partir des observations partielles haute-résolution de la SSH (cf. Figure 3a). Les opérateurs de convolution locaux sont calibrés sur des voisinages spatio-temporels correspondants à ± 10 jours et $\pm 1°$. L'approche par décomposition non-négative des opérateurs de convolution permet ici de réaliser des calibrations plus locales avec un nombre réduit de données.

L'approche proposée est évaluée à partir de simulations numériques d'un modèle de circulation océanique pour un cas d'étude en Méditerranée. Nous rapportons la distribution de l'erreur de reconstruction de la SSH à haute résolution pour un modèle convolutionnel global, ainsi que pour des décompositions localement ajustées du modèle global, en considérant différents types de contraintes : orthogonalité (PCA) [1], parcimonie (K-SVD) [6] et non-négativité (NN). Les résultats obtenus mettent en évidence l'intérêt de la décomposition non-négative pour améliorer la résolution spatiale des champs de SSH.

4 Conclusion

Nous avons proposé une généralisation de l'approche de décomposition non-négative introduite dans [5] à la décomposition d'opérateurs localement linéaires. Nous avons exploré des applications pour la prédiction et caractérisation de dynamiques géophysiques et la super-résolution de champs géophysiques. L'approche proposée exploite une formulation basée sur l'apprentissage de dictionnaires à partir d'opérateurs linéaires localement estimés. Ce formulation a une plus grande flexibilité et adaptabilité. Outre les contraintes de non-négativité, cette formulation peut exploiter d'autres contraintes, telle que la parcimonie des dictionnaires ou des coefficients de mélange, en exploitant des techniques adéquates d'apprentissage de dictionnaires [6]. Les travaux futurs exploreront et évalueront ces approches alternatives basées sur l'apprentissage de dictionnaires pour différentes applications de modélisation et reconstruction de dynamiques géophysiques.

Références

[1] A. Hannachi, I.T. Jolliffe, and D.B. Stephenson. Empirical orthogonal functions and related techniques in atmospheric science : A review. *International Journal of Climatology*, 27(9) :1119–1152, Jul 2007.

[2] P. Comon and C. Jutten. *Handbook of Blind Source Separation : Independent Component Analysis and Applications*. Academic Press, Oxford, 1st edition, 2010.

[3] D.D. Lee and H.S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755) :788–791, October 1999.

[4] F. Abrard and Y. Deville. A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. *Signal Processing*, 85(7) :1389–1403, July 2005.

[5] M. López-Radcenco, A. Aïssa-El-Bey, P. Ailliot, P. Tandéo, and R. Fablet. Non-negative decomposition of linear relationships : Application to multi-source ocean remote sensing data. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4179–4183, March 2016.

[6] M. Aharon, M. Elad, and A. Bruckstein. K-SVD : An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11) :4311–4322, Nov 2006.

[7] P. L. Combettes and J.C. Pesquet. Proximal Splitting Methods in Signal Processing. In H.H. Bauschke, R.S. Burachik, P.L. Combettes, V. Elser, D.R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer, 2011.

[8] Predictability : a problem partly solved. In *Seminar on Predictability, 4-8 September 1995*, volume 1, pages 1–18. ECMWF, 1995.

[9] R. Lguensat, P. Tandéo, P. Ailliot, Pulido M., and R. Fablet. The analog data assimilation. In *Monthly Weather Review*, Submitted.

[10] R. Fablet, J. Verron, B. Mourre, B. Chapron, and A. Pascual. Improving mesoscale altimetric data from a multi-tracer convolutional processing of standard satellite-derived products. Working paper or preprint, October 2016.

[11] E. Agustsson, R. Timofte, and L. Van Gool. Regressor basis learning for anchored super-resolution. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 3850–3855, Dec 2016.