

Réseaux de neurones artificiels pour la prédiction et la reconstruction de dynamiques océanographiques

M. Said Ouala¹

M. Cédric Herzet²

M. Ronan Fablet¹

¹ IMT Atlantique ; Lab-STICC, Brest, France

² INRIA Bretagne-Atlantique, Fluminance, Rennes, France

said.ouala@imt-atlantique.fr

1 Résumé

En raison de la disponibilité d'importantes bases de données d'observations à grande échelle acquises lors des différentes missions spatiales d'altimétrie, les modèles basés données apparaissent comme des représentations pertinentes de systèmes géophysiques pour un large éventail d'application. Dans ce travail, nous étudions la pertinence des représentations basées réseaux de neurones, pour la prédiction et la reconstruction de champs géophysiques à partir de données de télédétection satellitaires. La validation de nos modèles est démontrée sur un champ de température de la mer au large de l'Afrique du Sud, une zone qui implique une dynamique océanique intense et complexe. Nos expériences numériques démontrent que les représentations à base de réseaux de neurones surpassent les autres-modèles, y compris les schémas analogues, en matière de prédiction et d'interpolation des données manquantes avec un gain relatif allant jusqu'à 50%.

2 Introduction

La prédiction et la reconstruction des grandeurs océanographiques reposent essentiellement sur la formulation d'un modèle dynamique décrivant l'évolution temporelle de nos grandeurs d'intérêt. La sélection et la paramétrisation de ce modèle par des méthodes analytiques étant un problème complexe [1] sujet à de nombreuses limitations [2], la disponibilité d'importantes bases de données d'observations et de simulation a motivé l'utilisation de techniques exploitant les données pour approcher nos modèles réels de façon paramétrique ou non paramétrique. On peut citer à la fois les schémas basés sur la recherche d'analogues [2] et les réseaux de neurones [3] comme deux exemples efficaces de formulation de modèles à partir des données.

Dans ce travail, nous étudions l'utilisation des réseaux de neurones artificiels pour l'identification de systèmes dynamiques. Nous nous concentrons sur la représentation récemment introduite [4] qui rend explicite la relation entre l'architecture du réseau et le système dynamique sous-jacent. Comme détaillé ci-après, cette représentation peut être considérée comme une implémentation d'un schéma d'intégration numérique d'un modèle dynamique. Dans l'ensemble, les principales contributions de ce travail sont : i) nous proposons une nouvelle architecture d'identification de dynamiques locales fondée sur l'utilisation d'une structure particulière de réseau de neurones. Cette représentation est appliquée à la composante haute résolution de la température de la mer, ii) le modèle dynamique identifié est ensuite utilisé dans un schéma d'assimilation stochastique de données pour reconstruire les composantes haute résolution de nos séries temporelles de température à partir d'observations satellitaires avec des données manquantes, iii) la pertinence du modèle proposé en matière de performances en prédiction et en reconstruction est démontrée pour un cas d'étude au large de l'Afrique du Sud.

3 Modèle proposé

3.1 Identification locale de la composante haute résolution la température de la mer

Etant donné un champ spatio-temporel de température de la mer X dont la composante basse résolution \bar{X} est supposée connue (pour les échelles horizontales au-dessus de 100 km). Cette composante basse résolution peut être par exemple calculée avec une méthode d'interpolation optimale [5]. La composante haute résolution Y de notre champ de température est alors définie selon l'équation : $Y = X - \bar{X}$.

Conformément à [6], nous considérons une représentation locale au niveau patch de notre composante haute résolution. Cela revient à décomposer Y en $P \times P$ patches, avec P la longueur et la largeur de notre zone. La variabilité spatiale

Ce travail a été financé par le projet GERONIMO de l'ANR, Labex Cominlabs, CNES, ainsi que par le MESR, FEDER, Région Bretagne, Conseil Général du Finistère, Brest Métropole et par l'Institut Mines Télécom dans le cadre du program VIGISAT géré par "Groupement Bretagne Télédétection" (BreTel).

de nos patches est ensuite encodée en appliquant une décomposition ACP. Foremnellement, pour une zone \mathcal{P}_s donnée centrée en un point s , une projection dans l'espace ACP est calculée :

$$Y(\mathcal{P}_s, t) = \sum_{k=1}^{N_E} \alpha_k(s, t) \beta_k \quad (1)$$

avec β_k le k^{ieme} vecteur de notre base ACP et $\alpha_k(s, t)$ le coefficient correspondant pour la zone \mathcal{P}_s au temps t . La projection de notre composante haute résolution $Y(\mathcal{P}_s, t)$ sur la base ACP est donc le vecteur des N_E coefficients $\alpha_k(s, t)$ noté $Z(\mathcal{P}_s, t)$. La matrice de projection \mathcal{B} de l'espace ACP $Z(\mathcal{P}_s, t)$ vers l'espace des patches $Y(\mathcal{P}_s, t)$ peut être déduite à partir de l'équation 1 en concaténant les vecteurs β_k de notre base.

$$Z(\mathcal{P}_s, t) = \begin{bmatrix} \alpha_1(s, t) \\ \alpha_2(s, t) \\ \vdots \\ \alpha_{N_E}(s, t) \end{bmatrix}, \mathcal{B} = \begin{bmatrix} \beta_1^t \\ \beta_2^t \\ \vdots \\ \beta_{N_E}^t \end{bmatrix} \quad (2)$$

Par conséquent, l'équation (1) peut être écrite sous une forme matricielle :

$$Y(\mathcal{P}_s, t) = Z(\mathcal{P}_s, t) \mathcal{B} \quad (3)$$

3.2 Identification de dynamiques locales par réseau de neurones artificiel

Suite à notre récente étude [4], on considère comme modèle paramétrique utilisé pour identifier la dynamique locale de nos composantes hautes résolutions un réseau de neurones résiduel à couches bilinéaires. Les principales caractéristiques de ce type de modèles sont les suivantes :

- Il rend explicite la relation entre notre modèle, l'opérateur dynamique sous-jacent et le schéma d'intégration numérique utilisé ;
- Il intègre des termes bilinéaires qui sont des caractéristiques intrinsèques de plusieurs systèmes dynamiques[7].

Foremnellement, on suppose que notre patch est gouverné par une EDO dans l'espace ACP :

$$\frac{dZ(\mathcal{P}_s, t)}{dt} = F(Z(\mathcal{P}_s, t), \theta) \quad (4)$$

Avec F l'opérateur dynamique inconnu et θ ses paramètres associés. Notre but est donc de réussir à identifier F en utilisant comme modèle notre réseau de neurones résiduel à couches bilinéaires.

L'aspect clé de ce réseau est l'intégration de représentations polynomiales de second ordre dans une structure de réseau de neurones complètement connectés approchant l'opérateur dynamique F . La représentation globale implique une architecture à quatre blocs partagés, qui correspond à un schéma d'intégration de Runge-Kutta-4 dont l'élément-clé est l'opérateur dynamique défini par la structure résiduelle bilinéaire.

4 Expériences numériques

Dans cette section, nous présentons les expériences numériques réalisées pour évaluer la performance du modèle proposé en matière de prédiction et d'assimilation de données. Nos expériences sont comparées avec une méthode de modélisation non paramétrique basée sur la recherche d'analogues [2].

4.1 cas-d'étude considéré

La base de données utilisée pour réaliser nos expériences est obtenue en exploitant le produit OSTIA [5] avec une résolution spatiale de 0.05° de janvier 2008 à décembre 2015 et une résolution temporelle $h = 1$ journée. Les données de 2008 à 2014 sont utilisées comme base d'apprentissage et l'année 2015 est utilisée comme base de test. La région considérée se situe au large de l'Afrique du Sud ($5^\circ E$ à $75^\circ E$, $25^\circ S$ à $55^\circ S$).

La taille des patches utilisés est $P = 20$ et la dimension de l'espace ACP est $N_E = 50$ (ce qui correspond à 95% de la variance totale). Quatre zones différentes ont été sélectionnées pour évaluer notre méthode.

4.2 Paramètres expérimentaux

On a utilisé un réseau de neurones résiduel à couches bilinéaires reproduisant un schéma d'intégration numérique du type Runge Kutta pour identifier nos dynamiques locales, deux configurations d'apprentissage ont été envisagées. :

- Apprentissage d’un seul bloc sur la dérivée temporelle de la composante haute résolution projetée dans l’espace ACP, ce bloc est ensuite dupliqué quatre fois pour reproduire le schéma d’intégration de Runge Kutta 4. Cette configuration est notée Bi-NN(1)-RK4.
- Apprentissage du modèle dynamique du réseau de neurones directement à partir de la projection de la série temporelle d’anomalie dans un schéma d’intégration de Runge Kutta 4. Cette configuration est notée Bi-NN(4)-SL.

Nous avons comparé nos résultats aux approches fondées sur la recherche d’analogues. Les modèles analogues ont montré d’excellents résultats comparés aux techniques classiques dans la reconstruction des champs de température et de courant [8, 6]. Deux modèles de prédiction analogue ont été considérés : la prédiction analogue locale (PAL) et la prédiction analogue globale (PAG). Nous laissons le lecteur se référer à [8, 6] pour plus de détails sur les paramètres des méthodes basées sur la recherche d’analogues.

4.3 Performances en prédiction et en reconstruction de champs

Les performances en prédiction à différents pas de temps sont rapportées dans le tableau 1. De même, nous rapportons les résultats de l’expérience d’assimilation dans le tableau 2. Les deux expériences soulignent la pertinence de la représentation proposée comparée aux approches fondées sur la recherche d’analogues.

TABLE 1 – *Performances en prédiction des modèles considérées sur quatre patches* : Racine quarée de l’erreur quadratique moyenne pour des prédiction à différents pas de temps pour les modèles : PAG (A), PAL (B), Bi-NN(1)-RK4 (C), Bi-NN(4)-SL (D).

Model		A	B	C	D
patch 1	$t_0 + h$	1.61	1.58	1.10	1.08
	$t_0 + 4h$	3.70	3.29	3.36	2.76
patch 2	$t_0 + h$	0.96	0.75	0.36	0.33
	$t_0 + 4h$	1.35	1.88	1.33	1.11
patch 3	$t_0 + h$	1.09	0.99	0.86	0.87
	$t_0 + 4h$	1.80	1.31	1.32	1.16
patch 4	$t_0 + h$	2.34	1.89	1.37	1.44
	$t_0 + 4h$	1.81	1.55	1.73	1.25

TABLE 2 – *Reconstruction de la dynamique des composantes haute résolution des champs de températures sur l’année 2015 à partir d’observations avec des données manquantes générées en utilisant le masques de nuages METOP* : Racine quarée de l’erreur quadratique moyenne des champs reconstruits.

Model	Patch1	Patch2	Patch3	patch4
Bi-NN(1)-RK	0.89	0.44	0.59	0.60
Bi-NN(4)-SL	0.89	0.42	0.39	0.60
AnDA-G	2.10	1.78	3.14	0.93
AnDA-L	0.98	0.44	0.73	0.78

5 Conclusion

Cette étude confirme la pertinence de nos architectures [4] dans la prédiction et la reconstruction de champs géophysiques de surface de la mer à partir d’observations satellitaires partielles. Ces représentations surpassent les modèles classiques en prédiction et en reconstruction et fournissent une interprétation explicite du modèle en termes d’opérateur dynamique. Les travaux futurs exploreront plus en avant ce type de représentation et leurs applications aux données de télédétection satellitaire.

Références

- [1] E. N. Lorenz, “Atmospheric predictability experiments with a large numerical model,” *Tellus*, vol. 34, no. 6, pp. 505–513, Dec. 1982.
- [2] R. Lguensat, P. Tandeo, P. Ailliot, M. Pulido, and R. Fablet, “The Analog Data Assimilation,” *Monthly Weather Review*, Aug. 2017.
- [3] A. Braakmann-Folgmann, R. Roscher, S. Wenzel, B. Uebbing, and J. Kusche, “Sea Level Anomaly Prediction using Recurrent Neural Networks,” *arXiv :1710.07099 [cs]*, Oct. 2017, arXiv : 1710.07099.
- [4] R. Fablet, S. Ouala, and C. Herzet, “Bilinear residual Neural Network for the identification and forecasting of dynamical systems,” *SciRate*, Dec. 2017.
- [5] C. J. Donlon, M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, “The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system,” *Remote Sensing of Environment*, vol. 116, no. Supplement C, pp. 140–158, Jan. 2012.
- [6] R. Fablet, P. H. Viet, and R. Lguensat, “Data-Driven Models for the Spatio-Temporal Interpolation of Satellite-Derived SST Fields,” *IEEE Transactions on Computational Imaging*, vol. 3, no. 4, pp. 647–657, Dec. 2017.
- [7] S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, Apr. 2016.
- [8] R. Lguensat, P. Huynh Viet, M. Sun, G. Chen, T. Fenglin, B. Chapron, and R. FABLET, “Data-driven Interpolation of Sea Level Anomalies using Analog Data Assimilation,” Oct. 2017.