

Etude de la variabilité spatio-temporelle du phytoplancton à partir des pigments secondaires estimés par les observations satellitaires.

Roy El Hourany^{1,2}, Marie Abboud-Abi Saab³, Ghaleb Faour², Julien Brajard¹, Michel Crépon¹, Sylvie Thiria¹

¹ Laboratoire d'Océanographie et du Climat : Expérimentations et Approches Numériques (LOCEAN), Paris, France

² Centre Nationale de Télédétection, CNRS-L, Beirut, Liban

³ Centre Nationale des Sciences Marines, CNRS-L, Batroun, Liban

Roy.elhourany@locean-ipsl.upmc.fr

1 Introduction

Le phytoplancton représente l'ensemble des organismes végétaux vivants en suspension dans l'eau. Il constitue 50 % de la production primaire à l'échelle mondiale et est à la base des réseaux trophiques océaniques. Il joue un rôle essentiel dans le rétrocontrôle du climat global, notamment en pompant le CO₂ (gaz à effet de serre) de l'air. Il présente une grande diversité de taille, de couleur selon son contenu en pigments et de forme et se répartit en plusieurs groupes : les algues bleues ou cyanobactéries, les diatomées, les coccolithophoridés, les chrysophycées, les chlorophycées et les dinoflagellés. La variabilité du phytoplancton peut être observée par satellite grâce aux capteurs multi-spectraux tels que SeaWiFS, MODIS, MERIS, VIIRS et OLCI qui permettent d'estimer la concentration en Chlorophylle-a (Chla) ainsi que les pigments photosynthétiques majeurs caractérisant le phytoplancton.

Récemment plusieurs travaux ont été réalisés afin d'estimer, à l'échelle régionale et à partir des mesures satellitaires, la concentration des pigments autres que la Chla [1], [2], [3], [5], [6]. Ces pigments secondaires habituellement mesurés par chromatographie en phase liquide à haute performance (HPLC), permettent de réaliser des études taxonomiques du phytoplancton. Ces méthodes basées sur les mesures HPLC ont principalement été utilisées pour étudier les différentes communautés phytoplanctoniques, en associant des concentrations de pigments mesurés *in-situ* à des groupes de phytoplancton. De là, il a été alors possible de définir des pigments biomarqueurs, capables de représenter des catégories taxonomiques ou fonctionnelles, et qui soient suffisamment spécifiques et discriminants d'une classe de phytoplancton.

Dans cette étude, nous proposons une méthode d'estimation de la Chla et des pigments secondaires à l'échelle du globe, en utilisant une base de données globale *in-situ* de concentrations pigmentaires mesurés par HPLC et collocalisées avec des mesures satellitaires de réflectance marine à plusieurs longueurs d'ondes ($Rrs(\lambda)$) et la température de surface (SST). Pour cela, on utilise des cartes topologiques de type SOM (*self organizing maps*, [4]). Ce sont des modèles neuronaux à apprentissage non-supervisé. Le but de ces cartes topologiques est de regrouper les données en fonction de leurs similarités. Etant donné les bonnes performances obtenues par les classifications neuronales, on peut les utiliser pour déterminer les relations existantes entre les données satellitaires et les observations *in-situ* et arriver à estimer ces dernières d'une manière efficace et moins contraignante.

2 Matériels et Méthodes

Cette méthodologie consiste à reconstruire la variabilité spatio-temporelle des pigments secondaires à travers les observations journalière de quatre $Rrs(\lambda)$ (412, 443, 490, 555 nm) et de la Chla(OC5) fournie par GlobColour (<http://www.globcolour.info/>), la SST AVHRR (<https://podaac.jpl.nasa.gov/AVHRR-Pathfinder>) et des mesures par HPLC de pigment *in-situ* ceci l'échelle globale. GlobColour est un projet lancé par l'Agence Spatiale Européenne (ESA) et consiste à fusionner les images quotidiennes de différents capteurs satellitaires mesurant la couleur de l'océan (<https://oceancolor.gsfc.nasa.gov/>).

2.1 Base de données

La base de données PIGM-SAT qui va nous servir à calibrer la méthode d'estimation des pigments phytoplanctoniques à partir des mesures satellitaires, est constituée des six mesures satellitaires ci-dessus et de dix pigments phytoplanctoniques mesurés par HPLC couvrant une vaste échelle spatio-temporelle entre 1997 et 2014 échantillonnées dans l'océan global. PIGM-SAT provient de plusieurs bases de données telles que MAREDAT, GEP&CO, SEABASS, NOMAD, et de plusieurs expéditions comme Polarstern, Tara Ocean Expedition... Les dix pigments phytoplanctoniques sont: la Fucoxanthine (Fuco), la Peridinine (Perid), l'Alloxanthine (Allo), la Zeaxanthine (Zea), la 19Hex-Fucoxanthine (19HF), le 19But-Fucoxanthine (19BF), la Chlorophylle-b (Chlb), la Divynil Chlorophylle-b (DVChlb), la Divynil Chlorophylle-a (DVChla) et finalement la Chla.

La base de données PIGM-SAT est donc formée de vecteurs ayant seize composantes, regroupant 15276 mesures HPLC limitées à la première profondeur optique qui est d'environ de 15 à 35 m. Ces mesures sont collocalisées avec les

mesures satellitaires de Chla(OC5), des 4 Rrs(λ) et de la SST en retenant la valeur du pixel le plus proche des coordonnées de la mesure *in-situ*.

2.2 Concept général des cartes auto-organisatrices SOM

L'algorithme SOM est une méthode de classification non linéaire et non supervisée. Cette méthode permet de classifier une base de données composée de vecteurs d'entrée multidimensionnels en classes représentées par un réseau de neurones. Ce réseau (ou carte topologique) est défini en général par une grille rectangulaire, de dimension $p \times q$ où à chaque nœud est associé un neurone. Chaque neurone est associé à un vecteur référent \mathbf{w} qui caractérise un sous-ensemble de la base d'apprentissage ayant en commun un certain nombre de propriétés statistiques. Les neurones du réseau sont connectés entre eux et déterminent une relation topologique (voisinage) qui lie les neurones similaires, relation caractérisée par une distance δ . SOM permet donc la partition d'une base de données initiale dans le sens où chaque classe de données est représentée par un vecteur synthétique multidimensionnelle (vecteur référent \mathbf{w}). Les vecteurs référents \mathbf{w} sont appris lors d'une phase d'apprentissage à partir de la base PIGM-SAT. Chaque vecteur \mathbf{z}_i de PIGM-SAT est associé au neurone caractérisé par le vecteur référent \mathbf{w} qui est le plus proche au sens de la Distance Euclidienne dans le cas présent.

Cette distance est représentée par:

$$J_{SOM}^T(\chi, W) = \sum_{z_i \in SOM} \sum_{c \in SOM} K^T(\delta(c, \chi(z_i))) \|z_i - w_c\|^2$$

Où $c \in SOM$ dénote un neurone de la carte SOM, χ est la fonction d'allocation qui attribue à chaque élément z_i de la base de données initiale un vecteur référent w_c , $\delta(c, \chi(z_i))$ est la distance discrète sur la carte entre un neurone c et l'observation z_i associée, et finalement K^T , fonction Kernel, conditionnée par 'une température T ' qui donne un poids à la distance δ sur la carte et diminue graduellement durant le processus de minimisation. Cette fonction de coût prend en considération l'inertie propre de la base de données initiale et assure la préservation de la topologie. Deux vecteurs proches et similaires de la base de données initiale sont associés à deux neurones adjacents suivant δ . Les vecteurs référents \mathbf{w} sont obtenus par une moyenne pondérée des vecteurs de la base de données initiale affectés à chaque neurone de la carte lors de l'apprentissage.

2.3 Protocole

Phase 1 : Apprentissage : A partir de la base PIGM-SAT, une carte SOM-Pigments de 200x100 neurones est construite où à chaque neurone est associé un vecteur référent composé de 16 variables (six variables satellitaires, dix pigments). A la fin de l'apprentissage, la qualité de la carte est évaluée par cross-validation.

Phase 2 : Cross-validation : La cross-validation consiste à partitionner, à chaque itération, la base de données en sous-ensembles complémentaires ; en général, 95% des données sont prises comme ensemble d'apprentissage et 5% sont prises comme test. A chaque itération, les vecteurs de l'ensemble de test sont projetés sur la carte SOM-Pigments et sont affectés aux neurones les plus proches suivant les 6 variables satellitaires (Chla(OC5), 4 Rrs(λ) et SST). Une erreur est calculée à la fin de chaque itération, entre les concentrations des 10 pigments de l'ensemble de test et celles des neurones associés. Plusieurs cycles de validation sont effectués à l'aide de différentes partitions. 20 cycles de validation sont effectués en assurant le passage de tous les sous-ensembles comme test. Finalement, une erreur moyenne est déduite à la fin des 20 itérations.

Phase 3 : Décodage des pigments : Dans cette phase, des images satellitaires quotidiennes de Chla(OC5), des 4 Rrs(λ) et de la SST sont utilisées pour constituer un nouveau vecteur d'entrée. Les dix variables représentant les pigments *in-situ*, qui sont associées aux six variables satellitaires, sont inconnues. Pour cela, chaque nouveau vecteur est projeté sur la carte SOM-Pigments et est affecté au neurone le plus proche au sens des 6 variables satellitaires. Les valeurs pigmentaires associées au référent du neurone stimulé sont ensuite attribuées à ce vecteur. De cette façon, la distribution spatiale des pigments est reconstruite sur les grilles des images satellitaires de résolution 4km.

3 Résultats et Discussion

Une fois l'apprentissage de la SOM-Pigments est achevé, la relation entre les pigments et les paramètres satellitaires est évaluée. Un gradient est établi entre les valeurs regroupées sur la SOM. Ce résultat dénote une relation cohérente entre les différentes variables. Le résultat de la cross-validation des 10 pigments dénote une bonne estimation des pigments, marquée par un coefficient de détermination (R^2) qui varie entre 0.84 et 0.91 avec un RMSE global de 0.016 mg.m⁻³.

Le résultat satisfaisant de la cross-validation permet d'affirmer la bonne performance de la SOM-Pigments. La figure 1 représente un exemple de reconstruction hebdomadaire de la Chla et 3 pigments secondaires, Fuco, Zea et la 19HF qui représentent respectivement 3 classes de phytoplancton: les Diatomées, les Cyanobactéries et les Haptophytes. Ces images à l'échelle globale dévoilent une distribution spatiale qui diffère d'un pigment à un autre et ceci marque la

variabilité des communautés phytoplanctoniques ; Les Diatomées marquées par des fortes concentrations en Fuco semblent être confinées aux zones côtières et les zones d'upwelling comme les côtes ouest de l'Afrique et de l'Amérique, là où l'apport en nutriment est important. Parallèlement, les Cyanobactéries occupent majoritairement la zone équatoriale, avec des fortes concentrations de Zea suivant le courant équatorial. Et finalement, des concentrations importantes de 19HF prédominent dans les zones subtropicales, marquant l'abondance des Haptophytes au niveau de ces zones-là.

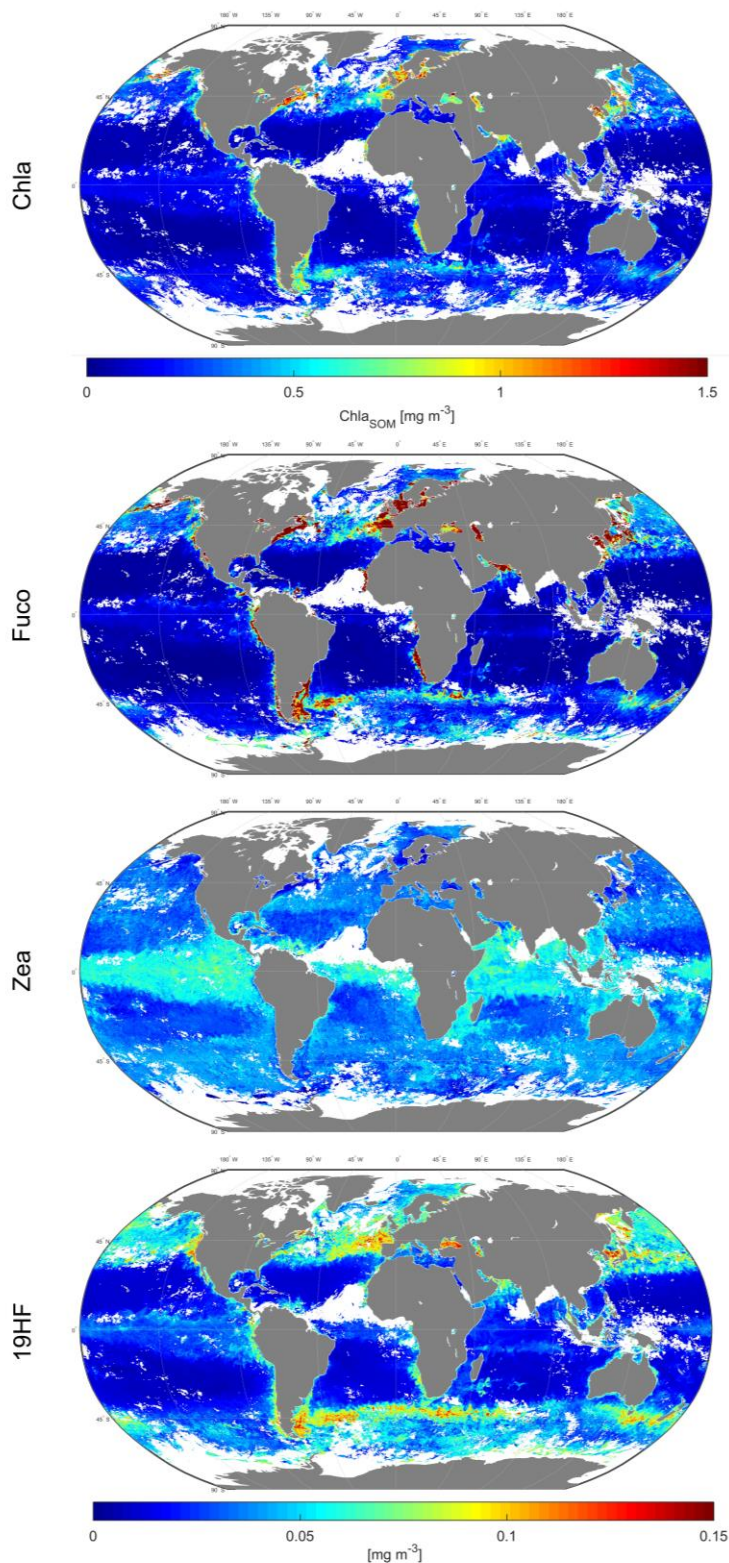


Figure 1: Reconstruction hebdomadaire de Chla, Fuco, Zea et 19HF (2-9 Avril 2017)

4 Conclusion

Le potentiel de reconstruction des pigments secondaires suivant les images satellitaires permet des études de la variabilité du phytoplancton à une fréquence journalière, hebdomadaire, mensuelle, saisonnière et interannuelle. Ceci évoque tout un champ d'application afin de comprendre la dynamique de l'écosystème phytoplanctonique à travers l'utilisation de la SOM-Pigments et à une échelle spatio-temporelle vaste.

5 Références

- [1] Bracher, A., Taylor, M., Taylor, B., Dinter, T., Roettgers, R., & Steinmetz, F.: *Using empirical orthogonal functions derived from remote sensing reflectance for the prediction of phytoplankton pigments concentrations*. Ocean Science, 11(1), 139-158, 2015
- [2] Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., Dall'Olmo, G., and Westberry, T. K.: *Decomposition of in situ particulate absorption spectra*, Methods Oceanogr., 7, 110–124, 2013
- [3] Craig, S. E., Jones, C. T., Li, W. K. W., Lazin, G., Horne, E., Caverhill, C., and Cullen, J. J.: *Deriving optical metrics of ecological variability from measurements of coastal ocean colour*, Remote Sens. Environ., 119, 72–83, 2012
- [4] Kohonen, T.: *Self-organizing maps* (3rd ed.). Springer, Berlin Heidelberg New York, 2001
- [5] Organelli, E., Bricaud, A., Antoine D., and Uitz, J.: *Multivariate approach for the retrieval of phytoplankton size structure from measured light absorption spectra in the Mediterranean Sea (BOUSSOLE site)*, Appl. Opt., 52, 2257–2273, 2013
- [6] Pan, X., Mannino, A., Russ, M. E., Hooker, S. B., and Harding, L. W.: *Remote sensing of phytoplankton pigment distribution in the United States northeast coast*, Remote Sens. Environ., 114, 2403–2416, 2010.