



**HAL**  
open science

# Prévision des ressources en eau à partir de séries temporelles d'images satellitaires à l'aide d'un réseau de neurones pour graphes

Corentin Dufourg, Charlotte Pelletier, Stéphane May, Sébastien Lefèvre

## ► To cite this version:

Corentin Dufourg, Charlotte Pelletier, Stéphane May, Sébastien Lefèvre. Prévision des ressources en eau à partir de séries temporelles d'images satellitaires à l'aide d'un réseau de neurones pour graphes. *Reconnaissance des Formes, Image, Apprentissage et Perception* 2024, Jul 2024, Lille, France. hal-04643544

**HAL Id: hal-04643544**

**<https://hal.science/hal-04643544v1>**

Submitted on 10 Jul 2024

**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.

# Prévision des ressources en eau à partir de séries temporelles d'images satellitaires à l'aide d'un réseau de neurones pour graphes

C. Dufourg<sup>1</sup>

C. Pelletier<sup>1</sup>

S. May<sup>2</sup>

S. Lefèvre<sup>1</sup>

<sup>1</sup> Université Bretagne Sud, IRISA, UMR CNRS 6074, Vannes, France

<sup>2</sup> Centre National d'Études Spatiales (CNES), Toulouse, France

(corentin.dufourg, charlotte.pelletier, sebastien.lefevre)@univ-ubs.fr  
stephane.may@cnes.fr

## 1 Introduction

L'augmentation du nombre de capteurs à bord des satellites génère de gros volumes de données pouvant contribuer à la gestion des territoires. En particulier, le taux de revisite élevé offert par certaines constellations de satellites permet le suivi régulier des dynamiques de l'ensemble des surfaces émergées. Par exemple, les deux satellites Sentinel-2 acquièrent des images tous les cinq jours au niveau de l'équateur avec une résolution spatiale allant jusqu'à 10 mètres. La pile d'images couvrant la même zone à différents moments est connue sous le nom de série temporelle d'images satellitaires (SITS). Ces images constituent des ressources précieuses pour des études à une échelle mondiale comme la cartographie de l'occupation des sols, le suivi de la déforestation, la détection des glissements de terrain ou encore la gestion des ressources naturelles.

Au cours de la dernière décennie, des progrès significatifs ont été réalisés dans le traitement automatique des SITS, en particulier avec l'adoption de techniques d'apprentissage profond, comme démontré par exemple par le benchmark BreizhCrops [16]. Les SITS forment des cubes de données structurés par leurs dimensions spatiales et temporelles, ce qui nécessite le développement d'architectures spécifiques. Dans le contexte de la segmentation sémantique des SITS (*i.e.*, une prédiction pour les séries temporelles extraites pour chaque pixel), la complémentarité des dimensions spatiales et temporelles a été exploitée à l'aide d'architectures à deux branches [5], de méthodes récurrentes ConvLSTM [17], de réseaux U-Net avec encodeur d'attention temporelle (TAE) [18], ou de variantes de Vision Transformers [19]. Des approches similaires peuvent également être employées pour la régression extrinsèque (*e.g.*, l'estimation du rendement) et les tâches de prévision [13].

Une autre approche étudiée pour l'analyse des SITS repose sur les graphes, un outil mathématique bien établi pour les données dont la structure n'est pas euclidienne, tels que les réseaux sociaux, les molécules chimiques ou les nuages de points. Les progrès récents de l'apprentissage basé sur les graphes, favorisés par les réseaux neuronaux pour graphes (GNN), offrent également une alternative convaincante aux techniques d'apprentissage profond basées sur les grilles pour certaines tâches liées à l'image [6], telles que l'interprétation de scènes. Les GNN tirent parti des connaissances sur les structures préétablies pour modéliser efficacement les relations et les interactions complexes entre les pixels ou les régions d'une image, en les traitant comme les nœuds d'un graphe. Cela permet aux GNN de capturer des informations structurelles dans les images, ce qui les rend appropriés pour des tâches où la compréhension du contexte et le raisonnement relationnel sont cruciaux.

Cet intérêt pour l'apprentissage basé sur les graphes appliqué aux données spatio-temporelles a explosé au cours des cinq dernières années dans les applications environnementales, et plus particulièrement pour la prédiction des conditions météorologiques [10, 8] et océaniques [2, 14]. Récemment, GraphCast [9] a réalisé une percée dans la prévision météorologique mondiale en publiant un modèle basé sur les graphes surpassant sur la plupart des variables d'étude les modèles physiques déterministes de prévision numérique du temps. Notons cependant que la plupart de ces approches ne sont pas basées sur des données SITS mais sur des ensembles de données dérivées comme ERA5, résultant d'une réanalyse de données multi-sources, provenant de mesures satellitaires, de stations météo et radar, de radiosondes et d'autres capteurs embarqués.

Concernant les données SITS acquises à moyenne ou haute résolution spatiale (inférieure à 50 mètres), les GNN sont couramment utilisés pour les applications de cartographie de l'occupation du sol [7, 3]. Par exemple, nos travaux antérieurs [4] ont démontré le potentiel des GNN pour effectuer une classification d'occupation des sols à partir de SITS (prédictions pour tous les pixels pour chaque date d'acquisition), par rapport aux méthodes de classification traditionnelles tels que les forêts aléatoires et les perceptrons

---

Les auteurs remercient le CNES et la région Bretagne pour leur financement. Ces travaux ont bénéficié d'un accès aux ressources en HPC/IA de l'IDRIS au travers de l'allocation de ressources 2024-AD011014108R1 attribuée par GENCI.

multicouches. L'étude a mis en évidence l'importance de prendre en compte les caractéristiques du voisinage, spatial comme temporel, de chaque région en plus de ses attributs intrinsèques, et ainsi comprendre son contexte et obtenir une meilleure représentation locale. Dans ce travail, nous proposons de faire le lien entre l'apprentissage basé sur les graphes, telle qu'utilisé en météorologie, et l'analyse des SITS en nous concentrant sur une tâche de prévision à partir de SITS à haute résolution spatiale. L'objectif est d'estimer les ressources en eau dans différents bassins hydrographiques en prédisant l'image suivante dans une séquence donnée d'images satellites. Plus précisément, les expériences sont menées sur le jeu de données SEN2DWATER utilisant l'indice d'eau par différence normalisé (NDWI) [12] comme indicateur des ressources en eau [11]. Pour cela, nous tirons parti des progrès récents dans l'utilisation des approches basées sur les graphes pour les prévisions météorologiques. En particulier, nous proposons une adaptation de GraphCast [9] pour des données SITS.

## 2 Méthode

En bref, GraphCast [9] est un modèle graphique autorégressif qui prédit les conditions météorologiques mondiales jusqu'à 10 jours avec un pas de temps de 6 heures et une résolution spatiale de 0.25 degré. En s'inspirant de simulateurs physiques basés sur les GNN [8, 15], l'architecture se compose d'un encodeur, d'un processeur et d'un décodeur. L'encodeur est alimenté par un historique des états météorologiques. Il projette les caractéristiques d'entrée (*i.e.*, 6 variables climatiques provenant de 37 niveaux atmosphériques, 5 variables de surface et certains états de forçage, y compris des informations temporelles et spatiales) d'une région locale (*i.e.*, un groupe de pixels) vers les nœuds d'un graphe à mailles multiples. Ce maillage modélise la géométrie sphérique de la Terre à l'aide d'un icosaèdre raffiné de manière itérative à différentes résolutions. C'est sur ce maillage que le processeur apprend et applique les interactions physiques existant entre les différentes régions du globe. Enfin, le décodeur reprojette les caractéristiques apprises sur le maillage vers la grille d'origine et prédit la prévision de l'état météorologique en tant que résidu, c'est-à-dire la différence par rapport à l'entrée la plus récente. Contrairement à GraphCast, qui vise à fournir des prévisions globales, notre étude se concentre sur des images spécifiques à un lieu et couvrant de petites zones. Cela a conduit à deux différences majeures avec GraphCast : (i) l'utilisation d'un maillage simple comme dans [8], et (ii) le changement de géométrie du maillage. Comme les images analysées sont de petite taille (cf. Section 3), nous faisons l'hypothèse qu'un seul maillage, c'est-à-dire une seule résolution, est suffisant pour modéliser les dépendances à longue distance. De plus, GraphCast fournit des prédictions à une échelle globale et utilise donc un maillage adapté pour approximer une sphère. Dans notre étude, les prédictions sont effectuées sur une étendue spatiale plus limitée, représentée à une résolution spatiale plus élevée, ce qui rend l'icosaèdre inapproprié. Nous proposons donc de nous appuyer sur une sur-segmentation spatiale de la SITS [1] pour créer un graphe d'adjacence de région utilisé comme maillage prédéfini. L'utilisation d'un maillage irrégulier peut permettre d'obtenir plus de détails dans des zones spécifiques et de limiter l'effet de lissage.

## 3 Données et Résultats préliminaires

Pour réaliser les expériences, nous avons utilisé le jeu de données SEN2DWATER [11]. L'objectif est de prévoir les ressources d'un bassin hydrographique, estimées par le NDWI, à une date donnée en utilisant un historique de six valeurs du NDWI. Le jeu de données se compose de 3682 séquences NDWI extraites d'images Sentinel-2 et recueillies dans 17 bassins différents entre juillet 2020 et décembre 2022. Chaque patch SITS a une taille de  $64 \times 64$  pixels et une résolution temporelle d'environ deux mois (seules les images les moins nuageuses sont traitées). Nous avons divisé les SITS en ensembles d'entraînement (70%), de validation (15%) et de test (15%). Contrairement à GraphCast utilisant une procédure d'entraînement adaptée à une prédiction autorégressive sur 12 pas de temps, nous nous focalisons dans nos expériences sur la prédiction de la septième image d'une série composée des six premières images des séquences NDWI. La stratégie proposée est ainsi comparée dans des conditions similaires au ConvLSTM utilisé dans le travail original [11]. Les deux réseaux sont entraînés sur 100 epochs avec l'optimiseur Adam et un pas d'apprentissage initiale de 0.0001. Pour évaluer les performances des modèles de prévision, nous calculons des mesures de qualité répandues, dont la moyenne est calculée sur l'ensemble des images tests. En particulier, nous utilisons une mesure basée sur la reconstruction, la racine de l'erreur quadratique moyenne (RMSE), et deux mesures basées sur la perception, en l'occurrence le rapport signal-bruit maximal (PSNR) et la mesure de l'indice de similarité structurelle (SSIM) [20].

Le tableau 1 montre des performances pour notre adaptation de GraphCast à la prévision de SITS similaires à ConvLSTM pour un nombre de paramètres moindre. Le transfert technologique d'architectures telles que GraphCast au-delà de son application météorologique initiale est ainsi prometteur.

TABLE 1 – Nombre de paramètres, RMSE, SSIM et PSNR pour les modèles ConvLSTM et notre adaptation de GraphCast. Les valeurs en gras indiquent la meilleure performance pour chaque mesure.

	Nb. Paramètres	RMSE $\searrow$	PSNR $\nearrow$	SSIM $\nearrow$
ConvLSTM	1 036 609	0.0617 $\pm$ 0.0020	30.21 $\pm$ 0.27	<b>0.8383</b> $\pm$ 0.0050
GraphCast modifié	211 649	<b>0.0611</b> $\pm$ 0.0006	<b>30.30</b> $\pm$ 0.08	0.8263 $\pm$ 0.0007

## Références

- [1] Radhakrishna ACHANTA et al. « SLIC superpixels compared to state-of-the-art superpixel methods ». In : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.11 (2012), p. 2274-2282.
- [2] Salva Rühling CACHAY et al. « The World as a Graph : Improving El Niño Forecasts with Graph Neural Networks ». In : *arXiv preprint arXiv :2104.05089* (2021).
- [3] Alessandro Michele CENSI et al. « Attentive spatial temporal graph CNN for land cover mapping from multi temporal remote sensing data ». In : *IEEE Access* 9 (2021), p. 23070-23082.
- [4] Corentin DUFOURG et al. « Graph Dynamic Earth Net : Spatio-Temporal Graph Benchmark for Satellite Image Time Series ». In : *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2023, p. 7164-7167.
- [5] Roberto INTERDONATO et al. « DuPLO : A DUal view Point deep Learning architecture for time series classificatiOn ». In : *ISPRS Journal of Photogrammetry and Remote Sensing* 149 (2019), p. 91-104.
- [6] Licheng JIAO et al. « Graph representation learning meets computer vision : A survey ». In : *IEEE Transactions on Artificial Intelligence* 4.1 (2022), p. 2-22.
- [7] Domen KAVRAN et al. « Graph Neural Network-Based Method of Spatiotemporal Land Cover Mapping Using Satellite Imagery ». In : *Sensors* 23.14 (2023), p. 6648.
- [8] Ryan KEISLER. « Forecasting global weather with graph neural networks ». In : *arXiv preprint arXiv :2202.07575* (2022).
- [9] Remi LAM et al. « Learning skillful medium-range global weather forecasting ». In : *Science* 382.6677 (2023), p. 1416-1421. eprint : <https://www.science.org/doi/pdf/10.1126/science.adi2336>.
- [10] Haitao LIN et al. « Conditional local convolution for spatio-temporal meteorological forecasting ». In : *Proceedings of the AAAI Conference on Artificial Intelligence*. T. 36. 7. 2022, p. 7470-7478.
- [11] Francesco MAURO et al. « SEN2DWATER : A Novel Multispectral and Multitemporal Dataset and Deep Learning Benchmark for Water Resources Analysis ». In : *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2023, p. 297-300.
- [12] Stuart K MCFEETERS. « The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features ». In : *International Journal of Remote Sensing* 17.7 (1996), p. 1425-1432.
- [13] Waytehad Rose MOSKOLAÏ et al. « Application of deep learning architectures for satellite image time series prediction : A review ». In : *Remote Sensing* 13.23 (2021), p. 4822.
- [14] Ding NING et al. « Harnessing the power of graph representation in climate forecasting : Predicting global monthly mean sea surface temperatures and anomalies ». In : *Earth and Space Science* 11.3 (2024), e2023EA003455.
- [15] Tobias PFAFF et al. « Learning Mesh-Based Simulation with Graph Networks ». In : *International Conference on Learning Representations*. 2020.
- [16] M RUSSWURM et al. « Breizhcrops : A time series dataset for crop type mapping ». In : *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 43 (2020).
- [17] Marc RUSSWURM et Marco KÖRNER. « Multi-temporal land cover classification with sequential recurrent encoders ». In : *ISPRS International Journal of Geo-Information* 7.4 (2018), p. 129.
- [18] Vivien SAINTE FARE GARNOT et Loic LANDRIEU. « Panoptic segmentation of satellite image time series with convolutional temporal attention networks ». In : *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021, p. 4872-4881.
- [19] Michail TARASIOU, Erik CHAVEZ et Stefanos ZAFEIRIOU. « ViTs for SITS : Vision Transformers for Satellite Image Time Series ». In : *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, p. 10418-10428.
- [20] Zhou WANG et al. « Image quality assessment : from error visibility to structural similarity ». In : *IEEE Transactions on Image Processing* 13.4 (2004), p. 600-612.