

Sujet de thèse

Optimisation par Machine Learning pour la PROPagation atmosphérique des ondes électromagnétiques au dessus de la surface de mer (MLPROP)

Ali Khenchaf* et Thomas Bonnafont†

Lab-STICC, CNRS UMR 6285, ENSTA Bretagne, 29806, Brest, France

1 Présentation de la thèse

1.1 Contexte et principaux objectifs

Cette thèse porte sur le développement d'une méthode de **Machine Learning (ML)** pour la modélisation de la **propagation** atmosphérique des **ondes électromagnétiques** sur de grandes distances au dessus de la **surface de mer**. Les objectifs sous-jacent sont nombreux. On peut citer par exemple la **prédiction de couverture d'antennes** (dans le contexte 4G et 5G par exemple) en fonction de l'environnement : *effets de la mer et des structures humaines* (comme les fermes solaires, les éoliennes et éoliennes off-shores) par exemple. Cela est aussi utile pour les problématiques de **détection** par mesures de champs électromagnétiques, par exemple pour **les nappes de pétroles sur la mer**. Enfin, la propagation des ondes électromagnétique en scénario de radio-occultation (RO) est également utilisée comme *thermomètre de l'atmosphère* par le GIEC [10, p. 476]. Ces applications sont ainsi d'intérêts majeurs actuellement pour la gestion de la santé des océans, mais aussi de l'environnement.

En effet, pour *déterminer une couverture* radar ou pour les couvertures des systèmes de communication (e.g. 4G et 5G), pour *détecter une nappe de pétrole* ou *prévoir la propagation en scénario de RO*, il est nécessaire de modéliser le plus fidèlement possible *le canal de propagation*. Dans ce cadre, nous nous intéressons à l'introduction d'une méthode de machine learning (ML) pour obtenir « le propagateur ». En effet, à partir d'une base de données construite artificiellement, à l'aide d'une méthode split-step wavelet (SSW) [15, 3, 4], le but est d'obtenir un algorithme de ML pouvant prédire la propagation (path loss) en fonction de données physiques (relief, sol – mer et/ou pétrole –, réfraction, construction humaine, etc.). L'avantage de l'utilisation du ML sera de réduire fortement le temps de calcul et de permettre des calculs en temps réel.

En s'appuyant sur la problématique qui combine à la fois des aspects physiques, mathématiques et numériques sur la propagation des ondes dans un environnement réel (principalement maritime), l'un des

*ali.khenchaf@ensta-bretagne.fr <https://www.ensta-bretagne.fr/khenchaf/fr/cv.php>

†thomas.bonnafont@ensta-bretagne.fr <https://www.ensta-bretagne.fr/bonnafont/>

objectifs consiste à obtenir une méthode numérique efficace en temps de calcul, taille mémoire et précision, basée sur le ML. Cette méthode sera ensuite utilisée pour les applications présentées, en particulier pour modéliser la propagation au dessus d’une mer polluée et en scénario de RO.

1.2 Pré-requis

Le sujet portant sur la propagation des ondes électromagnétiques et le machine learning, des compétences en phénomènes ondulatoires, électromagnétiques, programmation, machine learning supervisé et non supervisé sont attendues. Il n’est pas demandé une parfaite maîtrise de l’intégralité de ces sujets mais il est indispensable que le-la candidat-e soit familier-ère avec certaines de ces notions.

1.3 Environnement scientifique

La thèse sera encadrée par Ali Khencahf (Pr) et Thomas Bonnafont de l’équipe PIM Laboratoire STICC (LabSTICC) et aura lieu au sein de l’ENSTA Bretagne.

Le-la candidat-e bénéficiera naturellement des expériences et des compétences de l’ensemble des chercheurs-ses du LabSTICC et en particulier de l’équipe REMS de l’ENSTA Bretagne, aux spécialités variées et complémentaires.

2 Sujet de thèse

2.1 Introduction succincte aux deux principaux domaines scientifiques investis

D’un côté, la **propagation des ondes électromagnétiques** décrit le comportement physique des ondes électromagnétiques se propageant dans un milieu donné. En effet, de nombreux phénomènes physiques modifient cette propagation. On peut citer en particulier, pour le domaine étudié, la *réfraction* (qui “courbe” les rayons et peut augmenter la portée d’un système), le *relief* (qui induit principalement des réflexions, de la diffraction et une zone d’ombre), et la *composition du sol* (sur lequel une partie de l’onde est réfléchi). Tous ces effets ont un impact important sur la portée d’un système et sur le placement des antennes. En termes de modélisation, ces phénomènes peuvent être étudiés par la résolution de *l’équation d’onde parabolique*, qui est une approximation de *l’équation de Helmholtz* en domaine fréquentiel où seule la propagation vers l’avant est calculée. L’objectif principal dans ce contexte est d’obtenir une méthode précise en un temps le plus faible et dont l’empreinte mémoire est limitée.

De l’autre côté, le **machine learning** qui est un type d’intelligence artificielle impliquant l’entraînement d’un modèle à l’aide de donnée pour faire de la prédiction sans explicitement coder le modèle. Ces méthodes se sont particulièrement développées depuis les années 1980 et ont permis la résolution de nombreux problèmes dans de nombreux domaines [5, 11]. Récemment, *les méthodes de régression à partir d’apprentissage supervisé* se sont développées pour la prédiction de la propagation (en acoustique ou électromagnétique) dans un environnement donné [6, 1, 12, 2]. La difficulté majeure de l’utilisation de ces méthodes est la création d’une base de données adaptée de façon à apprendre le bon modèle.

Ainsi, combiner une méthode de machine learning avec une méthode précise et rapide de prédiction de la propagation des ondes électromagnétiques semblent un outil efficace et prometteur pour prédire le comportement des ondes électromagnétiques dans des environnements complexes en “temps réel”.

2.2 Résumé du sujet de thèse

La thèse proposée se place dans deux domaines scientifiques : les *méthodes numériques pour l'électromagnétisme* et le *machine learning*. Elle vise ainsi, en associant les deux, à proposer une méthode rapide et précise de calcul de la propagation troposphérique des ondes électromagnétiques (EM) sur de grandes distances.

De nos jours, les technologies de communication évoluent (5G) et l'environnement autour des antennes également (*fermes solaires, éoliennes*). De plus, la propagation des ondes électromagnétiques peut être utilisée pour **mesurer** de *nombreux phénomènes environnementaux*, comme la *température atmosphérique* à l'aide de la radio-occultation [10, p. 476], ou pour de la détection, comme celle des *nappes de pétrole sur la mer* [13]. Ainsi, pour ces applications, il est nécessaire de modéliser précisément la propagation haute fréquence (bande VHF à bande X) dans des conditions données. Dans ces applications, l'une des difficultés majeures est de calculer la propagation depuis l'antenne sur de longues distances en un **temps réduit**, tout en gardant une bonne précision et une empreinte mémoire réduite.

Dans ce contexte, de nombreuses méthodes ont été développées. Au vu de la taille des scénarios de propagation, les méthodes rigoureuses sont inenvisageable. C'est pourquoi, en général des méthodes asymptotiques sont utilisées. En particulier, pour la propagation troposphérique longue distance, la résolution de l'équation d'onde parabolique [8] par une méthode split-step (split-step Fourier [7]) a donné de très bons résultats. En effet, elle permet en 2D d'obtenir la zone de couverture d'une antenne en un temps court et avec une bonne précision en tenant compte de nombreux phénomènes physiques (réfraction, relief et composition du sol). Cependant, en 3D, ses performances en termes de temps de calculs et taille mémoire la rende inenvisageable [14].

C'est pourquoi récemment, une méthode basée sur les ondelettes a été développée : split-step wavelet (SSW) [15, 3]. Celle-ci permet d'obtenir des résultats équivalents en termes de précision, tout en réduisant fortement la taille mémoire et la complexité de la méthode. Cependant, en 3D, les domaines de calculs restent limités pour obtenir des résultats en un temps convenable [4].

Le but cette thèse est d'obtenir une méthode « temps réel » du calcul de la propagation. Pour cela, basé sur l'état de l'art actuel [6, 1, 12, 2], le-la candidat-e travaillera sur le développement d'une méthode basée sur le machine learning (ML) et SSW pour obtenir un « DeepSSW ». En effet, récemment les algorithmes de ML ont démontré leurs compétences à apprendre des opérateurs complexes, en particulier en propagation [6, 1], permettant de réduire fortement le temps de calcul de ces méthodes. De plus, SSW [3] permet d'obtenir des cartes de couvertures pour un scénario donné en un temps faible. Nous nous placerons dans un premier temps dans un cas simplifié, comme dans [12], où seule la propagation à hauteur de l'antenne est à modéliser, avant de chercher à obtenir une cartographie du champ sur tout le domaine 2D.

Ici, la méthode SSW sera utilisée pour créer la base de données nécessaire à l'apprentissage, permettant d'éviter le recours à des campagnes de mesures coûteuses. Dans ce cadre, il sera nécessaire d'échantillonner de manière intelligente les phénomènes physiques, de façon à faire apprendre à l'algorithme de ML les bonnes données, lui permettant de généraliser « le propagateur » tout en gardant une empreinte mémoire faible. Ainsi, une étude sur la forme et la taille de la base de données sera menée. Ensuite, l'architecture de l'algorithme de ML devra être étudiée. Une première idée se base sur un *encodeur-décodeur* en apprentissage supervisé, voir [2]. Une seconde idée est d'utiliser un *auto-encodeur* en *apprentissage non-supervisé*. Dans un premier temps la fonction coût considéré sera l'erreur quadratique moyenne, et une optimisation par descente de gradient stochastique pourra être effectuée.

Après avoir obtenu des résultats avec les premiers algorithmes de ML, une étude sera faite pour voir comment optimiser la méthode DeepSSW, soit en améliorant la base de données, soit l'architecture général de l'algorithme de ML. En effet, à l'aide de SSW la base de données pourra être améliorée ou agrandit de façon synthétique en un temps faible. De plus, une étude de la fonction perte et de son optimisation sera également nécessaire.

Cette méthode sera ensuite comparée aux méthodes utilisées dans le milieu : SSF et SSW. Cela permettra de voir les avantages et inconvénients de la méthode DeepSSW. La comparaison concernera en particulier la précision, le temps de calcul et la taille mémoire. De plus, des applications à des cas concrets sont prévues. Par exemple, nous étudierons la modélisation de la propagation au dessus d'une mer polluée dans un but de détection, ou l'optimisation d'antenne sur navire en fonction de la surface de mer.

Enfin, après le développement de la méthode en 2D, les travaux se focaliseront sur sa généralisation à la 3D, et sur quelle partie de la propagation peut être accéléré avec du ML. Ce point représente une difficulté majeure. En effet, en 3D peu de méthodes permettent de tenir compte du relief sur de la propagation longue distance, en étant précise et rapide. Il faudra donc voir comment et où placer l'apprentissage supervisé pour la généralisation à la 3D. Cependant, le développement d'une méthode comme DeepSSW-3D serait un atout pour la prédiction de la propagation EM. Le-la candidat-e pourra s'inspirer des réseaux de neurones convolutifs en nombres complexes comme [9]. Le-la candidat-e pourra également réfléchir à l'introduction d'algorithmes d'apprentissage par renforcement [1] dans ce domaine.

Références

- [1] A. Seretis and C. D. Sarris. An Overview of Machine Learning Techniques for Radiowave Propagation Modeling. *IEEE Transactions on Antennas and Propagation*, 70(6) :3970–3985, June 2022.
- [2] S. Bakirtzis, K. Qiu, J. Zhang, and I. Wassell. Deepray : Deep learning meets ray-tracing. In *2022 16th European Conference on Antennas and Propagation (EuCAP)*, pages 1–5. IEEE, 2022.
- [3] T. Bonnafont, R. Douvenot, and A. Chabory. A local split-step wavelet method for the long range propagation simulation in 2d. *Radio science*, 56(2) :1–11, 2021.
- [4] T. Bonnafont, R. Douvenot, and A. Chabory. Split-step wavelet with local operators for the 3d long-range propagation. In *2021 15th European Conference on Antennas and Propagation (EUCAP)*, pages 1–5. IEEE, 2021.
- [5] T. G. Dietterich. Machine learning. *Annual review of computer science*, 4(1) :255–306, 1990.
- [6] C. R. Hart, D. K. Wilson, C. L. Pettit, and E. T. Nykaza. Machine-learning of long-range sound propagation through simulated atmospheric turbulence. *The Journal of the Acoustical Society of America*, 149(6) :4384–4395, June 2021.
- [7] J. R. Kuttler and G. D. Dockery. Theoretical description of the parabolic approximation/fourier split-step method of representing electromagnetic propagation in the troposphere. *Radio science*, 26(2) :381–393, 1991.
- [8] M. Levy. *Parabolic equation methods for electromagnetic wave propagation*. Number 45. IET, 2000.
- [9] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui. Deepnis : Deep neural network for nonlinear electromagnetic inverse scattering. *IEEE Transactions on Antennas and Propagation*, 67(3) :1819–1825, 2018.

- [10] V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, et al. Climate change 2021 : the physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, 2, 2021.
- [11] T. M. Mitchell and T. M. Mitchell. *Machine learning*, volume 1. McGraw-hill New York, 1997.
- [12] K. Yang, X. Guo, Z. Wu, J. Wu, T. Wu, K. Zhao, T. Qu, and L. Linghu. Using multi-source real landform data to predict and analyze intercity remote interference of 5g communication with ducting and troposcatter effects. *Remote Sensing*, 14(18) :4515, 2022.
- [13] H. Zheng, C. Gou, A. Khenchaf, Y. Wang, and Y. Zhang. Retrieving oil-water mixture ratios of marine oil spills from l-band sar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 60 :1–11, 2022.
- [14] H. Zhou, A. Chabory, and R. Douvenot. A 3-d split-step fourier algorithm based on a discrete spectral representation of the propagation equation. *IEEE transactions on antennas and propagation*, 65(4) :1988–1995, 2017.
- [15] H. Zhou, R. Douvenot, and A. Chabory. Modeling the long-range wave propagation by a split-step wavelet method. *Journal of Computational Physics*, 402 :109042, 2020.