

## Mémoire

**Auteur** : Libert, Aurélien

**Promoteur(s)** : Alvera Azcarate, Aida; Barth, Alexander

**Faculté** : Faculté des Sciences

**Diplôme** : Master en océanographie, à finalité approfondie

**Année académique** : 2021-2022

**URI/URL** : <http://hdl.handle.net/2268.2/16167>

---

### *Avertissement à l'attention des usagers :*

*Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.*

*Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.*

---



# Analysis of spatial scales in satellite data reconstructed using a neural network

Analyse des échelles spatiales dans des  
données satellitaires reconstruites à l'aide  
d'un réseau neuronal

Mémoire rédigé par Aurélien Libert en vue de l'obtention du titre de  
Master en Océanographie, à finalité approfondie

Promoteur : BARTH Alexander  
Co-promotrice : ALVERA AZCARATE Aida  
Année académique : 2021-2022

**Université de Liège**

Conformément aux règles imposées à la rédaction, ce mémoire ne doit pas dépasser 50 pages, rédigées en Times 12 ou équivalent.

# Table des matières

<b>Table des matières</b>	<b>iii</b>
<b>Remerciements</b>	<b>iv</b>
<b>Résumé</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Liste des abréviations</b>	<b>vii</b>
<b>Liste des figures</b>	<b>vii</b>
<b>Liste des tableaux</b>	<b>xi</b>
<b>Introduction</b>	<b>1</b>
La mer d'Alboran	2
L'intérêt de la température de surface des océans	4
Les réseaux neuronaux	7
Apprentissage automatique	7
Perceptron multicouche	8
La fonction de perte	9
La rétropropagation du gradient	9
Réseaux de neurones convolutifs	9
Auto-encodeur	11
U-Net	11
DINCAE	12
Reconstruction	13
La densité spectrale de puissance	18
<b>Résultats</b>	<b>19</b>
Les essais	19
Analyse d'échelle	21
Etude saisonnière	28
<b>Exemples de reconstructions</b>	<b>37</b>
<b>Conclusion</b>	<b>40</b>
<b>Bibliographie</b>	<b>41</b>
<b>Annexes</b>	<b>44</b>
Tableau complet des résultats	44
Loss-Weight refine	47

# Remerciements

Tout d'abord, je souhaite remercier mon promoteur, Alexander Barth, qui m'a aidé tout au long de ce travail, autant pour la mise en place du code utilisé que pour la rédaction. Il aura toujours su être là lorsque j'ai eu besoin d'aide ou d'explications et je lui en suis très reconnaissant.

Je souhaite aussi remercier ma co-promotrice Aide Alvera Azcarate pour ses conseils ainsi que mon père, Alain Libert, pour avoir revu et corrigé ce travail.

Enfin, je souhaite remercier les membres du jury, Arthur Capet, Jean-Marie Beckers et Loïc Michel de prendre de leur temps pour lire ce mémoire.

# Résumé

La température de surface de l'eau (SST) est une variable très importante pour évaluer de nombreux phénomènes physiques et biologiques, et notamment l'impact de l'océan sur le réchauffement climatique. C'est aussi un paramètre qui peut être mesuré très régulièrement et sur toute la surface des mers et des océans grâce aux satellites. Le problème majeur de ces mesures est le manque de données dues à la présence de nuages. Plusieurs méthodes permettent de reconstruire les données sous les nuages grâce à différentes méthodes. Une de ces méthodes est d'utiliser un réseau neuronal avec une architecture U-Net (Siddique et al, 2021) appelée DINCAE (Data-Interpolating Convolutional Auto-Encoder) (Barth et al., 2020 et 2022).

Ce travail a pour objectif d'aider à améliorer la fiabilité des reconstructions faites par DINCAE sur différentes échelles, et particulièrement sur les petites échelles, avec 5 ou 6 couches de neurones. Pour cela, il va se concentrer particulièrement sur l'utilisation de skip-connections entre l'entrée et la sortie de différentes couches de neurones du réseau. L'évaluation de la fiabilité des reconstructions est faite par deux méthodes. La première est une évaluation de l'erreur RMS des données reconstruites par rapport aux données initiales, en incluant des données volontairement enlevées puis reconstruites. La seconde est une évaluation de la quantité de variance gardée lors de la reconstruction car les programmes de reconstructions d'images satellitaires sont connus pour la sous-estimer. L'évaluation de la variance est faite annuellement ainsi que saisonnièrement. La zone d'étude choisie pour ce travail est la mer d'Alboran, la partie la plus à l'est de la mer Méditerranée.

Les résultats obtenus lors de ce travail montre que les skip-connections à utiliser sont dépendantes des échelles et de la période étudiées. Aucune configuration ne permettant de minimiser l'erreur RMS tout en maximisant la variance n'a été trouvée, il est donc nécessaire de trouver un équilibre entre les deux. Cependant, une configuration avec 6 couches de neurones et des skip-connections sur les 5 premières couches a été montrée comme minimisant l'erreur RMS et une configuration avec 6 couches de neurones et une skip-connection sur la seconde couche uniquement a été montrée comme maximisant la variance gardée sur la majorité des échelles sur une année entière, mais pas sur chaque saison individuellement.

# Abstract

Sea surface temperature (SST) is a very important variable to assess numerous physical and biological phenomena, most notably the ocean's impact on the recent climatic changes. It's also a parameter that can be measured regularly and over the entire surface of the oceans and seas thanks to satellites. The major issue with those measurements are data breaches due to clouds. Numerous softwares can approximate the lacking data with various methods, one of those being the use of a neural network with a U-Net architecture (Siddique et al, 2021) named DINCAE (Data-Interpolating Convolutional Auto-Encoder) (Barth et al., 2020 et 2022).

This report aims to improve the reliability of the reconstructions made by DINCAE at differing scales, most notably small scales, with 5 or 6 neural layers. To achieve this, it will focus on the use of skip-connections between the input and the output of different layers. The reliability assessment will be done in 2 ways. +The first one is to evaluate the RMS error of the reconstructed data compared to the initial data, including purposefully removed data then restored. The second one is to assess the amount of variance kept by the software as reconstruction softwares are known to under-estimate it. This assessment is done annually and seasonally. The study zone chosen for this report is the Alboran sea, the most western part of the Mediterranean sea.

The work's result shows the most optimal skip-connections to use are dependent on the scales and period studied. No configuration could be found to both minimize the RMS error as well as to maximize the variance, we thus need to find a trade-off between the two. However, a 6 layers configuration with a skip-connection on each of the first 5 layers has been found to minimize the RMS error, while a 6 layers configuration with a skip-connection only on the second layer has been found to keep the most variance in the system over the entire year at most scales, although it's not the best over each and every one of the season.

# Liste des abréviations

SST : Sea Surface Temperature, température de surface

DINCAE : Data INterpolating Convolutional Auto-Encoder

MODIS : Moderate-resolution Imaging Spectroradiometer

NASA : National Aeronautics and Space Administration

CPU : Central Processing Unit

DINEOF : Data Interpolating Empirical Orthogonal Functions

PO.DAAC : Physical Oceanography Distributed Active Archive Center

RMS : Root Mean Square Error

EMODnet : European Marine Observation and Data Network

# Liste des figures

<b>Figure 1</b> : Représentation schématique des grands courants et gyres de la mer d’Alboran, repris de Renaul, et al., 2012 .....	2
<b>Figure 2</b> : Représentation schématique des grands courants et gyres de la mer Méditerranée, repris de Hernanz, et al., 2018 .....	3
<b>Figure 3</b> : Représentation des grands courants et gyres de la mer d’Alboran grâce à la SST (en degré Celsius).....	3
<b>Figure 4</b> : Représentation schématique simplifiée de la circulation thermohaline océanique globale, repris de Molina-Kesche, 2014.....	5
<b>Figure 5</b> : Représentation graphique de la SST en Mer D'alboran le 19/01/2021.....	6
<b>Figure 6</b> : Schéma d'un perceptron multicouche, repris de Abirami et Chitra, 2020.....	8
<b>Figure 7</b> : Fonctionnement d'une opération de convolution avec un tensor 5x5 et un kernel 3x3, repris de Yamashita et al., 2018.....	10
<b>Figure 8</b> : Architecture d'un réseau de neurones convolutifs pour la classification, repris de Hands on Machine Learning, Géron, 2019.....	11
<b>Figure 9</b> : Schéma de l'architecture U-Net. La partie gauche correspond à la partie de contraction et la partie droite, la partie d’expansion. Repris de Siddique et al., 2021.....	12
<b>Figure 10</b> : Exemple de reconstruction du 12 mai 2002 faite avec DINCAE et les paramètres initiaux décrits dans le tableau 1, première ligne. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.....	16
<b>Figure 11</b> : Exemple de reconstruction du 12 mai 2002 faite avec DINCAE et les paramètres optimaux décrits dans le tableau 1, seconde ligne. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.....	17
<b>Figure 12</b> : Densité spectrale de puissance des tests 0, 6, 10, 11 et 19 du tableau 2 et des données originelles non-reconstruites.....	21

<b>Figure 13</b> : Zoom sur les plus petites échelles des densités spectrales de puissance des tests 0, 6, 10, 11 et 19 du tableau 2 et des données originelles non-reconstruites.....	22
<b>Figure 14</b> : Comparaison des densités spectrales de puissances des tests 0, 6, 10, 11 et 19 du tableau 2 avec les données originelles non-reconstruites.....	24
<b>Figure 15</b> : Zoom sur les plus petites échelles des densités spectrales de puissance des tests avec une seule skip-connection sur chaque couche, ainsi que des données sans skip-connection et des données originelles.....	25
<b>Figure 16</b> : Zoom sur les plus petites échelles des densités spectrales de puissance des tests 10 et 19, ainsi que du test avec une seule skip-connection sur la seconde couche, des données sans skip-connections et des données originelles.....	27
<b>Figure 17</b> : Courbes saisonnières de la densité spectrale de puissance sans skip-connection.	29
<b>Figure 18</b> : Courbes saisonnières de la densité spectrale de puissance avec des skip-connections sur les couches 1, 2, 3, 4 et 5.....	30
<b>Figure 19</b> : Courbes saisonnières de la densité spectrale de puissance avec une skip-connection sur la seconde couche.....	31
<b>Figure 20</b> : Courbes en hiver des tests 0, 6, 19 et avec une skip-connection seulement sur la seconde couche.....	33
<b>Figure 21</b> : Courbes au printemps des tests 0, 6, 19 et avec une skip-connection seulement sur la seconde couche.....	34
<b>Figure 22</b> : Courbes en été des tests 0, 6, 19 et avec une skip-connection seulement sur la seconde couche.....	35

<b>Figure 23</b> : Courbes en automne des tests 0, 6, 19 et avec une skip-connection seulement sur la seconde couche.....	36
<b>Figure 24</b> : Exemple de reconstruction du 22 avril 2001 faite avec DINCAE et les paramètres décrits dans le tableau 3. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.....	38
<b>Figure 25</b> : Exemple de reconstruction du 29 novembre 2002 faite avec DINCAE et les paramètres décrits dans le tableau 3. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.....	39

# Liste des tableaux

<b>Tableau 1</b> : Paramètres initiaux et optimaux pour les reconstructions DINCAE réalisées lors de ce travail.....	15
<b>Tableau 2</b> : Tableau reprenant tous les essais faits lors de ce travail ainsi que leurs erreurs RMS.....	19
<b>Tableau 3</b> : Tableau reprenant les paramètres de la meilleure reconstruction.....	37

# Introduction

La première partie du travail sera consacrée à exposer les différents concepts utilisés et nécessaires à la compréhension de cet écrit. Cela commencera donc par une brève introduction de la mer d'Alboran et de ses courants, suivie par l'importance des mesures de la SST et enfin une introduction aux réseaux neuronaux et en particulier à DINCAE, le programme de reconstruction d'images utilisé lors de ce travail. Enfin, la densité de puissance spectrale sera utilisée en tant que variable afin de caractériser la variabilité des reconstructions.

La seconde partie du travail sera consacrée aux résultats. C'est-à-dire aux évolutions de la variabilité des reconstructions faites lorsque l'on modifie les paramètres d'entrée de DINCAE, et particulièrement lorsque l'on modifie les skip-connections entre les couches neuronales. Dans un premier temps, nous nous attacherons à analyser les résultats sur une année entière, afin de voir quelle série de paramètres permet la meilleure fiabilité des mesures de SST dans la mer d'Alboran. Le deuxième temps sera consacré aux résultats saisonniers des différentes configurations découvertes dans la première partie afin de découvrir et de comprendre les différences de variabilité dues aux saisons.

Pour terminer, la troisième et dernière partie du travail sera consacrée à observer les reconstructions avec les paramètres optimaux obtenus dans la partie résultat du travail.

## La mer d'Alboran

La mer d'Alboran est la partie la plus à l'ouest de la mer Méditerranée. Elle est directement reliée au détroit de Gibraltar et à l'océan Atlantique. La mer se situe entre  $35^{\circ}$  N et  $36^{\circ} 83'$  N et entre  $1^{\circ}12'$  O et  $5^{\circ} 18'$  O (International Hydrographic Organization, 1953).

Au niveau des courants, la mer d'Alboran est caractérisée par deux grands gyres. Le premier est un grand gyre anticyclonique, qui se situe à l'ouest de la mer et le second est un gyre cyclonique situé plus au centre de la mer. Enfin, un dernier grand gyre anticyclonique est présent dans l'est de la mer, mais seulement en été et en automne (Renault, et al., 2012). La figure 1 représente les grands courants de la mer d'Alboran et notamment ses grands gyres.

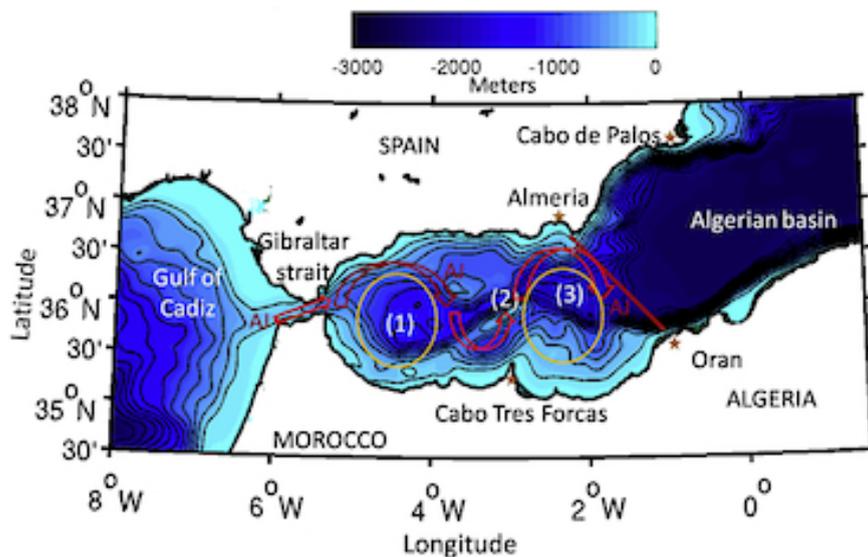


Figure 1 : Représentation schématique des grands courants et gyres de la mer d'Alboran, repris de Renault, et al., 2012.

La figure 2 représente les grands courants dans la mer Méditerranée. On peut y voir l'importance de la mer d'Alboran, puisqu'elle est le point d'entrée de l'eau venant de l'océan Atlantique. Ainsi, l'eau entre par la mer d'Alboran, suit les deux grands gyres de celles-ci et continue vers l'est dans le bassin ouest de la mer méditerranée. Après un passage dans le bassin est, l'eau va revenir dans le bassin ouest par la mer Tyrrhénienne et revenir dans la mer d'Alboran par les côtes italiennes, françaises et espagnoles.

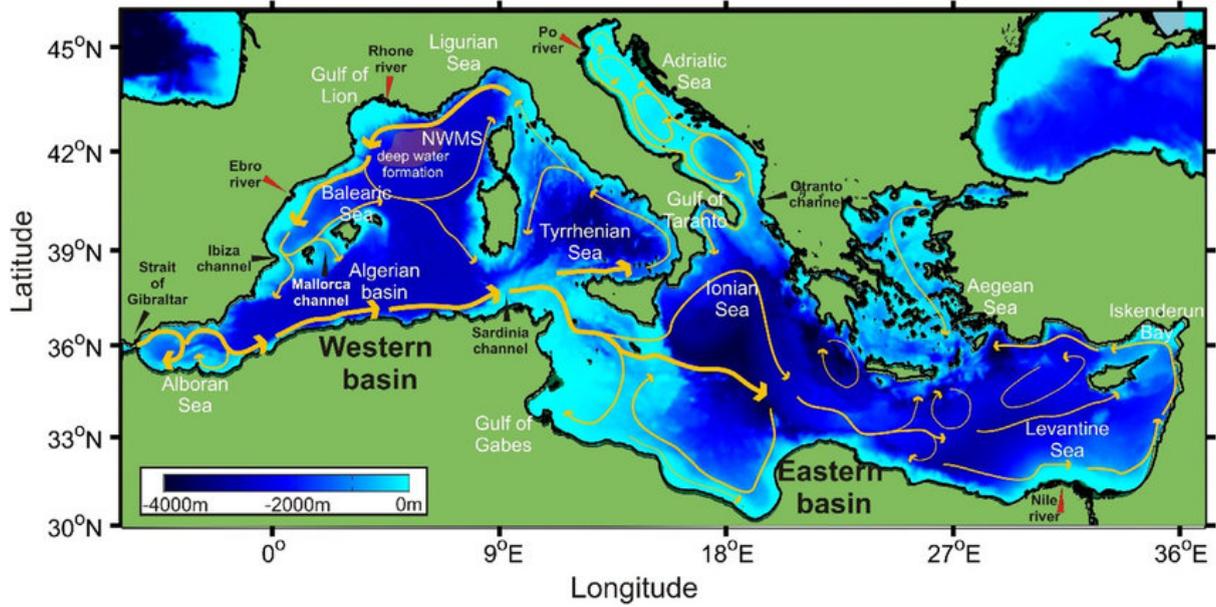


Figure 2 : Représentation schématique des grands courants et gyres de la mer Méditerranée, repris de Hernanz, et al., 2018.

La figure 3 montre les grands gyres de la mer d'Alboran vus grâce à la SST. Sur cette figure, on peut aussi voir l'entrée d'eau de l'océan atlantique.

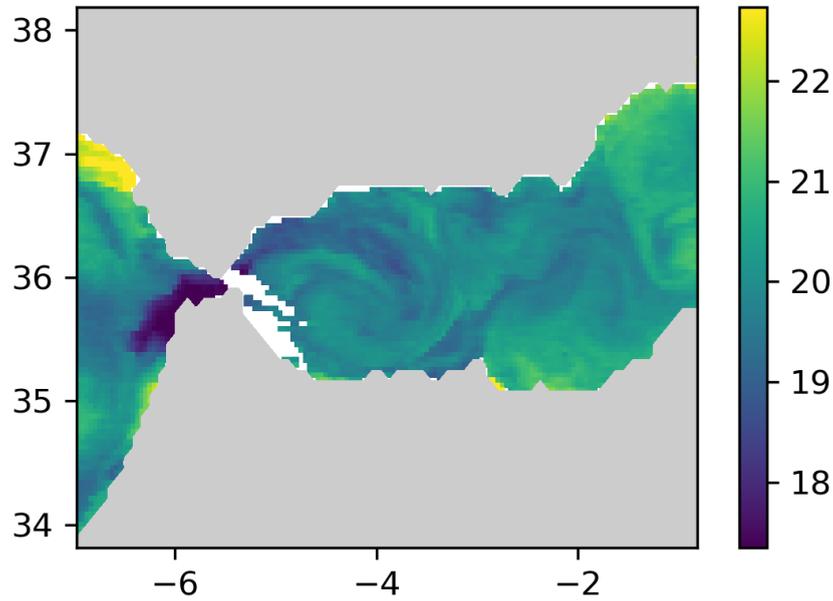


Figure 3 : Représentation des grands courants et gyres de la mer d'Alboran grâce à la SST (en degré Celsius).

## L'intérêt de la température de surface des océans

Tout d'abord, il est essentiel d'avoir des mesures de la SST fiables et d'en comprendre la raison. Une des grandes influences de la SST se fait sur les courants marins. Puisque la température est généralement assez homogène le long des courants marins, les images représentant la SST permettent de les visualiser. C'est notamment vrai pour les gyres et les upwellings. Ces deux phénomènes sont extrêmement importants en océanographie car ils permettent d'en expliquer beaucoup d'autres. Les upwellings sont aussi importants pour les populations locales car ils augmentent la production des eaux côtières.

Les phénomènes d'El Niño et de La Niña, grands acteurs de la production océanique en Amérique du Sud, sont aussi faciles à visualiser et à étudier grâce à la SST. Une SST haute est associée à un effet El Niño fort et une SST faible à un phénomène de El Niño inverse voire à un phénomène de La Niña (Meehl & Geralt, 1987).

La température de surface de l'eau a aussi de nombreux effets météorologiques. Tout d'abord, l'océan est très important au niveau climatique car l'eau a une capacité thermique plus élevée que celle de l'air. L'océan est donc capable de contenir beaucoup plus de chaleur que l'atmosphère. Cette chaleur provient de la radiation solaire, qui va être absorbée par les premiers centimètres d'eau de l'océan (Donner & Large, 2008). C'est donc la SST qui sera importante pour évaluer ce flux de chaleur. Cette capacité de l'océan à retenir la chaleur est très importante car elle permet à celui-ci de mitiger les effets du réchauffement climatique sur l'atmosphère. La SST est aussi utile pour évaluer les flux de chaleur latentes, sensibles et le flux infrarouge.

Cette chaleur accumulée par l'océan est ensuite transférée des tropiques vers les pôles et le reste du globe par la circulation thermohaline océanique, représentée en figure 4. Ces transferts de chaleur dans les océans sont une des causes majeures du climat sur terre.

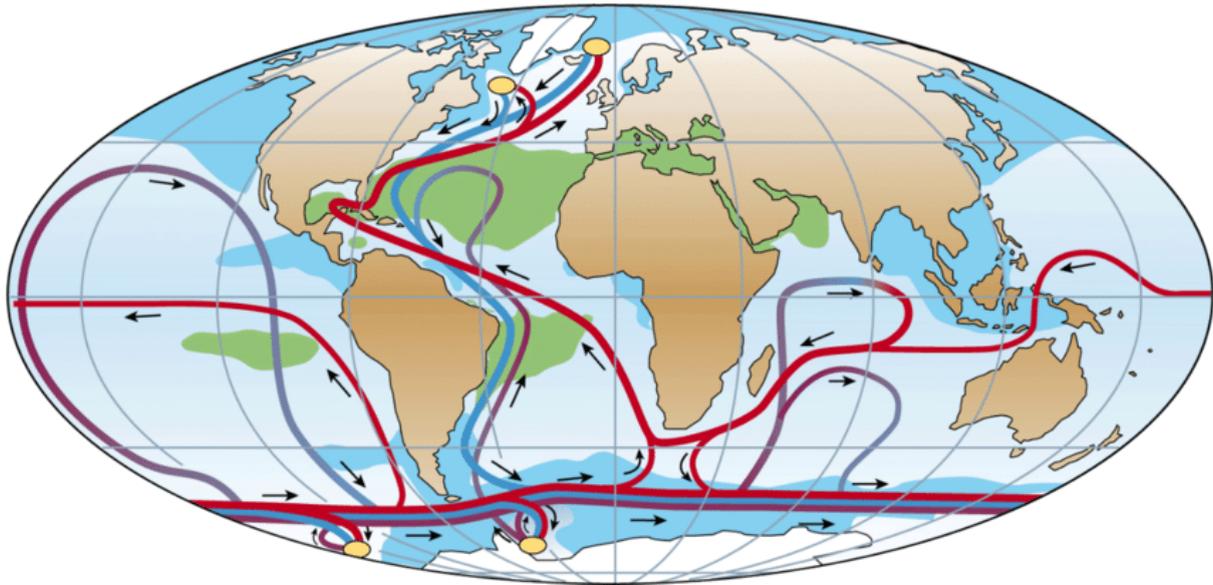


Figure 4 : Représentation schématique simplifiée de la circulation thermohaline océanique globale, repris de Molina-Kesche, 2014.

Ensuite, elle est une bonne indicatrice du potentiel de création d'ouragans. En effet, les ouragans sont créés par de grandes différences de température verticale de l'air. On peut considérer que l'air au niveau de la mer a une température égale à celle de la SST car il y a de nombreux échanges entre les deux milieux, notamment thermiques. En connaissant cette dernière, on a donc un des deux éléments permettant de prédire les formations d'ouragans, l'autre étant le gradient vertical de température de l'air (Palmen, 1948).

Au niveau météorologique, une autre conséquence des changements de la SST sont les moussons, notamment en Inde. En effet, lorsque la SST diminue fortement, on peut remarquer que les vents de l'océan indien sont plus forts et que les précipitations sur les régions touchées par ceux-ci sont aussi plus hautes. La mousson est donc plus forte lorsque la SST diminue. L'inverse est également vrai, une SST particulièrement haute va diminuer la force de la mousson (Zhu & Houghton, 1996).

Enfin, une autre utilisation de la SST est l'étude des populations. La température de l'eau est un facteur écologique très important pour tous les êtres vivants. Connaître cette dernière permet donc de pouvoir évaluer l'évolution de la population de certaines espèces dans des endroits spécifiques ainsi que de pouvoir évaluer le potentiel de survie de celles-ci au réchauffement climatique des océans. Un lien a ainsi été montré entre la SST et les densités

de populations de nombreuses espèces, telles que le dauphin commun (*Delphinus delphis*) dans la mer d'Alboran (Cañadas & Vázquez, 2017), l'espèce de sardine *Sardinops melanostictus* dans l'extension du Kuroshio, au Japon (Noto & Yasuda, 1999) ou encore l'espèce de corail *Balanophyllia europaea*, en mer méditerranée (Goffredo et al., 2008).

Les changements de la température de surface de la mer ont donc des effets sur de nombreux phénomènes, dont certains qui touchent directement à notre société et à la vie terrestre. Il est donc important de pouvoir étudier ceux-ci de manière efficace et donc de savoir les mesurer de manière fiable à plusieurs échelles.

Au niveau de la mer d'Alboran en elle-même, tous ces effets n'y sont pas présents, mais l'étude de la SST dans cette région en particulier reste importante pour y étudier les courants, qui affectent directement la mer méditerranée et le détroit de Gibraltar, ainsi que les populations comme les dauphins précédemment cités. Par exemple, la figure 5 représente la SST en mer d'Alboran le 19 Janvier 2021. On peut y voir un gyre au Nord-Ouest ainsi que de nombreux filaments au Sud-Ouest.

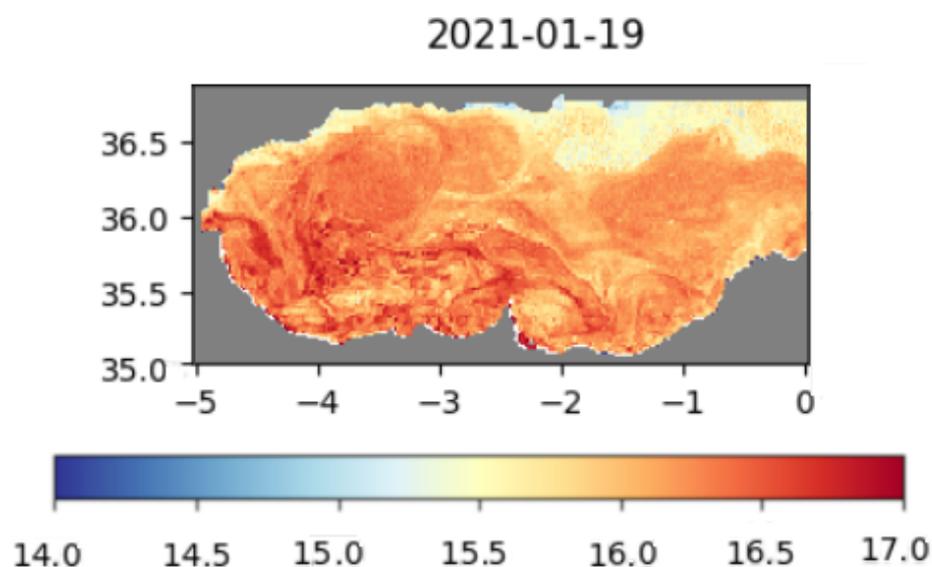


Figure 5 : Représentation graphique de la SST en Mer D'alboran le 19/01/2021.

Le but de ce travail est d'étudier la fiabilité des reconstructions satellitaires de la température de surface des mers et océans, ou SST à plusieurs échelles, et en particulier aux petites échelles. Pour cela, ce travail va étudier des reconstructions faites grâce au programme DINCAE sur la mer d'Alboran. Il est nécessaire de faire des reconstructions des images

satellites car ces derniers utilisent des capteurs pour mesurer la SST. Ces capteurs vont mesurer le rayonnement infrarouge de la terre. Hors, ce rayonnement peut-être bloqué avant d'atteindre le satellite. Les obstacles, qui bloquent les rayons, sont généralement des nuages, très nombreux au-dessus des océans. Les données satellitaires de SST ont donc toujours des fractions manquantes et les reconstructions permettent de combler ces lacunes.

## Les réseaux neuronaux

Il existe plusieurs méthodes pour reconstruire les images satellites. Les méthodes principales sont l'interpolation optimale, DINEOF (Data Interpolating Empirical Orthogonal Functions) (Alvera Azcarate, et al., 2005 et Beckers & Rixen, 2003) et DINCAE. Dans ce travail, nous avons décidé d'utiliser ce dernier car, de par sa nature de réseau neuronal, il est non linéaire, au contraire des deux autres méthodes, l'interpolation optimale et DINEOF. Il est donc bien mieux équipé pour représenter les éléments non linéaires, comme les mélanges et les séparations de différentes masses d'eau.

## Apprentissage automatique

L'apprentissage automatique est la capacité d'une machine à apprendre par elle-même. C'est une partie de l'intelligence artificielle qui permet à la machine d'apprendre et d'augmenter sa précision comme le ferait un être humain. On distingue souvent l'apprentissage automatique surveillé, et l'apprentissage automatique non-surveillé.

L'apprentissage automatique surveillé nécessite des données étiquetées, en entrée comme en sortie. C'est à dire qu'il faut une intervention humaine qui, lors du processus d'apprentissage, va apprendre à la machine à reconnaître les différents types de données qu'elle peut rencontrer. L'apprentissage automatique surveillé est surtout utilisé pour des problèmes de classification et des modèles de régression. Dans notre cas, avec DINCAE, nous n'utilisons pas de données étiquetées car nous n'en avons pas.

Au contraire, l'apprentissage automatique non-surveillé ne nécessite pas de données étiquetées. Ce type d'apprentissage est surtout utilisé afin de trouver des patterns cachés dans un ensemble de données. Il est donc particulièrement utile dans des problèmes de regroupement ou d'association des données ou dans des problèmes de réduction de la dimension des données.

## Perceptron multicouche

Les réseaux neuronaux sont des techniques d'apprentissages automatiques inspirées des systèmes neuronaux biologiques. On retrouve ainsi dans les réseaux neuronaux des couches de nœuds, représentant les neurones, qui vont communiquer entre elles.

Le perceptron multicouche est un réseau neuronal de type feed-forward. L'information ne voyage que dans un seul sens, de l'entrée vers la sortie. Comme son nom l'indique, il y a plusieurs couches de nœuds, avec trois types de couches différents : la couche d'entrée, la couche de sortie et un nombre variable de couches intermédiaires déterminé empiriquement. La figure 6 représente schématiquement le fonctionnement d'un perceptron multicouche. On peut y voir que l'information, représentée par les flèches en noires, rentre par la couche d'entrée (input layer) et continue jusqu'à la couche de sortie (output layer). On peut aussi y voir que tous les neurones d'une couche sont reliés avec tous les neurones de la couche suivante. Un neurone reçoit donc des informations de tous les neurones de la couche précédente. A cette information, un poids est associé permettant de modifier l'importance de l'apport de chaque neurone. En pratique, ces poids sont repris dans une matrice appelée la matrice  $\mathbf{W}$  et dans un vecteur appelé vecteur  $\mathbf{b}$ .

Le perceptron multicouche est surtout utilisé dans des problèmes de classification de patterns, de reconnaissance, de prédiction et d'approximation. Ils servent ainsi à approximer une fonction continue arbitraire.

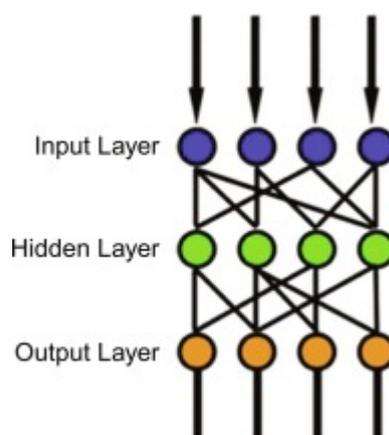


Figure 6 : Schéma d'un perceptron multicouche, repris de Abirami et Chitra, 2020.

## La fonction de perte

La fonction de perte est la fonction qui sert à évaluer la précision d'une solution lors d'un apprentissage automatique. Elle permet donc de quantifier l'erreur du système. C'est une fonction dont on cherche à diminuer la valeur autant que possible. Il faut donc faire une optimisation du système où l'on va chercher à diminuer la valeur moyenne des erreurs. Cette fonction de perte est généralement une différence entre la solution obtenue et la vraie solution.

## La rétropropagation du gradient

La rétropropagation du gradient est un algorithme utilisé pour entraîner les réseaux feed-forward. Le but est de trouver un gradient de la fonction de perte qui va permettre de diminuer l'erreur dans le système. Cet algorithme est fait de manière itérative en modifiant les poids synaptiques à chaque itération afin de les amener aux poids optimaux.

## Réseaux de neurones convolutifs

Les réseaux de neurones convolutifs sont une classe de réseaux neuronaux. Ils sont utilisés pour étudier des données sous forme de tableaux à deux ou trois dimensions, comme par exemple des images, et sont composés de 3 types de couches de neurones.

Tout d'abord, nous retrouvons la couche de neurones convolutifs. Ces neurones servent à l'opération de convolution, c'est-à-dire une opération linéaire spécialisée. Pour faire celle-ci, une petite grille, appelée le kernel, va être appliquée à une partie de la grille des données appelée tensor, comme illustré dans la figure 6. Une fois le kernel appliqué à toutes les parties du tensor, on obtient une nouvelle grille appelée la carte des caractéristiques (feature map) qui va représenter une caractéristique de nos données. Comme on peut le voir sur la figure 7, cette carte est plus petite que le tensor initial. Pour éviter cela, on va ajouter des 0 tout autour du tensor avant l'application du kernel. Cette technique d'ajout s'appelle le zero-padding (Yamashita et al., 2018). Une couche de convolution est un cas particulier d'un perceptron puisque cette opération peut aussi être représentée par une multiplication matricielle.

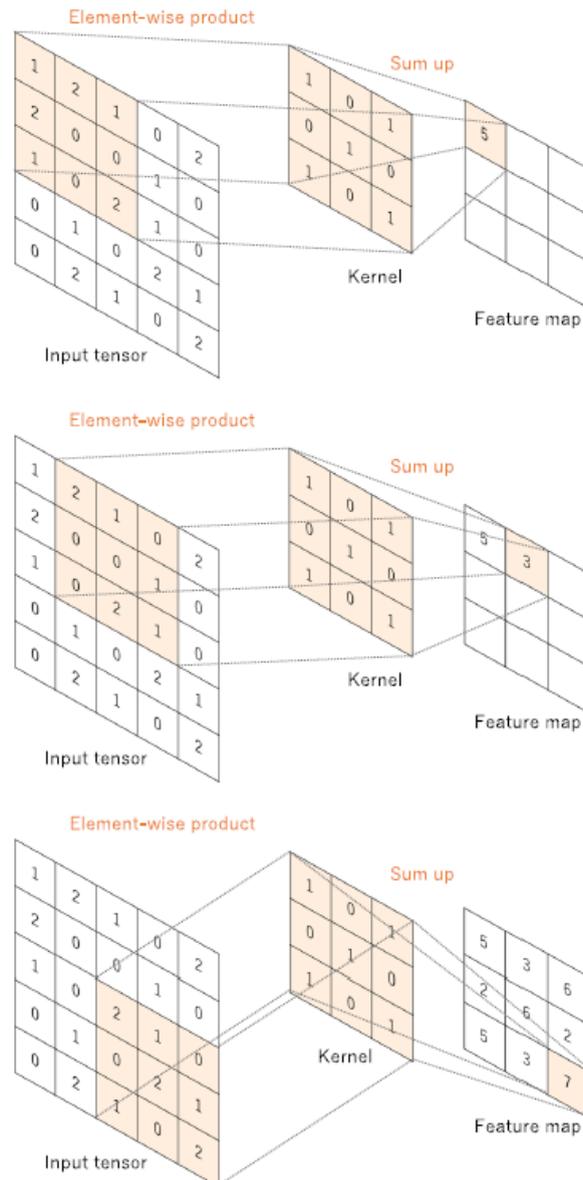


Figure 7 : Fonctionnement d'une opération de convolution avec un tensor 5x5 et un kernel 3x3, repris de Yamashita et al., 2018.

La deuxième catégorie de couche est la couche de neurones de pooling. Ces neurones servent à réduire la taille de la feature map. Pour ce faire, on va soit garder une seule zone de la feature map, on parlera alors de max pooling, ou faire une moyenne sur celle-ci, on parlera alors d'average pooling.

Une fois les feature maps passées à travers ces deux couches, elles sont amenées vers les couches entièrement connectées. Cette couche est notre troisième type de couche. Cette dernière couche va activer les feature maps grâce à une fonction d'activation non-linéaire, souvent une unité linéaire rectifiée. Cette activation va permettre d'obtenir les valeurs finales

des features maps et de calculer la fonction de perte. La figure 8 illustre l'architecture d'un réseau de neurones convolutifs avec les 3 types de couches. Cette architecture est spécifique pour la classification. Dans les autres cas, comme dans celui de DINCAE, il n'y a pas de couche fully connected.

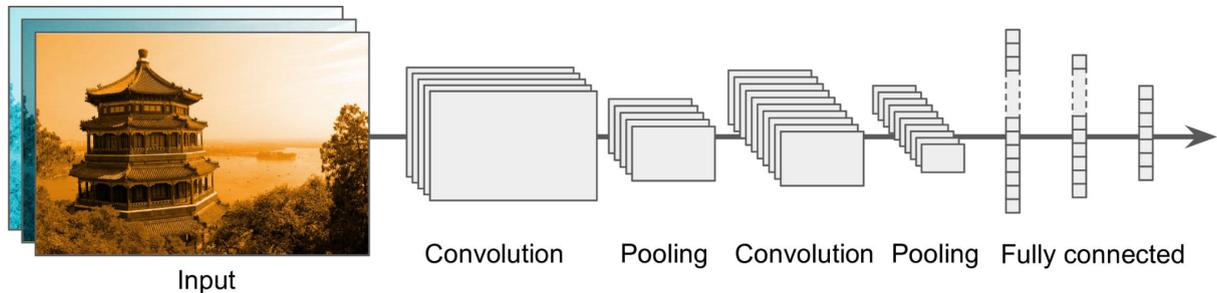


Figure 8 : Architecture d'un réseau de neurones convolutifs pour la classification, repris de *Hands on Machine Learning, Géron, 2019*.

Dans un réseau à neurones convolutifs, on utilise un poids unique dans tous les neurones d'un même noyau de convolution, au contraire des perceptrons multicouches. Cela permet de meilleures performances pour le traitement des images.

## Auto-encodeur

Un auto-encodeur est un réseau neuronal non-surveillé. Il est utilisé pour apprendre une représentation d'un set de données en ignorant les bruits. L'auto-encodeur a une architecture très proche de celle des perceptrons multicouches avec une couche d'entrée, l'encodeur, une couche de sortie, le décodeur, et une ou plusieurs couches cachées. L'encodeur et le décodeur ont généralement le même nombre de neurones et donc de dimensions. Au contraire, la ou les couches cachées ont un nombre de dimensions moins élevées.

## U-Net

U-net est une architecture de réseau neuronal. Elle sert principalement à la segmentation d'images, c'est-à-dire qu'elle va attribuer à chaque pixel une classe différente et les classifier, tout en gardant les informations relatives à chaque pixel. Elle est composée de 2 grandes parties. La première est un réseau de neurones convolutifs. La seconde partie est la partie d'expansion, qui combine convolution et concaténation pour apprendre au réseau la classification spatiale de chaque pixel. Comme on peut le voir dans la figure 9, U-Net est une

architecture en forme de U, d'où elle tire son nom. On peut aussi voir qu'il y a des connexions directes entre les deux parties. Ces connexions sont des "skip-connexions" et elles permettent aux informations de passer directement d'une partie à l'autre. On peut ainsi passer des informations directement des entrées aux sorties de même niveaux. Ces skip-connexions sont très importantes car elles permettent de combler, au moins en partie, les pertes créées par les fonctions de convolution et d'expansion (Siddique et al, 2021).

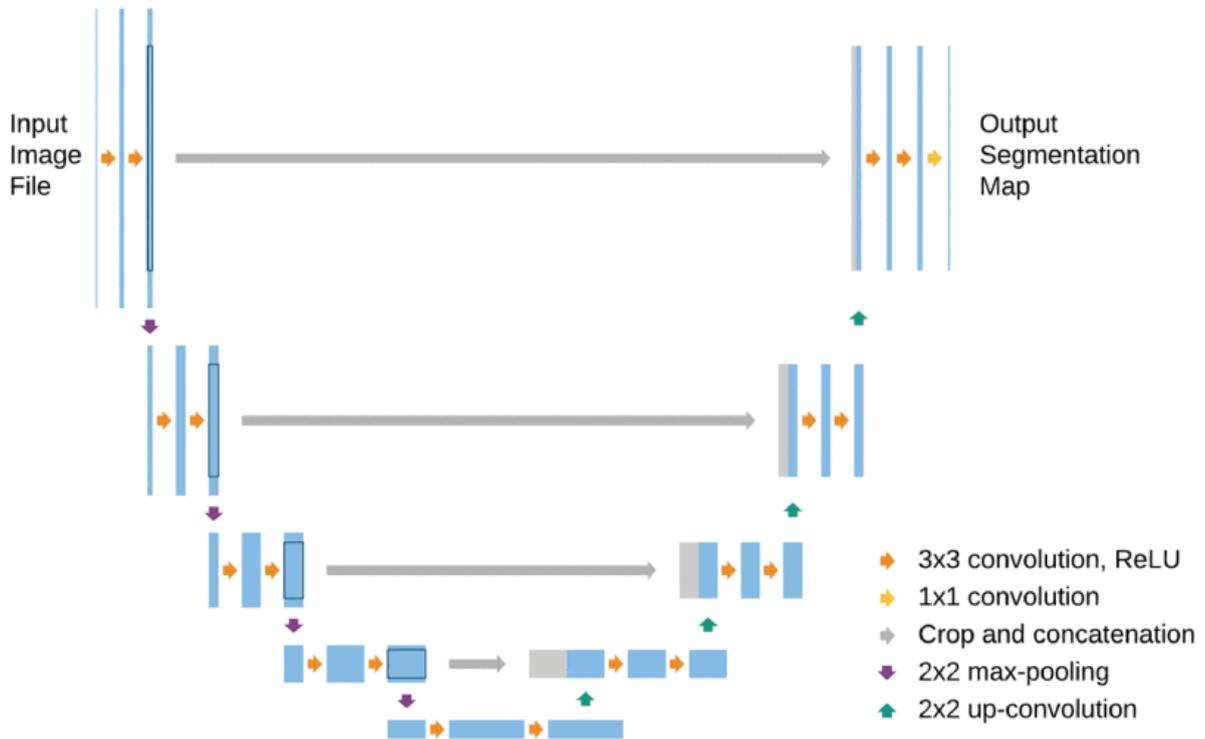


Figure 9 : Schéma de l'architecture U-Net. La partie gauche correspond à la partie de contraction et la partie droite, à la partie d'expansion. Repris de Siddique et al., 2021.

## DINCAE

DINCAE (Data Interpolating Convolutional Auto-Encoder) est le réseau neuronal qui servira de base à ce travail. DINCAE est un auto-encodeur avec une structure U-Net classique. DINCAE utilise donc les caractéristiques d'un réseau U-Net comme les skip-connexions, les couches de neurones de convolutions et de pooling dans la partie de réseau convolutif et les couches de neurones de convolutions et d'interpolations dans la partie d'expansion. Ces différentes parties permettent au programme d'obtenir une image reconstruite sans données manquantes à partir de l'image satellite fournie en entrée. En sortie, en plus de l'image reconstruite, le programme donne aussi une estimation de l'erreur (Barth et al., 2020 et 2022).

## Reconstruction

Les reconstructions de ce travail sont faites grâce au réseau neuronal DINCAE. En entrée du programme, on utilise les données du satellite MODIS (Moderate-resolution Imaging Spectroradiometer) de la NASA (National Aeronautics and Space Administration). Les données sont des données de températures de surface avec une résolution spatiale de 4,63 km et sont publiées par le Physical Oceanography Distributed Active Archive Center (PO.DAAC). Le travail utilise les données allant de -7 degrés à -0,8 degrés de longitude et de 33,8 à 38,2 degrés de latitude. La période étudiée va du 25 février 2000 au 31 décembre 2020, avec 7924 reconstructions pendant cette période.

Chaque reconstruction prend environ 5 à 6 heures de calcul. L'ordinateur utilisé pour les reconstructions est équipé d'un processeur Intel i7-7700 CPU et d'une carte graphique GeForce GTX 1080. Le temps de reconstruction peut donc changer avec un matériel différent. Il est cependant nécessaire d'avoir, dans tous les cas, une carte graphique avec au moins 8 GB de mémoire de marque Nvidia car DINCAE ne fonctionne pas avec les cartes graphiques d'autres marques.

Une des étapes importantes dans le traitement des données est l'ajout de "nuages". Ces "nuages" sont appelés des points de cross-validation et servent à la validation de la reconstruction. Ils entrent aussi en jeu dans le calcul des erreurs car on peut comparer directement des données connues avec des données reconstruites. Pour l'ajout de ces points, le programme va trouver l'instant où on a le plus de données, c'est-à-dire l'image satellite la plus propre, avec le moins de nuages présents. Il va ensuite récupérer le masque de nuage, c'est-à-dire la position des nuages, d'un autre instant choisi aléatoirement et appliquer ce masque sur le moment le plus propre retenu précédemment. Enfin, il va calculer le pourcentage de données manquantes sur la nouvelle image. Si ce pourcentage est en dessous de dix pourcents des pixels totaux des images, il va ajouter un autre masque de nuage aléatoirement choisi à un autre instant. Une fois les dix pourcents atteints, la série d'images est prête à être utilisée pour la reconstruction. Cette étape de cross-validation est faite pour chaque instant.

De nombreux paramètres servent à améliorer la reconstruction et à adapter les réseaux à la zone et à la variable choisie. Lors de ce travail, plusieurs paramètres ont ainsi été modifiés

afin de trouver la combinaison permettant de diminuer l'erreur RMS (Root Mean Square Error, ou racine carré de l'erreur moyenne). Cette erreur RMS est calculée grâce aux points de cross-validation ajoutés précédemment. Pour la calculer, le programme va faire une différence entre la valeur reconstruite sur ces points et la valeur réelle, connue et initiale de ceux-ci. Enfin, on va calculer la racine carrée des carrés de ces différences pour les rendre toutes positives, et les additionner pour trouver l'erreur RMS finale de la reconstruction.

Le tableau complet des essais fait est disponible en annexe A. Les paramètres étudiés lors de ce travail sont les suivants;

- Le nombre d'époch que le programme va faire. Ce paramètre (epochs) représente le nombre d'itérations à travers toutes les données d'entraînement.
- Le nombre de pas de temps qui influence une seule reconstruction (ntime\_win)
- Le poids de chaque couche lors du raffinement de la fonction de perte (loss\_weights\_refine). Ce paramètre est expliqué plus en détail dans l'annexe B.
- Le nombre de couches de neurones et le nombre de neurones à chaque couche (enc\_nfilter\_internal)
- Les niveaux de couches où on inclut une skip connection (skipconnections) (He, et al., 2015)
- La vitesse d'apprentissage de l'optimiseur d'Adam (learning\_rate) (Kingma & Ba, 2017)

Le tableau 1 montre les valeurs par défauts de ces paramètres ainsi que les valeurs optimales trouvées lors des essais. On peut remarquer que les paramètres par défauts sont généralement déjà bons et que la diminution de l'erreur RMS obtenue est faible.

	RMS error	epochs	ntime_win	loss_weight_refine	enc_nfilter_internal	skipconnections	learning_rate
Paramètres initiaux	0.5981	1000	3	0.3-0.7	(32,64,128,256,512)	(1,2,3,4)	0.001
Paramètres optimaux	0.5704	500	3	Pas de raffinement	(32,64,128,256,512)	(1,2,3,4)	0.001

*Tableau 1 : Paramètres initiaux et optimaux pour les reconstructions DINCAE réalisées lors de ce travail.*

Les paramètres `enc_nfilter_internal` et `skip-connections` étant les paramètres touchant directement aux changements de résolution, de par les opérations de concaténation et de convolution, ce sont les deux paramètres qui seront étudiés plus en détails dans la suite de ce travail. Les autres paramètres resteront aux valeurs optimales écrites dans le tableau 1 pour le reste du travail.

Les figures 10 et 11 montrent des exemples de reconstruction faites en utilisant DINCAE. La figure 10 présente les paramètres initiaux, avec les données initiales (a), les données avec les nuages ajoutés (b), les données reconstruites (c) et l'erreur estimée (d). La figure 11 présente ces mêmes graphiques, mais avec les paramètres optimaux trouvés. Pour une meilleure comparaison, les deux figures présentent des données de la même date; le 12 mai 2002.

Date: 2002-05-12

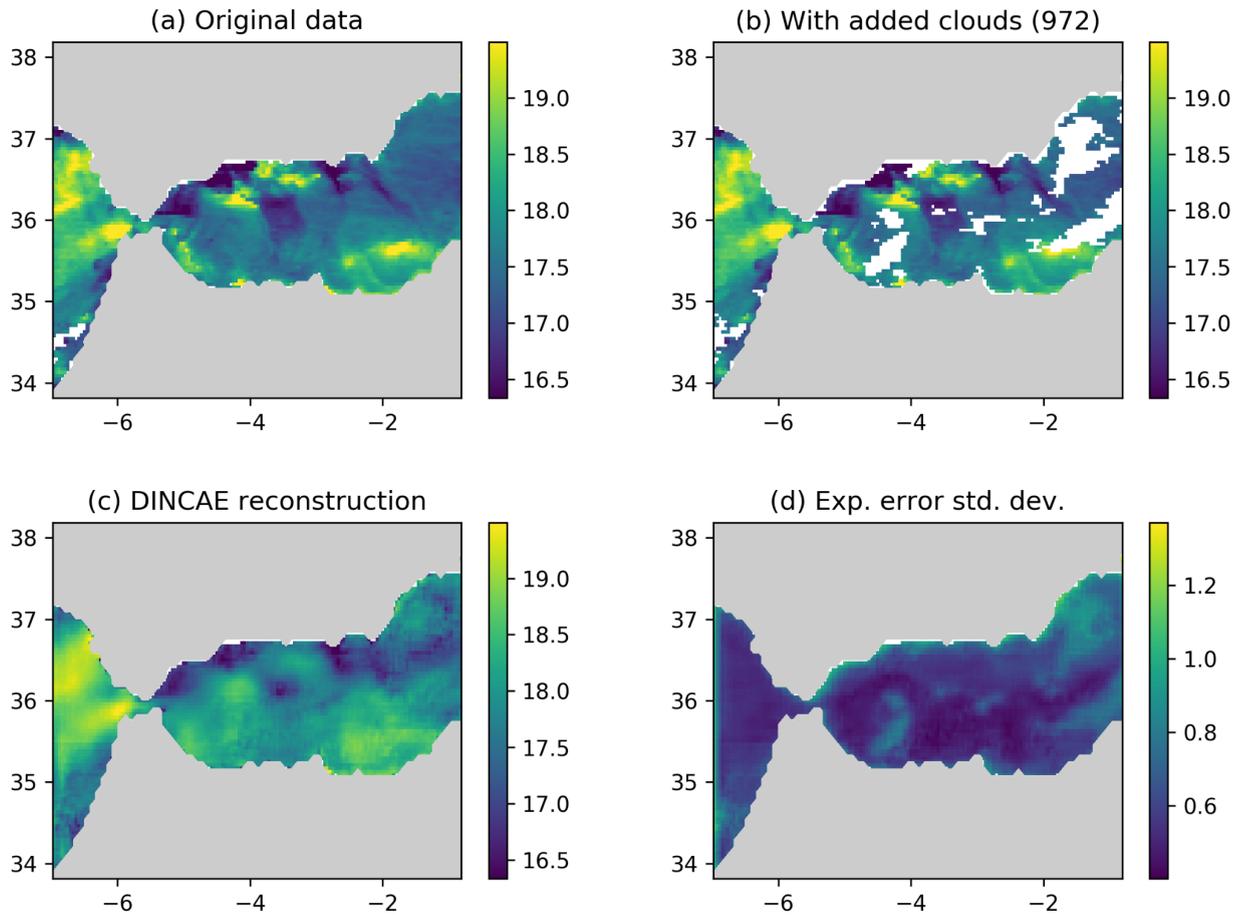


Figure 10 : Exemple de reconstruction du 12 mai 2002 faite avec DINCAE et les paramètres initiaux décrits dans le tableau 1, première ligne. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.

Date: 2002-05-12

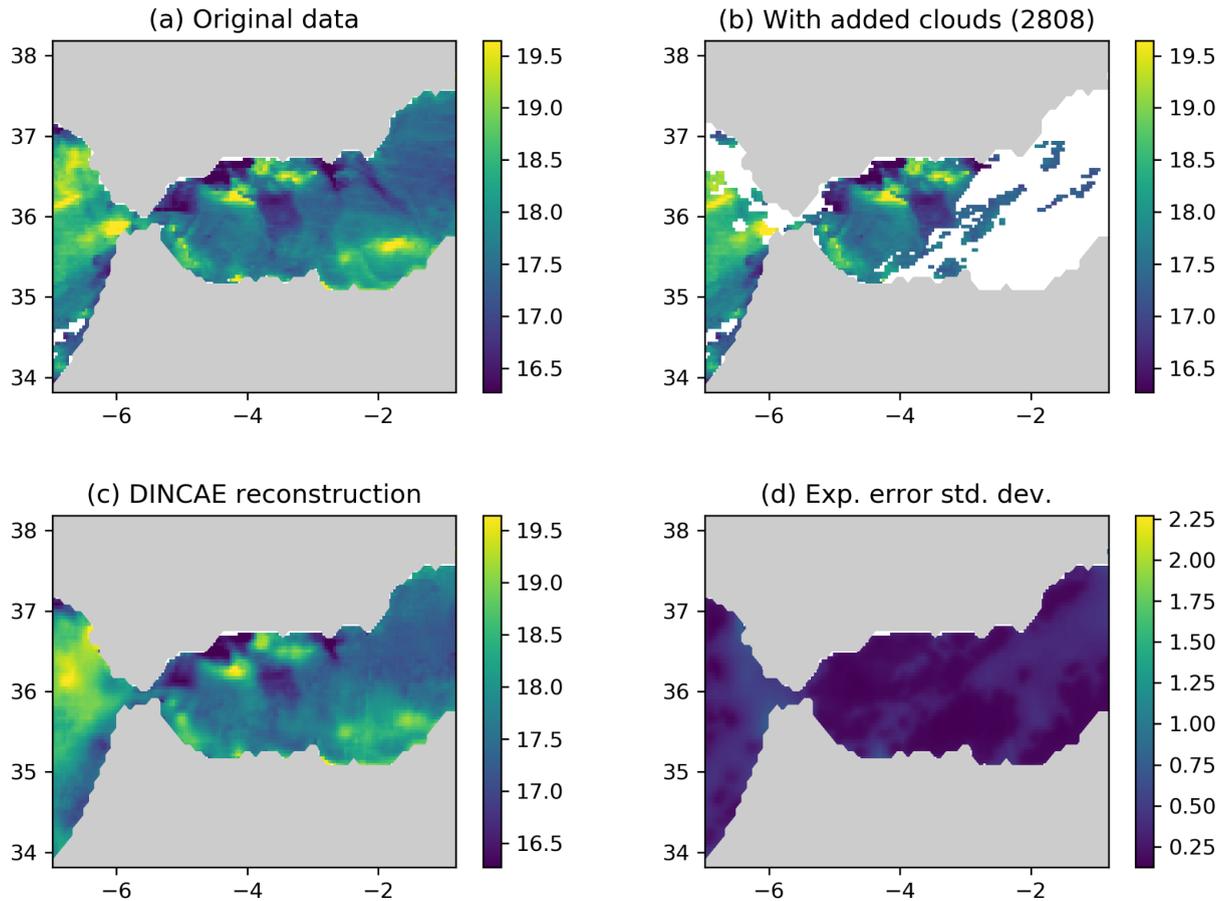


Figure 11 : Exemple de reconstruction du 12 mai 2002 faite avec DINCAE et les paramètres optimaux décrits dans le tableau 1, seconde ligne. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.

Sur ces figures, on peut voir que la reconstruction avec les paramètres optimaux est mieux réussie que celle avec les paramètres initiaux. On peut, en effet, voir les zones plus chaudes du nord de la mer d'Alboran dans la reconstruction de la figure 10, là où elles ne sont pas visibles dans la reconstruction de la figure 11. L'erreur attendue est aussi bien plus faible dans la reconstruction faite avec les paramètres optimaux que dans celle faite avec les paramètres initiaux. Ces deux figures ne restent cependant qu'un exemple, et les différences d'erreurs y sont particulièrement marquées. La différence d'erreur moyenne n'étant que de 0,028 °C sur la totalité des données, comme calculé dans le tableau 1. Il est aussi à noter que, pour ces deux figures, la quantité et la position des pixels de nuages ajoutés sont différentes, et que la comparaison n'est donc pas parfaite. Dans la suite du travail, ces deux paramètres garderont

une valeur fixe. Le masque de nuages ajouté pour la cross-validation sera donc le même pour chaque instant étudié.

## La densité spectrale de puissance

Afin de pouvoir comparer la fiabilité des différents tests aux différentes échelles, il est intéressant de pouvoir comparer la conservation de la variance dans le système. Plus la quantité de variance gardée dans le système est haute, plus les données du système sont fiables. Pour cela, on va utiliser le spectre de puissance aux différentes échelles pour voir quelle quantité de variance est gardée à chaque échelle. En général, les systèmes de reconstructions des données sous-estiment la variance car elle n'est pas prise en compte. En vérifiant si notre système garde de la variance, on améliore donc un des principaux points faible des reconstructions, particulièrement important pour les gradients de température, définis par la variance.

Afin de trouver ce spectre de puissance aux différentes échelles, ou densité spectrale de puissance, nous utilisons la transformation de fourier rapide qui permet de transformer une série de données exprimées en temps vers une série de données exprimées en fréquence. La transformation rapide est choisie plutôt que la transformation discrète car elle est moins complexe en implémentation pour l'ordinateur et demande donc moins de temps et de ressources informatiques (Donnelly & Rust, 2005).

Pour que cette transformation de Fourier soit réalisable, il faut cependant que les données utilisées soient complètes. Or, dans notre cas, nous utilisons des données incomplètes puisque nous utilisons intentionnellement des images satellites avec des nuages et des côtes. Pour résoudre ce problème, il existe plusieurs solutions, comme celle de combler les trous avec des zéros ou celle de les combler avec la moyenne des données. Nous avons décidé de ne pas utiliser ces solutions car elles ajouteraient un biais aux données. A la place, nous avons décidé d'utiliser des séries de tracks. Ces tracks sont pris aléatoirement sur les images et sont donc tous complets. Pour rendre les résultats des transformations rapides de fourier plus comparables entre eux, l'aléatoire dans le choix des tracks a été enlevé en choisissant une graine aléatoire fixe et identique pour chaque test sur chaque figure. Une fois la transformation rapide de Fourier effectuée, nous obtenons, dans notre cas, un vecteur qui est ensuite utilisé pour visualiser la densité spectrale de puissance.

# Résultats

## Les essais

Une fois les premiers paramètres fixés, une autre série de tests a été faite. Cette fois-ci, seul le nombre de couches de neurones et le nombre de skip-connections ont été modifiés. Pour cette nouvelle série de tests, la quantité et la position des points manquants ont aussi été fixées afin que les résultats soient le plus comparable possible. Le masque obtenu lors de notre premier essai est donc gardé et utilisé à nouveau dans tous les essais suivants.

Le tableau 2 reprend tous les tests faits. L'erreur RMS a, encore une fois, été calculée pour ces essais. Les essais avec 5 couches de neurones ont rapidement montré une erreur RMS plus élevée que ceux à 6 couches de neurones. Il a donc été décidé de se concentrer plus intensément sur des essais avec six couches de neurones plutôt que cinq.

Le test 0 est un test sans skip-connections, il sert ainsi de contrôle négatif pour comparer et vérifier que les skip-connections aident à augmenter la fiabilité des résultats à petites échelles. On peut déjà remarquer sur les résultats du tableau 2 que ce test 0 a une RMS error plus élevée que beaucoup de tests avec des skip-connections. Les skip-connections ont donc bien un effet positif sur l'erreur du modèle.

	RMS error	Nombre de couches de neurones et nombres de neurones par couches	Skip-connections
<b>test 0</b>	<b>0.5957</b>	<b>(32,64,128,256,512,1024)</b>	<b>()</b>
test 1	0.5957	(32,64,128,256,512)	(1,2,3,4)
test 2	0.5992	(32,64,128,256,512)	(2,3,4)
test 3	0.5871	(32,64,128,256,512)	(2,3,4,5)
test 4	0.5921	(32,64,128,256,512)	(3,4,5)
test 5	0.5878	(32,64,128,256,512,1024)	(1,2,3,4)
<b>test 6</b>	<b>0.5857</b>	<b>(32,64,128,256,512,1024)</b>	<b>(1,2,3,4,5)</b>

test 7	0.5924	(32,64,128,256,512,1024)	(1,2,3,4,5,6)
test 8	0.5973	(32,64,128,256,512,1024)	(2,3,4,5,6)
test 9	0.5916	(32,64,128,256,512,1024)	(2,3,4,5)
<b>test 10</b>	<b>0.5928</b>	<b>(32,64,128,256,512,1024)</b>	<b>(3,4,5,6)</b>
<b>test 11</b>	<b>0.588</b>	<b>(32,64,128,256,512,1024)</b>	<b>(3,4,5)</b>
test 12	0.5957	(32,64,128,256,512,1024)	(1,2,3)
test 13	0.5918	(32,64,128,256,512,1024)	(1,2,4)
test 14	0.5926	(32,64,128,256,512,1024)	(1,2,5)
test 15	0.597	(32,64,128,256,512,1024)	(1,2,6)
test 16	0.5916	(32,64,128,256,512,1024)	(2,3,4)
test 17	0.5899	(32,64,128,256,512,1024)	(2,4,5,6)
test 18	0.5929	(32,64,128,256,512,1024)	(2,3,5,6)
<b>test 19</b>	<b>0.587</b>	<b>(32,64,128,256,512,1024)</b>	<b>(2,3,4,6)</b>

*Tableau 2 : Tableau reprenant tous les essais faits lors de ce travail ainsi que leurs erreurs RMS.*

Les essais où l'erreur est la plus basse, repris en gras dans le tableau 2, ont été retenus car la faible erreur RMS montre une plus grande fiabilité.

## Analyse d'échelle

La figure 12 montre la densité spectrale de puissance des tests avec la plus haute variance sur les petites échelles du tableau 2. Sur cette figure, en plus des tests à l'erreur la plus basse, nous avons repris le contrôle négatif (test 0) ainsi que la densité spectrale de puissance des données non-reconstruites.

Sur cette figure, l'axe des abscisses représente le nombre d'onde en 1/km. Ce nombre est donc l'inverse de la longueur d'onde. Nous retrouvons ainsi les grandes échelles à la gauche de l'axe, et les petites échelles, à la droite de l'axe. Les courbes sur cet axe démarrent à environ 0,05 de nombre d'onde par km jusqu'à environ 0,8 de nombre d'onde par km. L'axe des ordonnées représente la puissance.

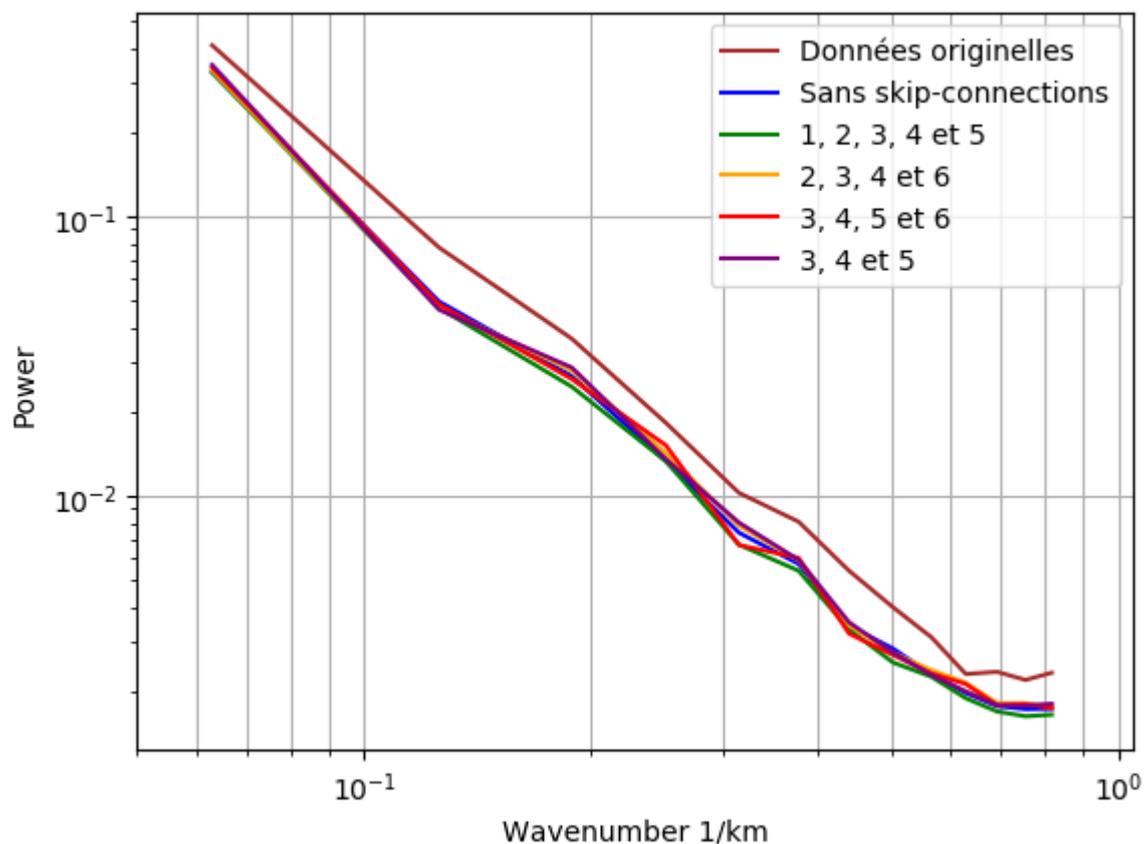


Figure 12 : Densité spectrale de puissance des tests 0, 6, 10, 11 et 19 du tableau 2 et des données originelles non-reconstruites.

La première chose que l'on peut remarquer sur cette figure est qu'une courbe est largement plus haute que les autres. Ce sont les données originelles, avant la reconstruction, qui sont en

effet bien plus fiables que les données reconstruites, peu importe les skip-connections utilisées. On peut cependant voir que cette courbe rejoint les autres courbes vers 0,5 de nombre d'ondes par km. Les différentes courbes des données reconstruites sont, quant à elles, toutes très semblables et à toutes les échelles. On ne peut pas vraiment les différencier sur la figure 12. Il est attendu que cette courbe sans reconstruction soit la plus fiable puisque ce sont les données satellitaires exactes. De plus, puisqu'on utilise que des tracks où les données sont complètes, il n'y a pas non plus de perte de fiabilité des données initiales à cause des nuages.

Pour mieux comparer les plus petites échelles, et les courbes des données reconstruites, la figure 13 montre un zoom sur celles-ci, en démarrant l'axe des abscisses à 0,5 de nombre d'ondes par km (2 km) jusqu'à 0,85 de nombre d'onde par km (1,17 km). Le nombre d'onde maximal possible avec la résolution spatiale de base des données (4,63 km) est de 0,74 par km.

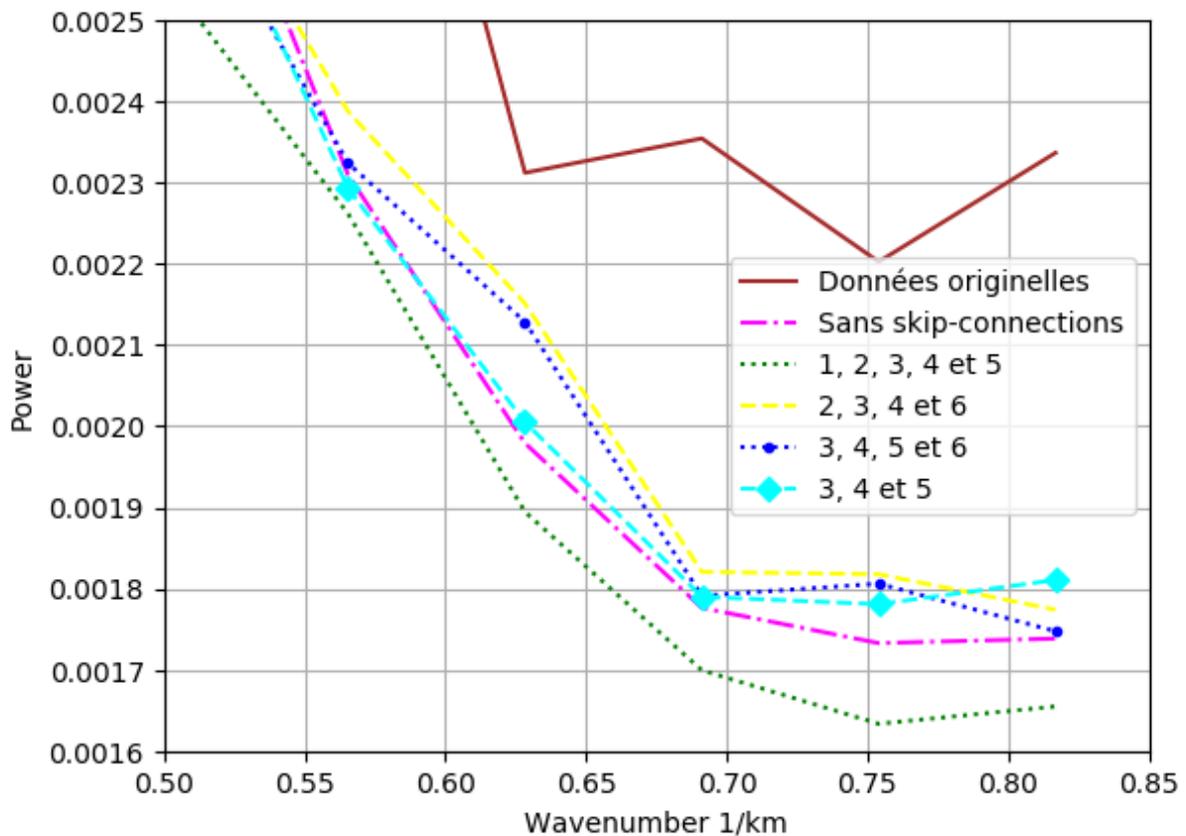


Figure 13 : Zoom sur les plus petites échelles des densités spectrales de puissance des tests 0, 6, 10, 11 et 19 du tableau 2 et des données originelles non-reconstruites.

Sur cette figure, on peut remarquer que les données originelles, la courbe brune solide, restent plus détaillées aux petites échelles que les données reconstruites à plus petites échelles. On peut donc en conclure que, bien que la reconstruction des images avec DINCAE permet d'avoir une vue plus globale de la température de surface dans un lieu choisi, cela se fait avec une perte de variance à toutes les échelles. Cette perte n'est pas compensée par l'ajout de skip-connections, et ce peu importe leurs nombres ou leurs placements sur les différentes couches de neurones. Cette perte est cependant logique et attendue, car il nous est impossible d'avoir les données cachées par les nuages aux petites échelles.

Ces dernières ont cependant un impact visible aux petites échelles. On peut en effet remarquer que les courbes reconstruites ont des densités spectrales de puissances différentes selon les skip-connections utilisées.

L'effet le plus intéressant sur cette figure se note sur la courbe en vert et pointillés, représentant le test 6 avec des skip-connections sur les couches 1, 2, 3, 4 et 5. C'est la courbe qui garde le moins de variance, à toutes les échelles étudiées. C'est aussi le seul test qui possède une skip-connection sur la première couche, celle qui amène directement des données de l'entrée vers la sortie. Hors, ces données de base gardent le plus de variance dans le système. On aurait donc pu s'attendre à une variance élevée pour ce test 6, qui avait aussi l'erreur RMS la plus basse. Il semblerait donc qu'une skip-connection sur la première couche ne soit pas la meilleure pour garder une bonne variance dans le système, et qu'il n'y a pas de proportionnalité entre l'erreur RMS et la capacité du système à garder un maximum de variance.

Au niveau des autres tests, les courbes sont très semblables. La courbe jaune, en traits, et représentant le test 19 avec des skip-connections sur les couches 2, 3, 4 et 6 est généralement la plus élevée des courbes.

Afin de pouvoir encore mieux comparer les différentes courbes avec des skip-connections, la figure 14 représente une comparaison directe de ces courbes avec les données originelles. Toutes les courbes représentent donc le ratio des courbes initiales sur la courbe brune en solide des données originelles. Cette courbe brune en solide est donc présente tout en haut du graphique avec une valeur de 1 sur toute la plage des longueurs d'ondes par km étudiées. Cette manière de présenter les courbes permet aussi de pouvoir différencier les courbes sur des échelles plus grandes que 0,5 de nombre d'ondes par km.

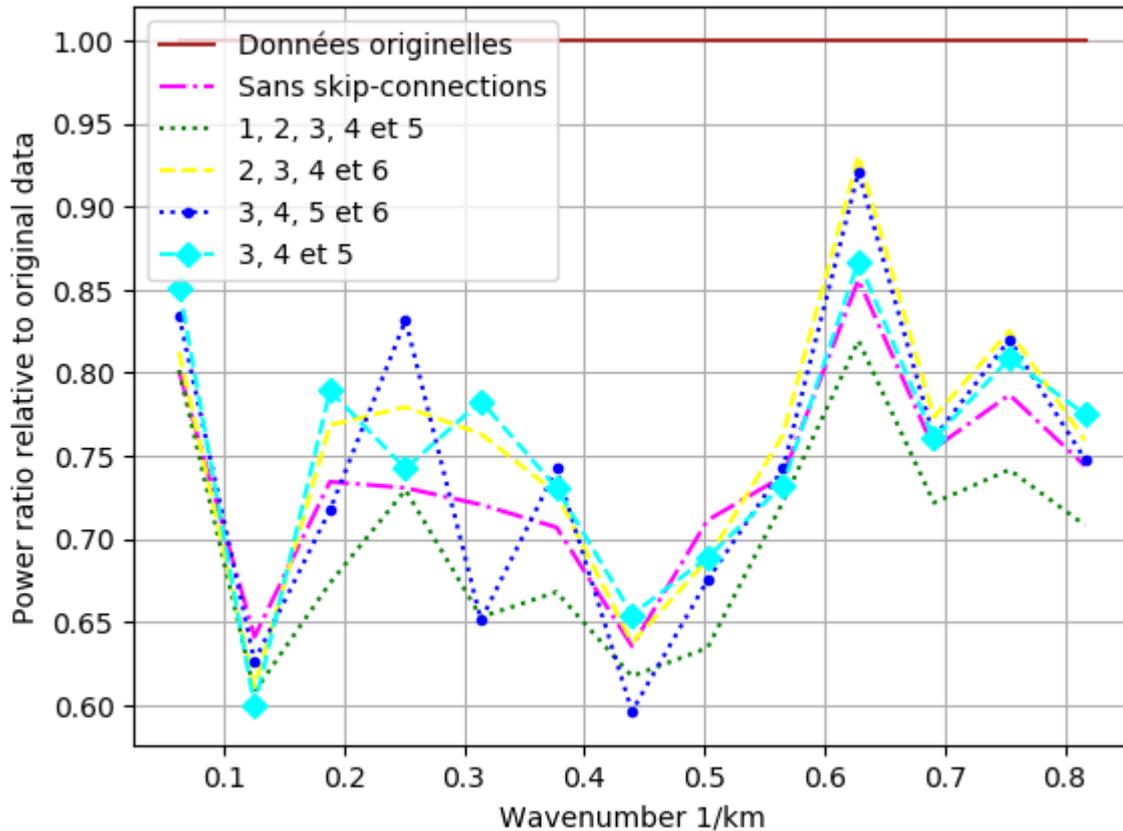


Figure 14 : Comparaison des densités spectrales de puissances des tests 0, 6, 10, 11 et 19 du tableau 2 avec les données originelles non-reconstruites.

Sur cette figure, on peut remarquer que bien que la présentation soit différente, les conclusions faites avec la figure 13 restent vraies pour tous les nombres d'ondes par km plus grands que 0,5. La courbe verte en pointillés, avec des skip-connections sur les couches 1, 2, 3, 4 et 5, est toujours la moins fiable, là où les courbes jaunes en traits et bleues en pointillés avec des gros points sont les plus fiables, sauf aux toutes petites échelles où la courbe violette devient plus fiable.

Aux échelles plus grandes que 0,5, par contre, on a des dynamiques assez différentes. La courbe verte en pointillés, avec des skip-connections sur les couches 1 à 5, reste la plus basse, et la conclusion des effets non-intéressants d'une skip-connection en première couche reste vraie. Mais la courbe cyan en traits et diamants, avec des skip-connections sur les couches 3, 4 et 5, se démarque maintenant à d'autres échelles que les plus petites échelles. Elle est, en effet, la plus efficace aux nombres d'ondes par km proches de 0,2, et aux nombres d'ondes par km d'environ 0,3 à 0,4. La courbe bleue en pointillés et des gros points, avec des

skip-connections sur les couches 3, 4, 5 et 6, pourtant la plus intéressante aux petites échelles, est dépassée par les courbes orange et violette sur une grande partie des grandes échelles. Elle présente cependant un très haut pic à environ 0,25 de nombre d'onde par km. Ce pic se situe juste entre les deux pics de la courbe cyan en pointillés et diamants, et est aussi présent sur la courbe verte en pointillés. Puisque les courbes bleues et vertes sont les deux seules courbes avec une présence de skip-connection en seconde couche, il est fort probable qu'il soit associé à cette dernière.

Les figures 15 et 16 présentent des tests où une seule skip-connection est utilisée, afin de voir si les patterns relevés jusqu'ici peuvent être attribués à l'ajout d'une skip-connection dans une couche en particulier, ou s'ils viennent plutôt des associations entre plusieurs skip-connections sur plusieurs couches. La courbe sans reconstruction et la courbe sans skip-connections sont aussi ajoutées pour faciliter la comparaison.

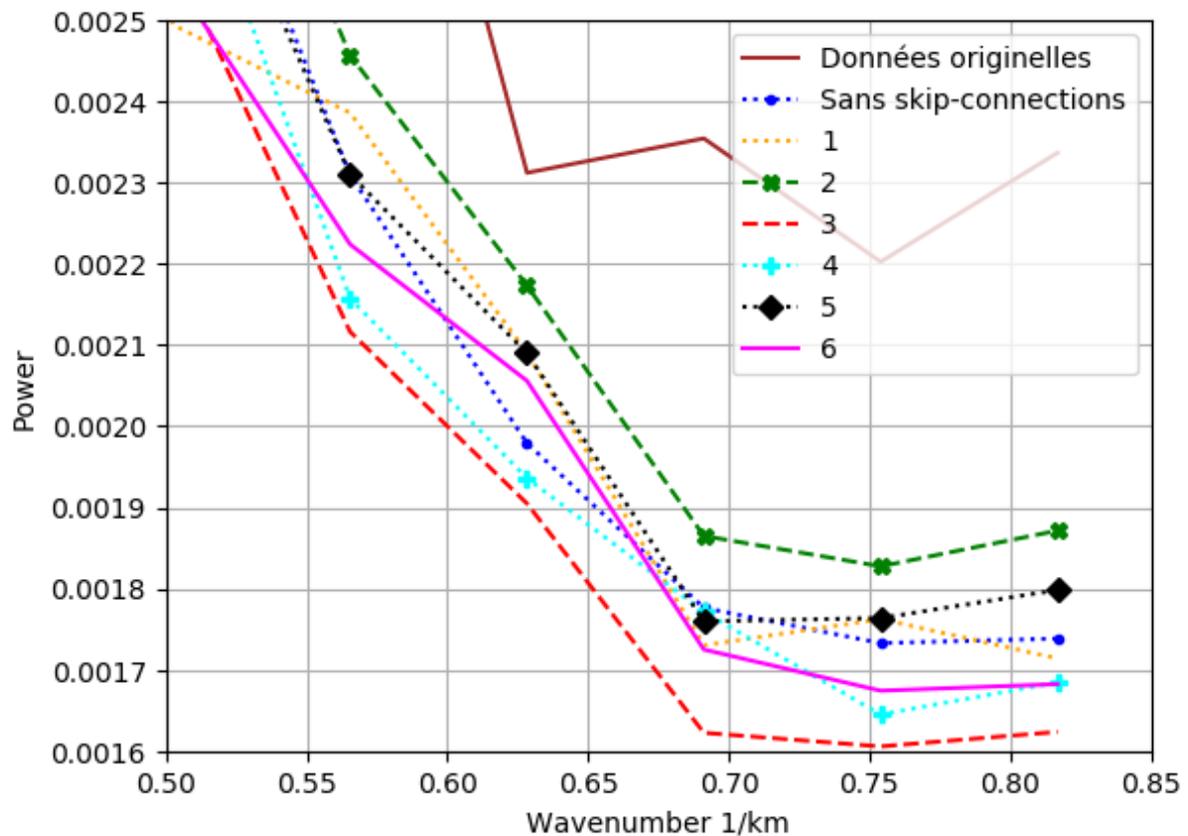


Figure 15 : Zoom sur les plus petites échelles des densités spectrales de puissance des tests avec une seule skip-connection sur chaque couche, ainsi que des données sans skip-connections et des données originelles.

Sur cette figure, on peut voir qu'une seule skip-connection n'est pas toujours suffisante pour avoir une densité spectrale de puissance plus élevée sur les petites échelles que sans skip-connection. En particulier, la courbe rouge en traits avec une skip-connection sur la 3ème couche seulement est particulièrement faible sur les plus petites échelles. Elle est en effet toujours la courbe la plus basse. Une skip-connection sur cette couche se trouve pourtant dans tous les meilleurs essais fait jusqu'ici. Ceci montre bien l'importance des combinaisons des skip-connections.

Une autre couche intéressante est la courbe jaune en pointillés, avec une skip-connection seulement sur la première couche. Cette skip-connection, pourtant jugée peu intéressante voire néfaste pour les résultats, est étonnamment élevée par rapport aux autres couches. On peut voir qu'elle est plus élevée que les, généralement plus élevée, courbes représentant les essais avec une skip-connection seulement sur les couches de neurones 4 et 6, respectivement les courbes en cyan pointillés avec des plus et en rose solide. Ces deux couches étaient pourtant présentes dans la plupart, voir tous, les tests avec les plus hautes variances sur les petites échelles. Elle est aussi évidemment plus élevée que la courbe rouge en traits.

Outre la fiabilité inattendue de cette courbe jaune en pointillés, on peut voir que la courbe la plus fiable sur les plus petites échelles est, et de loin, la courbe verte en traits et croix, avec une skip-connection seulement sur la deuxième couche. Cette courbe dépasse les autres courbes à toutes les échelles possibles. La deuxième couche n'était pourtant pas la couche la plus utilisée dans les courbes avec les plus hautes variances gardées trouvées jusqu'ici, bien que l'ajout d'une skip-connection sur cette couche ait montré des capacités à améliorer la fiabilité à certaines échelles, et notamment à environ 0,25 de nombre d'ondes par km.

Au final, on peut voir qu'il n'y a pas de vraie corrélation forte entre les skip-connection individuelle et leur présence ou absence sur les courbes à plusieurs skip-connections sur différentes couches. Il n'est donc pas possible de trouver une courbe idéale à plusieurs couches grâce à la figure 15.

Cependant, on voit bien sur cette figure une courbe avec une densité spectrale de puissances généralement très haute. C'est-à-dire la courbe verte en pointillés et croix, avec une skip-connection unique sur la seconde couche de neurones. La figure 16 montre cette courbe

ainsi que les courbes avec les plus hautes variances aux petites échelles afin de pouvoir comparer leurs fiabilités respectives.

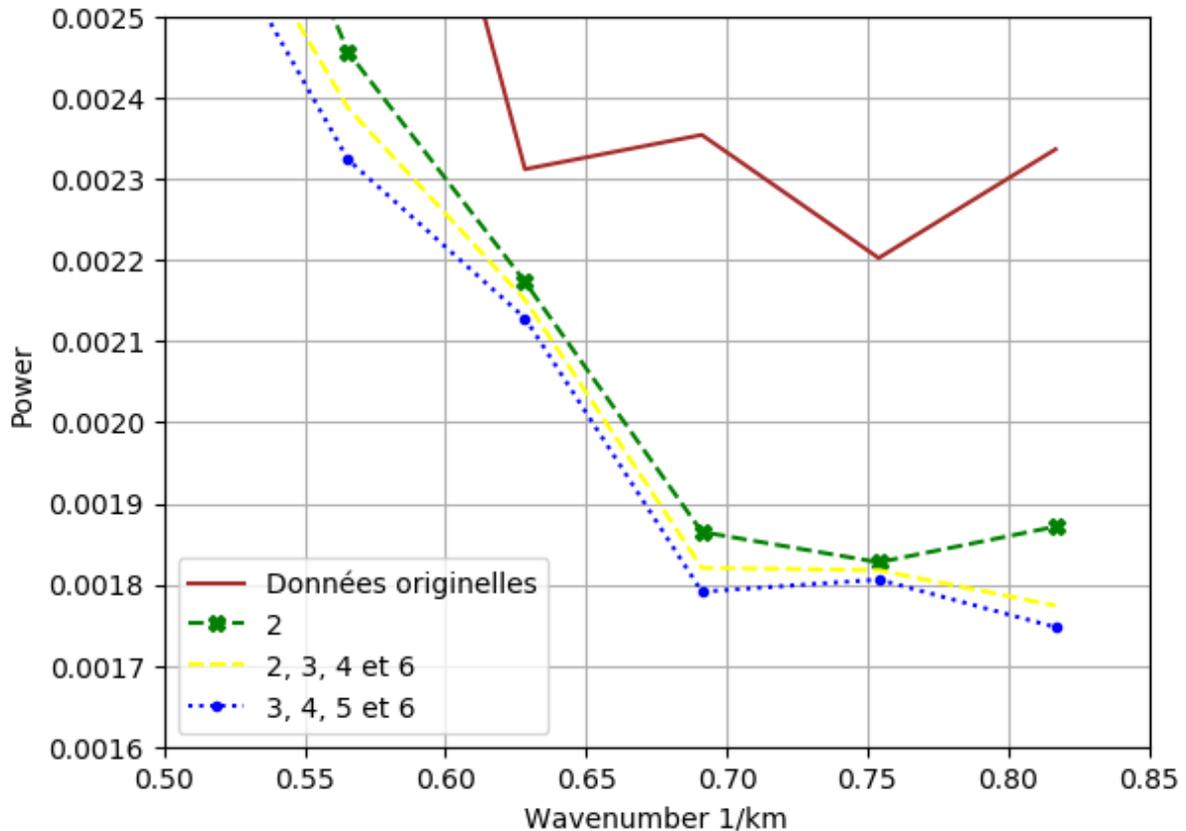


Figure 16 : Zoom sur les plus petites échelles des densités spectrales de puissance des tests 10 et 19, ainsi que du test avec une seule skip-connection sur la seconde couche, des données sans skip-connections et des données originelles.

On peut voir que la courbe verte en traits et croix, avec une skip-connection sur la deuxième couche uniquement, est plus fiable que les autres courbes du graphique, courbe brune solide des données originelles mise à part. Aux plus petites échelles, il est donc plus fiable de ne posséder qu'une seule skip-connection sur la seconde couche.

La configuration la plus fiable pour les plus petites échelles est donc dépendante des échelles que l'on veut étudier en particulier, mais sur la majorité des petites échelles, c'est la configuration avec une seule skip-connection sur la seconde couche qui garde le plus de variance. Elle n'a cependant pas l'erreur RMS la plus basse, avec 0,5907 contre 0,5857 pour le test 6 avec une skip-connection sur les cinq premières couches. Il est donc nécessaire de

choisir si l'on préfère garder de la variance ou si l'on préfère avoir une erreur RMS la plus basse possible.

Une autre étude possible sur les avantages des différents placements de skip-connections sur les couches de neurones est de regarder leur fiabilité aux différentes saisons.

## Etude saisonnière

Une autre possibilité intéressante est d'étudier l'évolution de la densité spectrale de puissance en fonction des saisons. Les figures 17, 18 et 19 représentent les courbes de densité spectrale de puissance pendant les différentes saisons ainsi que la courbe annuelle pour comparaison.

Les périodes des différentes saisons ont été choisies grâce aux définitions des saisons de EMODnet, European Marine Observation and Data Network, qui considère que l'hiver en Mer Méditerranée démarre le premier janvier jusqu'au 31 mars, le printemps du premier avril jusqu'au 30 juin, l'été du premier juillet jusqu'au 30 septembre et enfin l'automne du premier octobre jusqu'au 31 décembre.

Ces trois figures présentent le test 0, sans skip-connections, le test 6, avec des skip-connections dans les cinq premières couches car ce test a la meilleure RMS error et était le test le moins fiable sur la figure 13, et, enfin le test avec une seule skip-connection sur la seconde couche car c'est le test avec la plus haute variance sur la majorité des échelles sur la figure 16.

Toutes ces figures sont présentées en axes logarithmiques sur les deux axes. L'axe des ordonnées représente la densité spectrale de puissance et l'axe des abscisses représente le nombre d'ondes par km. Ce nombre est égal à 1 divisé par la longueur d'onde. Nous retrouvons donc les petites échelles, celles qui nous intéressent particulièrement, à droite, dans les grands nombres d'ondes par km et les grandes échelles, à gauche, dans les petits nombres d'ondes par km. Nous avons ainsi un axe qui montre les résolutions allant de 0,06 nombre d'onde par km jusqu'à  $10^0$  ou 1 nombre d'onde par km.

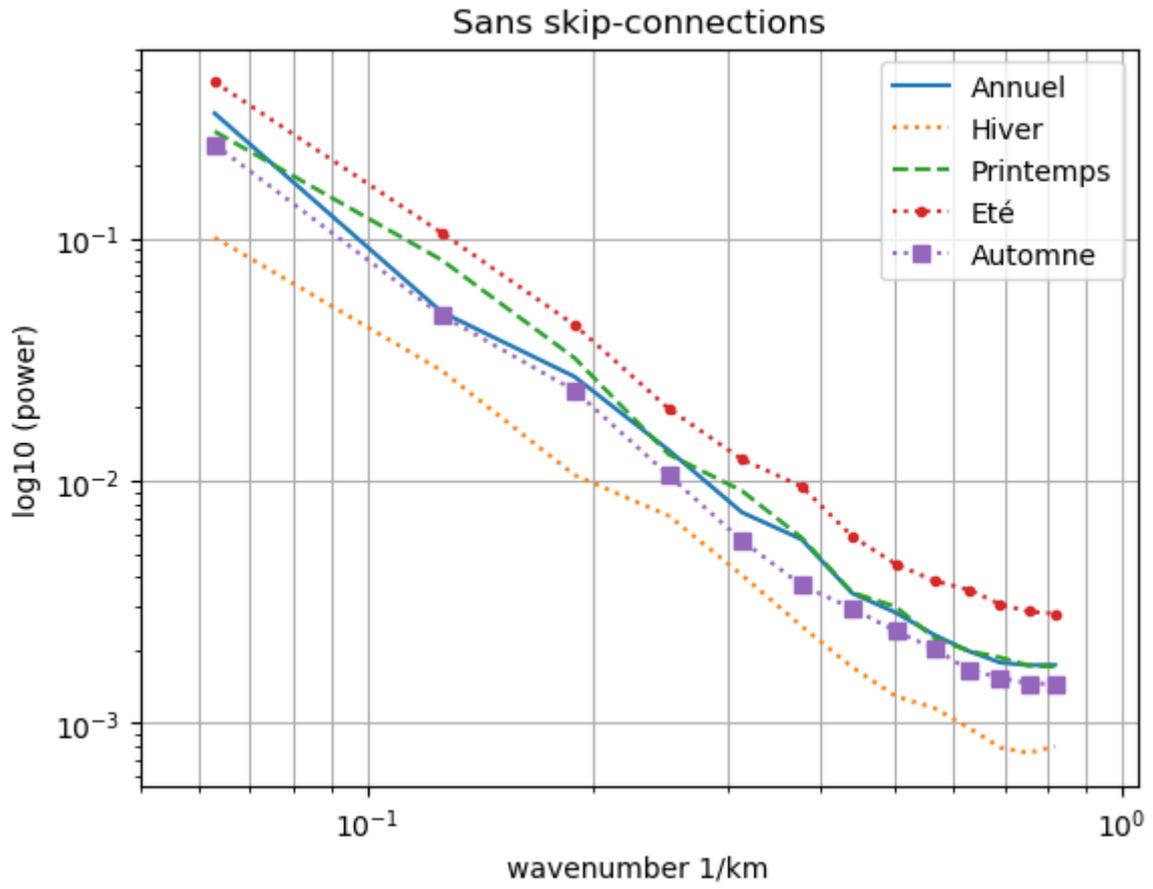


Figure 17 : Courbes saisonnières de la densité spectrale de puissance sans skip-connections.

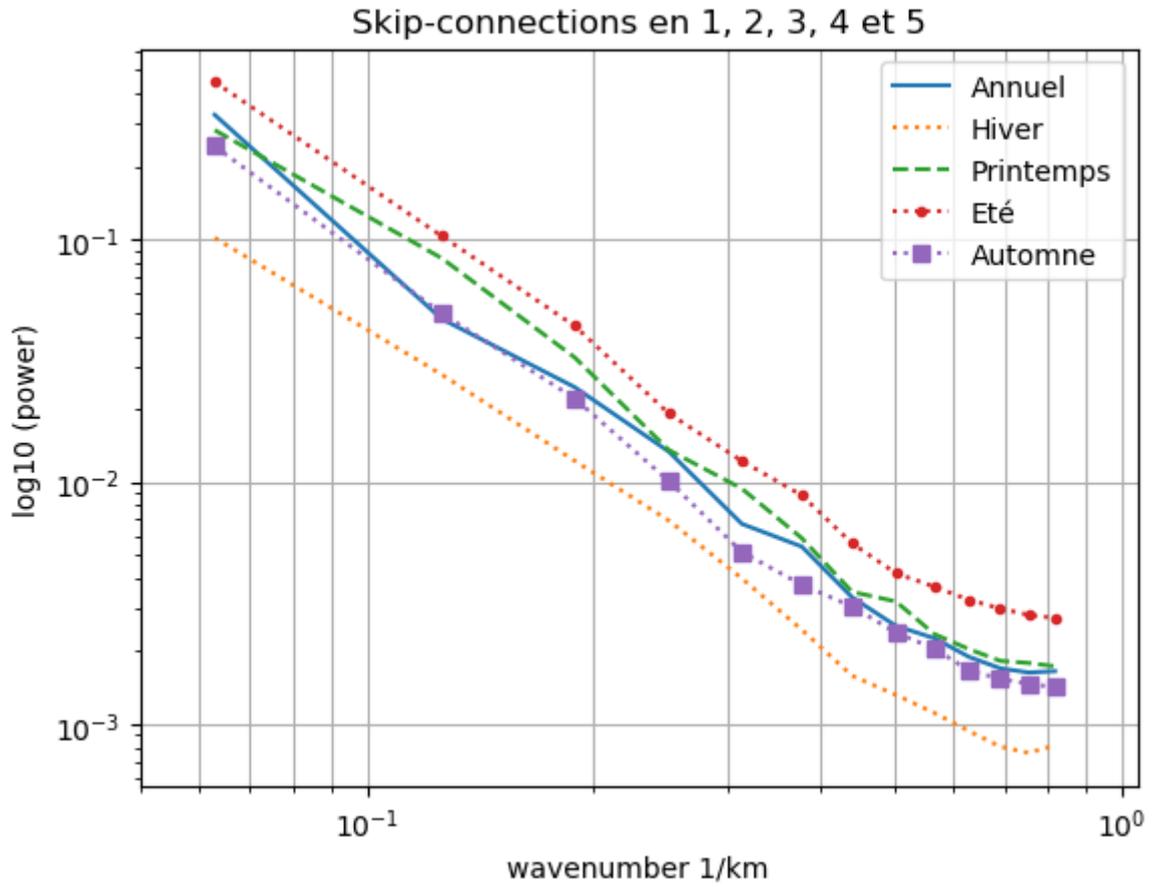


Figure 18 : Courbes saisonnières de la densité spectrale de puissance avec des skip-connections sur les couches 1, 2, 3, 4 et 5.

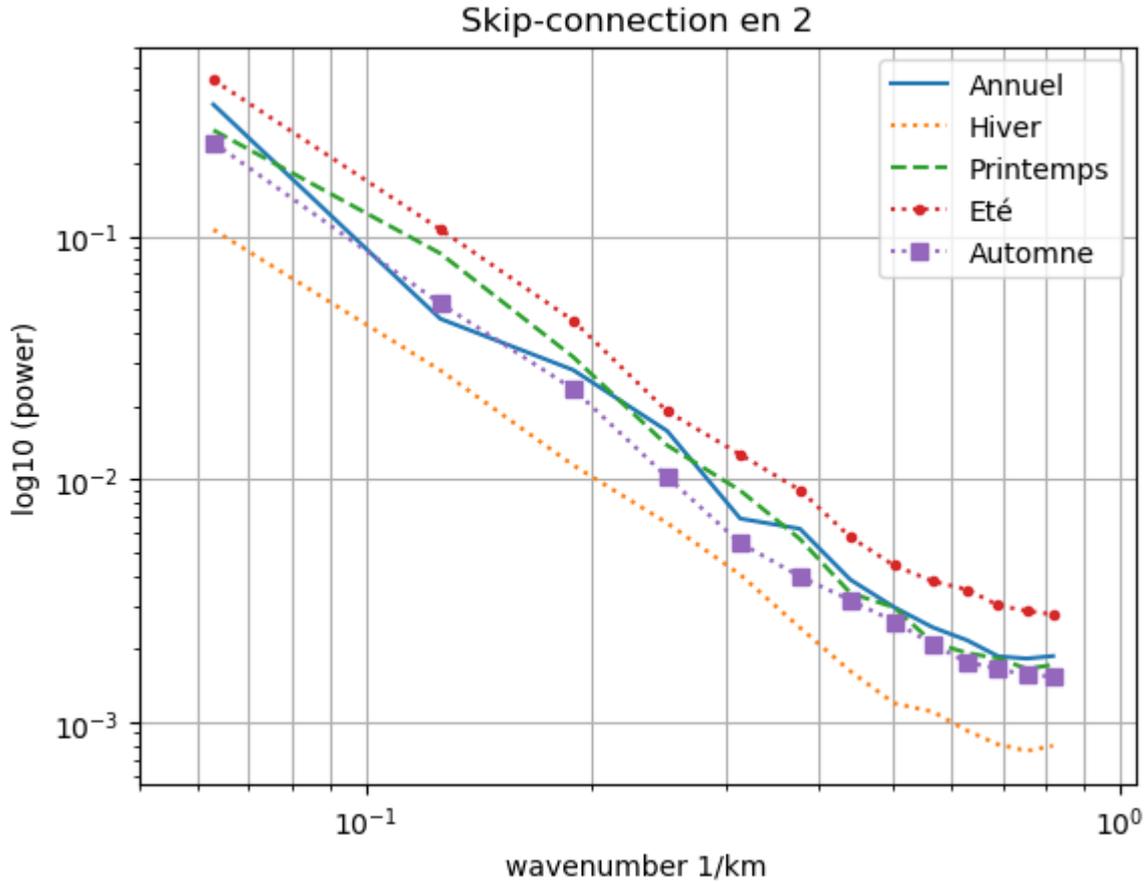


Figure 19 : Courbes saisonnières de la densité spectrale de puissance avec une skip-connection sur la seconde couche.

Sur ces trois figures, on peut voir que la courbe en été est celle qui garde le mieux la variance à toutes les échelles. Elle est aussi fortement linéaire. Au contraire, la courbe en hiver garde toujours le moins de variance, et généralement moins linéaire. Ces tendances sont probablement dues aux changements de températures, puisqu'on peut voir que l'ordre des courbes correspond à l'ordre saisons classées de la plus chaude (été) à la plus froide (hiver). Ces changements de température sont aussi accompagnés d'une plus forte variabilité des températures de surface. En effet, Shaltout et Ornsted (2014) ont calculé un coefficient de variation des différentes saisons en calculant le ratio de l'écart-type sur les valeurs moyennes de SST des différentes saisons et ont trouvé qu'en mer Méditerranée, ce coefficient était de 14,4% en hiver contre seulement 4,8% en été. Ces différences peuvent expliquer les changements de linéarité des courbes, et ont probablement aussi un rôle à jouer dans les différences visibles sur les courbes des densités spectrales de puissance.

Les courbes en automne et au printemps sont, elles, beaucoup plus proches des courbes sur l'année complète. La courbe du printemps est légèrement plus haute que celle de l'automne, mais les deux courbes sont très proches, de par leur petite différence de températures moyennes.

Une autre possibilité de regarder les différents tests aux différentes saisons est de comparer directement les courbes de densité spectrale de puissance des différents tests à chaque saison pour voir si un test en particulier montre des performances inattendues sur une saison particulière. Les figures 20, 21, 22 et 23 montrent donc les courbes des tests étudiés auparavant, c'est-à-dire le test sans skip-connection, le test 6 avec des skip-connections sur les cinq premières couches et, enfin, le test avec une seule skip-connection sur la seconde couche, aux différentes saisons, respectivement en hiver, au printemps, en été et en automne.

Il est cependant à noter que, comme remarqué précédemment, les différentes saisons ont des valeurs sur l'axe des ordonnées assez différentes. Ces axes ne sont donc pas les mêmes dans les figures 20, 21, 22 et 23 qui ne sont pas faites pour être comparées entre elles. La valeur de densité spectrale la plus haute sur la figure 20, en hiver, est plus petite que la valeur la plus haute sur la figure 23, en été.

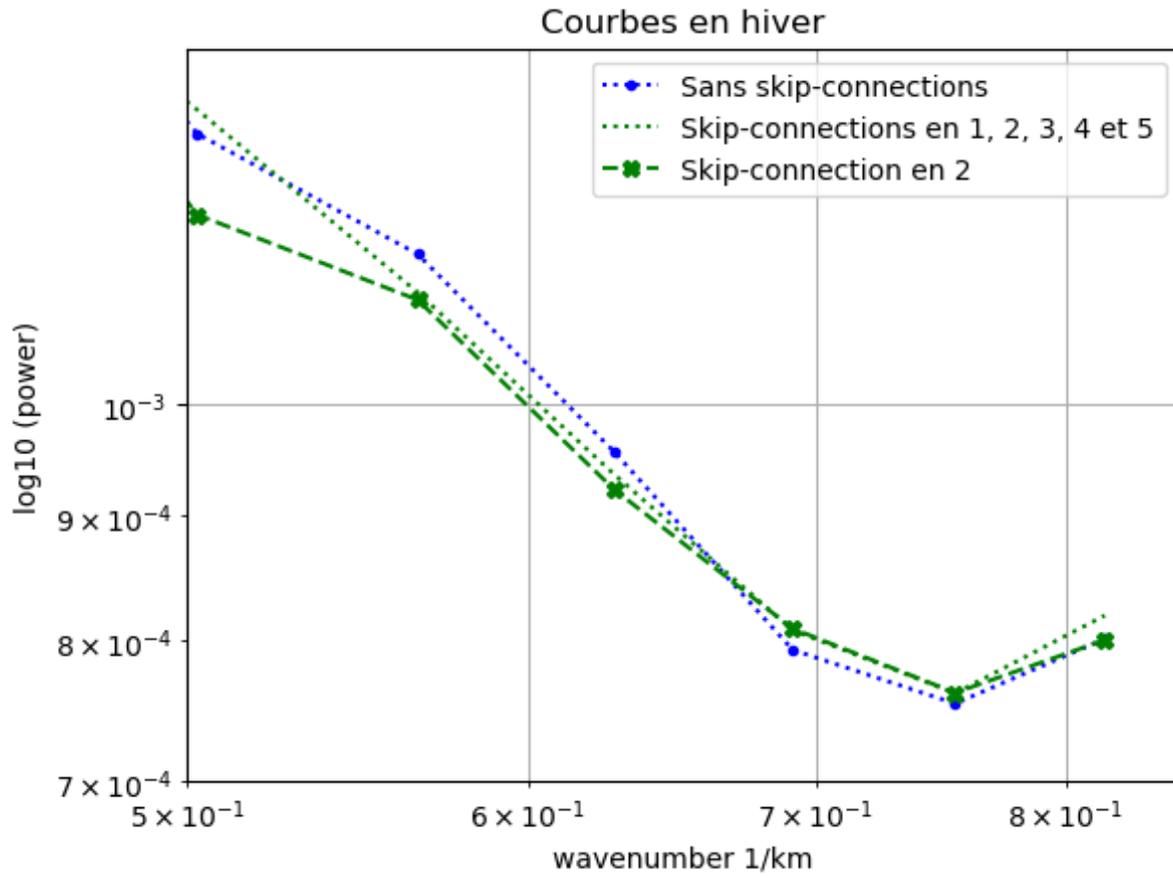


Figure 20 : Courbes en hiver des tests 0, 6 et avec une skip-connection seulement sur la seconde couche.

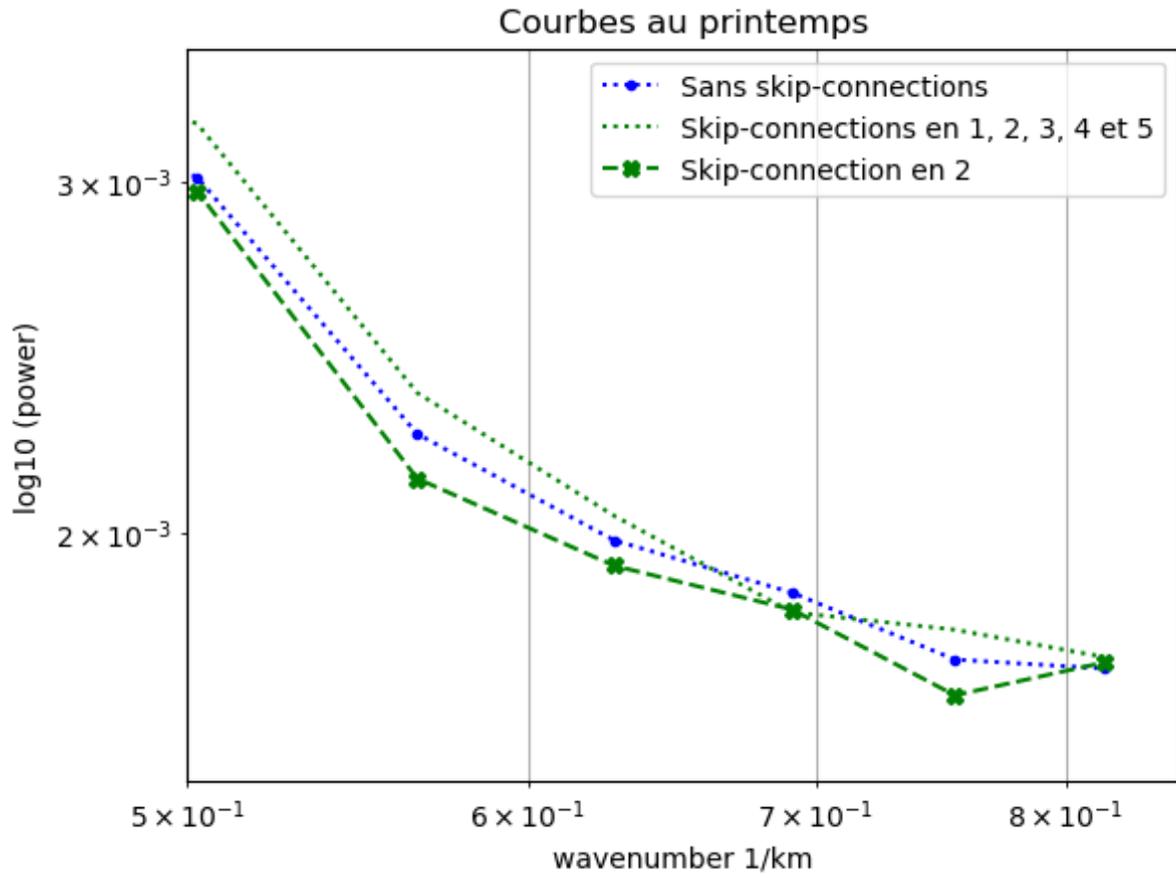


Figure 21 : Courbes au printemps des tests 0, 6 et avec une skip-connection seulement sur la seconde couche.

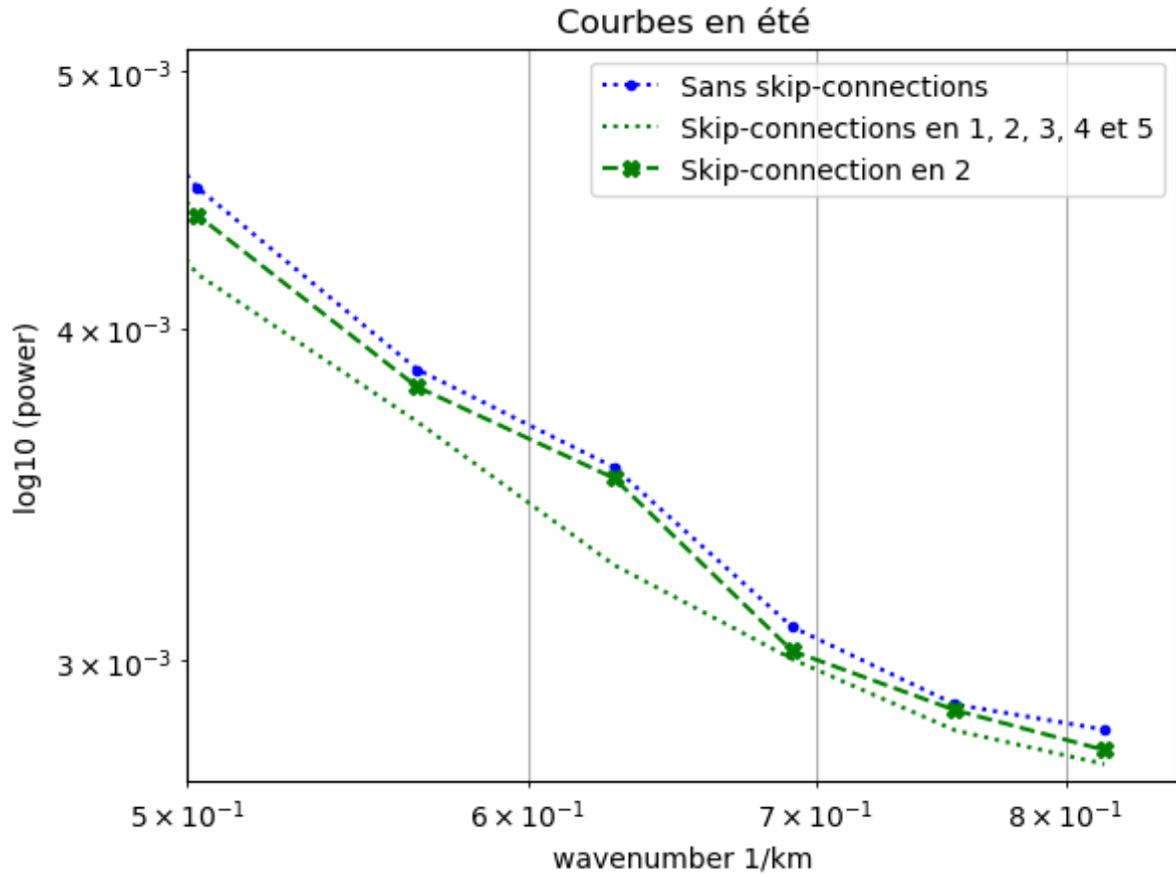


Figure 22 : Courbes en été des tests 0, 6 et avec une skip-connection seulement sur la seconde couche.

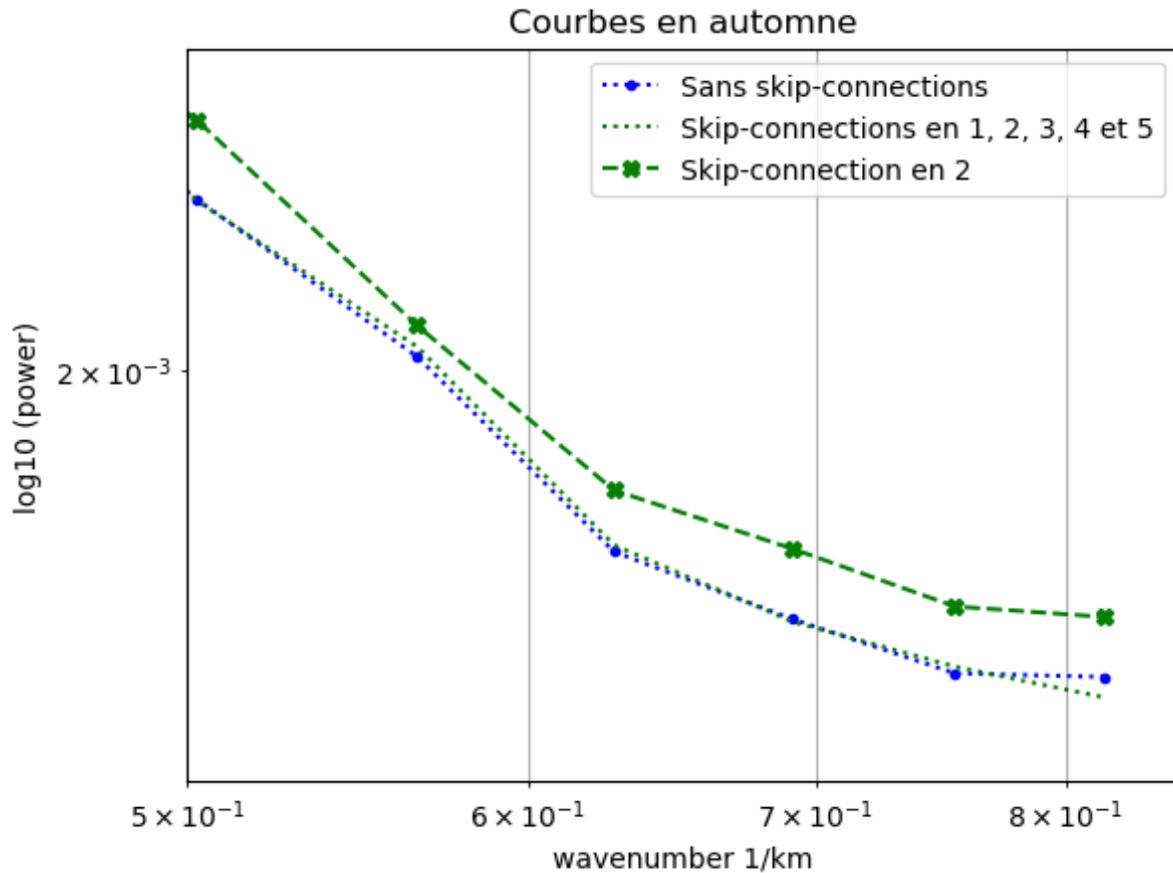


Figure 23 : Courbes en automne des tests 0, 6 et avec une skip-connection seulement sur la seconde couche.

Grâce aux figures 20, 21, 22 et 23, on peut remarquer que les différents tests ne présentent pas de hiérarchie de fiabilité identique selon les saisons. Dans le cas où on essaierait d'étudier la SST sur une saison spécifique comme, par exemple, si on veut étudier le gyre anticyclonique à l'est de la mer d'Alboran, présent seulement en été et en automne, il peut être plus intéressant de choisir des skip-connections sur des couches différentes que celles qui montrent de meilleures performances sur une année entière. En effet, le test avec une seule skip-connection sur la seconde couche était généralement le plus fiable lorsqu'on étudiait une année entière sur la figure 16. Il est pourtant ici le moins fiable en hiver et au printemps sur les figures 20 et 21.

## Exemples de reconstructions

Maintenant que l'on a trouvé une configuration au niveau des skip-connections assez intéressante, nous pouvons regarder les reconstructions. Le tableau 3 reprend les paramètres optimaux utilisés pour ces reconstructions.

	RMS error	epochs	ntime_win	loss_weight_refine	enc_nfilter_ternal	skipconnections	learning_rate
Paramètres optimaux	0.5907	500	3	Pas de raffinement	(32,64,128,256,512)	(2)	0.001

Tableau 3 : Tableau reprenant les paramètres de la meilleure reconstruction.

Comme expliqué auparavant, les erreurs des reconstructions ne sont pas les mêmes d'un instant à un autre. Les figures 24 et 25 montrent donc les deux extrêmes reconstructions possibles faites avec ces paramètres optimaux. La figure 24 montre en effet une reconstruction avec une erreur assez haute et beaucoup de perte d'éléments, là où la figure 25 montre une reconstruction mieux réussie avec des éléments importants qui ont pu être bien reconstruits.

Date: 2001-04-22

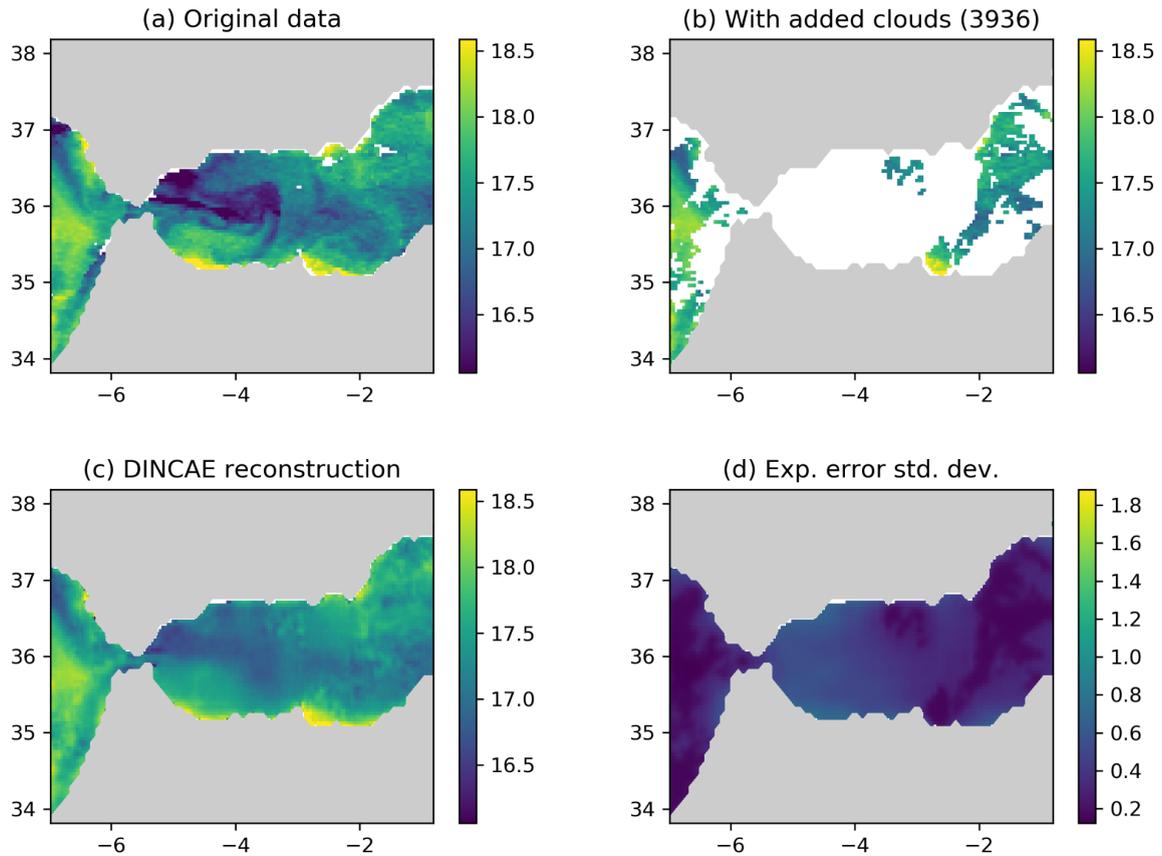


Figure 24 : Exemple de reconstruction du 22 avril 2001 faite avec DINCAE et les paramètres décrits dans le tableau 3. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.

Sur cette figure 24, on peut voir, sur le panel b, que le programme a enlevé beaucoup de données surtout sur toute la partie ouest de la mer d'Alboran. Les données originelles, sur le panel a, montrent à l'ouest des courants plus froids. Ces parties plus froides ne sont presque plus visibles après la reconstruction, où on ne retrouve qu'une légère partie plus froide à environ 17°C au lieu des moins de 16.5°C initiaux. Une particularité intéressante de DINCAE, reste sa capacité à lier des éléments pourtant non connectés. On peut en effet voir sur le panel a des données originales deux parcelles plus chaudes au sud. Ces parcelles se retrouvent aussi sur les données reconstruites, alors que toutes les données du sud-ouest de la mer sont éliminées dans le panel b. Sur le panel d, on peut voir que l'erreur pour les données reconstruites se situe environ à 0,7 ou 0,8. On est donc bien au-dessus de l'erreur moyenne de 0,5907 calculée dans le tableau 3.

Date: 2002-11-29

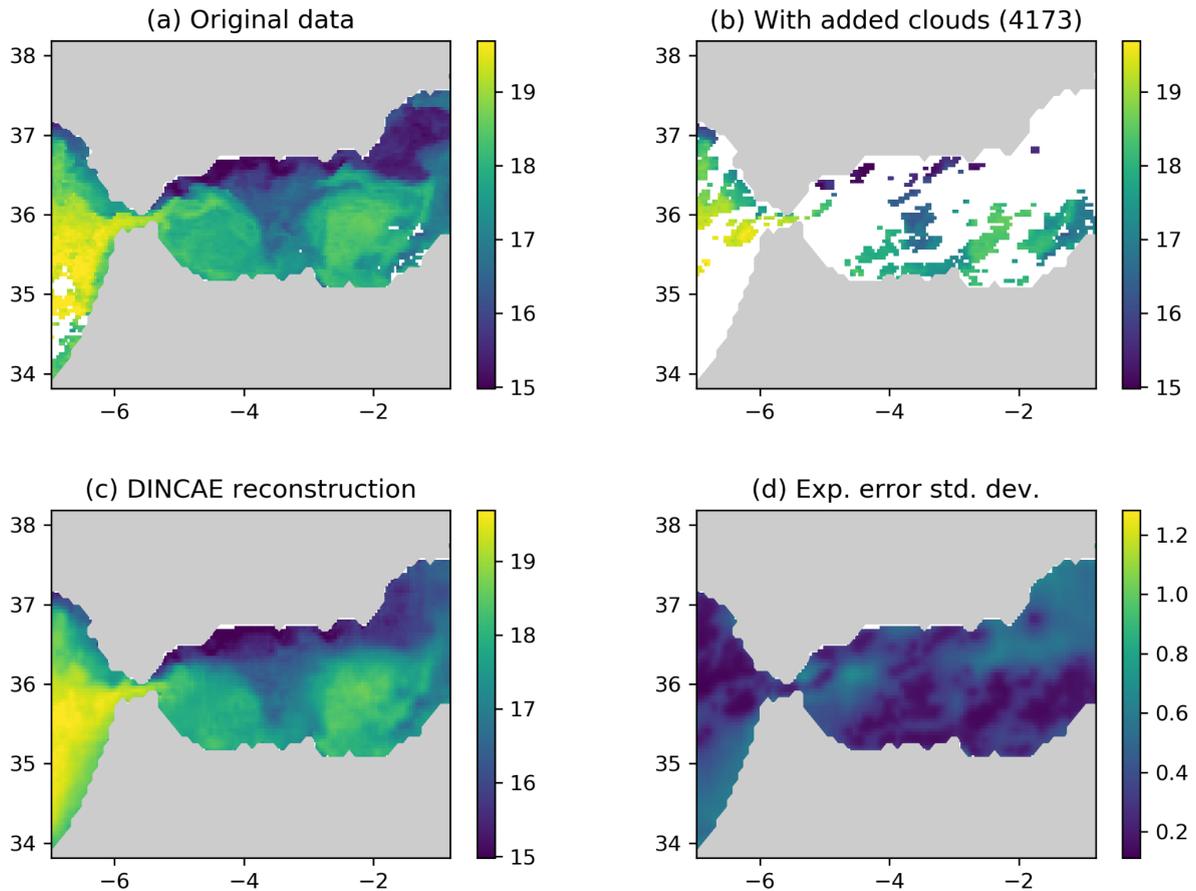


Figure 25 : Exemple de reconstruction du 29 novembre 2002 faite avec DINCAE et les paramètres décrits dans le tableau 3. (a) : données originales, (b) : données avec nuages ajoutés, le nombre entre parenthèses correspond au nombre de pixels cachés, (c) : données reconstruites et (d) : erreurs attendues.

Grâce à la figure 25, on peut voir sur le panel b, que le 29/11/2002, le programme a, à nouveau, enlevé beaucoup de données. Cependant, en comparant les panels a et c, on peut voir que les données reconstruites sont beaucoup plus proches que sur la figure 25. On peut, en effet, retrouver les deux gyres plus chaud au sud de la mer d'Alboran. Ce sont les deux gyres majeurs de la mer d'Alboran que l'on peut voir sur la figure 1. La partie au nord est de la mer n'est pas aussi bien reconstruite, mais on peut néanmoins retrouver la même forme présente. Sur le panel d, on peut voir que l'erreur des pixels reconstruits tourne généralement vers 0,4 ou 0,5, ce qui est proche de la moyenne d'erreur de 0,5907 voir un peu en dessous de celle ci.

## Conclusion

En conclusion, ce travail a permis de montrer que l'ajout de skip-connections lors d'une reconstruction avec un réseau neuronal permet d'améliorer la fiabilité des reconstructions sur les petites échelles. Sur un réseau à six couches de neurones, la meilleure configuration de celles-ci pour garder de la variance est la configuration avec une seule skip-connection sur la seconde couche lorsque l'on regarde une année entière et l'ensemble des échelles de moins de 20 km. Cette configuration ne garde cependant pas le plus de variabilités sur toutes les échelles dans cet intervalle et d'autres peuvent être plus intéressantes à utiliser à certaines échelles précises. Elle n'est aussi pas la configuration avec la meilleure erreur RMS, qui est la configuration avec une skip-connection sur les cinq premières couches.

Ce travail a aussi montré qu'il n'y a pas de configurations plus fiables que toutes les autres sur les quatre saisons. Le choix des skip-connections est donc dépendant de la période étudiée. La configuration la plus fiable trouvée pour l'hiver est celle avec des skip-connections sur les couches de neurones 2, 3, 4 et 6. Pour le printemps, une configuration avec des skip-connections sur les cinq premières couches de neurones est la plus fiable sur les échelles étudiées. Pour l'été, la configuration la plus fiable est une configuration sans skip-connection. Enfin, pour l'hiver, une seule skip-connection sur la seconde couche donne les résultats les plus fiables.

Enfin, il est important de noter que l'amélioration de la fiabilité des résultats des reconstructions à petites échelles par l'ajout de skip-connections reste assez faible. Il pourrait donc être intéressant de continuer à étudier les skip-connections pour voir s'il est possible d'encore augmenter ses gains.

Il serait aussi intéressant de pouvoir automatiser le choix des skip-connections lors de l'entraînement des réseaux neuronaux. On pourrait imaginer une équation de type  $C_{out} = C_{in} + \alpha * C$ , où  $C$  représenterait les données et  $\alpha$  représenterait les skip-connections. Le but serait alors de trouver un  $\alpha$  qui minimiserait l'erreur et donc la différence entre  $C_{out}$  et  $C_{in}$ .

# Bibliographie

Meehl, Gerald A. « The Annual Cycle and Interannual Variability in the Tropical Pacific and Indian Ocean Regions ». *Monthly Weather Review* 115, n° 1, 1 janvier 1987, 27-50. [https://doi.org/10.1175/1520-0493\(1987\)115<0027:TACAIV>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<0027:TACAIV>2.0.CO;2).

Palmen, E. « On the formation and structure of tropical hurricanes. » *Geophysica* 3, 1948, 26-38.

Zhu, Y., et D.D. Houghton. « The Impact of Indian Ocean SST on the Large-Scale Asian Summer Monsoon and the Hydrological Cycle ». *International Journal of Climatology* 16, n° 6, 1996, 617-32. [https://doi.org/10.1002/\(SICI\)1097-0088\(199606\)16:6<617::AID-JOC32>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1097-0088(199606)16:6<617::AID-JOC32>3.0.CO;2-I).

Cañadas, A., et J.A. Vázquez. « Common Dolphins in the Alboran Sea: Facing a Reduction in Their Suitable Habitat Due to an Increase in Sea Surface Temperature ». *Deep-Sea Research Part II: Topical Studies in Oceanography* 141, 2017, 306-18. <https://doi.org/10.1016/j.dsr2.2017.03.006>.

Noto, M., et I. Yasuda. « Population decline of the Japanese sardine, *Sardinops melanostictus*, in relation to sea surface temperature in the Kuroshio Extension ». *Canadian Journal of Fisheries and Aquatic Sciences* 56, n° 6, juin 1999, 973-83. <https://doi.org/10.1139/f99-028>.

Goffredo, S., E. Caroselli, G. Mattioli, E. Pignotti, et F. Zaccanti. « Relationships between Growth, Population Structure and Sea Surface Temperature in the Temperate Solitary Coral *Balanophyllia europaea* (Scleractinia, Dendrophylliidae) ». *Coral Reefs* 27, n° 3, 1 septembre 2008, 623-32. <https://doi.org/10.1007/s00338-008-0362-y>.

Abirami, S., et P. Chitra. « Chapter Fourteen - Energy-Efficient Edge Based Real-Time Healthcare Support System ». In *Advances in Computers*, édité par Pethuru Raj et Preetha Evangeline, 117:339-68. The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases. Elsevier, 2020. <https://doi.org/10.1016/bs.adcom.2019.09.007>.

Yamashita, R., M. Nishio, R. Kinh Gian Do, et K. Togashi. « Convolutional Neural Networks: An Overview and Application in Radiology ». *Insights into Imaging* 9, n° 4, 1 août 2018, 611-29. <https://doi.org/10.1007/s13244-018-0639-9>.

Siddique, N., Sidike P, C. P. Elkin, et V. Devabhaktuni. « U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications ». *IEEE Access* 9 (2021): 82031-57. <https://doi.org/10.1109/ACCESS.2021.3086020>.

Barth, A., Alvera-Azcárate, A., Licer, M., and Beckers, J.-M.: DINCAE 1.0: a convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations, *Geosci. Model Dev.*, 13, 1609–1622, <https://doi.org/10.5194/gmd-13-1609-2020>, 2020

Barth, A., Alvera-Azcárate, A., Troupin, C., and Beckers, J.-M.: DINCAE 2.0: multivariate convolutional neural network with error estimates to reconstruct sea surface temperature satellite and altimetry observations, *Geosci. Model Dev.*, 15, 2183–2196, <https://doi.org/10.5194/gmd-15-2183-2022>, 2022.

« Data Products | Emodnet Chemistry ». Consulté le 22 juin 2022. <https://www.emodnet-chemistry.eu/products>.

Shaltout, M., et A. Omstedt. « Recent sea surface temperature trends and future scenarios for the Mediterranean Sea ». *Oceanologia* 56, n° 3, 1 juin 2014, 411-43. <https://doi.org/10.5697/oc.56-3.411>.

Kingma, Diederik P., et J. Ba. « Adam: A Method for Stochastic Optimization ». arXiv, 29 janvier 2017. <https://doi.org/10.48550/arXiv.1412.6980>.

Renault, L., T. Oguz, A. Pascual, G. Vizoso, et J. Tintoré. « Surface circulation in the Alboran Sea (Western Mediterranean) inferred from remotely sensed data ». *Journal of Geophysical Research* 117, 2012 : 8009. <https://doi.org/10.1029/2011JC007659>.

Donner, L.J., et W.G. Large. « Climate Modeling ». *Annual Review of Environment and Resources* 33, 2008: 1-17. <https://doi.org/10.1146/annurev.enviro.33.020707.160752>.

Molina-Kescher, M. « Deep-water circulation and detrital provenance in the South Pacific, from the present day until 240 000 years ago. Evidence from Nd, Sr and Pb isotopes and Rare Earth Elements », 2014. <https://doi.org/10.13140/2.1.4675.8089>.

Clarke, A., J. Church, et J. Gould. « Chapter 1.2 Ocean Processes and Climate Phenomena ». In *International Geophysics*, édité par Gerold Siedler, John Church, et John Gould, 77:11-30. Ocean Circulation and Climate. Academic Press, 2001. [https://doi.org/10.1016/S0074-6142\(01\)80109-1](https://doi.org/10.1016/S0074-6142(01)80109-1).

Donnelly, D., et B. Rust. « The Fast Fourier Transform for Experimentalists ». *Computing in Science & Engineering* 7 1 juin 2005, 71. <https://doi.org/10.1109/MCSE.2005.60>.

intro\_spectral2d.m. « 2D Spectral Analysis », 19 juin 2017. [https://wiki.math.uwaterloo.ca/sheets/matlab/html/intro\\_spectral2d.html#8](https://wiki.math.uwaterloo.ca/sheets/matlab/html/intro_spectral2d.html#8).

« Limits of Oceans and Seas », International Hydrographic Organization (IHO), International Hydrographic Organization., Bremerhaven, PANGAEA, <https://epic.awi.de/id/eprint/29772/>, 1953

Hernanz, P. M., M.-F. Racault, J. Font-Muñoz, et G. Basterretxea. « Trends in phytoplankton phenology in the Mediterranean Sea based on ocean-colour remote sensing ». *Remote Sensing of Environment* 221, 2018, 50-64. <https://doi.org/10.1016/j.rse.2018.10.036>.

Alvera Azcarate, A., A. Barth, M. Rixen, et J.-M. Beckers. « Reconstruction of Incomplete Oceanographic Data Sets Using Empirical Orthogonal Functions: Application to the Adriatic Sea Surface Temperature ». *Ocean Modelling* 9, n° 4, 2005. <https://doi.org/10.1016/j.ocemod.2004.08.001>.

Beckers, J.-M., et M. Rixen. « EOF Calculations and Data Filling from Incomplete Oceanographic Datasets ». *Journal of Atmospheric and Oceanic Technology* 20, n° 12, 2003. [https://doi.org/10.1175/1520-0426\(2003\)020<1839:ECADFF>2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)020<1839:ECADFF>2.0.CO;2).

He, K., X. Zhang, S. Ren, et J. Sun. « Deep Residual Learning for Image Recognition ». arXiv, 10 décembre 2015. <https://doi.org/10.48550/arXiv.1512.03385>.

# Annexes

## A. Tableau complet des résultats

	RMS error	epochs	ntime_win	loss_weight_refine	enc_nfilter_internal	skipconnections	learning_rate
test 1	0.5981	1000	3	0.3-0.7	(32,64,128,256,512)	(1,2,3,4)	0.001
test 2	0.5849	500	3	0.3-0.7	(32,64,128,256,512)	(1,2,3,4)	0.001
test 3	0.6012	300	3	0.3-0.7	(32,64,128,256,512)	(1,2,3,4)	0.001
test 4	0.5891	500	5	0.3-0.7	(32,64,128,256,512)	(1,2,3,4)	0.001
test 5	0.5867	500	7	0.3-0.7	(32,64,128,256,512)	(1,2,3,4)	0.001
test 6	0.586	500	3	0.5-0.5	(32,64,128,256,512)	(1,2,3,4)	0.001
test 7	0.5704	500	3	Pas de raffinement	(32,64,128,256,512)	(1,2,3,4)	0.001
test 8	0.6	500	3	0.9-0.1	(32,64,128,256,512)	(1,2,3,4)	0.001
test 9	0.5807	500	3	0.1-0.9	(32,64,128,256,512)	(1,2,3,4)	0.001
test 10	0.5851	500	3	0.9-0.1	(32,64,128,256,512)	(1,2,3,4)	0.001
test 11	0.5999	500	3	0.05-0.95	(32,64,128,256,512)	(1,2,3,4)	0.001
test 12	0.6103	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(1,2,3,4)	0.001
test 13	0.6104	500	3	Pas de raffinement	(32,64,128,256,512)	(1,2,3)	0.001
test 14	0.5996	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(1,2,3,4,5)	0.001

test 15	0.60 27	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(1,2,3,4,5,6)	0.001
test 16	0.60 12	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(1,2,3)	0.001
test 17	0.59 85	500	3	Pas de raffinement	(32,64,128,256,512)	(1,2,3,4)	0.002
test 18	0.59 21	500	3	Pas de raffinement	(32,64,128,256,512)	(1,2,3,4)	0.0005
test 19	0.59 79	500	3	Pas de raffinement	(32,64,128,256,512)	(1,2,3,4)	0.00075
test 20	0.60 39	500	3	Pas de raffinement	(32,64,128,256,512)	(2,3,4)	0.001
test 21	0.58 2	500	3	Pas de raffinement	(32,64,128,256,512)	(2,3,4,5)	0.001
test 22	0.58 77	500	3	Pas de raffinement	(32,64,128,256,512)	(3,4,5)	0.001
test 23	0.59 73	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(2,3,4,5)	0.001
test 24	0.58 55	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(2,3,4,5,6)	0.001
test 25	0.57 85	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(3,4,5,6)	0.001
test 26	0.59 87	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(3,4,5,6)	0.00075
test 27	0.59 8	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(3,4,5,6)	0.002
test 28	0.58 45	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(4,5,6)	0.001
test 29	0.60 11	500	3	Pas de raffinement	(32,64,128,256,512, 1024)	(3,4,5)	0.001
test 30	0.59 77	500	3	Pas de raffinement	(32,64,128,256,512, 1024,2048)	(3,4,5,6,7)	0.001

test 31	0.58 52	500	3	Pas de raffinement	(32,64,128,256 ,512)	(2,4,5)	0.001
test 32	0.61 82	500	3	Pas de raffinement	(32,64,128,256 ,512)	(2,3,5)	0.001

## B. Loss-Weight refine

Lors de l'étude avec un réseau U-Net, il peut être intéressant d'utiliser deux réseaux U-Net, un réseau intermédiaire et un réseau final, l'un à la suite de l'autre. Le paramètre `loss_weight_refine` permet de voir quel poids est donné à chaque réseau. Par exemple, un `loss_weight_refine` égal à 0.3-0.7 indique un poids de 30% mis sur le réseau intermédiaire et un poids de 70% mis sur le réseau final.

L'intérêt d'utiliser deux réseaux est de pouvoir augmenter le nombre de couches de neurones pour augmenter la capacité du réseau à représenter des relations complexes. On utilise cette méthode car un réseau trop long aurait plus de mal lors de l'apprentissage, qui se fait grâce à la rétropropagation du gradient. Cette méthode d'apprentissage démarre à la fin, et privilégie avec de plus hauts poids les dernières couches. Si le réseau a de trop nombreuses couches, l'apprentissage est moins bon et les premières couches reçoivent des poids très faibles. Ce problème est appelé le problème de vanishing gradient. L'utilisation de deux réseaux U-Net permet donc de l'éviter, tout en augmentant le nombre de couches du réseau.