
Climate Model Downscaling

Jeremy Qin¹ Santiago Gomez² Emulie Chhor³ Diego Gonzalez⁴
Francisco Lopez⁵

¹Université de Montréal

Abstract

Ce rapport étudie l'impact de l'ajout de nouvelles variables, telles que la couverture du sol(landcover), dans le processus de downscaling climatique pour améliorer la précision des prédictions à des échelles plus petites. Nous explorons l'utilisation de l'intelligence artificielle, des sciences des données et des compétences en informatique pour augmenter la résolution des modèles climatiques, en passant d'une échelle nationale à une échelle plutôt régionale. En intégrant des données géographiques mondiales, nous cherchons à comprendre comment les caractéristiques géographiques influencent le climat et comment leur prise en compte peut améliorer les estimations des modèles. Les résultats de cette recherche peuvent contribuer à la mise au point d'outils précis et fiables pour la prévision météorologique, la planification d'événements, l'agriculture et l'économie, en tenant compte des spécificités locales et régionales.

<https://github.com/jeremy-qin/Climate>

1 Introduction

Commençons par le début : Que sont les modèles climatiques ? D'après le National Centre for Atmospheric Science du Royaume Uni (8), les modèles climatiques sont des outils computationnels qui servent à simuler le système climatique de la Terre. Cela inclut de nombreuses prédictions sur plusieurs facteurs, tels que la température atmosphérique, la pression, le vent ou l'humidité entre autres. La question subséquente est naturellement : Quelle est leur utilité pour l'humanité ? Il se peut en effet que ces modèles n'ont jamais été si importants que maintenant. Alors que leur but de base était d'évaluer la capacité de prédiction des scientifiques, l'utilisation actuelle la plus en vogue cherche à évaluer les impacts du réchauffement climatique sur la planète. C'est le cas du plus récent rapport du *Intergovernmental Panel on Climate Change* (5), lequel explore les effets d'une augmentation des températures globales de 1,5°C grâce à l'utilisation d'un modèle climatique

Abordons maintenant le deuxième sujet important de cette introduction : la réduction de dimensionnalité ou *downscaling*. Comme le nom l'indique, le downscaling cherche à réduire l'échelle de l'output des données climatiques. Dit dans d'autres mots, cet outil permet d'avoir une meilleure résolution des prédictions faites par nos modèles climatiques. Actuellement, nous avons accès à des modèles climatiques mondiaux, mais ils sont basés sur des mesures à grande échelle. À des échelles plus petites, les données peuvent être peu fiables, conduisant à des conclusions erronées. Pour illustrer: imaginons qu'un jour donnée on regarde la météo et on voit uniquement la température moyenne au Canada: en moyenne, il fait froid ce jour là. On pourrait donc penser qu'il fera froid à Vancouver, même si en réalité il fait très chaud! Le *downscaling* est particulièrement pratique lorsque des organismes locaux, tels que des

gouvernements ou organisations régionales, ont besoin de données de qualité se rapportant uniquement sur l'espace où ils peuvent exercer une influence.

Sachant cela, notre projet consiste à utiliser *l'IA* et les sciences des données, ainsi que nos compétences en informatique, pour améliorer les estimations à des échelles plus petites (comme passer d'une échelle nationale à une régionale), en augmentant la résolution dans des *échelles réduites*! Nous explorerons en particulier la possibilité d'améliorer les estimations d'IA en utilisant des données géographiques mondiales: le climat est fortement lié aux structures géographiques et on soupçonne que les prendre en compte améliorera nos estimations.

2 Revue de la littérature

Pitfalls. Les modèles à nos jours ont souvent des problèmes à bien prédire des *anomalies climatiques* (tempêtes, ouragans) et le climat des *frontières des sections géographiques* (côtes, montagnes).

ClimAlign. Cet article, cherche à utiliser les normalizing flows comme modèles génératifs et combiner avec une fonction de perte discriminateur. Un normalizing flow est essentiellement un modèle permettant d'aller d'une distribution simple à une distribution plus complexe par le biais de transformations inversibles et bijectives. ClimAlign traite le problème de réduction de modèles comme un problème de *domain alignment* dont le but est d'apprendre un mapping entre deux variables aléatoires (1). ClimAlign est intéressant dans le fait que c'est une méthode non-supervisée pour la réduction d'échelle. Contrairement à un GAN, normalizing flows a l'avantage d'être plus simple à entraîner. **Une idée intéressante à considérer est l'utilisation potentielle d'un temporal normalizing flow qui tient en compte l'aspect de la temporalité lors du calcul des distributions.**

Configuration and intercomparison of DL Models. Pour mieux mesurer l'impact des modèles en apprentissage machine, établir des **benchmarks** est crucial. Ce document présente les résultats de modèles linéaires généralisés (GLM) et de réseaux de neurones convolutionnels (CNN) pour prédire la température et la précipitation d'une région locale selon la méthodologie VALUE. Les résultats suggèrent que le **CNN avec 3 couches de 50 neurones et un filtre 3x3** est meilleur qu'un CNN avec plus de couches et un filtre plus grand (5,7 couches respectivement), ce qui montre que l'ajout de paramètres n'est pas nécessaire lorsque le modèle peut représenter toute la non-linéarité (comme on voit souvent avec le problème de **dégradation**) (6). Pour les **GLM, utiliser les 4 grilles les plus proches** pour la prédiction donne des résultats satisfaisants.

Explainability. Cet article propose l'utilisation de techniques du XAI (**explainable AI**) pour offrir une meilleure compréhension interne de nos modèles d'inférence des modèles climatiques. Les conclusions des chercheurs affirment que l'utilisation du XAI permet une meilleure évaluation de nos modèles grâce à des outils comme les **saliency maps** (7). Ces derniers vérifient que le modèle ne néglige pas les features considérées comme importantes. On observe également des **différents erreurs** systématiques au hémisphère **nord** vs **sud**. On constate aussi que **la pression atmosphérique est un des attributs plus significatifs** dans la tâche d'inférence des données climatiques.

Physics-Constrained DL. On met en évidence l'importance de **l'ajout de contraintes par rapport aux lois physiques** dans le contexte de l'apprentissage. En effet, ces contraintes **accélèrent et améliorent l'entraînement** des modèles d'apprentissage profond; on obtient

aussi des **estimations plus précis et corrects**. L'article présente aussi des possibles **implémentations** de l'ajout de ces contraintes, ainsi que leurs différences. On met en évidence que ces contraintes sont utiles dans une multitude de domaines, **notamment dans le downscaling des modèles climatiques!**

VALUE framework. Comment évaluer la performance des modèles de réduction d'échelle? L'article nous présente VALUE, un "open framework" développé pour **évaluer et comparer diverses techniques de downscaling**. VALUE vise à fournir une **base de données de validation** complète des méthodes de pointe de réduction d'échelle, ce qui nous permettra de mieux comprendre **quelles méthodes fonctionnent le mieux pour différents types de données climatiques** (2). Nous pouvons utiliser VALUE ou les méthodes et raisonnements présentés dans l'article pour évaluer, mesurer et comparer les modèles à développer et ainsi avoir une échelle égale pour tester la performance des différentes méthodes à utiliser.

Climate4R framework. L'acquisition de données est un défi courant en apprentissage machine, particulièrement pour le problème de la réduction d'échelle des modèles climatiques. **Les organisations mesurent les données différemment**, ce qui nécessite un **pré-traitement** qu'introduit un **bias systémique** et **complique la reproductibilité**. Ce document présente des méthodologies, telles que le *local-scaling* et le *empirical quantile mapping (EQM)*, pour réduire ce biais (9). Le *local-scaling* pondère la valeur locale par un facteur additif/multiplicatif calculé à partir de la différence et du ratio entre la prédiction du modèle et la valeur observée, tandis que le *EQM* corrige la fonction de densité en extrapolant les résultats à partir des quantiles. Le document discute des résultats de ces méthodologies sur des **ensembles de données, tels que ERA-Interim, CORDEX, ESGF et CMIP**, utilisés dans les documents précédemment discutés.

3 Méthodes

CNN pour le downscaling

Le développement de l'apprentissage profond a permis de nombreuses percées quant à la modélisation pour plusieurs tâches. L'architecture la plus commune est notamment celle des réseaux de neurones à propagation ou perceptron multicouches. Ces réseaux sont notamment composés de couches entièrement connectées dont une couche d'entrée, une couche de sortie et plusieurs couches cachées. Ces réseaux ont de très bonnes performances pour plusieurs tâches, mais n'est pas optimale pour le traitement d'images. En effet, pour traiter une image, le perceptron multicouche a besoin d'un nombre de neurones en entrée équivalant aux nombres de pixels de l'image. Si l'image est de haute résolution, nous avons donc besoin d'un réseau très grand. Ainsi, nous voyons qu'ils ne sont pas optimales sur le côté computationnel. Cependant, les CNN constituent une solution intéressante qui règle ce problème et apporte plusieurs autres avantages. Les réseaux de neurones convolutifs (CNN) utilisent la convolution pour capturer des relations spatiales complexes dans les données. La manière dont la convolution fonctionne pour approximer et représenter un signal à un point spécifique consiste à utiliser une somme pondérée des mesures voisines avec plus de poids donnés aux mesures récentes. Différents poids associés à un même voisinage pourraient faire ressortir différentes informations. Ces poids forment ce qu'on appelle un feature detector, ou un noyau. Ainsi, les CNNs sont utiles pour la réduction d'échelle de modèles climatiques car ils peuvent capturer les relations entre les caractéristiques des modèles à grande échelle et celles des modèles à échelle locale.

ResNet

ResNet est un type de CNN qui est l'état de l'art pour la classification d'images, la détection d'objets et la segmentation d'images. Il résout le problème de la disparition du gradient dans les réseaux de neurones de grande taille en utilisant des connections résiduelles, permettant à l'information de sauter d'une couche à une autre, sans avoir à passer par des couches intermédiaires, pour accélérer l'entraînement du modèle. En effet, dans des réseaux de neurones très profonds, les gradients ont tendance soit à avoir de très petites valeurs, et donc contribue à peine lors de la descente de gradients, soit à avoir de très grandes valeurs ce qui nuit à la convergence du modèle vers un minimum global. En résumé, ResNet est une architecture de réseau de neurones profond qui utilise des raccourcis résiduels pour permettre la transmission directe des caractéristiques entre les couches. Cette technique permet de construire des réseaux très profonds tout en maintenant une bonne performance en termes de précision.

UNet

UNet est un type de CNN qui a été utilisé en premier pour la segmentation d'images biomédicales, et qui est principalement composé d'un encodeur et un décodeur. L'encodeur réduit l'image en représentations de plus petites dimensions et capture les caractéristiques de l'image à l'aide de convolutions et de max pooling, tandis que le décodeur reconstruit l'image en utilisant des convolutions transposées et de l'upsampling. L'architecture d'un UNet est caractérisée par son utilisation d'une structure de "U" symétrique, qui permet d'associer les caractéristiques des couches de l'encodeur aux couches correspondantes du décodeur. Cette structure est obtenue en ajoutant des couches de déconvolution successives (convolution transposée) dans le décodeur, qui ont la même taille que les couches correspondantes de l'encodeur. En résumé, UNet est une architecture de réseau de neurone convolutifs qui utilise un encodeur et décodeur pour apprendre les différentes relations dans les images.

4 Données

ERA5. L'ensemble de données ERA5 est un produit de réanalyse atmosphérique global produit par le Centre européen pour les prévisions météorologiques à moyen terme (CEP-MMT). Il fournit des informations sur les variables météorologiques telles que la couverture nuageuse, la température, la pression, l'humidité, le vent et les précipitations à une résolution spatiale de 0,25 degré et une résolution temporelle horaire pour la période 1979 à présent. Ces données sont produites en assimilant de nombreuses observations provenant de différentes sources, telles que les satellites, les radiosondages et les stations au sol, dans un modèle numérique de prévision du temps.

ESRI Sentinel-2. Le jeu de données de couverture du sol ESRI Sentinel-2 est basé sur les images multispectrales acquises par le satellite Sentinel-2 de l'Agence spatiale européenne. Il fournit des informations sur la couverture du sol à une résolution spatiale de 10 à 20 mètres pour le monde entier. La classification des types de couverture du sol inclut les forêts, les zones humides, les surfaces artificielles, les terres cultivées et les zones aquatiques. Les données Sentinel-2 sont particulièrement utiles pour surveiller les changements de la couverture terrestre et sont largement utilisées dans des applications telles que l'agriculture, la foresterie et la gestion des ressources en eau.

Redimension des données ESRI

La base de données ESRI fournit les données pour le landcover de certaines régions seulement, il a donc fallu reconstruire la carte mondiale par nous-mêmes en concaténant chacune des tuiles ensemble. Nous avons rencontré quelques difficultés lors de la reconstruction de la carte. Premièrement, la résolution de chaque tuile est gigantesque et nécessite autour de 5 GB de mémoire RAM pour sa lecture. Il a donc fallu modifier la librairie 'PILLOW' afin d'accomoder la dimension. Deuxièmement, la plupart des données océaniques étaient naturellement indisponibles (pas beaucoup de *landcover* là où il n'y a pas de *land*). Il a donc fallu substituer les tuiles manquantes par les tuiles océaniques disponibles les plus proches pour proprement construire une carte du monde sans trous. Nous nous avons préoccupé qu'étant donné que la Terre est ronde, la carte générée sans ajustement serait ovale, et nous avons donc considéré d'appliquer la formule d'Haversine pour rectifier ce problème. Cependant, nous avons rapidement constaté que les images fournis par ESRI étaient nativement bien adaptées pour éviter ce problème. Nous avons donc simplement concaténé l'ensemble des images pour reconstruire notre carte.

Pour faire cette concaténation nous avons réduit la dimensions de tous nos images à 15x15, en gardant la taille de chaque image pour les reconstruire lors de la concaténation. Nous avons après concaténée nos images et construit la carte qui suit. Les défauts dans les frontières des images locales seront négligeables au moment de réduire encore plus la résolution d'image pour nos modèles, donc on a continué en les ignorant.



Carte globale du landcover

Union des données. Les données ERA5 et ESRI prétraitées ont ensuite été fusionnées (concaténées) pour créer des cartes d'entrée complètes. Ces cartes d'entrée ont été utilisées pour entraîner les modèles ResNet et U-Net et produire des prévisions climatiques à plus haute résolution (en rajoutant une dimensions aux images, ceci dit, un canal au modèle pour ces entrées).

5 Résultats

Afin d'évaluer les performances de notre jeu de donnée fusionné, nous avons évalué les performances de quatre variantes du jeu de données sur leur capacité à prédire la couverture totale des nuages ainsi que la température à 2 mètres, ceci grâce à trois métriques différentes: la racine de l'erreur quadratique moyenne, le coefficient corrélation de Pearson et l'erreur de biais moyen. Voici un tableau qui résume les résultats obtenus pour ResNet:

| Expérience | RMSE | Coeff. Pearson | Mean bias error |
|---------------------------------|-------|----------------|------------------------|
| Merged Resnet Total Cloud Cover | 0,319 | 0,482 | $-3,72 \times 10^{-2}$ |
| Baseline Total Cloud Cover | 0,235 | 0,760 | $2,22 \times 10^{-3}$ |
| Merged Resnet 2m Temp | 7,887 | 0,935 | 2,52 |
| Baseline 2m Temp 4 | 2,152 | 0,995 | $-1,4 \times 10^{-1}$ |

Voici le tableau qui résume les résultats pour UNet:

| Expérience | RMSE | Coeff. Pearson | Mean bias error |
|-------------------------------|-------|----------------|------------------------|
| Merged UNet Total Cloud Cover | 0,318 | 0,477 | $3,22 \times 10^{-4}$ |
| Baseline Total Cloud Cover | 0,227 | 0,778 | $2,92 \times 10^{-5}$ |
| Merged UNet 2m Temp | 7,419 | 0,935 | $8,26 \times 10^{-1}$ |
| Baseline 2m Temp 4 | 1.898 | 0,996 | $-6,29 \times 10^{-2}$ |

De manière surprenante, nous n'avons observé que peu de changement dans les performances des modèles après avoir fusionné les données de landcover aux ensembles de données de base ERA5. Dans les métriques utilisées, telles que la corrélation de Pearson, l'erreur quadratique moyenne et la perte d'entraînement, les modèles avec landcover ont obtenu de moins bons résultats que les modèles de base avec U-Net et ResNet. Cette différence est peu remarquable sur des métriques comme le *RMSE* sur le Total Cloud Cover, où la variation pour ResNet et UNet est de 31,8% et 40%, respectivement. En revanche, la différence devient pas du tout négligeable pour le *mean bias error*, où la variation est supérieure à un ordre de grandeur de 10 pour UNet et ResNet sur Total Cloud Cover ainsi que sur 2m Temperature.

6 Conclusion

Nous avons combiné les informations sur l'occupation des sols d'ESRI avec les données climatiques ERA5 et appliqué divers modèles d'apprentissage automatique pour étudier les améliorations potentielles de la modélisation du climat. Cependant, nos expériences n'ont pas produit de meilleurs résultats qu'avec seulement les données ERA5.

Retour sur les résultats

La fusion des données du landcover d'ESRI avec les données climatiques ERA5 n'a pas conduit à de meilleurs résultats dans nos expériences. Nous soupçonnons trois raisons principales à cela :

- Apprentissage implicite du landcover : Le modèle a peut-être déjà déduit les informations sur l'occupation des sols à partir des données climatiques, surtout depuis que toutes les images Climate-Learn ont la même orientation. Cela pourrait être confirmé ou réfuté en testant avec différentes données géographiques.
- Impact négligeable aux faibles résolutions : Les avantages de l'incorporation des données d'occupation des sols ne deviennent peut-être apparents qu'à des résolutions plus élevées, qui n'ont pas été utilisées dans notre étude.
- Problèmes méthodologiques : Il pourrait y avoir des problèmes dans le prétraitement des données qui ont eu un impact négatif sur les résultats.

Limitations

Notre étude comporte plusieurs limitations. D'une part, nous avons assumé que les données de chacune des régions étaient disponibles, alors que ce n'est pas toujours le cas si on avait choisi une autre donnée géospatiale. En effet, si on avait plutôt choisi des données d'imagerie satellite tel que le "Sentinel", nous n'aurions pas eu accès aux lieux ayant des bases militaires ou restreint d'accès au public. D'autre part, étant donné que le terrain reste relativement le même sur une période de 15 à 20 ans et que nous avons entraîné nos modèles de downscaling sur cette même période, nous avons supposé que les données de la couverture du sol reste la même au fil du temps, ce qui ne sera pas toujours vrai dans le futur. À cause du changement climatique, on peut s'attendre à plus de glissements de terrain ou d'inondations des terres et donc, il faudra probablement prendre en compte l'aspect temporel de la donnée géospatiale introduite dans le futur. Additionnellement, notre modèle s'avère être difficile à interpréter. La carte générée nous indique la température (ou le cloud cover) pour une certaine sous-région, par contre, il n'est pas toujours évident de savoir comment utiliser cette information pour guider les politiciens et démographes à prendre de meilleure décision.

Projets futurs

Les projets futurs que l'on propose cherche à palier certaines limitations de notre étude. Au niveau des données, il serait intéressant de tester le modèle en intégrant une autre donnée géospatiale autre que la couverture du sol pour voir si nous avons des résultats similaires. Évidemment, puisque nous n'avons pas considéré la temporalité des données géospatiales lors de l'entraînement de nos modèles, il serait intéressant de voir si on obtient des résultats similaires avec des données qui ne varient pas beaucoup avec le temps comme les données d'élévation du terrain (3) ou bien avec des données qui subies une plus grande variations dans le temps comme les terres cultivées (*cropland*). Si on veut encore assumer l'invariabilité temporelle, on peut aussi essayer de moyenner les données géospatiales sur quelques années. Toujours par rapport à l'aspect temporel, on peut essayer d'introduire la temporalité lorsqu'on entraîne le modèle et considérer non pas les données géospatiales pour une année, mais bien sur toute la période étudié pour voir si on a vraiment des gains considérables. En faisant cette expérience, on sera capable de déduire si le modèle est empiriquement Markovien, c'est-à-dire qu'on n'a pas besoin de connaître le passé pour prédire le futur.

En ce qui attrait l'architecture des modèles, nous pourrions réessayer cette expérience avec des modèles ayant plus de capacité tel que des visions transformers (ViT) ou bien des modèles non-supervisés comme des réseaux génératifs antagonistes (GANs) ou alors des autoencodeurs variationnels (VAEs).

Finalement, pour essayer faciliter l'interprétation du modèle, il serait intéressant de construire une carte de saillance (*saliency map*) pour identifier les biais et les faiblesses du modèle. Étant donné que la carte de saillance nous montre directement le gradient de la variable cible (dans notre cas, la température ou le cloud cover), on peut voir directement les régions qui ont subi le plus de changement et avoir une meilleure idée de l'impact des politiques et stratégies mises en place par les autorités.

7 Contributions des membres

La contribution de nos membres s'est faite de la façon suivante:

- Jeremy: gestion des modèles et leur développement; chargé de l'exécution des expériences et développeur principal du code utilisé. Développeur auxiliaire au pré-traitement. Contribué

au rapport et à la présentation. Recherche des méthodes de pré-traitement, modèles, et des datasets candidats.

- Santiago: pré-traitement des données. Développeur auxiliaire des modèles et leur développement. Contribué au rapport et à la présentation. Recherche des méthodes de pré-traitement, modèles, et des datasets candidats.
- Diego: Développeur auxiliaire au pré-traitement. Contribué au rapport et à la présentation. Recherche des méthodes de pré-traitement, modèles, et des datasets candidats.
- Francisco: Développeur auxiliaire au pré-traitement. Contribué au rapport et à la présentation. Recherche des méthodes de pré-traitement, modèles, et des datasets candidats.
- Émulie: Développeur auxiliaire au pré-traitement. Développeur auxiliaire des modèles. Contribué au rapport et à la présentation. Recherche des méthodes de pré-traitement, modèles, et des datasets candidats.

Remerciements

Merci à Mélisandre pour les précieux conseils et à Alex pour le template latex (4) !

References

- [1] Claire Monteleoni Brian Groenke, Luke Madaus. Climalign: Unsupervised statistical downscaling of climate variables via normalizing flows. 2020. URL: <https://arxiv.org/abs/2008.04679>.
- [2] JM Gutiérrez Sven Kotlarski Douglas Maraun, Martin Widmann. Value - a framework to validate downscaling approaches for climate change studies. 2014. URL: https://www.researchgate.net/publication/269339815_VALUE_-_A_Framework_to_Validate_Downscaling_Approaches_for_Climate_Change_Studies.
- [3] Google. Earth engine data catalog, 2023. URL: <https://developers.google.com/earth-engine/datasets/>.
- [4] Alex Hernandez-Garcia. mais-latex, 2022. URL: <https://github.com/alexhernandezgarcia/mais-latex>.
- [5] IPCC. Special report - global warming of 1.5 °c. 2023. URL: <https://www.ipcc.ch/sr15/>.
- [6] José Manuel Gutiérrez Jorge Baño-Medina, Rodrigo Manzananas. Configuration and intercomparison of deep learning neural models for statistical downscaling. 2020. URL: <https://gmd.copernicus.org/articles/13/2109/2020/>.
- [7] José Manuel Gutiérrez Jose Gonzalez-Abad, Jorge Baño-Medina1. Using explainability to inform statistical downscaling based on deep learning beyond standard validation approaches. 2023. URL: <https://arxiv.org/pdf/2302.01771.pdf>.
- [8] NCAS. What is a climate model? 2020. URL: <https://ncas.ac.uk/learn/what-is-a-climate-model/>.

- [9] SantanderMetGroup. The r-based climate4r open framework for reproducible climate data access and post-processing. 2019. URL: <https://www.sciencedirect.com/science/article/pii/S1364815218303049>, <https://github.com/SantanderMetGroup/climate4R>.