

Stage présenté par

Paolo CHEVALIER

Élève-ingénieur de l'INSA Rennes

Spécialité MA

2024-2025

PRÉDICTION DE
VAGUES DE CHALEUR
AVEC DES BASES DE
DONNÉES DE SIMULATIONS ET
ÉVALUATION DU « TRANSFER
LEARNING » POUR DES
RÉANALYSES.

Lieu du Stage

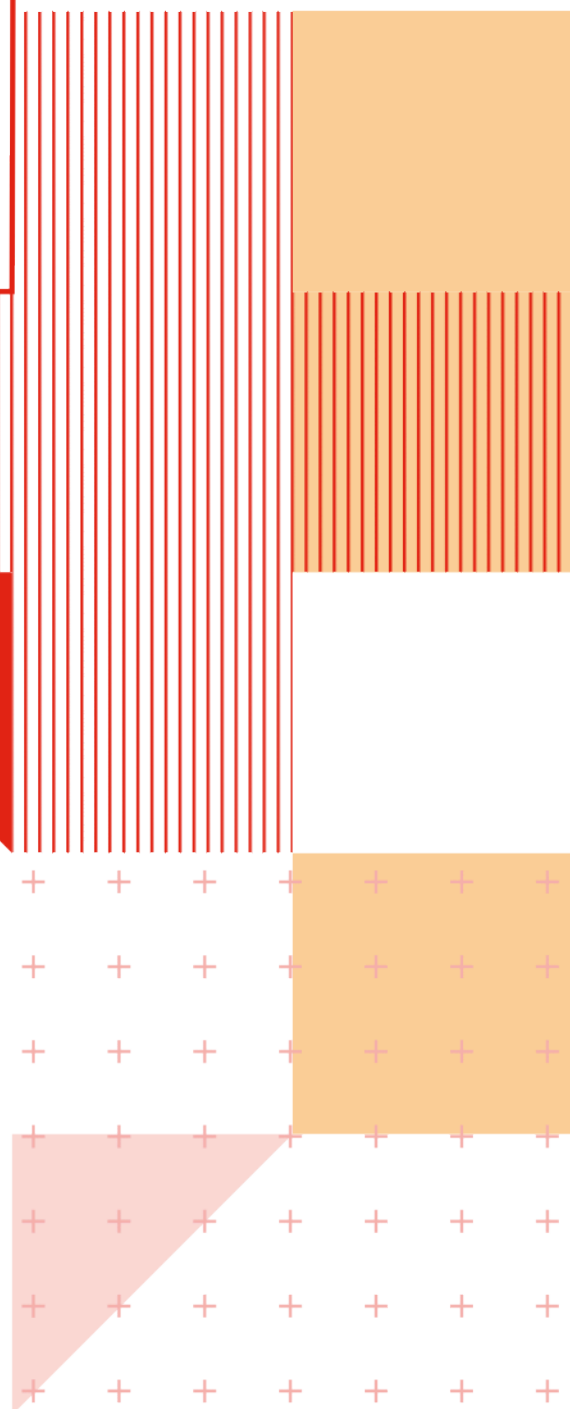
Laboratoire des Sciences du Climat et de
l'Environnement, UMR 8212 CEA-CNRS-
UVSQ Université Paris-Saclay & IPSL,
CEA Saclay l'Orme des Merisiers, 91191
Gif-sur-Yvette France

Maître de Stage

Pascal YIOU

Correspondant pédagogique INSA

Simon POSTEC



REMERCIEMENTS

Je tiens tout d'abord à remercier Pascal Yiou, mon maître de stage, qui m'a offert l'opportunité de réaliser ce stage sur des thématiques qui me parlent au LSCE et pour son encadrement toujours bienveillant.

Je remercie aussi Hannah avec qui j'ai partagé un bureau et Gilles pour leur gentillesse.

Je remercie aussi toutes les autres personnes passionnées que j'ai rencontrées au cours de mon stage et avec qui j'ai eu l'occasion d'échanger.

Pour le traitement des données, cette étude a bénéficié du Centre de données et de calcul de l'IPSL ESPRI, soutenu par le CNRS, Sorbonne Université, le CNES et l'École Polytechnique. Je remercie les équipes fournissant les données CMIP6 (disponibles dans /bdd/CMIP6) et ERA5, ainsi que le Centre de données et de calcul de l'IPSL ESPRI pour leur soutien dans l'accès à ces données.

Ce stage a été financé par le Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA).

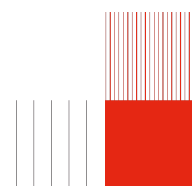
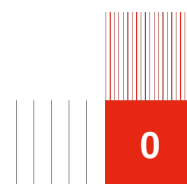


TABLE DES MATIÈRES

| | | |
|-------|--|----|
| 1 | Introduction | 1 |
| 2 | Environnement de travail et données | 3 |
| 3 | Éléments théoriques et méthodes | 6 |
| 3.1 | Fonction committor | 6 |
| 3.2 | Analogues de circulation | 6 |
| 3.3 | Estimation de la fonction committor par les analogues | 7 |
| 3.4 | Modèle UNET | 8 |
| 4 | Étude exploratoire | 8 |
| 4.1 | Cartes composites | 9 |
| 4.2 | Typicalité des événements de température extrêmes dans IPSL-CM6A-LR | 9 |
| 4.3 | RMSE et corrélations spatiales entre évènements extrêmes | 10 |
| 4.4 | Cartes de l'évolution | 12 |
| 4.5 | Conclusion | 13 |
| 5 | Résultats | 13 |
| 5.1 | Analogues | 13 |
| 5.1.1 | Qualité globale des analogues | 13 |
| 5.1.2 | Prévisions, estimation de la fonction committor et « transfer learning » | 14 |
| 5.2 | Machine learning | 16 |
| 5.2.1 | Fonction de transfert au jour même dans IPSL-CM6A-LR | 16 |
| 5.2.2 | Fonction de transfert avec $\tau > 0$ jour d'avance dans IPSL-CM6A-LR | 16 |
| 5.2.3 | « transfer learning » | 18 |
| 6 | Conclusion | 19 |
| | Code et Données | 22 |
| | Glossaire | 22 |
| | Bibliographie | 22 |
| | Annexes | 24 |
| -A | Considérations numériques pour le calcul des analogues | 24 |
| -B | Éléments théoriques supplémentaires | 26 |
| -C | Résultats des UNET dans IPSL | 28 |



1 INTRODUCTION

Étudiant en quatrième année à l'INSA de Rennes en Mathématiques Appliquées, j'ai eu l'opportunité, à travers mon cursus, de réaliser un stage durant l'été 2025. De ce fait, j'ai effectué durant 15 semaines un stage au Laboratoire des Sciences du Climat et de l'Environnement (LSCE) basé à Gif-sur-Yvette. J'ai, grâce à ce stage, pu travailler sur des sujets liés au climat auxquels je suis sensible. En particulier, je me suis intéressé aux vagues de chaleur extrêmes qui sont considérées comme parmi les pires conséquences du changement climatique anthropique (IPCC, 2023a; Russo et al., 2015) et face auxquelles la société devra s'adapter tant leur intensité et durée vont augmenter selon l'évolution actuelle du climat (Figure 1). Selon le 6ème rapport du GIEC (IPCC, 2023b), un réchauffement de 1.5°C causera au moins 30 000 décès par été en Europe, touchant particulièrement les populations les plus vulnérables. À 3°C de réchauffement planétaire il y aura des risques additionnels sur la santé qui seront inévitables et conséquemment une augmentation des inégalités climatiques, d'autant plus que l'Europe est le continent qui se réchauffe le plus rapidement.

Le laboratoire de Recherche dans lequel j'ai effectué mon stage a ses quartiers sur le site du Commissariat à l'énergie Atomique et aux Énergies Renouvelables (CEA) de Saclay à l'Orme des Merisiers. Le LSCE est une unité mixte de recherche (UMR) entre le CEA, le Centre National de la Recherche Scientifique (CNRS) et l'Université Versailles Saint-Quentin (UVSQ) et fait partie de l'Institut Pierre Simon Laplace (IPSL) et l'université Paris Saclay. Je ne m'attarderai pas ici sur la présentation complète de l'IPSL ni du CEA ou de l'UVSQ mais plutôt à celle du LSCE. Le LSCE regroupe plus de 350 chercheurs, ingénieurs, techniciens et personnels administratifs dont plus de 150 personnels permanents issus des 3 tutelles et plusieurs dizaines de doctorants. Le LSCE est organisé en trois thèmes scientifiques. Ils sont composés chacun de 4 à 6 équipes, pour un total de 16 équipes scientifiques. Les thèmes de recherche sont les suivants: « Archives et Traceurs » dont l'objectif est de comprendre la dynamique et la variabilité naturelle du climat sur différentes échelles de temps et d'espace; « Cycles biogéochimiques et transferts dans l'Environnement » dont l'objectif est de caractériser et quantifier les cycles biogéochimiques de

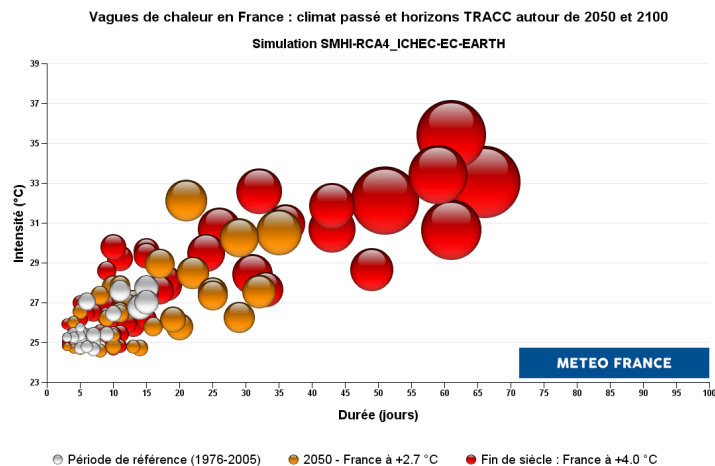


Figure 1 : Vagues de chaleur sur la période de référence et sur les horizons TRACC¹ d'une France à +2.7°C et +4.0°C, simulées par le modèle climatique SMHI-RCA4_ICHEC-EC-EARTH. En abscisses figure la durée de l'épisode. En ordonnées figure son intensité, correspondant à la température moyenne du jour le plus chaud de l'épisode. La taille des bulles correspondent à la sévérité, indicatrice de la quantité de chaleur perçue durant tout l'épisode. Les couleurs des bulles correspondent aux périodes passées ou futures.

¹<https://www.drias-climat.fr/accompagnement/sections/417>

certains composés (en particulier les gaz à effet de serre) et les transferts de matière dans et entre les différents compartiments de l'environnement (interface air-eau, air-surface continentales, eau-surfaces continentales); et enfin « Climat et Cycles – Modélisation de leurs variabilités et de leurs interactions » dans lequel j'ai réalisé mon stage. Les différentes équipes de ce dernier thème s'intéressent respectivement à la variabilité et aux interactions entre les différentes composantes du système Terre (MERMAID), à la dynamique intégrée du climat sur de multiples échelles de temps (CLIM), à l'infrastructure de modélisation numérique (CALCULS) et aux méthodes statistiques pour étudier la variabilité, les extrêmes ou les descentes d'échelle (ESTIMR). L'équipe que j'ai rejointe est l'équipe ESTIMR. L'équipe ESTIMR a pour objectifs principaux : la compréhension et la modélisation de la variabilité climatique et environnementale à différentes échelles spatiales – depuis les très grandes structures liées à la dynamique atmosphérique jusqu'à des phénomènes très locaux – et à différentes échelles temporelles – pour l'étude de climats passés, des processus présents et des évolutions futures. Dans ce cadre très vaste, l'une des forces de l'équipe ESTIMR repose sur l'utilisation et le développement de modèles statistiques de l'état-de-l'art adaptés aux problématiques climatiques, grâce à une interaction multidisciplinaire soutenue entre climatologie, modélisation, physique, statistiques et intelligence artificielle. Le labo est un environnement assez vivant et convivial.

Le stage s'étant principalement déroulé en été, j'ai le plus souvent travaillé seul, avec des directives données par mon maître de stage.

Dans un contexte où Google (Lam et al., 2023) ou Huawei (Bi et al., 2023) ont développé des modèles (GraphCast et PanguWeather) météorologiques qui utilisent de l'IA nous pouvons nous poser certaines questions. Ces modèles fonctionnent très bien mais n'utilisent que des réanalyses (Hersbach et al., 2020) pour l'apprentissage. Cette approche bien que très performante (plus efficace que les modèles déterministes) possède tout de même des limites. Ces modèles ne prennent pas en compte le changement climatique (dans 50 ans des mêmes conditions initiales pourraient ne pas mener vers des états similaires) et pourraient aussi juste bénéficier de plus de données d'apprentissage pour être d'autant plus performants. Il est quand même important de préciser que ces modèles bien que plus performants que les modèles usuels de manière générale restent moins bons quand il s'agit d'événements extrêmes (Zhang et al., 2025), notamment pour la prédiction d'événements encore jamais observés. Cette problématique est celle qui va nous intéresser.

Nous nous posons alors la question suivante, est-il viable d'apprendre (au sens de l'IA) sur des ensembles de données de simulations à grande échelle CMIP6 (Eyring et al., 2016) pour tirer des conclusions ou faire des prédictions dans des conditions réelles. Autrement dit, nous cherchons à évaluer la possibilité de « transfer learning » entre des ensembles de simulations réalisés pour le GIEC et des réanalyses (proches de la réalité).

Dans ce cadre, le travail de ce stage s'articule autour de deux axes :

1. L'estimation de « fonctions committor » à l'aide d'analogues de circulation atmosphérique. Prédire la probabilité d'occurrence d'une vague de chaleur étant donné l'état actuel du système climatique.
2. La reconstruction et la prévision de champs de températures à l'aide d'algorithmes de deep learning. Par exemple dans (Cariou et al., 2025), l'auteur arrive à reconstruire la température en fonction de la circulation atmosphérique.

Ce stage était aussi l'occasion pour moi de me familiariser avec l'utilisation de serveurs de calcul et les bonnes pratiques qui vont avec (Git, bash, ssh etc).

Ce rapport fera état des connaissances et compétences acquises durant le stage en suivant la forme globale d'un article, présentation des données, méthodes et éléments théoriques puis résultats.

Je commencerai par utiliser une grande base de données de simulations numériques du climat (CMIP6) pour en détecter les plus grandes vagues de chaleur en Europe. Ce travail d'exploration sera accompli par « data mining » de l'archive CMIP6 disponible sur les serveurs de l'IPSL. Ensuite, j'évaluerai les conditions météorologiques (en termes de circulation atmosphérique, humidité, température) qui mènent à ces vagues de chaleur, avec des échéances de 1 jour à 2 semaines à l'avance. Je travaillerai au calcul d'analogues et j'entraînerai un algorithme d'intelligence artificielle (de type U-net) à partir de ces données de simulations, pour calculer la probabilité d'observer une vague de chaleur et reconstruire des champs de températures, en partant de conditions initiales arbitraires. Puis j'évaluerai les performances de ces méthodes à partir de réanalyses (e.g. ERA5) et observations récentes, pour déterminer le potentiel de « transfer learning » de cet apprentissage. Pour chaque méthode, je m'attellerai à ce qu'elle marche au mieux dans le modèle de l'IPSL et ERA5 avant de tester le « transfer learning ».

2 ENVIRONNEMENT DE TRAVAIL ET DONNÉES

Dans le cadre de ce stage, j'utiliserai python et principalement les bibliothèques `xarray` (Hoyer et Hamman, 2017), `pandas` (McKinney, 2010), `matplotlib` (Hunter, 2007), `dask` (Dask Development Team, 2016), `xskillscore` (Bell et al., 2021), `numpy` (Harris et al., 2020) et `cartopy` (Met Office, 2010) sur un serveur distant de l'IPSL basé à l'école Polytechnique pour accéder directement aux données qui y sont déjà stockées et avoir plus de puissance de calcul (les données sont volumineuses).

Avant de présenter les données, il est nécessaire d'expliquer ce qu'est un modèle climatique et pourquoi nous les utilisons. Un modèle climatique est une représentation numérique du système climatique terrestre qui combine les lois fondamentales de la physique (dynamique des fluides, thermodynamique, transferts radiatifs) avec la description des processus biogéochimiques (cycle du carbone, aérosols, végétation). Ces modèles couplés couvrent le globe entier en combinant des modèles pour les composantes suivantes: atmosphère, océan, surface et cryosphère. Ainsi, les équations aux dérivées partielles, discrétisées dans l'espace et le temps, sont résolues numériquement car il n'existe pas de solutions de forme fermée à certaines de ces équations (Navier-Stokes). Le calcul numérique de ces solutions (les simulations) est très coûteux en termes de ressources de calcul et peut prendre plusieurs mois sur des serveurs. L'initiative CMIP6 (Coupled Model Intercomparison Project phase 6) (Eyring et al., 2016) fournit un cadre standardisé permettant de comparer et d'évaluer différents modèles développés par des centres de recherche, en utilisant des scénarios communs d'évolution des forçages naturels et anthropiques comme les gaz à effet de serre. L'objectif est d'analyser la variabilité passée et présente du climat et de produire des projections cohérentes du climat futur. Ces modèles sont déterministes mais chaotiques. En effet, si un même modèle de ce type est lancé deux fois de suite avec les mêmes conditions initiales, le résultat sera exactement le même. Cependant, si on part de deux états initiaux proches mais avec une variation même minime, les résultats seront rapidement différents. D'où l'intérêt lorsque ces modèles sont lancés de faire plusieurs simulations (runs). Cet aspect chaotique connu sous le nom d'effet papillon explique la raison pour laquelle on étudie ces données comme des distributions de probabilités. Si le système climatique n'était pas chaotique nous n'aurions aucun intérêt à travailler sur toutes les thématiques liées aux probabilités et au climat.

Je travaillerai avec les données du modèle IPSL-CM6A-LR (Boucher et al., 2020) pour la phase 6 du CMIP. Ces données sont des simulations historiques de 1850 à 2014 (total de $N = 15180$ jours). Une simulation historique est une simulation qui démarre en général en 1850 ou 1950 et qui s'étend jusqu'à nos jours, les différents forçages anthropiques comme ceux des gaz à effet de serre suivent l'évolution réelle observée. Dans ce modèle, les données issues des simulations numériques sont sous la forme d'un maillage de la planète, illustré par Figure 2.

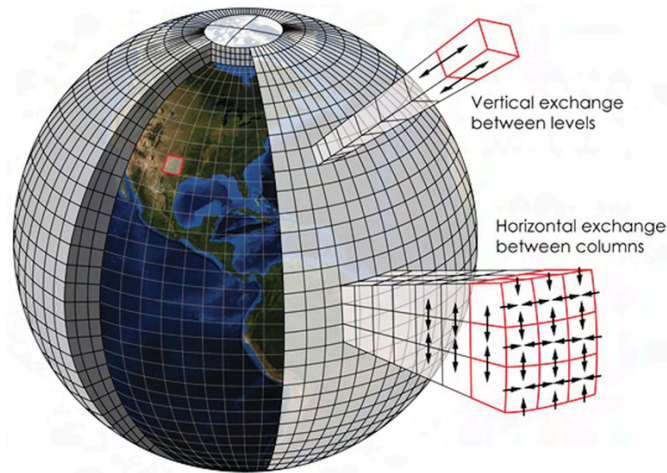


Figure 2 : Représentation du fonctionnement d'un modèle climatique (Schwertfeger, 2022). La manière dont le système terre est discrétisé varie selon les modèles et ne ressemble pas forcément à une grille en 3d comme ici (par exemple certains utilisent des méthodes spectrales) mais les données en sortie sont toujours sous cette forme

Je travaillerai aussi avec des données de réanalyses du centre européen (ECMWF) ERA5 (Hersbach et al., 2020), qui sont les données les plus complètes et précises de la météo et du climat passé récent. Les réanalyses sont un mélange entre observations et prévisions à court terme du passé avec des modèles météorologiques modernes. Ces données sont donc « sans trous », i.e. continues dans le temps et l'espace. Pour comprendre le changement climatique et les extrêmes, il est essentiel d'avoir des observations les plus lointaines possibles. Cependant, un modèle basé uniquement sur des observations passées serait incomplet et imprécis. Les observations ont toujours été distribuées non uniformément, elles peuvent être faussées et même à l'ère des satellites elles ne suffisent pas à fournir une représentation complète de l'état du climat à un instant t . Les réanalyses pallient ces problèmes. Concrètement, les réanalyses combinent des prévisions à court terme du passé avec des observations par assimilation de données (Bouttier et Courtier, 2002). L'objectif est de produire des prévisions quotidiennes, utilisant une à chaque fois l'état courant du système comme conditions initiales, puis en y assimilant les observations réelles disponibles.

Les variables journalières d'intérêt que j'étudierai sont présentées Tableau 1 ci-dessous.

Tableau 1 : Variables utilisées

| Variable | Notation | Unité | Description |
|----------|---------------|-------------------|---|
| tas | \mathcal{T} | K | Température moyenne proche de la surface (2m) par jour, en Kelvin (K) (je la passerai en °C lors des visualisations dans un souci de simplicité d'interprétation). |
| zg500 | \mathcal{Z} | m | Hauteur géopotentielle à 500hPa, correspond à l'altitude à laquelle on atteint 500 hPa par rapport au niveau moyen des mers ² . Cette mesure permet donc d'identifier les niveaux de pression constante. |
| mrso | \mathcal{M} | kg/m ² | Humidité du sol, masse totale d'eau dans le sol de la maille divisée par la surface de cette dernière. |
| pr10 | \mathcal{P} | mm | Cumul de précipitation lors des 10, 30 ou 90 derniers jours ³ . |
| psl | \mathcal{P} | Pa | Pression au niveau de la mer, c'est la pression que l'on obtiendrait si la maille était « déplacée » au niveau de la mer. |

²La pression décroît avec l'altitude

Dans un souci de visualisation, pour la plupart des variables, je privilégierai de travailler avec des anomalies plutôt qu'avec des observations. La notion d'anomalie en climatologie est assez intuitive. Soit $\mathcal{F}(\varphi, \lambda, t)$ le champ associé à une variable évalué à la latitude φ , la longitude λ et au temps t . L'anomalie est alors $\mathcal{F}_{\text{anom}}(\varphi, \lambda, t) = \mathcal{F}(\varphi, \lambda, t) - \mathbb{E}_t[\mathcal{F}(\varphi, \lambda, t)]$. Concrètement, je centre les variables en 0 pour chaque maille par rapport à la valeur moyenne climatologique (moyenne par jour de l'année dans ce travail). La manière dont la dimension temporelle est encodée est décrite en Annexe -B.

Pour identifier les événements de chaleur, j'isole 5 événements de températures très extrêmes lors de la période Juin, Juillet et Août (JJA). Dans ce cadre, afin d'éviter une dépendance entre les événements, il faut s'assurer que tous les extrêmes sont séparés d'au moins 10 jours. Je fais l'hypothèse que deux événements très extrêmes, s'ils sont séparés par au moins 10 jours consécutifs sont indépendants. C'est une procédure similaire à celle utilisée en théorie des valeurs extrêmes (TVE) pour fit des excès au-dessus d'un seuil à une distribution de Pareto en renforçant l'hypothèse d'indépendance des observations. La seule différence étant qu'ici nous n'avons pas de seuil donc que nous cherchons juste des maximas et non des clusters et nous choisissons un nombre d'extrêmes plutôt qu'un seuil. Une représentation graphique de la procédure est donnée dans Fig. 3.

Il n'y a pas de convention internationale pour la définition d'une vague de chaleur, mais le terme est toujours utilisé pour des températures extrêmes. En fonction des régions et pays, des seuils différents sont choisis en fonction des impacts locaux, qui considèrent notamment les impacts sociétaux. Par exemple dans le rapport européen sur l'état du climat 2024 elles sont définies comme « trois jours ou plus avec une température maximale dépassant le quantile 0.98 de la période de référence 1961-1990 et dépassant 28 °C » (Copernicus Climate Change Service (C3S) and World Meteorological Organization (WMO), 2025). Le quantile utilisé dans toute cette étude sera alors de 98%.

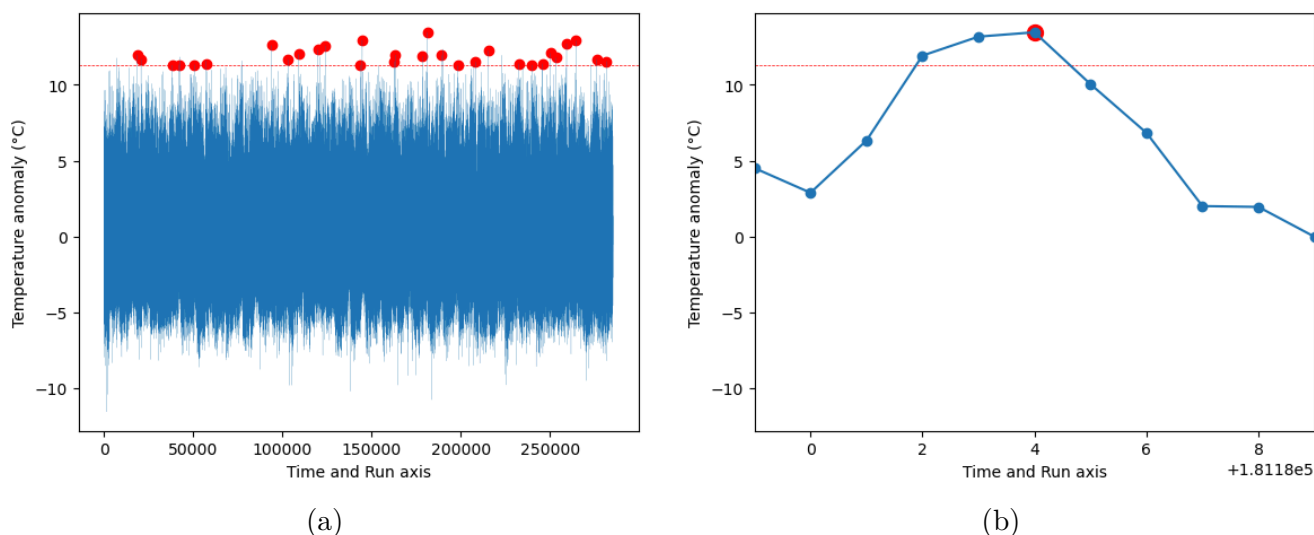


Fig. 3 : Représentations de la température moyenne par jour en fonction du temps (les différents runs sont agrégés à la suite, c'est ici juste l'index de chaque pas de temps, ~ 280000 jours) pour la période 1950-2014 à différentes échelles. (a) en bleu la série, en pointillés rouges la valeur du plus petit maxima choisi et en points rouges les maximas de température extrêmes ($n = 30$), (b) zoom sur un maxima qui permet de mieux comprendre l'intérêt de s'assurer de l'indépendance. Ici, trois valeurs clairement corrélées auraient été prises si nous n'avions pas imposé cette condition de 10 jours d'écart.

³Initialement, nous n'utilisons pas cette variable mais nous nous sommes rendus compte par la suite que l'humidité du sol dans CMIP6 n'était pas une donnée intéressante alors qu'elle devrait l'être intuitivement. Le cumul de précipitation permet alors d'émuler l'humidité du sol

3 ÉLÉMENTS THÉORIQUES ET MÉTHODES

3.1 Fonction committor

Un élément principal de mon travail sera les « fonctions committor ».

Introduisons ici dans le cadre d'un exemple ce qu'elles sont. Soit $\{\mathcal{X}(t)\}_{t \geq 0}$ un processus dynamique représentant l'état du système (prenons le climat à un point donné, Paris par exemple) à valeurs dans Ω dont l'une des composantes est la température T . Soit un ensemble cible $A = \{\mathcal{X}(t) : T > T_{\text{extreme}}\} \subset \Omega$ pour un certain T_{extreme} (quantile 98% par exemple), soit $B \subset \Omega$ l'ensemble des états initiaux possibles. $A \cap B \neq \emptyset$. On définit la fonction committor suivante

$$q(t, \tau) = \mathbb{P}(\mathcal{X}(t) \in A \mid \mathcal{X}(t - \tau) \in B), \quad \tau \geq 0. \quad (1)$$

Cette fonction représente la probabilité de passer dans un état A (chaleur extrême) sachant que nous étions dans un état B arbitraire τ unités de temps avant.

Le paramètre τ sera appelé le lag et quantifie le délai de conditionnement.

Un défi consiste à calculer q . Cette fonction est généralement calculé empiriquement à partir d'un grand nombre de trajectoires.

3.2 Analogues de circulation

Définissons ce que l'on appelle un analogue comme dans Fig. 4 et pourquoi on s'intéresse à ces derniers. Un analogue de circulation atmosphérique d'un jour t_0 est, dit simplement, un jour \tilde{t}_0 dont la configuration atmosphérique est similaire à celle observée à t_0 . Ces analogues permettent d'identifier des situations météorologiques comparables dans le passé. Ils sont utiles pour de nombreuses applications, par exemple pour reconstruire un champ de température à partir d'observations historiques, ou encore pour estimer une fonction committor.

Mathématiquement, on considère un tuple de champs (anomalies) qu'on a normalisé $\mathcal{X}(t) = \{\bar{\mathcal{Z}}(t, \varphi, \lambda), \bar{\mathcal{P}}(t, \varphi, \lambda), \bar{\mathcal{P}\mathcal{P}}(t, \varphi, \lambda)\}$, un jour $t_0 \in \llbracket 1, N \rrbracket$ et on définit⁴

$$d(\mathcal{X}(t_0), \mathcal{X}(t)) = \sqrt{\sum_i \frac{\alpha_i}{\dim(\mathcal{X}^i)} \sum_{\varphi, \lambda} (\mathcal{X}^i(t_0, \varphi, \lambda) - \mathcal{X}^i(t, \varphi, \lambda))^2}$$

Les analogues de t_0 sont alors

$$\tilde{t}_0^{(k)} = \left(\underset{t \neq t_0 \text{ et } |t-t_0| > 5 \text{ jours}}{\text{Argmin}_K} d(\mathcal{X}(t_0), \mathcal{X}(t)) \right)_k$$

ou Argmin_K représente les K jours dont la distance avec t_0 est la plus petite.

Lorsque les analogues seront calculés sur un seul champ, ils ne seront pas normalisés et lorsqu'ils seront calculés sur plusieurs champs ils seront normalisés « par rang », ie $\bar{\mathcal{Z}}(t, \varphi, \lambda) = \text{quantile}_t(Z(t, \varphi, \lambda))$. Le choix de cette normalisation vient du fait que les données ne suivent pas des lois normales, ce qui n'est pas étonnant : les données climatiques ont souvent des distributions à queues longues, plus ou moins lourdes selon la variable, et parfois asymétriques. Une normalisation par rang permet donc de rendre les variables comparables sans hypothèse de distribution sous-jacente. Pour le choix des poids $\{\alpha_i\}_i$ j'utiliserai une grille de recherche.

⁴Cette définition ressemble à celle de la RMSE. Dans la littérature la somme sur l'espace est parfois écrite comme une intégrale.

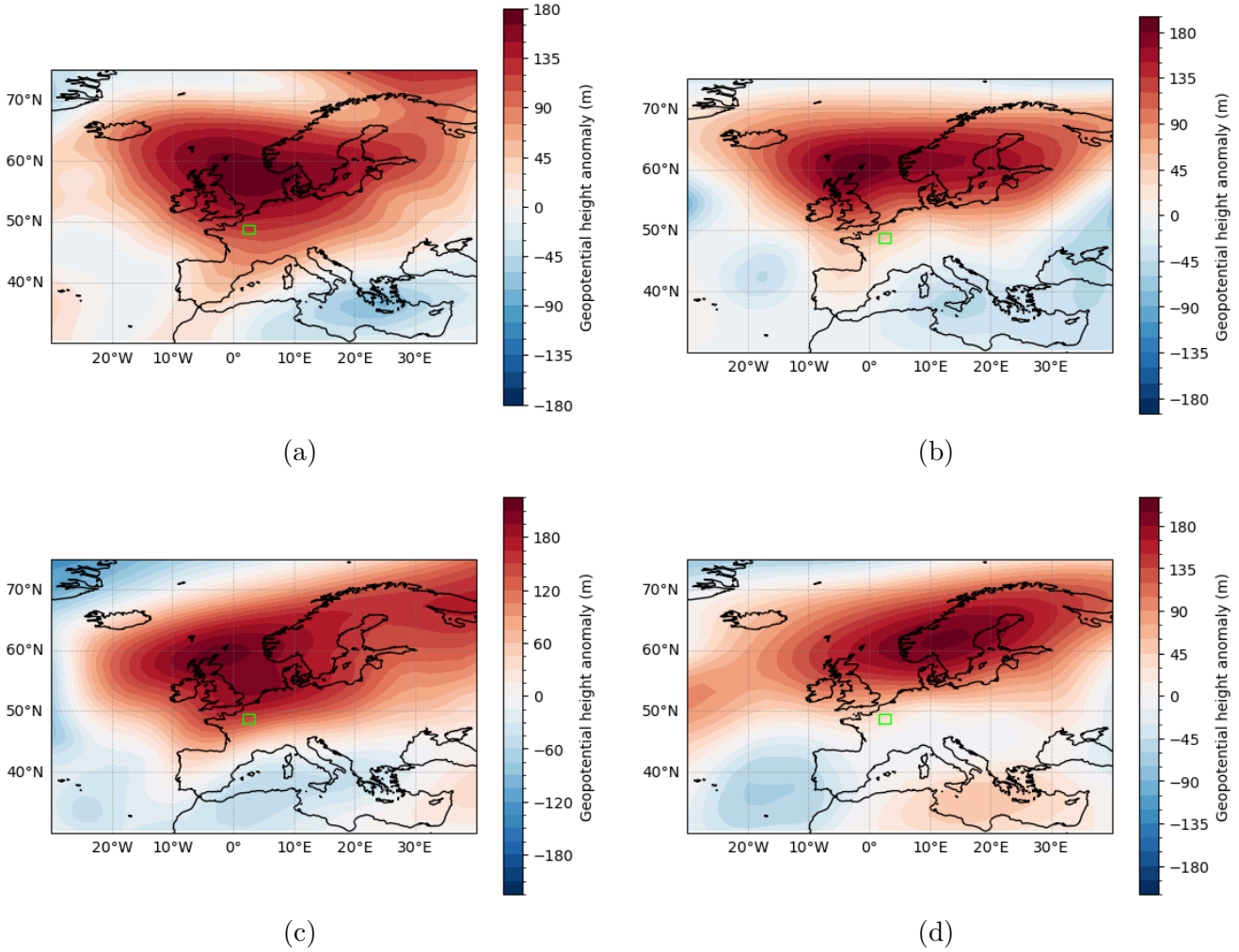


Fig. 4 : (a) les anomalies de géopotiel pour un évènement de température extrême (b)(c)(d) les anomalies de géopotiel pour des analogues de (a). Les corrélations spatiales de rang avec (a) sont respectivement de 0.80, 0.87 et 0.71. On constate aussi bien de part les corrélations que les tracés que les conditions atmosphériques sont bien similaires entre un jour et ses analogues.

3.3 Estimation de la fonction committor par les analogues

La procédure pour estimer de manière empirique la fonction committor (représentée dans la Figure 5) est la suivante;

1. Choisir un jour t_0 qui correspond à une vague de chaleur extrême.
2. Revenir dans le temps de τ jours jusqu'à $t_{-\tau}$.
3. Calculer les analogues de circulation atmosphérique via les données $\mathcal{X}[t_{-\tau}]$ pour ce jour (obtention des analogues $\{\tilde{t}_{-\tau}^{(k)}\}_{k \in [1, 20]}$).
4. Repartir de ces analogues $\{\tilde{t}_{-\tau}^{(k)}\}_{k \in [1, 20]}$ pour avancer dans le temps de τ jusqu'à $\{\tilde{t}_0^{(k)}\}_{k \in [1, 20]}$ jours et observer si oui ou non à $\{\tilde{t}_0^{(k)}\}_{k \in [1, 20]}$ on se retrouve dans des évènements de chaleurs extrêmes ou au moins de chaleurs conséquentes.

Les détails numériques du calcul de la RMSE et des analogues sont donnés dans Annexe -A. Il est important d'indiquer cependant que pour le calcul d'analogues il faut la même grille pour tous les modèles, les données d'ERA5 sont donc interpolées sur la grille de IPSL-CM6A-LR (downsampling). La fonction committor (1) est ensuite estimée comme le ratio entre le nombre de jours $\tilde{t}_0^{(k)}$ dont la température est supérieure au seuil et le nombre d'analogues calculés, ie

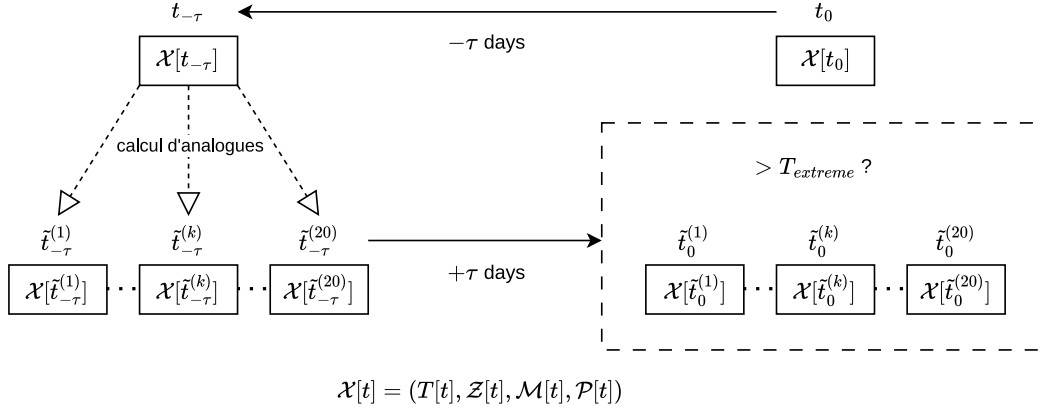


Figure 5 : Schéma de la procédure pour estimer la fonction committor avec les analogues

$$\forall \tau, \hat{q}(t_0, \tau) = \hat{\mathbb{P}}(T[t_0] > T_{\text{extreme}} \mid \mathcal{X}(t_0 - \tau), \tau) = \frac{1}{K} \sum_{k \in [1, K]} \mathbb{1}_{\{T[\tilde{t}_0^{(k)}] > T_s\}} \quad (2)$$

3.4 Modèle UNET

Le modèle U-Net (Ronneberger et al., 2015) a été initialement proposé pour la segmentation d'images biomédicales. Il s'agit d'un réseau de neurones convolutifs structuré en trois parties principales : un encodage (encodeur), un goulot d'étranglement (bottleneck) et un décodage (décodeur). L'encodeur extrait progressivement des représentations de plus haut niveau en réduisant la résolution spatiale, tandis que le décodeur reconstruit une carte de type voulu en restaurant la résolution initiale. Les connexions de type skip entre couches symétriques de l'encodeur et du décodeur permettent de combiner l'information contextuelle extraite en profondeur avec les détails fins préservés en surface, améliorant ainsi la précision en sortie.

Dans notre implémentation (Figure 6), le réseau est défini avec une profondeur de 4 niveaux et utilise 64 filtres convolutifs au premier niveau. Cette configuration assure un compromis entre richesse de représentation et coût computationnel.

Dans ce cadre, j'interpolerai les données de IPSL-CM6A-LR de sorte à ce que nos dimensions passent de 29×36 à 32×48 . C'est nécessaire pour que le maxpool 2×2 qui réduit les dimensions spatiales par 2 en prenant le max des sous matrices 2×2 ne cause pas de problèmes de dimensions non entières. Ainsi, il faut que les deux dimensions spatiales soient de tailles divisibles par 2^4 car le UNET présenté est de profondeur 4. Les données d'ERA5 seront interpolées sur la même grille. Les modèles UNET seront entraînés à l'aide de l'optimiseur Adam (Adaptive Moment Estimation (Kingma et Ba, 2017)) en minimisant comme fonction de perte l'erreur quadratique moyenne (Mean Squared Error, MSE). L'apprentissage sera régularisé par early stopping, interrompant le processus lorsque la performance sur l'ensemble de validation ne s'améliore plus, et par une réduction adaptative du taux d'apprentissage, divisé par deux lorsque la fonction de perte atteint un plateau, pour faciliter la convergence.

4 ÉTUDE EXPLORATOIRE

Cette partie est dédiée à l'exploration des variables présentées dans le Tableau 1. L'objectif est d'identifier des similarités entre les événements extrêmes et les conditions qui les précèdent, d'abord sur une simulation (r11f1p1) puis sur les 33 disponibles dans IPSL-CM6A-LR. Les résultats présentés ici seront ceux pour les 33 runs dans un souci de généralité. En fait, nous cherchons en quelque sorte à observer \mathcal{X}_0 dans 1.

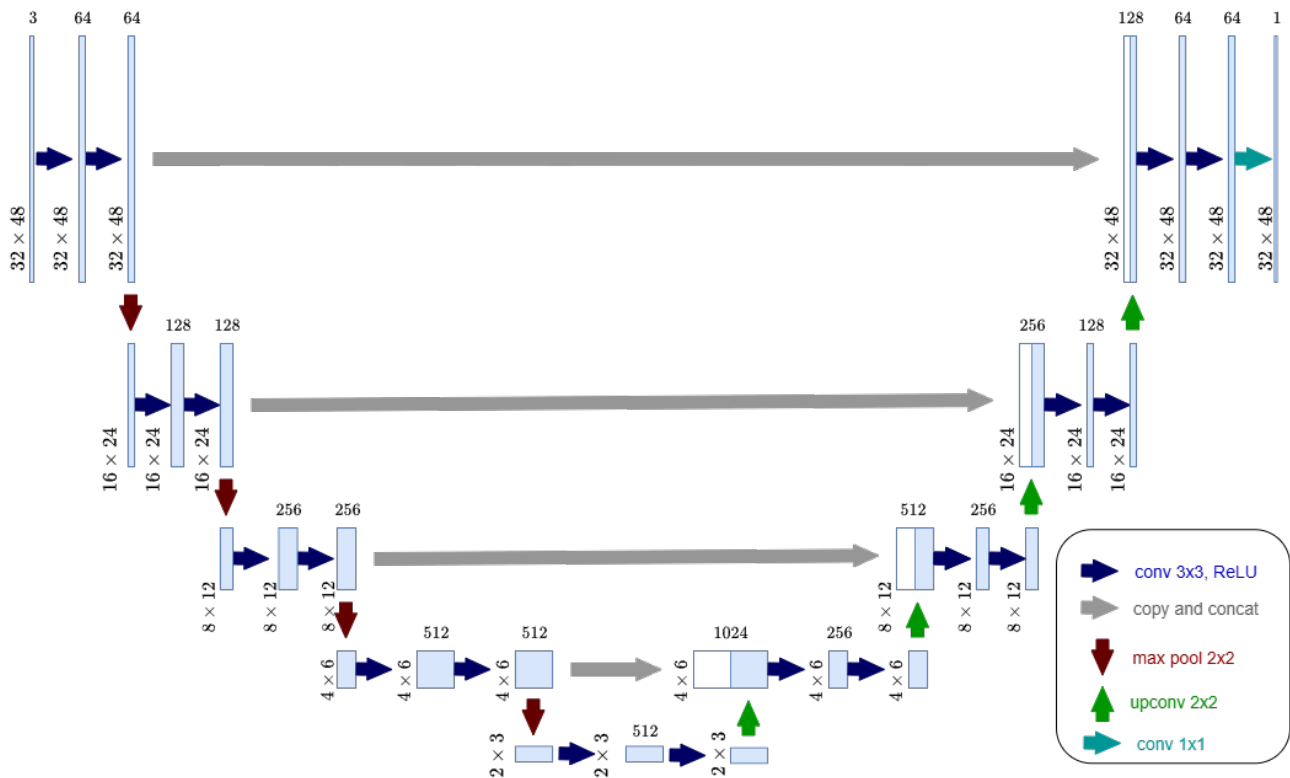


Figure 6 : Réseau U-net utilisé dans le stage. L’encodeur (gauche) combine convolutions 3×3 avec ReLU (flèches bleues) et sous-échantillonnage par maxpooling 2×2 (flèches rouges). Le décodeur (droite) reconstruit la carte via upconvolutions 2×2 (flèches vertes) et concaténation avec les cartes correspondantes du chemin contractant (flèches grises). La sortie est produite par une convolution 1×1 (flèche turquoise).

4.1 Cartes composites

Afin d’identifier des similarités sur les autres champs de A , je trace des cartes composites. En climatologie, une carte composite (souvent appelée composite map) est une représentation moyenne de plusieurs cartes climatiques correspondant à des périodes ou événements similaires. Elles servent à dégager les conditions typiques associées à un phénomène précis (ici les vagues de chaleur extrême). Elles synthétisent aussi des dizaines de cartes pour chaque variable et chaque extrême, cependant elles peuvent dissimuler la diversité des cas individuels.

4.2 Typicalité des événements de température extrêmes dans IPSL-CM6A-LR

Il a été montré (Noyelle et al., 2023) que la variance entre les trajectoires amenant à des événements extrêmes diminuait avec le niveau d’extrêmes. Autrement dit, plus l’extrême de température est intense plus il y a de chance que les conditions y ayant mené se ressemblent.

Ces conditions que nous observons bien sur Fig. 7 sont caractérisées par de faibles précipitations et des dépressions bloquées à l’Ouest de la France par des très hautes anomalies de géopotential sur la France que l’on appelle blocage en oméga (dû à leur forme particulière qui ressemble à la lettre grecque Ω). Ce phénomène de blocage en oméga Figure 8 fait remonter l’air chaud, et empêche l’air plus frais de progresser.

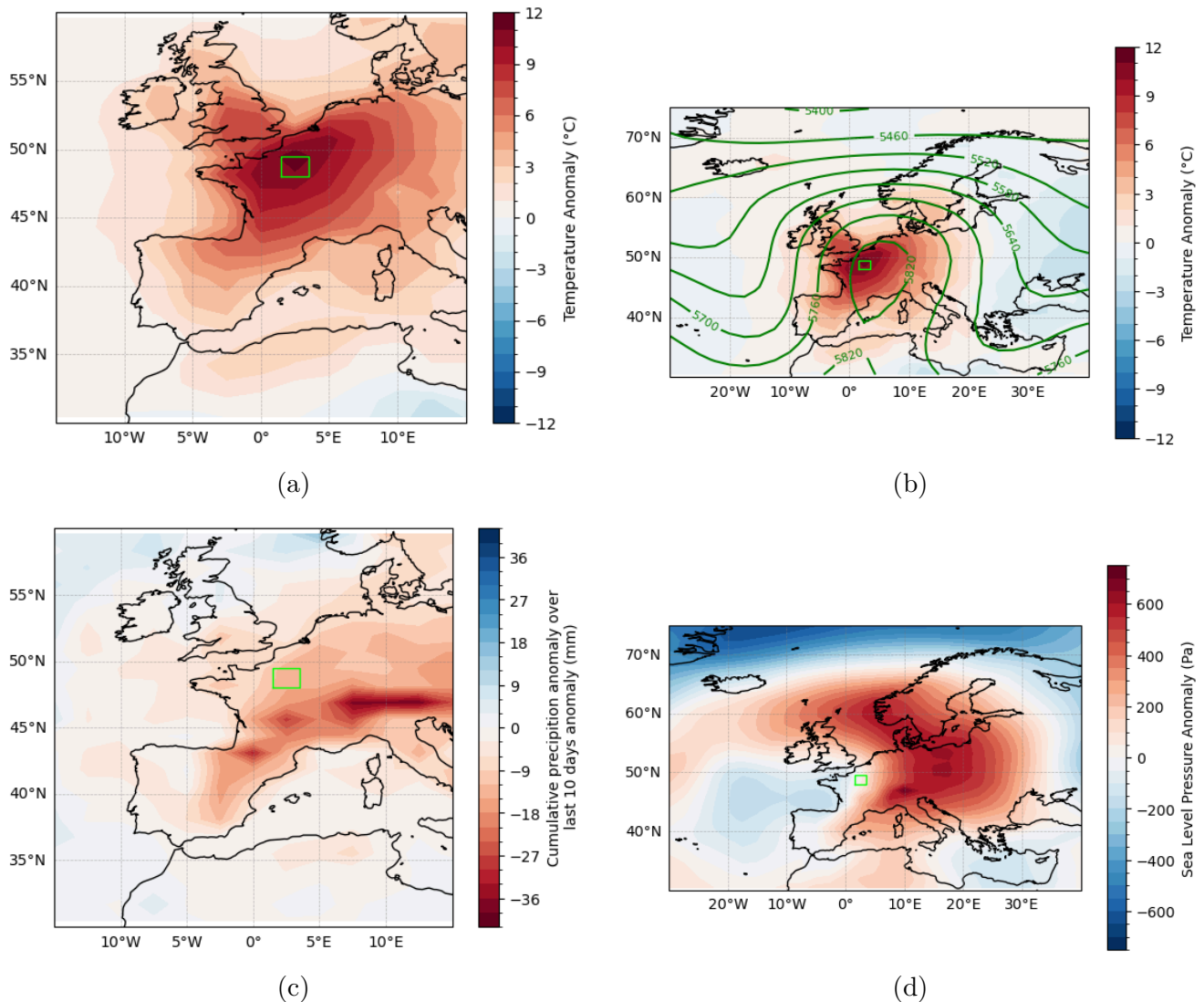


Fig. 7 : Cartes composites des anomalies de différents champs dans pour les 5 plus grands extrêmes (a) Anomalies de température (b) Anomalies de température en couleurs et hauteur géopotentielle à 500hPa (m) en vert. Le choix d'afficher la hauteur géopotentielle et non ses anomalies est dû à la volonté d'afficher le phénomène de blocage Oméga (c) Anomalie du cumul de précipitations (d) Anomalie de la pression au niveau de la mer.

4.3 RMSE et corrélations spatiales entre évènements extrêmes

Je souhaite désormais identifier si les différents champs sont similaires lors d'un évènement de température extrême. Cela passe par le calcul de RMSEs et corrélations spatiales, travaillant sur des extrêmes il est plus adapté de s'attarder sur une corrélation de rang plutôt que de valeurs pour identifier des patterns quand la variance n'est pas la même partout.

En fonction du champ considéré, la RMSE et les corrélations ne sont pas calculés sur la même zone donc sur des grilles différentes, le choix de la zone se fait par rapport aux cartes composites, typiquement une zone réduite uniquement sur la France est choisie pour les précipitations et l'humidité du sol. Afin d'avoir quelque chose avec quoi comparer nos résultats, en plus de calculer la RMSE moyenne entre les extrêmes, je la calcule entre des évènements aléatoires parmi nos données.

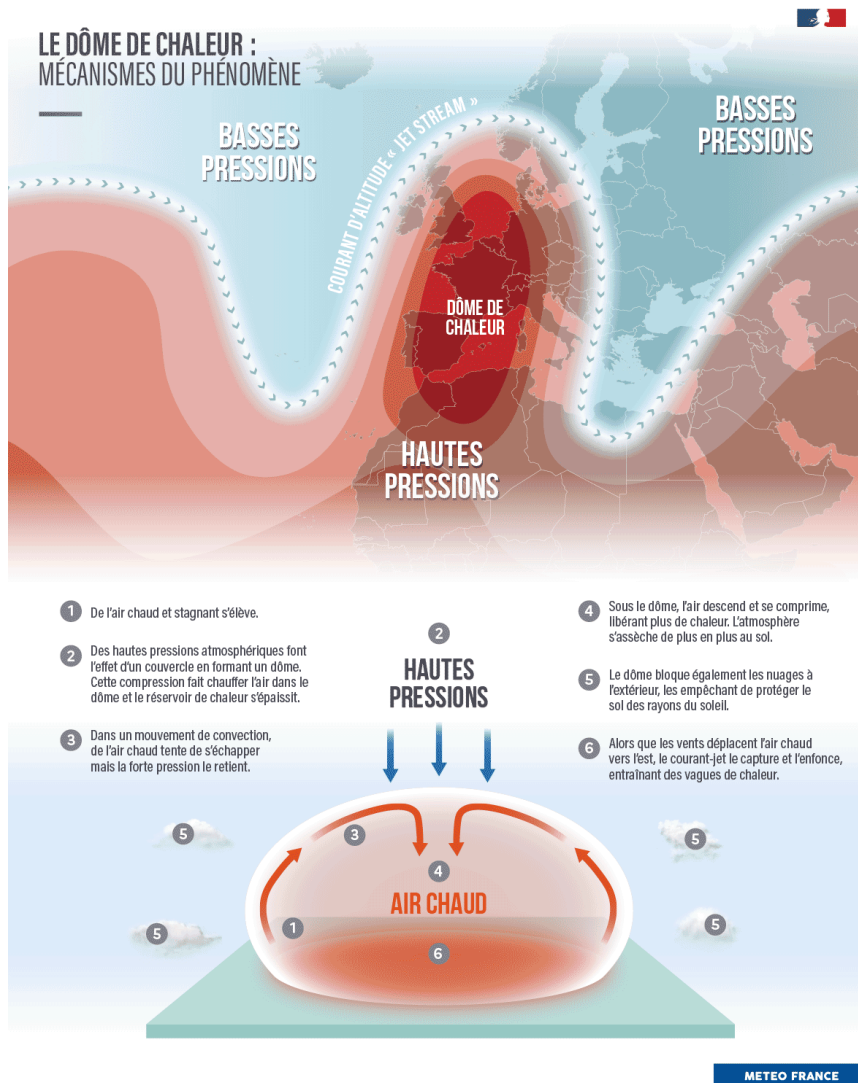


Figure 8 : Infographie d'explication du phénomène de blocage en oméga, principal responsable des vagues de chaleur (Une vague de chaleur exceptionnelle par sa durée en début d'été | Météo-France).

Les résultats (Figure 10) à $\tau = 0$ nous indiquent que la variabilité des données est réduite pour les événements extrêmes: c'est ce que nous espérons constater.

Les résultats (Figure 9) à $\tau = 0$ sont prometteurs (la corrélation globale, entre des événements aléatoires pour les anomalies est de 0) pour la suite, ils rassurent sur le fait qu'il y a une similarité entre les extrêmes.

Maintenant que nous avons identifié nos événements extrêmes, nous pouvons commencer à « revenir en arrière » dans le temps pour voir comment évolue le système lors de ces derniers.

Pour avoir une idée de combien de jours τ avant les événements extrêmes une dynamique se dégage nous traçons l'évolution de la corrélation et des RMSE. Nous observons des corrélations (Figure 9) qui baissent et des RMSE (Figure 10) qui augmentent au fur et à mesure qu'on s'éloigne, ce qui est tout à fait normal, c'est dû à la nature chaotique du climat, plus on s'éloigne de ces événements atypiques plus la corrélation se rapproche de celle d'événements complètement indépendants. Nous remarquons tout de même qu'à partir de 4-5 jours avant il devient quasiment impossible d'identifier des similarités dans les trajectoires menant à des événements de vagues de chaleur extrême (corrélations inférieures à 0.2). Cependant, cela n'est pas forcément très alarmant. En effet les vagues de chaleur extrême étant des

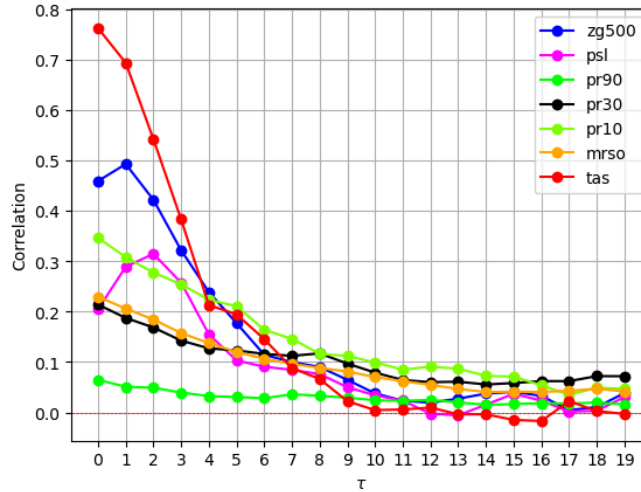


Figure 9 : Évolution de la corrélation spatiale de rang pour les anomalies lorsque l'on s'éloigne de τ jours avant les extrêmes. Les pointillés rouges correspondent à la corrélation « globale », calculée entre des évènements aléatoirement choisis.

évènements qui sont induits ou influencés par tous ces champs; nous pouvons espérer que l'utilisation de plusieurs champs en même temps avec une pondération sagement choisie améliorera nos résultats.

Les résultats de Figure 9 sont cependant décevants quant à l'utilisation de l'« humidité du sol \mathcal{M} , c'est probablement lié à la difficulté de modéliser les interactions du sol dans les modèles. Nous observons cependant Figure 11 à titre de comparaison que dans ERA5 cette variable est logiquement bien une variable intéressante, les vagues de chaleur ont principalement lieu quand le sol est sec. Au vu de ces résultats, il ne sera pas possible de prendre cette information en compte dans IPSL-CM6A-LR, nous utiliserons plutôt $pr(10)$, qui semble plus pertinente.

4.4 Cartes de l'évolution

Afin d'essayer d'identifier « à la main » des dynamiques similaires du système menant à des vagues de chaleur extrêmes nous traçons l'évolution du système sur les 12 jours précédant l'évènement. J'en ai fait des gifs disponibles [en ligne](#).

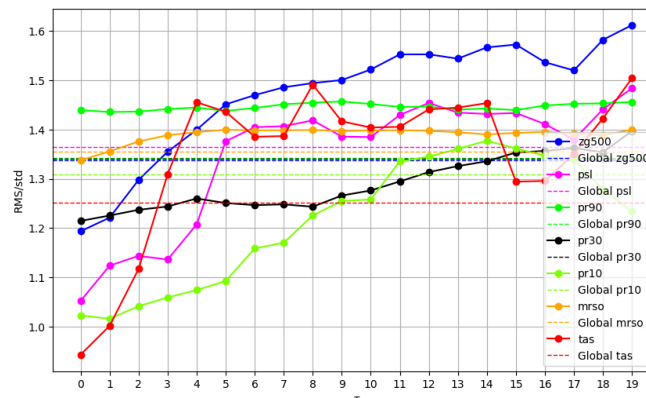


Figure 10 : Évolution de la RMSE normalisée par la variance sur la zone de calcul pour les anomalies lorsque que l'on s'éloigne de τ jours avant les extrêmes. Les pointillés de différentes couleurs correspondent à la RMSE « globale », calculée entre des évènements aléatoirement choisis.

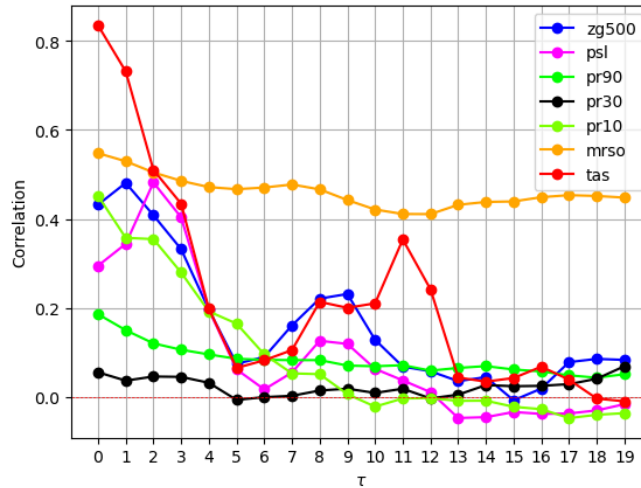


Figure 11 : Évolution de la corrélation pour les anomalies lorsque l'on s'éloigne de τ jours avant les extrêmes dans ERA5. Les pointillés rouges correspondent à la corrélation « globale », calculée entre des événements aléatoirement choisis. Le champ \mathcal{M} n'est pas soumis aux mêmes conventions dans les deux modèles. Dans ERA5 l'humidité volumétrique du sol (mm^3/mm^3) est divisée en 4 couches verticales de 0–7cm, 7–28cm, 28–100cm et 100–289cm alors que dans IPSL-CM6A-LR on a accès uniquement à l'humidité du sol (kg/m^2) dans les 10 premiers cm ainsi que sur toute la profondeur, on sait cependant (Boucher et al., 2020) qu'un modèle hydrologique en 11 couches sur 2m est utilisé. En faisant l'hypothèse que l'humidité du sol est répartie uniformément selon les couches on peut convertir les données d'ERA5 en des données comparables à celle du modèle de l'IPSL. Il suffit de sommer les 4 couches en les multipliant par $1000 \times \text{profondeur (m)}$ où 1000 correspond à la masse volumique de l'eau pour les passer en (kg/m^2) puis de multiplier par 200/289.

Nous retrouvons les conditions évoquées dans Chapitre 4.2 avec ces animations, l'humidité du sol qui augmente progressivement jusqu'à l'extrême de température et des conditions atmosphériques progressant vers des blocages.

4.5 Conclusion

Nous pouvons conclure que le modèle IPSL-CM6A-LR contient bien des vagues de chaleur s'apparentant à ce que nous pouvons observer dans ERA5, ce qui pouvait être attendu étant donné que ces modèles modélisent la physique du climat. Nous avons observé qu'il y a des similarités entre les événements extrêmes et les trajectoires y amenant. Cependant, nous avons quand même observé des différences fondamentales entre les modèles de l'IPSL et ERA5, notamment sur l'humidité du sol que nous émulerons donc par les cumuls de précipitations.

5 RÉSULTATS

5.1 Analogues

Dans toute cette partie, les analogues de circulation seront toujours calculés sur l'ensemble des données, sans quelconque séparation entraînement/validation car cela ne s'y prête pas vraiment.

5.1.1 Qualité globale des analogues

La Fig. 4 nous permet déjà de vérifier visuellement de la qualité des analogues calculés. Lorsque les analogues sont calculés sur une seule variable, on peut vérifier la corrélation de rang entre un jour et

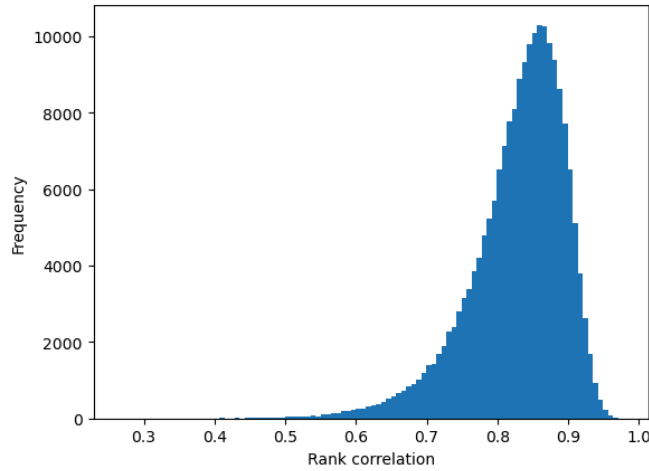


Figure 12 : Histogramme des corrélations de rang entre les jours dans ERA5 et leurs analogues dans ERA5 et IPSL-CM6A-LR calculés sur la variable \mathcal{P} . 95% des corrélations sont supérieures à 0.7.

ses analogues (Figure 12). Nous observons des corrélations qui sont très bonnes et nous rassurent pour la suite.

5.1.2 Prévisions, estimation de la fonction committor et « transfer learning »

Nous souhaitons maintenant observer à la main ce qu'il se passe si nous réalisons la procédure illustrée dans la Figure 5. La procédure est réalisée pour les 5 extrêmes les plus intenses.

Avec des analogues calculés dans ERA5 sur les champs \mathcal{Z} , \mathcal{P} et ρr , \mathcal{T} choisis respectivement avec des pondérations de 0.4, 0.4, 0.1 et 0.1 nous calculons la moyenne des anomalies de températures obtenues en $\left\{ \tau_0^{(k)} \right\}_{k \in \llbracket 0, 20 \rrbracket}$ au point le plus proche de Paris. Le choix de ces poids est fait après avoir cherché la combinaison qui minimise la RMSE entre le champ original de température et sa reconstruction sur une grille de pas 0.05 pour chaque variable.

Nous observons (Figure 13) que les anomalies obtenues sont supérieures à 1°C jusqu'à environ 4 jours à l'avance si l'on calcule les analogues uniquement dans ERA5 et jusqu'à 8 jours à l'avance si l'on calcule dans ERA5 et IPSL-CM6A-LR. Pour s'assurer de la validité de nos résultats nous réalisons aussi la même chose pour 200 jours aléatoirement choisis parmi ceux d'ERA5 afin d'observer un comportement plus global qui assez logiquement prévoit une anomalie de 0°C degré.

Grâce à ces analogues nous pouvons calculer une première estimation de la fonction committor pour les 5 extrêmes de température les plus intenses en calculant (2) avec T_{extreme} le quantile de niveau 98% des anomalies de température en JJA dans ERA5. De la même manière que juste avant nous calculons aussi cette même estimation pour 200 jours parmi ceux d'ERA5 afin d'observer un comportement plus global, nous observons que la fonction committor bien que proche de 0 n'y est pas exactement égale, il y a donc un bruit. Jusqu'à 3 jours à l'avance d'un extrême nous sommes capables de dire que la probabilité pour qu'un évènement de température extrême ait lieu est au moins deux fois supérieure au bruit si les analogues sont calculés dans ERA5. Si les analogues sont calculés cette fois dans les deux modèles on arrive à 5 jours. Cela signifie que grâce au « transfer learning », au lieu d'alerter 3 jours à l'avance de l'occurrence d'une vague de chaleur, nous pouvons alerter 5 jours à l'avance (voire potentiellement plus, la fonction committor moyenne diminue à $\tau = 6$ dans (Figure 14) mais est sinon 5 fois plus grande que le bruit jusqu'à 8 jours d'avance).

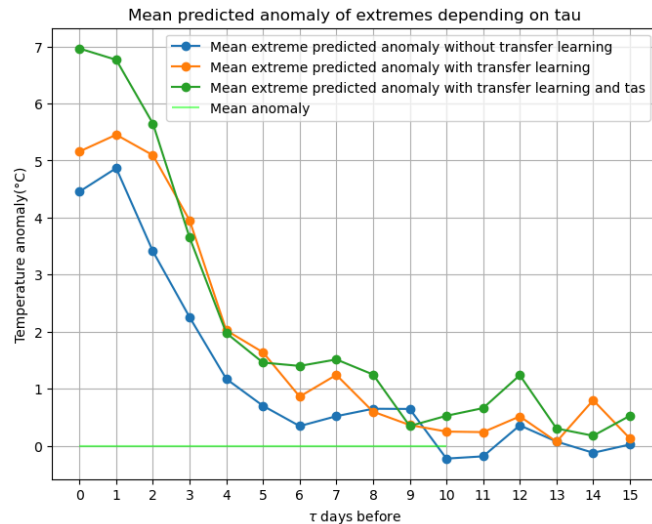


Figure 13 : Évolution de la température moyenne au point le plus proche de Paris des 20 analogues obtenus en fonction du nombre de jours τ imposé dans ERA5. En vert clair l'anomalie que l'on prédit pour 200 jours choisis aléatoirement.

La Figure 14 nous a permis d'évaluer la qualité de notre méthode à estimer les vrais positifs, à estimer des grandes probabilités quand elles doivent l'être. Nous souhaitons maintenant aussi évaluer si notre modèle est précis (ne produit pas trop de faux positifs, de probabilités grandes alors que la température ne l'est pas). Pour cela, nous construisons l'histogramme de l'ensemble des températures observées au point le plus proche de Paris, puis nous y superposons l'histogramme des températures correspondant aux instants où la fonction committor dépasse un seuil prédéfini pour la journée considérée. Ce seuil est choisi arbitrairement comme significatif (idéalement plus grand que le bruit, cf Figure 14).

Nous observons sur la Fig. 15 que de manière générale quand la fonction committor est grande (ici ≥ 0.3) alors la température réelle est effectivement au-dessus de la moyenne, avec ou sans « transfer learning ». La seule différence notable que nous constatons dans la Fig. 15 entre les histogrammes c'est que l'ajout du transfer learning permet d'identifier plus de jours pour lesquels il y a de grandes chances que la température dépasse T_{extreme} et de manière précise.

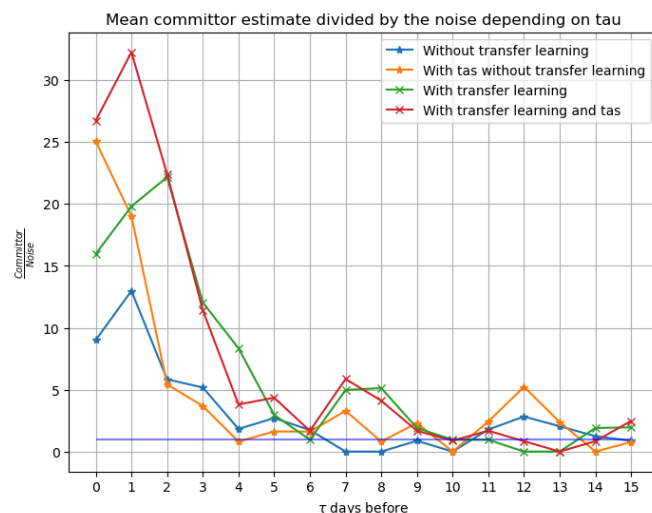


Figure 14 : Évolution de la fonction committor moyenne divisée par le bruit obtenue en fonction du nombre de jours τ imposé dans ERA5. Plus le ratio se rapproche de 1 (abscisse bleue), moins l'estimation est bonne.

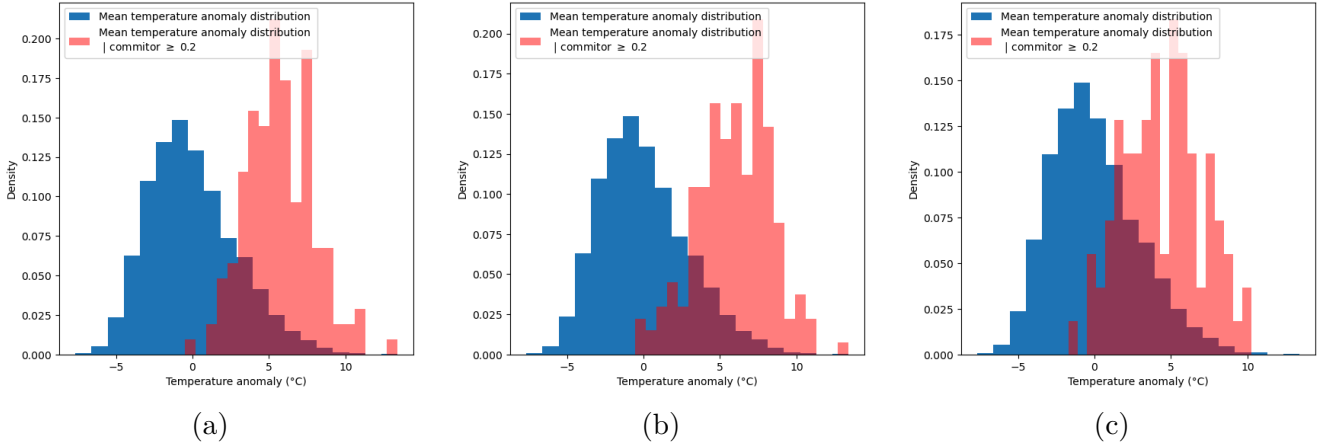


Fig. 15 : Histogrammes des températures et des températures sachant que la fonction committor est supérieure à 0.3 à $\tau = 2$ jours d’avance. (a) sans « transfer learning », 38 jours sont représentés en rouge (b) avec « transfer learning », 68 jours sont représentés en rouge (c) avec « transfer learning » et la température comme variable dans le calcul des analogues, 138 jours sont représentés en rouge

Nous concluons que les résultats obtenus sont encourageants. Grâce à l’utilisation des données du modèle CMIP6 en plus de celles d’ERA5, nous améliorons les résultats de prévision de plusieurs jours. Ces résultats sont prometteurs pour le « transfer learning » en utilisant un modèle plus complexe type U-NET.

5.2 Machine learning

Cette partie est consacrée à l’apprentissage et l’évaluation de modèles de deep learning. Plus précisément aux UNET, dont l’utilisation pour estimer une fonction de transfert $\mathcal{T} = F(\mathcal{P})$ (Cariou et al., 2025) dans le modèle MIROC donne de très bons résultats, c’est ce qui motive le choix de ce modèle. S’intéresser à une fonction de transfert pourrait permettre d’estimer, la fonction committor (en observant la température prédite au point donné, en ayant perturbé légèrement les conditions initiales).

5.2.1 Fonction de transfert au jour même dans IPSL-CM6A-LR

Pour commencer, nous évaluons l’utilisation d’un UNET pour la même tâche que (Cariou et al., 2025) mais dans IPSL-CM6A-LR. Nous obtenons des résultats très bons (Figure 16), améliorés aussi par l’ajout en entrée de deux nouvelles variables pour modéliser $\mathcal{T} = F(\mathcal{P}, \mathcal{Z}, \boldsymbol{pr})$.

Les résultats (Figure 17, Figure 18) sont très bons. Le réseau UNET reconstruit la température en fonction des 3 variables \mathcal{P}, \mathcal{Z} et \boldsymbol{pr} de manière très efficace. Sur l’Europe, aucun point ne descend en dessous des 0.95 de corrélation et le ratio des variances est à 0.9 pour une grande partie de l’Europe ce qui signifie que le réseau UNET arrive bien à reproduire la variabilité temporelle des anomalies de températures. Ce qui est aussi intéressant est que lorsque l’anomalie de température est extrême le modèle arrive bien à reconstruire une anomalie très grande (malgré une différence d’environ un °C pour les évènements les plus extrêmes).

5.2.2 Fonction de transfert avec $\tau > 0$ jour d’avance dans IPSL-CM6A-LR

Désormais, nous introduisons un décalage entre la reconstruction(prédiction) et les données en entrée, de la même manière qu’avec les analogues mais cette fois avec comme variable d’entrée supplémentaire \mathcal{T} . Concrètement, nous essayons de prédire le champ de température dans 2 jours par exemple connaissant $\mathcal{T}, \mathcal{P}, \mathcal{Z}$ et \boldsymbol{pr} .

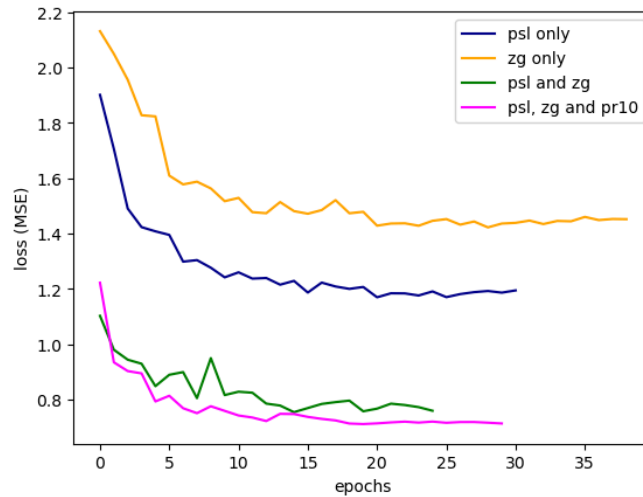


Figure 16 : Évolution de la perte (Erreur Quadratique Moyenne) du set de validation en fonction des epochs (nombre de fois ou le modèle a « vu » l'entièreté des données d'entraînement). Ajouter des variables augmente bien les résultats.

Nous observons directement grâce au tracé de la perte (Figure 19), aux figures Annexe -C et (Figure 20) qu'après 4 jours les résultats sont globalement décevants, avec des MSE bien trop grandes. Pour

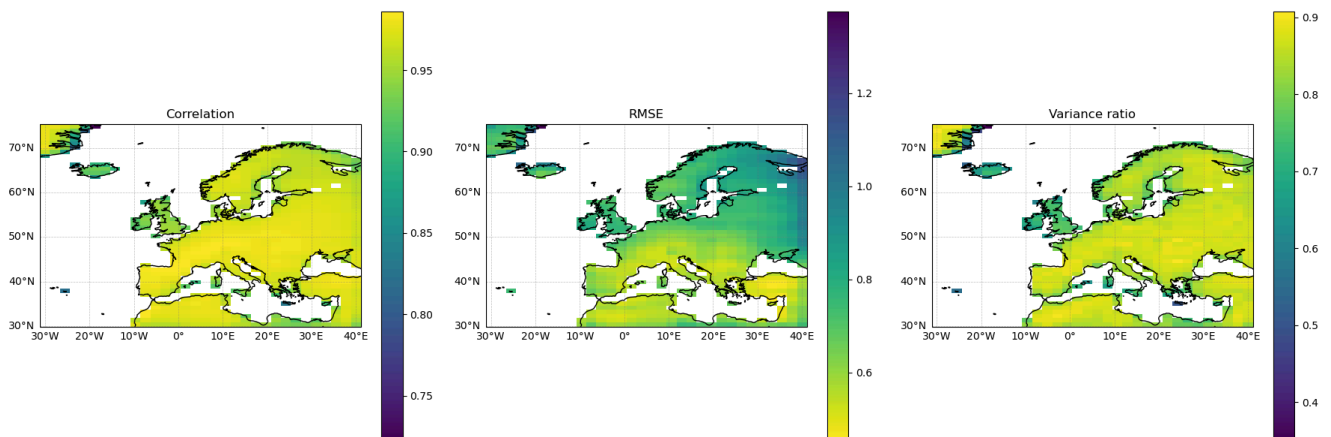


Figure 17 : Corrélations(rang), RMSE et ratio des variances pour des reconstructions d'anomalies de températures journalières pour le set de validation. Les valeurs moyennes sur toute la zone considérée (pas uniquement la terre ferme) sont respectivement de 0.88, 0.8 et 0.65

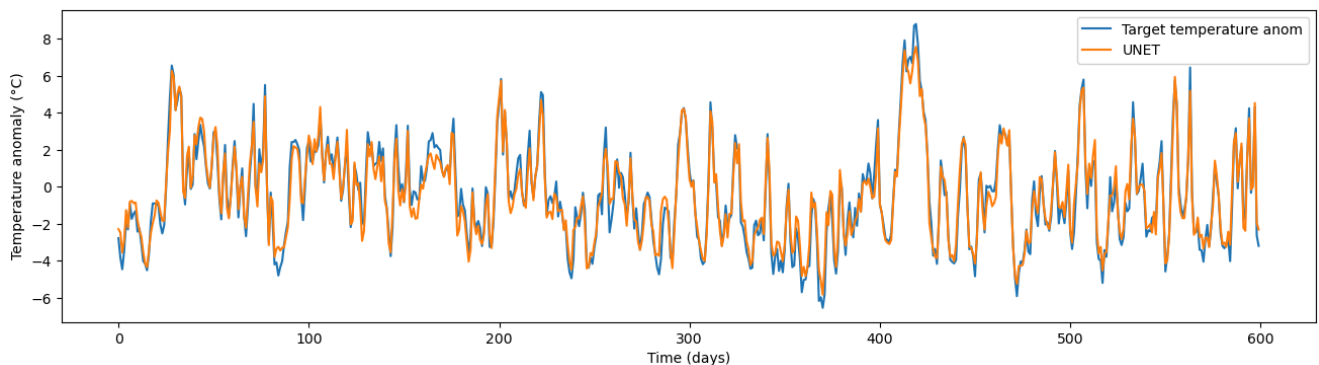


Figure 18 : Anomalies de températures pour Paris reconstruites (orange) et anomalies réelles (bleu). Le CCC entre ces deux séries est de 0.97

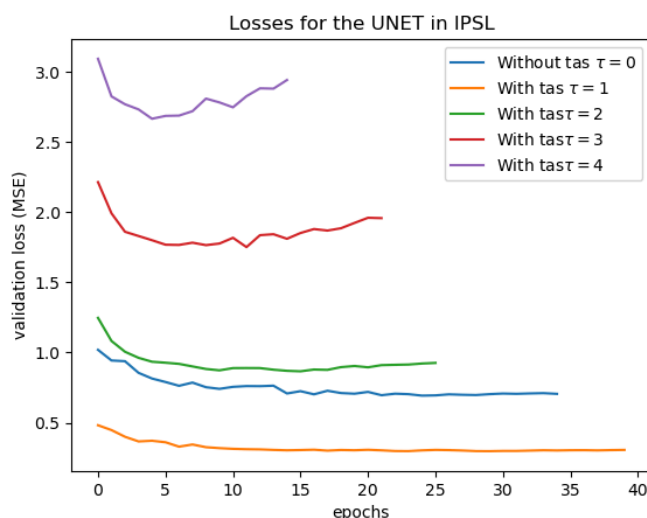


Figure 19 : Évolution de la fonction de perte du set de validation en fonction des epochs pour différents lags.

les prédictions de 1 à 3 jours cependant, nous observons quand même des corrélations sur l'Europe assez grandes (>0.95 à $\tau = 1$, >0.9 à $\tau = 2$ et >0.75 à $\tau = 3$); les 2 autres métriques sont elles aussi relativement bonnes.

Nous observons (Figure 20) une difficulté du réseau UNET à prévoir les anomalies les plus éloignées de 0 à mesure que l'on ajoute du lag. Ceci paraît normal. Le modèle est tout de même capable de prédire de très grandes anomalies même à $\tau = 3$, qui diminuent avec τ .

5.2.3 « transfer learning »

Nous souhaitons désormais voir si nous gagnons en performances dans ERA5 avec du « transfer learning ». Concrètement, le « transfer learning » revient à utiliser ce que le modèle a appris lors de la réalisation d'une tâche pour la réalisation d'une autre tâche similaire, ici nous souhaitons entraîner un réseau UNET dans CMIP-CM6A-LR puis le re-entraîner dans ERA5 en espérant obtenir un gain de performances. Nous observons aussi ce qu'il se passe si on donne simplement des données d'ERA5 en entrée du modèle entraîné sur l'IPSL sans le re-entraîner.

Nous utilisons les modèles à $\tau = 0$, avec une profondeur de 3 (diminuée de 1 pour un entraînement plus rapide afin de faire des tests) et 64 filtres à la base. On notera

$\mathcal{M}_{\text{control}}$ Le modèle entraîné sur le set d'entraînement d'ERA5

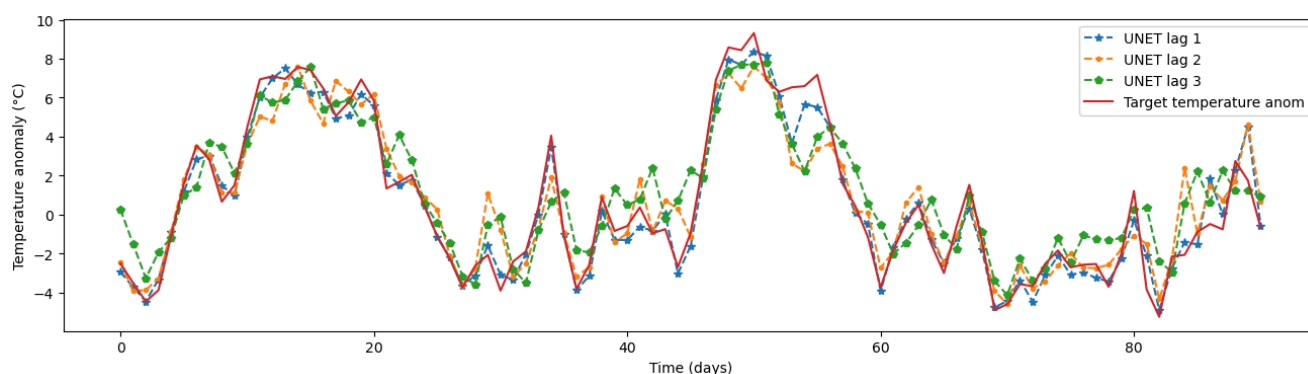


Figure 20 : Anomalies de températures (bleu) $\tau = 1$ (orange) $\tau = 2$ (vert) $\tau = 3$ pour Paris reconstruites et anomalies réelles (rouge). Les CCC entre les deux séries sont respectivement de 0.96, 0.91 et 0.74.

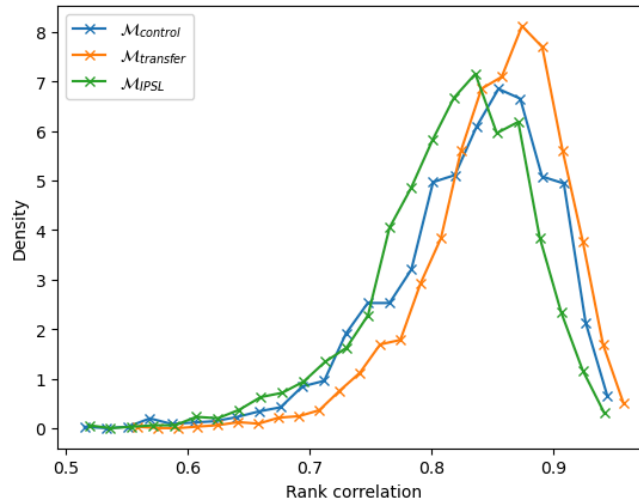


Figure 21 : Histogrammes de la corrélation de rang pour les trois modèles. Chaque point est placé au milieu des barres habituelles dans un souci de visualisation.

\mathcal{M}_{IPSL} Le modèle entraîné sur le set d’entraînement de IPSL-CM6A-LR (on lui donnera en entrée des données provenant du set de validation d’ERA5 à titre de comparaison)

$\mathcal{M}_{transfer}$ Le modèle entraîné sur le set de d’entraînement d’IPSL-CM6A-LR puis re entraîné sur celui d’ERA5

Les résultats obtenus sont nuancés. D’un côté les corrélations Figure 21 pour $\mathcal{M}_{transfer}$ sont meilleures, il est visible que $\mathcal{M}_{transfer}$ produit moins de reconstructions dont la corrélation est inférieure à 0.7.

Quand nous regardons les séries temporelles de températures au point le plus proche de Paris (Figure 22) il est difficile d’identifier un modèle vraiment meilleur qu’un autre entre $\mathcal{M}_{control}$ et $\mathcal{M}_{transfer}$. Seul \mathcal{M}_{IPSL} se démarque comme étant moins bon. Nous pouvons confirmer cela en calculant certaines statistiques sur les 3 séries pour les comparer (Tableau 2) où nous observons observation est que les séries sont très similaires et que $\mathcal{M}_{control}$ est meilleur en terme de CCC.

Notre objectif premier était de travailler sur les vagues de chaleur. Avec ces 3 modèles, si nous isolons les évènements les plus extrêmes en terme de chaleur dans le set de validation d’ERA5, nous retrouvons des évènements très chauds mais bien moins précis que lorsque nous travaillions dans IPSL-CM6A-LR avec ici environ 2.5 degrés de différence avec la vraie température.

Nous concluons donc par rapport au « transfer learning » pour les algorithmes de deep learning que c’est une méthode qui pourrait être intéressante mais dépend de beaucoup de variables dont principalement la manière dont le « transfer learning » est effectué et la qualité des données du GCM utilisé. Bien que les 3 modèles soient suffisamment fonctionnels, il apparaît que les données de IPSL-CM6A-LR ont encore du mal à suffisamment bien reproduire la variabilité présente dans ERA5 pour qu’elles soient utilisées comme données d’entraînement supplémentaires. Ou alors il faudrait passer par des méthodes de « transfer learning » plus sophistiquées. Mais la très légère amélioration que nous avons observée sur $\mathcal{M}_{transfer}$ nous indique que c’est un axe à continuer d’explorer, potentiellement avec d’autres GCMs aussi (notamment à plus grande résolution). Les résultats obtenus à $\tau = 0$ se retrouvent à $\tau > 0$ avec une baisse de qualité des résultats comme dans IPSL-CM6A-LR

6 CONCLUSION

L’objectif de ce stage était double : premièrement, de se familiariser avec les calculs d’analogues et les réseaux UNETs (et toute la théorie nécessaire en deep learning, car le cours est proposé en 5A, via

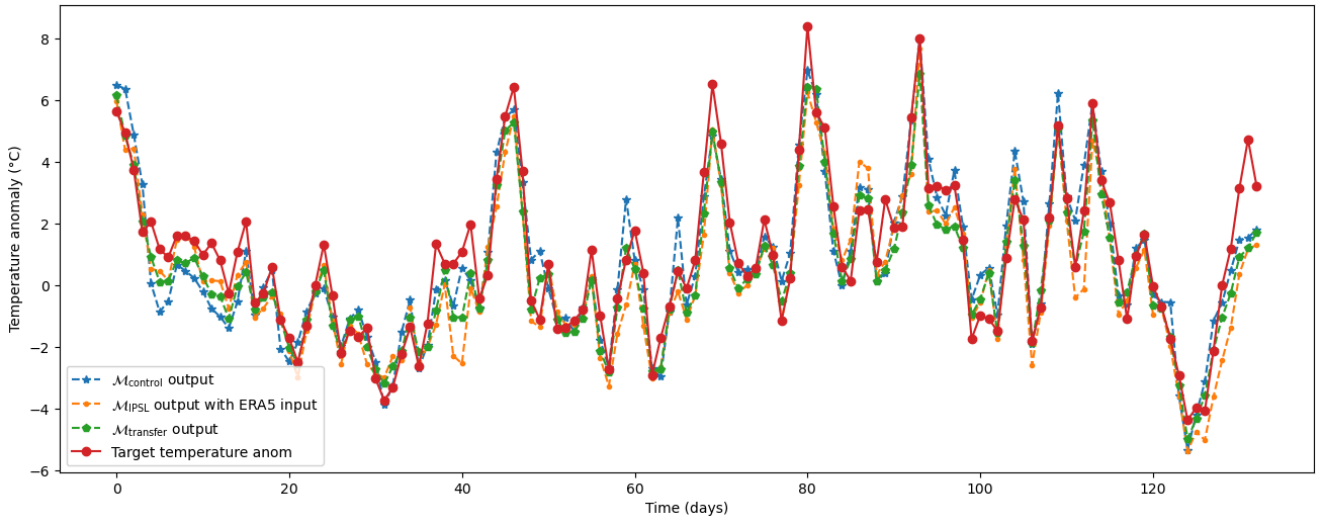


Figure 22 : Séries temporelles de la température au point le plus proche de Paris pour un été dans ERA5. En rouge la série réelle, en bleu $\mathcal{M}_{\text{control}}$, en orange $\mathcal{M}_{\text{IPSL}}$ et en vert $\mathcal{M}_{\text{transfer}}$

Tableau 2 : Différentes métriques sur les séries temporelles extraites pour Paris des ensembles de validation. En gras la meilleure valeur pour chaque métrique.

| Métrique | $\mathcal{M}_{\text{control}}$ | $\mathcal{M}_{\text{IPSL}}$ | $\mathcal{M}_{\text{transfer}}$ |
|---------------------|--------------------------------|-----------------------------|---------------------------------|
| Corrélation de rang | 0.93 | 0.93 | 0.95 |
| Corrélation | 0.94 | 0.94 | 0.96 |
| CCC | 0.94 | 0.9 | 0.93 |
| RMSE | 1 | 1.23 | 0.99 |

les documentations de TensorFlow/Keras et le très bon livre ISLP) pour les implémenter en Python de manière optimisée sur des données climatiques du modèle de l'IPSL et des réanalyses ERA5, dont les formats sont différents des données avec lesquelles nous avons l'habitude de travailler (données spatio-temporelles sous format .nc). Le deuxième objectif, que nous pouvons considérer comme l'objectif « scientifique » auquel nous essayons de répondre, était d'évaluer le « transfer learning » entre les données du modèle IPSL-CM6A-LR et ERA5 pour les deux méthodes implémentées, c'est-à-dire évaluer la viabilité de tirer des conclusions pour les données d'ERA5 à partir des données d'IPSL-CM6A-LR.

Premièrement, en conclusion du travail effectué, je pense qu'il est important de noter l'importance d'un code propre, réutilisable et optimisé. Pour des raisons aussi bien pratiques qu'éthiques, cela permet à d'autres personnes de reprendre le travail effectué si besoin et à limiter le temps de calcul⁵ et l'utilisation de ressources des serveurs qui ont besoin d'être climatisés. Le travail qui était prévu a été réalisé et nous avons pu conclure sur les questions posées au début du stage en réalisant de nombreuses versions plus ou moins bonnes.

Puis, en conclusion plus scientifique, le « transfer learning » s'avère être une option intéressante pour les modèles climatiques basés sur l'IA, qui ne sont pas des modèles de circulation générale et qui ont besoin de données en entrée pour apprendre. Nous avons vu avec les analogues que la qualité des analogues était améliorée et ainsi le nombre de jours avant où la prédiction est intéressante. Avec les modèles UNET nous avons montré que les résultats obtenus dans (Cariou et al., 2025) s'appliquent aussi dans IPSL-CM6A-LR et que nous pouvons utiliser un modèle entraîné dans le modèle de l'IPSL pour inférer des

⁵Cela a aussi comme avantage d'éviter de méditer trop longtemps en contemplant la lente agonie d'une barre de chargement.

résultats sur ERA5 (en utilisant directement un modèle entraîné dans IPSL-CM6A-LR ou mieux un modèle que nous avons ensuite re entraîné dans ERA5) même plusieurs jours à l'avance en ayant des résultats qualitatifs pour un modèle relativement simple dans une étude pilote. Ces derniers résultats avec les réseaux UNETs étaient cependant nuancés, car l'amélioration était minime, le « transfer learning » a donc un potentiel important mais il faut trouver comment bien l'utiliser en ML (il n'y a pas de méthode générale).

Il reste cependant énormément de choses qui pourraient être implémentées. Dans les premières semaines de mon stage j'ai par exemple implémenté un brouillon de générateur stochastique pour générer des trajectoires plus précises (on prend un analogue au hasard d'un jour donné pour aller au suivant et on répète ainsi de suite, voir (Miloshevich et al., 2024) pour plus de détails) et en plus grand nombre mais ce n'était pas particulièrement l'objectif ici. Cependant c'est quelque chose à regarder potentiellement dans un contexte de « transfer learning » ou de prévision basée sur les analogues à plus long terme. Comparer les analogues au ML sur des tâches similaires aurait aussi pu être intéressant. Un autre point à étudier est évidemment l'utilisation de modèles de deep learning plus avancés, par exemple rendre le modèle UNET récurrent via des cellules LSTM (Long Short Term Memory) et donner en entrée plusieurs jours de données en espérant mieux prévoir l'évolution future, ou mieux utiliser directement des modèles de l'état-de-l'art comme AIFS (Lang et al., 2024) seraient une vraie validation dans un contexte réel de cette étude. Nous pourrions aussi prévoir de la même manière qu'un générateur stochastique de manière récurrente en utilisant le jour précédent pour estimer le suivant, afin d'étudier si les trajectoires obtenues seraient réalistes et contiendraient assez de variabilité pour générer des extrêmes.

Ce travail pourra possiblement être utilisé par d'autres étudiants du même niveau (master) ou même chercheurs comme support ou bien à poursuivre. Il y a encore de nombreuses perspectives à explorer qui viennent d'être décrites. En particulier les tests des modèles AIFS dans ce contexte, un projet sera dédié à cet effet l'année prochaine au LSCE.

En termes de bilan personnel et d'apports du stage, ce stage a été une expérience formatrice et enrichissante pour une découverte du milieu des sciences du climat et en particulier des maths appliquées à ces dernières. J'ai eu l'opportunité de rencontrer de nombreuses personnes toutes aussi passionnées et intéressantes les unes que les autres. Ce stage m'a aussi permis de travailler dans un laboratoire et sur des sujets qui me parlent et auxquels je suis sensible. Il me paraît essentiel de nos jours de travailler dans un domaine où les thématiques nous intéressent plutôt que de faire de l'argent pour faire de l'argent. C'est une chance que nous avons d'avoir ce choix en tant qu'élèves ingénieurs en mathématiques appliquées et il faut la prendre.

J'ai aussi pu continuer à développer un certain état d'esprit analytique et critique en m'ouvrant à des données provenant d'un domaine que nous abordons peu dans notre formation. Cette expérience a non seulement enrichi mes compétences techniques mais aussi amélioré ma capacité à aborder les problèmes et la difficulté.

Dans le futur, après cette expérience, j'aimerais potentiellement rester dans ce domaine, que ce soit dans le public (recherche et/ou conseil) ou le privé en assurance dans la mesure où idéalement le projet reste respectueux socialement. Cette dualité entre recherche académique et industrie est une question que je me pose souvent et à laquelle je n'ai pas encore de réponse, peut-être par exemple qu'une thèse (aussi bien académique que CIFRE) liée à des thématiques comme celles des extrêmes et/ou du ML dans le contexte des sciences du climat serait une option intéressante en fin d'études.

CODE ET DONNÉES

Tout mon code et les données utilisées sont disponibles sur [github](#).

GLOSSAIRE

INSA Institut National des Sciences Appliquées

CEA Commissariat à l'Énergie Atomique et aux Énergies Renouvelables

CNRS Centre National de la Recherche Scientifique

LSCE Laboratoire des Sciences du Climat et de l'Environnement

IPSL Institut Pierre Simon Laplace

UVSQ Université Versailles Saint-Quentin-en-Yvelines

CMIP6 Coupled Model Intercomparison Project Phase 6

ERA5 Fifth generation ECMWF reanalysis of the global climate

TRACC Trajectoire de réchauffement de référence pour l'adaptation au changement climatique

GIEC Groupe d'experts intergouvernemental sur l'évolution du climat

IPCC Intergovernmental Panel on Climate Change (GIEC en anglais)

ANR Agence Nationale de la Recherche

ESPRI Équipe de Service pour la Recherche et l'Innovation (Centre de données et de calcul de l'IPSL)

CNES Centre National d'Études Spatiales

UMR Unité Mixte de Recherche

Git Système de contrôle de version

bash Bourne Again SHell (interpréteur de commandes)

ssh Secure Shell

MIROC Model for Interdisciplinary Research on Climate

MSE Mean Squared Error (erreur quadratique moyenne, sans racine)

JJA Juin, Juillet, Août

TVE Théorie des valeurs extrêmes

RMSE Root Mean Squared Error (erreur quadratique moyenne)

CCC Coefficient de Corrélacion de Concordance de Lin

U-NET Réseau convolutif pour segmentation et reconstruction (encodeur–bottleneck–décodeur avec skip connections)

CNN Convolutional Neural Network

LSTM Long Short-Term Memory (réseau récurrent)

GCM General Circulation Model (Modèle de circulation générale)

AIFS Artificial Intelligence Forecasting System (Système de prévision à base d'intelligence artificielle)

CIFRE Convention industrielle de formation par la recherche

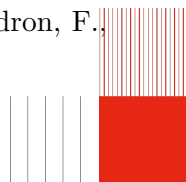
BIBLIOGRAPHIE

Une vague de chaleur exceptionnelle par sa durée en début d'été | Météo-France: <https://meteofrance.com/actualites-et-dossiers/actualites/une-vague-de-chaleur-exceptionnelle-par-sa-duree-en-debut-dete>.

Bell, R., Spring, A., Brady, R., Huang, A., Squire, D., Blackwood, Z., Sitter, M. C., et Chegini, T.: xskillscore: Metrics for verifying forecasts, , <https://doi.org/10.5281/zenodo.5173153>, 2021.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., et Tian, Q.: Accurate medium-range global weather forecasting with 3D neural networks, Nature, 619, 1-6, <https://doi.org/10.1038/s41586-023-06185-3>, 2023.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F.,



Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., Lavergne, C. de, Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, E., Lionel, Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., et Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002010, <https://doi.org/https://doi.org/10.1029/2019MS002010>, 2020.

Bouët, F. et Courtillot, P.: Data Assimilation Concepts and Methods, Meteorological Training Course Lecture Series, 2002.

Cariou, E., Cattiaux, J., Qasmi, S., Ribes, A., Cassou, C., et Doury, A.: Linking European Temperature Variations to Atmospheric Circulation With a Neural Network: A Pilot Study in a Climate Model, *Geophysical Research Letters*, 52, e2024GL113540, <https://doi.org/https://doi.org/10.1029/2024GL113540>, 2025.

Conover, W. J.: Practical nonparametric statistics, 3. ed., Wiley, New York, NY [u.a.], VIII,584 pp., 1999.

Copernicus Climate Change Service (C3S) and World Meteorological Organization (WMO): European State of the Climate 2024, <https://doi.org/10.24381/14j9-s541>, 2025.

Dask Development Team: Dask: Library for dynamic task scheduling,, 2016.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., et Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, 9, 1937-1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.

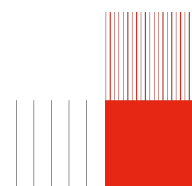
Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., et Oliphant, T. E.: Array programming with NumPy, *Nature*, 585, 357-362, <https://doi.org/10.1038/s41586-020-2649-2>, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. de, Rozum, I., Vamborg, F., Villaume, S., et Thépaut, J.-N.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999-2049, <https://doi.org/https://doi.org/10.1002/qj.3803>, 2020.

Hoyer, S. et Hamman, J.: xarray: N-D labeled arrays and datasets in Python, *Journal of Open Research Software*, 5, <https://doi.org/10.5334/jors.148>, 2017.

Hunter, J. D.: Matplotlib: A 2D graphics environment, *Computing in Science & Engineering*, 9, 90-95, <https://doi.org/10.1109/MCSE.2007.55>, 2007.

IPCC: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, édité par: Core Writing Team,



H. L. et Romero, J., IPCC, Geneva, Switzerland, <https://doi.org/10.59327/IPCC/AR6-9789291691647>, 2023b.

IPCC: Weather and Climate Extreme Events in a Changing Climate, in: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 1513-1766, 2023a.

Kingma, D. P. et Ba, J.: Adam: A Method for Stochastic Optimization, 2017.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., et Battaglia, P.: Learning skillful medium-range global weather forecasting, *Science*, 382, 1416-1421, <https://doi.org/10.1126/science.adi2336>, 2023.

Lang, S., Alexe, M., Chantry, M., Dramsch, J., Pinault, F., Raoult, B., Clare, M. C. A., Lessig, C., Maier-Gerber, M., Magnusson, L., Bouallègue, Z. B., Nemesio, A. P., Dueben, P. D., Brown, A., Pappenberger, F., et Rabier, F.: AIFS – ECMWF's data-driven forecasting system, 2024.

Lin, L. I.-K.: A Concordance Correlation Coefficient to Evaluate Reproducibility, *Biometrics*, 45, 255-268, 1989.

McKinney, W.: Data Structures for Statistical Computing in Python, in: Proceedings of the 9th Python in Science Conference, 56-61, <https://doi.org/10.25080/Majora-92bf1922-00a>, 2010.

Met Office: Cartopy: a cartographic python library with a Matplotlib interface,, 2010.

Miloshevich, G., Lucente, D., Yiou, P., et Bouchet, F.: Extreme heat wave sampling and prediction with analog Markov chain and comparisons with deep learning, *Environmental Data Science*, 3, e9, <https://doi.org/10.1017/eds.2024.7>, 2024.

Noyelle, R., Yiou, P., et Faranda, D.: Investigating the typicality of the dynamics leading to extreme temperatures in the IPSL-CM6A-LR model, *Climate Dynamics*, 62, 1329-1357, <https://doi.org/10.1007/s00382-023-06967-5>, 2023.

Ronneberger, O., Fischer, P., et Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, *CoRR*, 2015.

Russo, S., Sillmann, J., et Fischer, E.: Top ten European heatwaves since 1950 and their occurrence in the coming decades, *ENVIRONMENTAL RESEARCH LETTERS*, 10, 124003, <https://doi.org/10.1088/1748-9326/10/12/124003>, 2015.

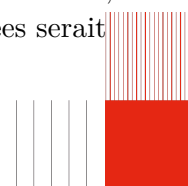
Schwertfeger, B.: The influence of bias corrections on variability, distribution, and correlation of temperatures in comparison to observed and modeled climate data in Europe, Doctoral dissertation, <https://doi.org/10.13140/RG.2.2.22363.23843>, 2022.

Zhang, Z., Fischer, E., Zscheischler, J., et Engelke, S.: Numerical models outperform AI weather forecasts of record-breaking extremes, 2025.

ANNEXES

-A Considérations numériques pour le calcul des analogues

Le calcul d'analogues de ERA5 ($N = 9975$) vers les 33 runs du modèle de l'IPSL et ERA5 ($N = 8645 \times 33 + 9975 = 295260$) demande de travailler avec des arrays de très grande taille (il y a aussi les deux dimensions spatiales), le calcul de tous les analogues représente près d'1.5 billion de différences à effectuer, il est donc primordial d'effectuer ces calculs de manière réfléchie, faire quatre boucles imbriquées serait



très long (plusieurs semaines). Il faut alors vectoriser pour prendre avantage de numpy, stocker les données sous forme de chunk pour éviter d'utiliser trop de RAM et paralléliser le calcul sur les différents cœurs disponibles (se fait avec dask automatiquement). Le moyen le plus rapide Figure 23 essayé est d'après avoir aplati les dimensions spatiales de décomposer la distance en:

$$\begin{aligned}
 d(\mathcal{X}(t_0), \mathcal{X}(t)) &= \sqrt{\sum_i \frac{\alpha_i}{\dim(\mathcal{X}^i)} \sum_r (\mathcal{X}^i(t_0, \mathbf{r}) - \mathcal{X}^i(t, \mathbf{r}))^2} \\
 &= \sqrt{\sum_i \frac{\alpha_i}{\dim(\mathcal{X}^i)} \left[\sum_r \mathcal{X}^i(t_0, \mathbf{r})^2 + \sum_r \mathcal{X}^i(t, \mathbf{r})^2 - 2 \underbrace{\sum_r \mathcal{X}^i(t_0, \mathbf{r}) \mathcal{X}^i(t, \mathbf{r})}_{\mathcal{X}^i \mathcal{X}^{iT}(t_0, t)} \right]} \\
 &= \sqrt{\sum_i \frac{\alpha_i}{\dim(\mathcal{X}^i)} \left[\sum_r \mathcal{X}^i(t_0, \mathbf{r})^2 + \sum_r \mathcal{X}^i(t, \mathbf{r})^2 - 2(\mathcal{X}^i \mathcal{X}^{iT})(t_0, t) \right]}
 \end{aligned}$$

Cette manière de calculer la distance permet de ne jamais calculer les différences paire à paire. Le calcul entier d'analogues dans ERA5 prend environ 3 secondes et 2min de ERA5 vers IPSL-CM6A-LR et ERA5.

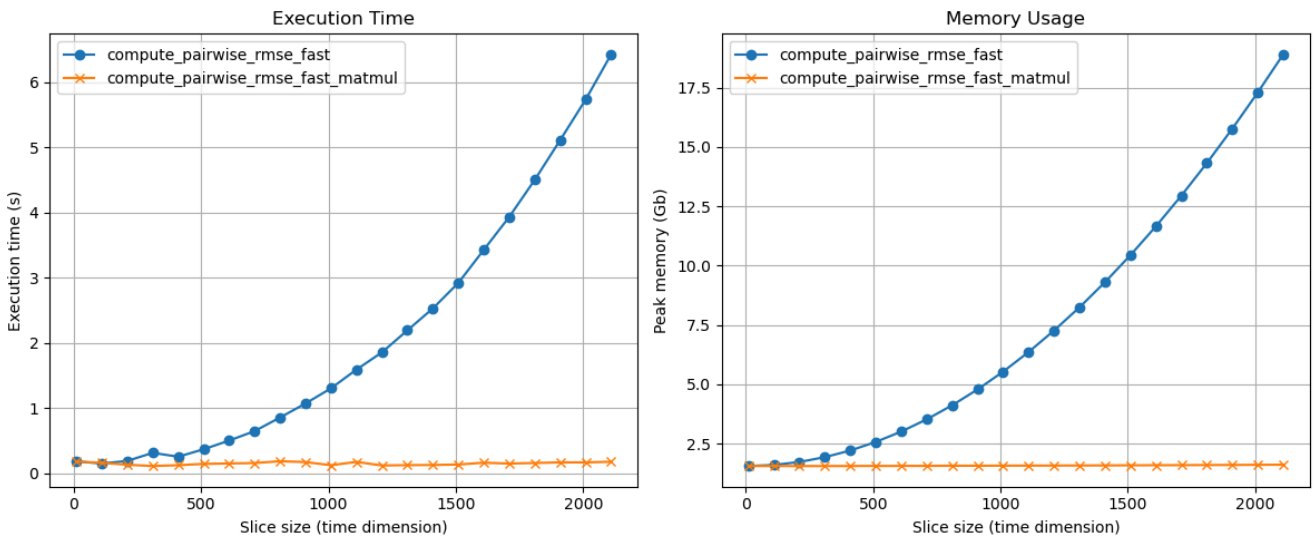
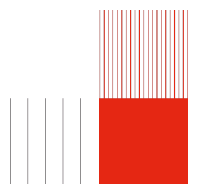


Figure 23 : Comparaison entre deux manières de calculer toutes les RMSE paire à paire entre tous les indices de temps. A gauche le temps d'exécution en fonction de la taille de la dimension temporelle et à droite l'utilisation de RAM en fonction de la taille de la dimension temporelle. En bleu une fonction déjà relativement rapide mais assez coûteuse en RAM car on vectorise tout via numpy pour calculer les différences paire à paire avant de sommer et passer à la racine (création d'un tenseur de dimension 4) alors qu'en orange on utilise l'astuce donné précédemment et la vectorisation pour le calcul des moyennes.

Bien que l'on dirait dans Figure 23 que la fonction orange ne croit pas, elle croit bien mais de manière moins agressive Figure 24, cela n'empêche donc pas de devoir utiliser dask pour chunk nos données et calculs.



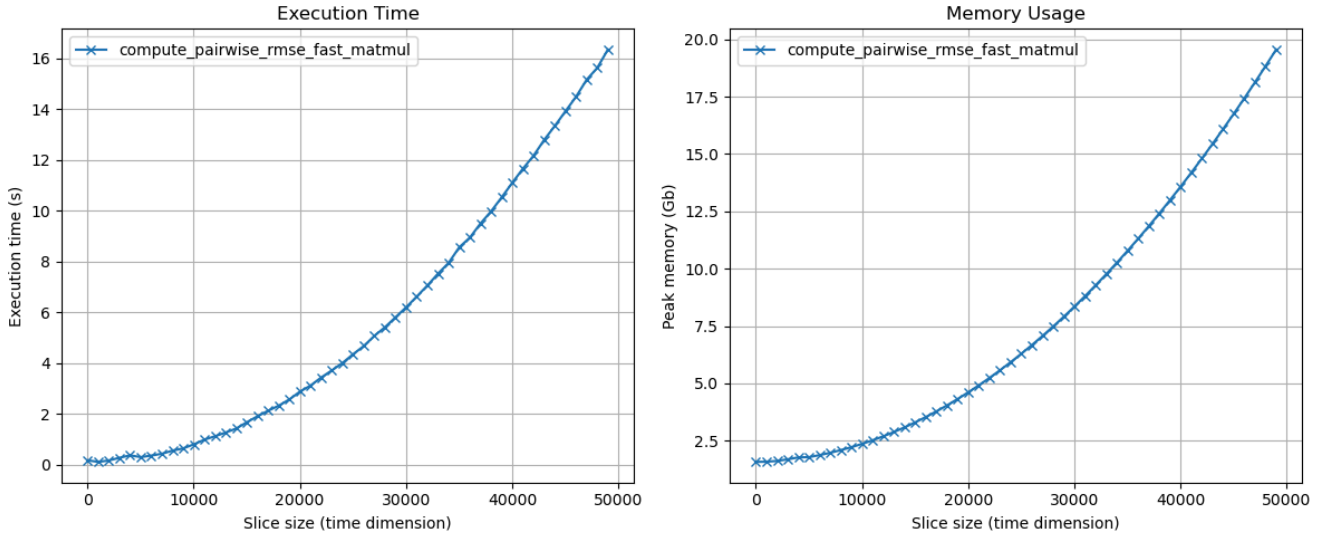


Figure 24 : Temps d'exécution et mémoire maximale utilisée pour la meilleure fonction dans Figure 23

-B Eléments théoriques supplémentaires

- Encodage de la dimension temporelle

Dans un souci de simplicité, nous ré-encodons la dimension temporelle de la forme suivante:

$$\text{time} = \text{run} \times 10^8 + \text{year} \times 10^4 + \text{month} \times 10^2 + \text{day}.$$

Ainsi, cela nous permet de concaténer tous les runs de IPSL-CM6A-LR sous une seule dimension. Pour accéder aux différentes composantes temporelles il suffit de jouer avec les divisions entières (`//`) et les modulus (`%`), ie

$$\begin{aligned} \text{run} &= \text{time} // 10^8 \\ \text{year} &= (\text{time} // 10^4) \% 10^4 \\ \text{month} &= (\text{time} // 10^2) \% 10^2 \\ \text{day} &= \text{time} \% 10^2 \end{aligned}$$

L'autre avantage de cet encodage est que nous encoderons les données d'ERA5 comme le run n°0.

- RMSE

La RMSE⁶ spatiale⁷ entre deux temps t_i et t_j pour un champ \mathcal{F}

$$\text{RMSE}(\mathcal{F}, t_i, t_j) = \sqrt{\frac{1}{\dim(\mathcal{F})} \sum_{\mathbf{r}} (\mathcal{F}(t_i, \mathbf{r}) - \mathcal{F}(t_j, \mathbf{r}))^2}$$

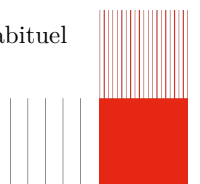
Où $\dim(\mathcal{F})$ correspond à la taille de grille sur laquelle nous considérons notre champ, $\dim(\mathcal{F}) = n_{\text{lat}} \times n_{\text{lon}}$.

- Corrélation de rang pour des champs en espace

La corrélation de Spearman (Conover, 1999) est une statistique non paramétrique définie comme la corrélation de Pearson (Conover, 1999) entre les variables de rang associées. On note ρ l'opérateur

⁶J'utiliserai aussi la MSE (pareil mais sans la racine)

⁷Il est aussi possible de calculer la RMSE temporelle entre deux séries temporelles et cette fois le calcul est habituel



associé à la corrélation de Pearson habituelle et $(X_i, Y_i)_{i=1, \dots, n}$ un échantillon de (X, Y) . La corrélation de Spearman est alors calculée par

$$r_s = \rho(\text{rg}(X), \text{rg}(Y)) = \frac{\text{cov}(\text{rg}(X), \text{rg}(Y))}{\sigma_{\text{rg}(X)} \sigma_{\text{rg}(Y)}}$$

ou les covariances et variances sont calculées à partir de l'échantillon. Elle est comprise entre -1 et 1 et est interprétée de la même manière qu'une corrélation linéaire.

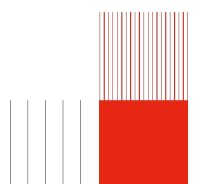
- Coefficient de Corrélation de Concordanance de Lin (CCC)

Le Coefficient de Corrélation de Concordanance de Lin (CCC), introduit par (Lin, 1989) quantifie l'accord entre deux observations X et Y . Il nous permet de quantifier aussi bien l'accord en terme de similarité qu'en terme de précision avec la différence des moyennes.

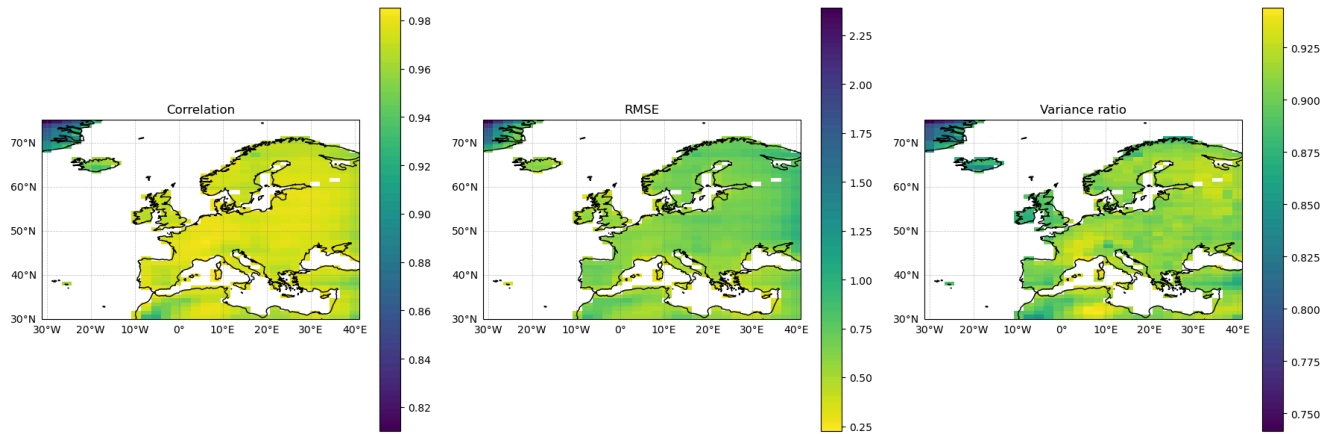
Il se calcule par

$$r_{\text{CCC}} = \frac{2\rho(X, Y)\sigma_X\sigma_Y}{(\bar{X} - \bar{Y})^2 + \sigma_X^2 + \sigma_Y^2} \in [-1, 1]$$

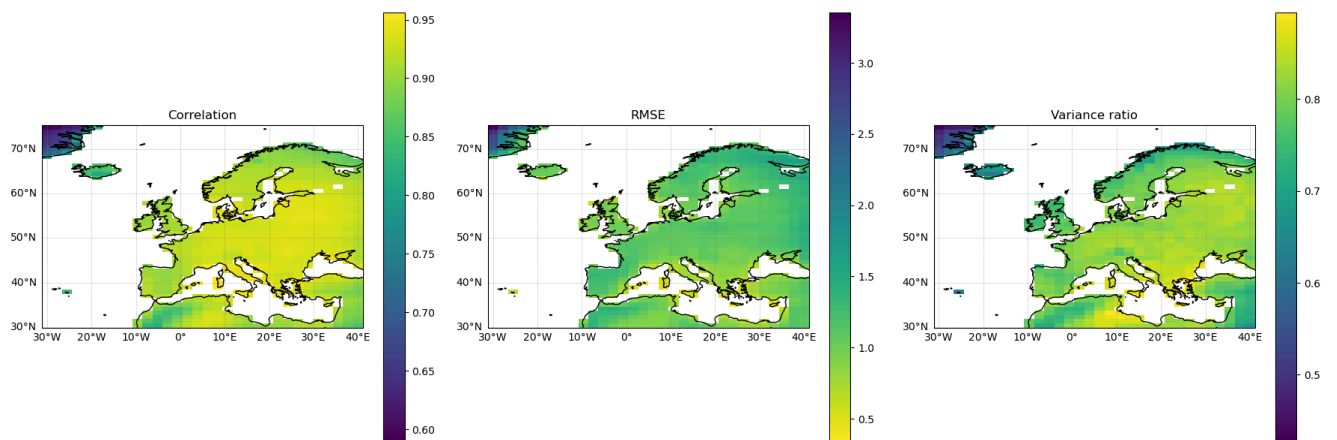
et s'interprète comme une corrélation linéaire.



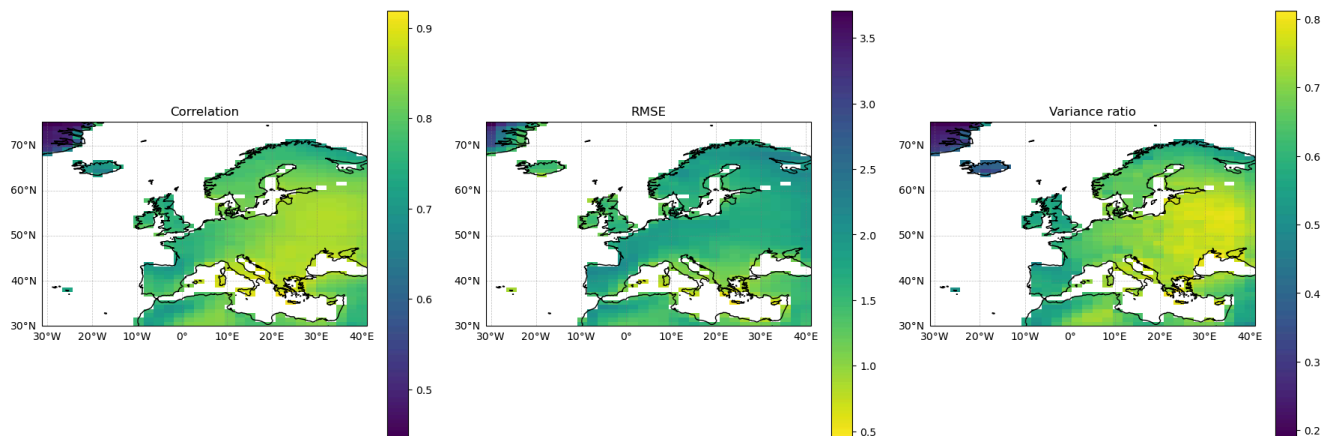
-C Résultats des UNET dans IPSL



(a) $\tau = 1$

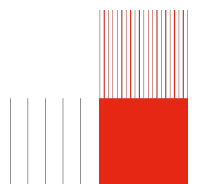


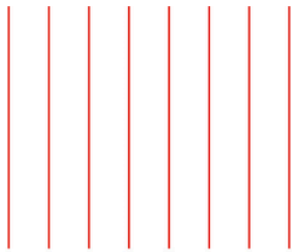
(b) $\tau = 2$



(c) $\tau = 3$

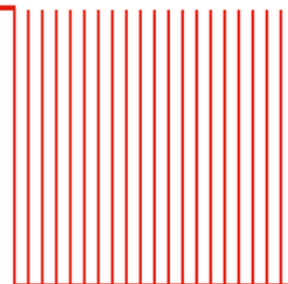
Fig. 25 : Corrélations(rang), RMSE et ratio des variances pour des reconstructions d'anomalies de températures journalières pour le set de validation avec lag 1,2 et 3. Les valeurs moyennes sur toute la zone considérée (pas uniquement la terre ferme) pour les corrélations sont respectivement de 0.97, 0.9 et 0.8. Pour les RMSE elles sont de 0.5, 0.87 et 1.24. Pour les ratios de variances elles sont de 0.89, 0.76 et 0.60. Le constat est que les résultats se détériorent avec τ , ce qui est normal.





A l'ère de l'IA et en particulier du deep learning, de nombreuses questions font apparition. Parmi elles la question de la disponibilité de données d'entraînements est cruciale. Les modèles de deep learning requièrent de grands ensembles de données, alors que les réanalyses climatiques restent limitées. Nous évaluons l'apport du «transfer learning» à partir de simulations à grande échelle pour améliorer la prédiction d'événements extrêmes car ces derniers sont ceux qui sont les plus impactants et avec lesquels les modèles IA peuvent avoir des difficultés. Deux approches sont considérées : la méthode des analogues pour estimer une fonction committor et des modèles UNET pour reconstruire des champs de température. Les résultats montrent que le transfer learning améliore la qualité des estimations et des reconstructions, bien que son efficacité dépende de la stratégie mise en œuvre et de la qualité des données provenant du modèle de simulation numérique.

In the era of AI, and particularly deep learning, many questions arise. Among them, the availability of training data is crucial. Deep learning models require large datasets, while climate reanalyses remain limited. We assess the contribution of transfer learning from large-scale simulations to improve the prediction of extreme events, since these are the most impactful and the most challenging for AI models. Two approaches are considered: the analogue method to estimate a committor function and UNET models to reconstruct temperature fields. The results show that transfer learning improves the quality of estimates and reconstructions, although its effectiveness depends on the strategy implemented and on the quality of the data from the numerical simulation model.



INSA Rennes

20 avenue des Buttes de Coësmes
CS 70839
35708 Rennes cedex 7
Tél : + 33 (0)2 23 23 82 00

www.insa-rennes.fr



INSA INSTITUT NATIONAL
DES SCIENCES
APPLIQUÉES
RENNES


MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR
ET DE LA RECHERCHE
*Liberté
Égalité
Fraternité*