



Titre: Apprentissage profond pour la réduction d'échelle des précipitations
Title: en changements climatiques

Auteur: Pierre-Louis Lemaire
Author:

Date: 2025

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Lemaire, P.-L. (2025). Apprentissage profond pour la réduction d'échelle des
Citation: précipitations en changements climatiques [Mémoire de maîtrise, Polytechnique
Montréal]. PolyPublie. <https://publications.polymtl.ca/68482/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/68482/>
PolyPublie URL:

**Directeurs de
recherche:** Julie Carreau, & Youssef Diouane
Advisors:

Programme: Maîtrise en mathématiques appliquées
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Apprentissage profond pour la réduction d'échelle des précipitations en
changements climatiques**

PIERRE-LOUIS LEMAIRE

Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Mathématiques

Août 2025

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Apprentissage profond pour la réduction d'échelle des précipitations en
changements climatiques**

présenté par **Pierre-Louis LEMAIRE**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Jonathan JALBERT, président

Julie CARREAU, membre et directrice de recherche

Youssef DIOUANE, membre et codirecteur de recherche

Richard LABIB, membre

DÉDICACE

À la différenciation automatique.

REMERCIEMENTS

Je tiens à remercier particulièrement mes co-superviseurs de recherche, Julie Carreau et Youssef Diouane, qui m'ont guidé et accompagné à travers un univers un peu flou au début ! Merci pour votre encadrement, votre disponibilité et vos conseils. J'ai énormément appris au cours de ces deux dernières années, et vous avez contribué à façonner une partie de qui je serai dans le futur.

Je remercie formellement Polytechnique Montréal et OURANOS pour leur confiance, et le soutien financier accordé à ce projet.

Ce mémoire n'aurait sûrement pas vu le jour sans le soutien et les nombreux encouragements de mes proches. Je tiens à profondément remercier ma famille et particulièrement mes parents, qui ont été présent tout au long de ce travail (malgré la distance), et qui se sont toujours trouvés à mes côtés pour m'écouter, pendant les moments de joie mais aussi de fatigue.

Je tiens également à faire un remerciement spécial à Manille, qui a été présente à côté de moi pendant la quasi-totalité de la rédaction de ce mémoire, et dont les ronronnements ont permis de grandement adoucir la charge de travail !

Enfin et surtout, merci mon cœur de croire en moi, merci pour ton soutien sans faille, et pour tes encouragements quotidiens.

RÉSUMÉ

L'accélération des changements climatiques entraîne une intensification des événements météorologiques extrêmes, avec des impacts majeurs sur les populations, les infrastructures et les écosystèmes. Les projections climatiques à haute résolution sont indispensables pour anticiper ces évolutions, mais les modèles climatiques globaux offrent une résolution spatiale insuffisante pour représenter les phénomènes de fine échelle, impliquant une sous-représentation des extrêmes, souvent très localisés. Les modèles régionaux, bien qu'offrant une meilleure résolution, restent coûteux et limitent la production d'ensembles de simulations à haute résolution.

L'apprentissage profond constitue une alternative prometteuse pour la réduction d'échelle des variables climatiques. Ce mémoire explore l'utilisation d'un réseau convolutif profond de type UNet pour effectuer de la réduction d'échelle de champs de précipitations journalières. Deux défis majeurs sont ciblés : (1) la sous-estimation des précipitations intenses et (2) la dégradation des performances en climat non stationnaire, par exemple dans un contexte de changements climatiques. Ces limites freinent l'application concrète de l'apprentissage profond pour cette tâche. Nous formulons l'hypothèse qu'elles seraient en partie dues au non-apprentissage de la dépendance des précipitations intenses à l'évolution des températures.

Deux contributions principales sont proposées. Premièrement, une contrainte douce inspirée de la relation de Clausius-Clapeyron est ajoutée à la fonction de perte pour inciter le modèle à produire des champs de précipitations dont l'évolution avec la température reste cohérente avec les données de référence. Deuxièmement, deux approches complémentaires intègrent au modèle des *a priori* issus de la théorie des valeurs extrêmes. La non-stationnarité liée à l'évolution des températures est ainsi encodée par les paramètres d'extensions de la loi de Pareto généralisée, améliorant la représentation des événements extrêmes.

Les méthodes sont évaluées sur des données issues d'un vaste ensemble de simulations climatiques régionales couvrant la période 1955–2099 et incluant notamment le sud du Québec. Les résultats mettent d'abord en évidence une dépendance avérée des extrêmes de précipitations à l'évolution des températures moyennes, tout en soulignant que les modèles d'apprentissage profond peinent à reproduire fidèlement cette relation. Ils montrent ensuite le potentiel des approches développées pour améliorer la prédiction des précipitations extrêmes et ouvrent des perspectives prometteuses pour de futurs travaux.

ABSTRACT

The acceleration of climate change is driving an intensification of extreme weather events, with major impacts on populations, infrastructure, and ecosystems. High-resolution climate projections are essential to anticipate these changes, but global climate models provide insufficient spatial resolution to capture fine-scale phenomena, leading to an underrepresentation of extremes, which are often highly localized. Regional climate models, while offering better resolution, remain costly and limit the production of large ensembles of high-resolution simulations.

Deep learning offers a promising alternative for the downscaling of climate variables. This thesis explores the use of a deep convolutional network (UNet) to perform downscaling of daily precipitation fields. Two main challenges are targeted: (1) the underestimation of high-intensity precipitation and (2) the degradation of performance under a non-stationary climate, as under climate change. These limitations hinder the practical application of such approaches. We hypothesize that they are partly due to the model’s failure to learn the dependence of extreme precipitation on temperature changes.

Two main contributions are proposed. First, a soft constraint inspired by the Clausius-Clapeyron relations is added to the loss function to encourage the model to produce precipitation fields whose evolution with temperature remains consistent with the reference data. Second, two complementary approaches integrate priors from extreme value theory into the model. The non-stationarity linked to temperature evolution is thus encoded through the parametrization of extensions of the generalized Pareto distribution, improving the representation of extreme events under future climate.

Proposed approaches are evaluated on data originating from a large ensemble of regional climate simulations covering years 1955 to 2099. The results first highlight a clear dependence of precipitation extremes on mean temperature changes, while underscoring that deep learning models struggle to faithfully reproduce this relationship. They then demonstrate the potential of the proposed approaches to improve the prediction of extreme precipitation and open promising avenues for future research.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vi
LISTE DES TABLEAUX	ix
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS	xv
LISTE DES ANNEXES	xvi
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 CADRE THÉORIQUE	4
2.1 Apprentissage profond	4
2.1.1 Réseaux convolutifs profonds	4
2.1.2 UNet	6
2.2 Théorie des valeurs extrêmes	7
2.2.1 Approche par maximum de blocs	8
2.2.2 Approche par dépassement de seuil	9
2.2.3 Extension de la loi de Pareto généralisée	10
2.3 Relation entre les précipitations et les températures	12
2.3.1 Clausius-Clapeyron	12
2.3.2 Estimation par méthode de regroupement	13
2.3.3 Estimation par la loi des valeurs extrêmes généralisée	14
CHAPITRE 3 REVUE DE LITTÉRATURE	16
3.1 Modélisation numérique du climat	16
3.2 Réduction d'échelle dynamique et statistique	17
3.3 Apprentissage profond pour la réduction d'échelle	19
3.3.1 Méthodes déterministes	19
3.3.2 Méthodes génératives	20
3.4 Apprentissage profond informé par la physique	21

CHAPITRE 4	DONNÉES ET ANALYSE EXPLORATOIRE	24
4.1	ClimEx : grand ensemble de simulations en réchauffements climatiques . . .	24
4.2	Analyse statistique exploratoire	25
CHAPITRE 5	APPRENTISSAGE PROFOND ET RÉDUCTION D'ÉCHELLE DES PRÉCIPITATIONS EN CHANGEMENTS CLIMATIQUES	29
5.1	UNet pour la réduction d'échelle	29
5.1.1	Présentation de l'architecture	29
5.1.2	Stratégie, <i>pré</i> et <i>post</i> traitement des données	32
5.2	Améliorations du UNet	36
5.3	Évaluation en changements climatiques	38
5.3.1	Estimation des TPSRs par la loi des valeurs extrêmes généralisée . .	40
CHAPITRE 6	INTÉGRATION DE LA RELATION TEMPÉRATURE-PRÉCIPITATIONS POUR LA RÉDUCTION D'ÉCHELLE	45
6.1	Contrainte douce de Clausius-Clapeyron par approche de regroupement . . .	45
6.1.1	Intégration de la relation précipitations-températures par contrainte douce	45
6.1.2	Limites de l'approche par regroupement	47
6.2	Intégration d' <i>a priori</i> statistiques issus de la théorie des valeurs extrêmes . .	48
6.2.1	Modélisation des précipitations par extension de la loi de Pareto géné- ralisée	48
6.2.2	Réduction d'échelle par prédiction des niveaux de quantile	54
6.2.3	Apprentissage par alignement des niveaux de quantile	57
CHAPITRE 7	ÉVALUATION	58
7.1	Métriques	58
7.2	Résultats	60
7.2.1	Contrainte douce par approche de regroupement	60
7.2.2	Intégration d' <i>a priori</i> statistiques issus de la théorie des valeurs extrêmes	64
CHAPITRE 8	CONCLUSION	69
8.1	Synthèse des travaux	69
8.2	Perspectives de recherche	70
RÉFÉRENCES	71
ANNEXES	83

LISTE DES TABLEAUX

Tableau 5.1	MAE (en mm/jour), proportions et la norme ℓ_2 des valeurs négatives, pour aucune transformation, T_1 ou T_2 . UNet entraîné sur la période 1960 – 2000, et évalué sur 2000 – 2010.	34
Tableau 5.2	UNet entraîné sur la période 1955 – 2010 et évalué sur la période 2010 – 2020, sur 50 époques. Comparaison des performance avec ou sans la contrainte de post-traitement (5.8). MAE générale (en mm/jour), et calculée au voisinage de certains quantiles (0.5, 0.90, 0.95, et 0.99). . .	35
Tableau 5.3	MAE pour différentes séquences de blocs par niveau de l’encodeur (décodeur symétrique avec un bloc supplémentaire par niveau). UNet entraîné sur 50 époques avec 64 échantillons par mini-lot.	36
Tableau 5.4	MAE pour différents types de convolutions et nombre de canaux initiaux. UNet entraîné sur 50 époques.	37
Tableau 5.5	MAE pour différentes tailles de filtres de convolution. UNet entraîné sur 50 époques.	37
Tableau 5.6	MAE avec certaines couches du bloc résiduel retirées : (1) seulement SiLU, et (2) SiLU et normalisation par groupe (GN). UNet entraîné sur 50 époques.	37
Tableau 5.7	Log-vraisemblance négative moyenne des lois de GEV du domaine sur la période 2020 – 2099 pour toutes les combinaisons de modèles des paramètres $\mu(T)$ et $\sigma(T)$	42
Tableau 5.8	Moyennes des TPSRs estimés à plusieurs niveaux de quantiles pour les données de référence, et les prédictions du UNet, pour la période de test 2020 – 2099.	43
Tableau 7.1	MAE globale et évaluée aux voisinages de plusieurs niveaux de quantiles pour l’interpolation bicubique sur la période de test.	60
Tableau 7.2	MAE globale et évaluée aux voisinages de plusieurs niveaux de quantiles, en fonction de la valeur du coefficient γ de la fonction de perte (6.1), pour $s = 0.0$ mm/jour (seuil de filtrage (6.4)).	61
Tableau 7.3	MAE globale et évaluée au voisinage de plusieurs niveaux de quantiles, en fonction de la valeur du coefficient γ de la fonction de perte (6.1), pour $s = 1.0$ mm/jour (seuil de filtrage (6.4)).	61

Tableau 7.4 MAE globale (en mm/jour) et évaluée au voisinage de plusieurs niveaux de quantiles, pour le UNet standard $\mathcal{L}_{\text{reco}}$ et les méthodes de prédictions et d'alignement sur les niveaux de quantile de la loi ExtGPD : $\mathcal{L}_{\text{ExtGPD}_1}$ (1) et $\mathcal{L}_{\text{ExtGPD}_2}$ (2), avec ou sans anomalies de températures T . 65

LISTE DES FIGURES

Figure 1.1	Illustration du cadre théorique de réduction d'échelle des précipitations par apprentissage profond.	3
Figure 2.1	Illustration du champ réceptif d'une cellule (couche 3 - orange), qui correspond à la portion de l'entrée à laquelle elle est sensible. À mesure que l'on progresse à travers les couches du réseau, le champ réceptif s'élargit (rouge à la couche 2 puis orange à la couche 3), permettant aux neurones de capter des caractéristiques de plus haut niveau. Dans cet exemple, les filtres sont de dimensions 3×3 , et le pas, correspondant au déplacement du filtre de convolution dans chaque direction, de 1. Le champ réceptif r_i de la couche i , pour un filtre de dimension k_i , et le pas de la couche précédente s_{i-1} est égal à : $r_i = r_{i-1} + (k_i - 1) \cdot s_{i-1}$	5
Figure 2.2	UNet composé de 3 niveaux, avec convolution transposée ou suréchantillonnage et convolution. L'entrée et la sortie sont de dimensions $64 \times 64 \times 3$, et l'espace latent $16 \times 16 \times 256$	7
Figure 2.3	Densités de probabilité pour différentes valeurs de ξ correspondant aux 3 familles incluent dans la GEV, avec $\mu = 0$, et $\sigma = 1$	9
Figure 2.4	Densités du modèle (1) pour des « petites » et « moyennes » valeurs (gauche) et des « grandes » valeurs (droite), en fonction de κ . La forme de la GPD est retrouvée lorsque $\kappa = 1$. La distribution Gamma avec comme paramètres (1.4, 1.4) est affichée en comparaison. Figure inspirée de [1].	12
Figure 4.1	Soient une période A allant de 1955 à 2030, et une période B chevauchant de cinq ans allant de 2025 à 2099. (a) Différence absolue, entre périodes B et A , de la médiane des températures moyennes journalières sur le domaine de ClimEx, et en pointillés noirs : le sous-domaine utilisé dans ce mémoire. (b) Médiane des précipitations journalières non-nulles sur la période A . (c) 95ème quantile des précipitations journalières non-nulles sur la période A . (d) Différence relative de la médiane des précipitations non-nulles. (e) Différence relative du 95ème quantile des précipitations non-nulles.	26

Figure 4.2	Afin de simuler des observations à basse résolution (LR), les précipitations journalières haute résolution (HR) issues de ClimEx sont agrégées en moyennant des blocs de 256 cellules (16×16). (À gauche) Les histogrammes des précipitations journalières, séparés selon les intensités ($0 - 5\text{mm/jour}$ en haut, et $\geq 5\text{mm/jour}$ en bas) comparant les périodes A et B , en haute et basse résolution. (À droite) Un champ de précipitations journalières sélectionné aléatoirement (en haut), et sa version agrégée (en bas).	27
Figure 5.1	Présentation de l'architecture du modèle UNet utilisé, et des modifications effectuées sur le bloc résiduel original [2].	31
Figure 5.2	Histogrammes des précipitations prédites par le UNet avec ou sans la contrainte (5.8) sur la période d'évaluation (2010 – 2020).	35
Figure 5.3	(Haut) MAE moyenne annuelle (en mm/jour) du UNet sur la période de test. (Bas) Évolution relative (en %) de la MAE annuelle, comparée à la période 2020–2025, calculée au voisinage des niveaux de quantile 0.5, 0.75, 0.90, 0.95 et 0.99. Courbes obtenues par moyenne mobile sur 10 valeurs.	39
Figure 5.4	Comparaison de l'évolution de l'erreur absolue moyenne (MAE) des prédictions de précipitations haute résolution du UNet entre les périodes 2020 – 2059 et 2060 – 2099, par rapport l'évolution des températures sur ces mêmes périodes.	40
Figure 5.5	Comparaison des TPSRs pour le niveau de quantile $q = 0.99$ estimés sur la période de test 2020 – 2099 à partir des données de référence, et des prédictions du UNet.	43
Figure 5.6	Boîtes à moustaches des TPSRs estimés à plusieurs quantiles pour les données de référence, et les prédictions du UNet, pour la période de test 2020 – 2099.	44
Figure 6.1	\mathcal{L}_{CC} pour la période 1955 – 2020 sur un domaine restreint en fonction de tailles croissantes de mini-lots (256, 512, 1024, 2048, et 4096). . . .	48
Figure 6.2	Log-vraisemblance négative obtenue après optimisation des paramètres définis en (6.22-6.24) de la loi ExtGPD, ajustée indépendamment sur chaque cellule du domaine à partir des données de référence couvrant la période 1955–2020. Les étoiles numérotées 1 à 4 identifient les cellules analysées aux figures 6.5 et 6.6.	52
Figure 6.3	Valeurs des paramètres $\exp(\alpha_0)$, α_1 , $\exp(\beta_0)$, β_1 pour chaque cellule du domaine après maximisation de la log-vraisemblance (6.17).	53

Figure 6.4	Valeurs du paramètre ξ après application de la transformation (6.24).	54
Figure 6.5	(Première ligne) Diagrammes quantile-quantile des données de la période 2020–2099 provenant des cellules 1 à 4 (voir figure 6.2), pour les modèles stationnaires ($\alpha_1 = \beta_1 = 0$) ajustés sur la période 1955–2020. (Deuxième ligne) Histogramme des quantiles obtenus.	55
Figure 6.6	(Première ligne) Diagrammes quantile-quantile des données de la période 2020 – 2099 provenant des cellules 1 à 4 (voir figure 6.2), pour les modèles non-stationnaires ajustés sur la période 1955 – 2020. (Deuxième ligne) Histogramme des quantiles obtenus.	55
Figure 7.1	MAE annuelle (en mm/jour) pour la perte MSE ($\gamma = 0$), et pour $s = 0$ ou $s = 1$ mm/jour avec $\gamma = 10$. L’erreur de l’interpolation cubique n’est pas affichée pour préserver la lisibilité de la figure. Lissage par moyenne mobile sur 10 valeurs.	62
Figure 7.2	Boîtes à moustaches des TPSRs (en $\%/^{\circ}C$) obtenus <i>via</i> la méthode définie à la section 6.1 pour chaque modèle. On sépare l’ensemble de test en deux périodes : 2020 – 2050, et 2050 – 2099, pour évaluer l’évolution temporelle.	62
Figure 7.3	Histogramme des précipitations pour chaque modèle, incluant l’ensemble des cellules du domaine sur la période de l’ensemble de test 2020 – 2099	63
Figure 7.4	PSD_{rel} pour chaque modèle par rapport à la densité spectrale de puissance des données de référence, calculées suivant la méthode détaillée à la section 7.1.	64
Figure 7.5	MAE (en mm/jour) annuelle pour le UNet standard (\mathcal{L}_{reco}) et les variantes stationnaire et non-stationnaire de la méthode de prédiction des niveaux de quantile (voir section 6.2.2) des lois ExtGPD ajustées précédemment : \mathcal{L}_{ExtGPD_1} et $\mathcal{L}_{ExtGPD_1}(T)$. La courbe est obtenue par moyenne mobile sur une fenêtre de 10 valeurs.	65
Figure 7.6	MAE (en mm/jour) annuelle pour le UNet standard (\mathcal{L}_{reco}) et les variantes stationnaire et non-stationnaire de la méthode d’alignement des niveaux de quantile (voir section 6.2.3) : \mathcal{L}_{ExtGPD_1} et $\mathcal{L}_{ExtGPD_1}(T)$. La courbe est obtenue par moyenne mobile sur une fenêtre de 10 valeurs.	66
Figure 7.7	Histogramme des précipitations pour chaque modèle, incluant l’ensemble des cellules du domaine sur la période de l’ensemble de test 2020 – 2099.	67

Figure 7.8	PSD _{rel} pour chaque modèle par rapport à la PSD des données de référence, calculées suivant la méthode détaillée à la section 7.1.	67
Figure 7.9	Champs de précipitations à basse résolution (interpolation 1nn), interpolés par méthode bicubique, prédits par chacun des modèles étudiés, et à haute résolution issus de ClimEx, pour deux dates sélectionnées aléatoirement sur la période de test.	68
Figure A.1	Log-vraisemblance négative à l'issue de l'optimisation des paramètres sous la combinaison $M_{\mu,4} + M_{\sigma,2}$ pour les données de référence (à gauche) et les prédictions du UNet (à droite).	83
Figure A.2	Paramètres μ_0 de $M_{\mu,4}$ pour les données de référence (à gauche) et les prédictions du UNet (à droite).	83
Figure A.3	Paramètres μ_1 de $M_{\mu,4}$ pour les données de référence (à gauche) et les prédictions du UNet (à droite).	84
Figure A.4	Paramètres σ_0 de $M_{\sigma,2}$ pour les données de référence (à gauche) et les prédictions du UNet (à droite).	84
Figure A.5	Paramètres σ_1 de $M_{\sigma,2}$ pour les données de référence (à gauche) et les prédictions du UNet (à droite).	85
Figure A.6	Paramètres ξ pour les données de référence (à gauche) et les prédictions du UNet (à droite).	85

LISTE DES SIGLES ET ABRÉVIATIONS

GCM	<i>Global Climate Model</i>
ESM	<i>Earth System Model</i>
RCM	<i>Regional Climate Model</i>
CMIP6	<i>Coupled Model Intercomparaison Project Phase 6</i>
CORDEX	<i>Coordinated Regional Climate Dowsaling Experiment</i>
GES	Gaz à Effet de Serre
TPSR	<i>Temperature-Precipitation Scaling Rates</i>
CNN	<i>Convolutional Neural Network</i>
GAN	<i>Generative Adversarial Network</i>
GEV	<i>Generalized Extreme Value distribution</i>
GPD	<i>Generalized Pareto Distribution</i>
ExtGPD	<i>Extended Generalized Pareto Distribution</i>

LISTE DES ANNEXES

Annexe A	Estimation des TPSRs par la loi des valeurs extrêmes généralisée . . .	83
----------	--	----

CHAPITRE 1 INTRODUCTION

L'accélération des changements climatiques renforce la vulnérabilité des populations, des écosystèmes et des infrastructures humaines. Elle s'accompagne de modifications profondes des régimes climatiques, caractérisées notamment par une intensification et une fréquence accrue des événements extrêmes tels que les vagues de chaleur, les inondations et les précipitations intenses. En particulier, la distribution des précipitations pourrait être significativement modifiée selon les régions, avec, dans certains cas, une diminution des précipitations annuelles moyennes, mais une augmentation de la fréquence et de l'intensité des événements extrêmes.

Pour anticiper au mieux ces évolutions, il est essentiel de produire des projections climatiques à fine échelle, afin de pouvoir modéliser les fortes disparités pour certaines variables atmosphériques, dans les zones « complexes » telles que les régions à forte orographie, les zones côtières ou, par exemple, celles soumises à des vents intenses. L'amélioration continue des modèles climatiques au cours des cinquante dernières années a permis de générer des projections globales avec une résolution spatiale d'environ $\approx 50 - 100$ kilomètres (km), une échelle toutefois insuffisante pour simuler certains processus de fine échelle, pourtant cruciaux pour la prévision des événements extrêmes. Les modèles climatiques régionaux constituent une solution partielle, offrant une résolution spatiale d'environ ≈ 10 km pour les modèles les plus performants, ce qui améliore significativement la représentation des précipitations, en particulier pour les fortes intensités. Cependant, le coût élevé en temps de calcul de ces modèles limite le nombre de simulations réalisables, surtout lorsqu'il s'agit de constituer de larges ensembles de projections, pourtant indispensables pour représenter correctement la variabilité naturelle du climat.

La quantité croissante d'observations atmosphériques (issues de stations terrestres, ballons, bouées, et satellites) et de simulations numériques, couplée aux récents progrès en apprentissage profond, ont permis d'ouvrir de nouveaux horizons en modélisation numérique du climat, avec l'essor des approches basées sur les données. Initialement limitées à des tâches relativement simples, les méthodes fondées sur les réseaux de neurones sont désormais capables d'émuler des modèles climatiques numériques à l'échelle globale, avec des performances comparables, tout en réduisant le coût de calcul de manière drastique. Par ailleurs, l'utilisation de l'apprentissage profond pour le raffinement de la résolution spatiale des simulations climatiques s'est fortement développée ces dernières années, offrant une alternative prometteuse aux approches traditionnelles de réduction d'échelle.

Dans ce mémoire, nous nous intéressons à l'augmentation de la résolution spatiale (appelée

aussi « réduction d'échelle ») des précipitations issues de simulations climatiques à l'aide de réseaux de neurones. Cette tâche, illustrée par la figure 1.1, consiste à estimer la fonction inconnue de réduction d'échelle h , liant les champs de précipitations basse résolution, à ceux haute résolution, à l'aide d'un réseau de neurones f_θ paramétré par l'ensemble de poids θ . Plus particulièrement, nous considérons les réseaux convolutifs profonds, qui exploitent l'opération de convolution discrète pour traiter efficacement des données multidimensionnelles (*e.g.*, images médicales ou champs de variables atmosphériques). Malgré leur potentiel, les réseaux de neurones présentent plusieurs limitations bien connues qui restreignent leur application à des contextes concrets. D'une part, en raison de phénomènes tels que l'apprentissage par raccourci et le surapprentissage, les modèles d'apprentissage profond peuvent encoder des relations artificielles ou physiquement incohérentes, qui ne se maintiennent pas hors de leur domaine et période d'entraînement, en particulier dans le contexte des changements climatiques. D'autre part, ces modèles peinent à représenter correctement les valeurs extrêmes, en raison de leur biais en faveur des signaux de basse fréquence. Dans le cadre de la réduction d'échelle des champs de précipitations, les conséquences de ces limites sont bien documentées et confirmées par nos expérimentations :

1. Une dégradation des performances au fur et à mesure que la période d'évaluation s'éloigne de celle d'entraînement, sous l'effet de l'intensification des changements climatiques ;
2. Des champs de précipitations à haute résolution présentant des structures floues et peu localisées, les modèles d'apprentissage profond ayant tendance à converger vers des solutions moyennes pour minimiser l'erreur, au détriment de la représentation des événements extrêmes.

L'objectif central de ce mémoire est d'atténuer ces deux principales limites. Nous faisons le choix de limiter la taille du modèle, et de fixer *a priori* une architecture, afin de concentrer la plus grande partie de nos efforts sur le développement de méthodes pouvant s'adapter à tout type de modèle d'apprentissage profond appliqué à la réduction d'échelle de champs de précipitations. Pour ce faire, nous introduisons des concepts issus des sciences atmosphériques, tels que la relation de Clausius-Clapeyron, qui quantifie l'évolution de la pression saturante d'un gaz (*e.g.*, vapeur d'eau dans l'atmosphère) par rapport à sa température, afin d'intégrer au modèle la dépendance des extrêmes de précipitations à l'évolution des températures en changements climatiques. Nous exploitons également des méthodes issues de la théorie des valeurs extrêmes pour intégrer des *a priori* statistiques sur la distribution des précipitations lors de l'entraînement des modèles, afin d'améliorer la représentation des extrêmes, tout en intégrant la non-stationnarité liée aux températures.

Ce mémoire est structuré comme suit : la section 2 présente le cadre théorique ainsi que les

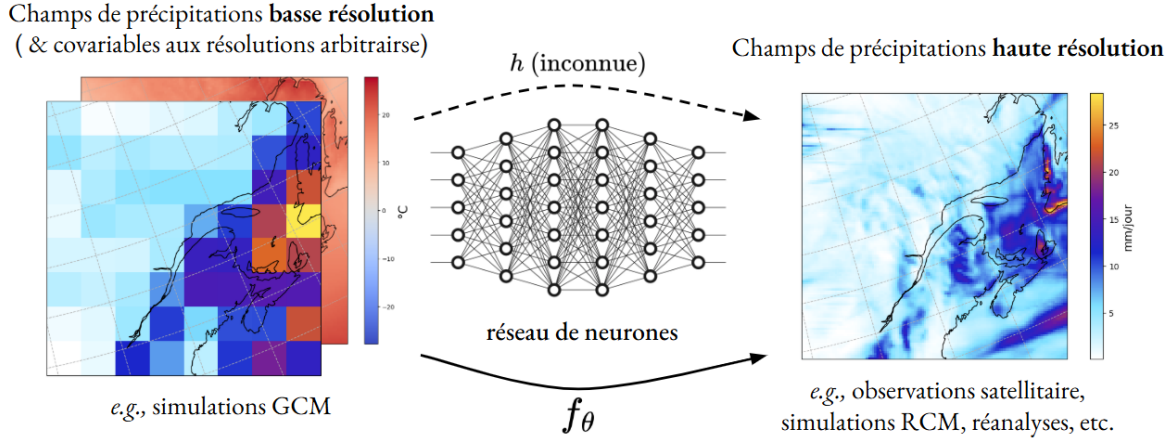


FIGURE 1.1 Illustration du cadre théorique de réduction d'échelle des précipitations par apprentissage profond.

méthodes mobilisées, suivie d'une revue de littérature en section 3. La section 4 décrit le jeu de données utilisé pour l'évaluation des méthodes développées et propose une analyse exploratoire afin de mettre en évidence certaines caractéristiques. L'architecture du modèle d'apprentissage profond, le traitement des données, et l'évaluation des performances du modèle en changements climatiques sont détaillés à la section 5. La section 6 développe les approches introduites pour mieux représenter les extrêmes et intégrer la dépendance aux températures. Les résultats des méthodes proposées sont présentés et analysés à la section 7. Enfin, la section 8 synthétise les principales contributions de ce mémoire, discute de leurs limites et propose de nouvelles perspectives de recherche.

CHAPITRE 2 CADRE THÉORIQUE

Ce chapitre a pour objectif d'introduire les bases théoriques des méthodes et approches utilisées à travers ce mémoire. Le corps méthodologique des travaux présentés appartenant à l'apprentissage profond, nous présentons la famille d'origine du modèle utilisé : les réseaux convolutifs profonds (CNN), ainsi que le modèle choisi pour notre projet. Un tour d'horizon de la théorie des valeurs extrêmes permettra de poser les bases théoriques sur lesquelles reposent certaines stratégies de réduction d'échelle développées dans ce mémoire. Finalement, nous introduirons la relation de Clausius-Clapeyron, ainsi que les différentes méthodes utilisées pour estimer les taux de dépendance entre les températures et les précipitations. Cela servira ensuite à évaluer la capacité des modèles d'apprentissage profond à apprendre des caractéristiques physiquement cohérentes afin d'améliorer leur robustesse aux changements introduits par les réchauffements climatiques.

2.1 Apprentissage profond

L'apprentissage profond est une branche de l'intelligence artificielle, consistant à entraîner des réseaux de plusieurs couches de paramètres, *via* l'algorithme de rétro-propagation, pour apprendre des représentations d'un jeu de données, et utiliser ces représentations pour résoudre des tâches, comme la reconnaissance d'image, ou la traduction de texte [3, 4].

2.1.1 Réseaux convolutifs profonds

Les réseaux convolutifs profonds (ou *Convolutional Neural Networks* en anglais, CNN) constituent une famille importante de modèles d'apprentissage profond conçus pour traiter majoritairement des données spatiales. Leur structure consiste en une succession de couches de convolutions, chacune composée d'un ou plusieurs filtres (ou noyaux) permettant d'extraire des motifs spatiaux caractéristiques des données, grâce à l'opération de convolution discrète.

Considérons un ensemble $C \in \mathbb{N}_+^*$ de champs bidimensionnels : $x \in \mathbb{R}^{C \times H \times W}$, de dimensions spatiales $H \in \mathbb{N}_+^*$ et $W \in \mathbb{N}_+^*$, et un filtre $k \in \mathbb{R}^{h \times w}$ de hauteur $h \in \mathbb{N}_+^*$ et de largeur $w \in \mathbb{N}_+^*$ avec $h \ll H$ et $w \ll W$. La valeur de la cellule (i, j) du champ bidimensionnel y , résultant de l'application de l'opération de convolution discrète aux champs d'entrée x , peut être exprimé en suivant la formule :

$$y_{i,j} = \sum_{c=1}^C \sum_{u=0}^h \sum_{v=0}^w k_{i+u,j+v} x_{c,i+u,j+v} \quad (2.1)$$

Les dimensions des filtres étant généralement bien inférieures à celles des données ($h \ll H$, et $w \ll W$), le nombre de paramètres par couche restant ainsi faible comparé à une couche dense traditionnelle, dans laquelle chaque variable d'entrée est reliée à chaque variable de sortie par un paramètre distinct), cela permet de construire des modèles profonds, avec de nombreuses couches et plusieurs filtres par couche. Le modèle CNN, consistant en une succession de couches de convolutions séparées par des opérateurs d'agrégation et de non-linéarité, détient la capacité d'apprendre des structures complexes représentant de son jeu de données d'entraînement, en exploitant à la fois les connexions locales et globales au sein des données, comme illustré à la figure 2.1.

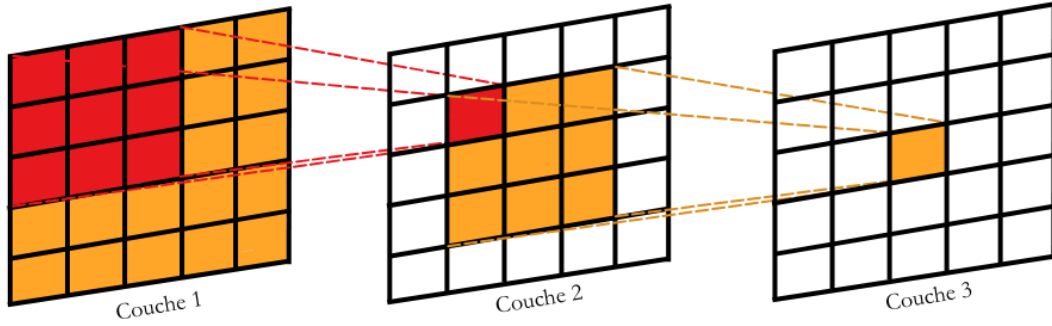


FIGURE 2.1 Illustration du champ réceptif d'une cellule (couche 3 - orange), qui correspond à la portion de l'entrée à laquelle elle est sensible. À mesure que l'on progresse à travers les couches du réseau, le champ réceptif s'élargit (rouge à la couche 2 puis orange à la couche 3), permettant aux neurones de capter des caractéristiques de plus haut niveau. Dans cet exemple, les filtres sont de dimensions 3×3 , et le pas, correspondant au déplacement du filtre de convolution dans chaque direction, de 1. Le champ réceptif r_i de la couche i , pour un filtre de dimension k_i , et le pas de la couche précédente s_{i-1} est égal à : $r_i = r_{i-1} + (k_i - 1) \cdot s_{i-1}$.

La non-linéarité, inhérente aux données complexes, est prise en compte par l'ajout de fonctions d'activation entre les couches de convolution, comme $\text{ReLU}(z) = \max(0, z)$ ou $\text{SiLU}(z) = z \cdot \text{sig}(z)$, où $\text{sig}(x) = 1/(1 + e^{-x})$ est la fonction sigmoïde. D'autres opérateurs, notamment d'agrégation par maximum, permettent de modéliser la non-linéarité des données.

Les paramètres des couches de convolution sont estimés par descente de gradient stochastique, *via* la rétro-propagation. Soit f_θ le CNN paramétré par les poids $\theta \in \mathbb{R}^d$, regroupés dans un vecteur de dimension d , destiné à modéliser $h : x \rightarrow y$, où x et y sont des tenseurs à valeurs réelles, de dimensions arbitraires. L'apprentissage des paramètres s'effectue par la minimisation d'une fonction de coût \mathcal{L} selon la règle :

$$\theta_{i+1} = \theta_i - \eta \cdot \nabla_{\theta_i} \mathcal{L}(f_{\theta_i}(x^{(ni:n(i+1))}), y^{(ni:n(i+1))}) \quad (2.2)$$

avec η le pas d'apprentissage, et $x^{(n:i:n(i+1))} \subset x$, $y^{(n:i:n(i+1))} \subset y$ les i -ème mini-lots de données, de n échantillons. Les paramètres η et n n'étant pas appris, on les fixe manuellement ou *via* des algorithmes d'optimisation d'hyperparamètres. La fonction de coût joue un rôle central en apprentissage profond, car elle définit l'espace dans lequel les paramètres θ sont optimisés, influençant ainsi fortement les performances et le comportement du modèle après l'apprentissage. L'impact de la fonction de coût pour la réduction d'échelle est exploré plus en détails à la section 3.3.1.

Les CNNs ont d'abord été développés et utilisés pour de la reconnaissance d'images, en utilisant les opérateurs d'agrégation pour réduire les dimensions spatiales tout en conservant les caractéristiques dominantes des données [5]. Ils ont rapidement été adaptés à des tâches dites *image vers image*, notamment en imagerie médicale, nécessitant une sortie pour chaque pixel de l'entrée. C'est dans ce contexte qu'a été introduite l'architecture UNet [6].

2.1.2 UNet

Initialement conçue pour de la segmentation d'images médicales, cette architecture s'est imposée à travers de multiples applications, et comme un modèle convolutif profond de référence. Le UNet est composé de deux chemins symétriques, contractif et expansif, fonctionnant tel un encodeur/décodeur :

- **L'encodeur** : agit comme un extracteur de caractéristiques $E : x \rightarrow z$. Contenant L niveaux, E transforme une entrée $x \in \mathbb{R}^{C \times H \times W}$ de grandes dimensions spatiales, en une représentation latente $z \in \mathbb{R}^{D \times H/2^L \times W/2^L}$ contractée, mais avec de nombreuses caractéristiques, aussi appelées canaux ($D \in \mathbb{R}_+^* \gg C$). Chaque niveau est composé d'au moins une couche de convolution, suivie d'une fonction d'activation, et d'une opération d'agrégation (par moyenne ou maximum) permettant de réduire la résolution spatiale.
- **Le décodeur** : permet de recomposer les détails de la sortie $y \in \mathbb{R}^{C \times H \times W}$ à partir des caractéristiques intermédiaires $E_l(x) \in \mathbb{R}^{d_l \times H/2^l \times W/2^l}$ issues du niveau $l \in [1, L]$ de l'encodeur et comprenant $C < d_l \in \mathbb{R}_+^* < D$ canaux, et de la sortie de l'encodeur : $D : (z, E_{1:L-1}(x)) \rightarrow y$. Chaque niveau utilise une couche de convolution transposée (ou un opérateur de sur-échantillonnage, suivi d'une couche de convolution) pour augmenter la résolution spatiale, en intégrant les caractéristiques apprises à chaque niveau de l'encodeur par des connexions de saut. Comme pour l'encodeur, une fonction d'activation est appliquée après chaque couche de convolution.

Les connexions de saut (*skip connecions* en anglais) sont cruciales pour le UNet, permettant de conserver l'information spatiale fine qui serait sinon perdue par la réduction de résolution

spatiale lors de la contraction vers l'espace latent. La figure 2.2 représente une architecture UNet standard avec 3 niveaux.

De nombreuses améliorations à l'architecture initiale ont été proposées, comme l'intégration de blocs résiduels, l'ajout de connexions denses entre niveaux, l'utilisation des mécanismes d'attention pour pondérer les régions de l'image, ou des combinaisons de ces améliorations [7–9]. Dans ce mémoire, nous utiliserons un modèle UNet résiduel issu de la littérature, introduit plus en détails dans la section 5.1.1.

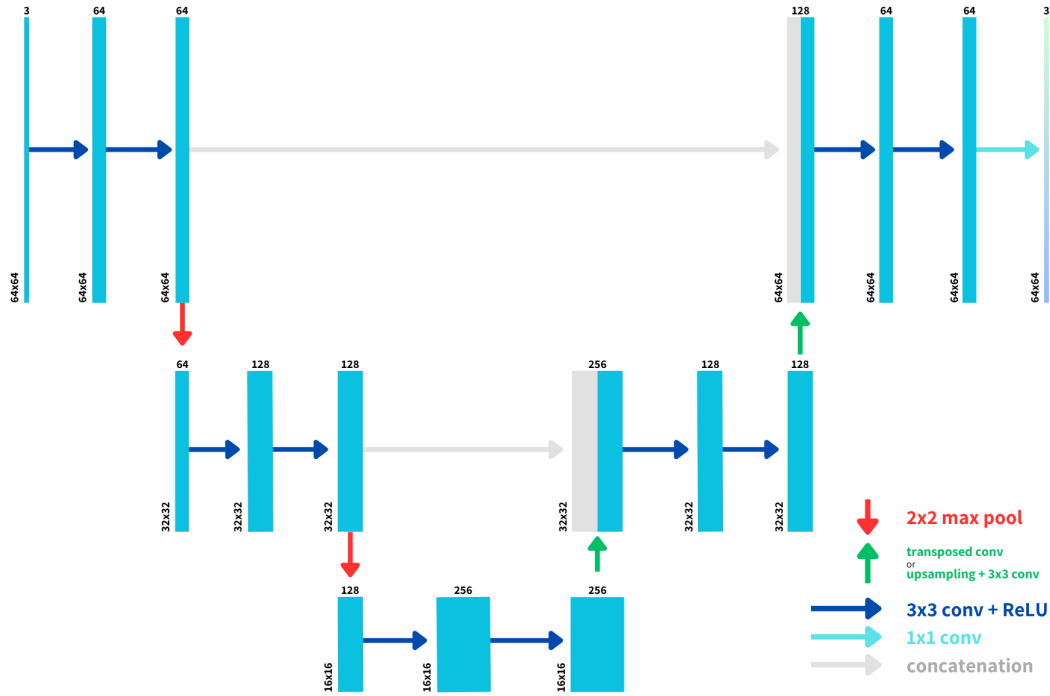


FIGURE 2.2 UNet composé de 3 niveaux, avec convolution transposée ou suréchantillonnage et convolution. L'entrée et la sortie sont de dimensions $64 \times 64 \times 3$, et l'espace latent $16 \times 16 \times 256$.

2.2 Théorie des valeurs extrêmes

La théorie des valeurs extrêmes est une branche des probabilités et des statistiques qui s'intéresse spécifiquement aux événements rares et aux queues des distributions. Cette théorie trouve de nombreuses applications dans des domaines variés tels que la météorologie ou la finance [10].

2.2.1 Approche par maximum de blocs

Soit une séquence de variables aléatoires indépendantes X_1, \dots, X_n (correspondant au « bloc »), suivant une même fonction de répartition F , la loi des valeurs extrêmes généralisée s'intéresse au comportement de la variable :

$$M_n = \max\{X_1, \dots, X_n\}. \quad (2.3)$$

Théoriquement, il est possible d'obtenir la densité de la variable aléatoire M_n , en utilisant l'indépendance des variables $(X_i)_{[1,n]}$:

$$\mathbb{P}(M_n \leq z) = \mathbb{P}(X_1 \leq z, \dots, X_n \leq z) = \prod_{i=1}^n \mathbb{P}(X_i \leq z) = \{F(z)\}^n. \quad (2.4)$$

Or, en pratique, F est inconnue et d'infimes erreurs d'estimation (inévitables) peuvent mener à de larges erreurs lorsque n est grand. Une autre approche consiste à approximer directement la distribution F^n lorsque $n \rightarrow \infty$.

Une propriété importante à noter, est que pour tout $z < z_+ = \inf\{z \in \mathbb{R} \mid F(z) = 1\}$, on a $\{F(z)\}^n \rightarrow 0$ lorsque $n \rightarrow \infty$. Autrement dit, la densité de probabilité de la variable M_n se concentre asymptotiquement au voisinage supérieur de la borne z_+ . Il est donc nécessaire d'effectuer une renormalisation de M_n pour stabiliser son emplacement et sa variance lorsque n augmente. Le théorème 1 donne, sous l'hypothèse de la renormalisation, une famille de lois permettant de modéliser F^n pour n assez grand.

Théorème 1 (Théorème de Fisher-Tippett-Gnedenko) *Soient X_1, \dots, X_n une séquence de variables aléatoires indépendantes et identiquement distribuées, et $M_n = \max\{X_1, \dots, X_n\}$. Pour deux suites appropriées $\{a_n > 0\}$ et $\{b_n\}$ telles que :*

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z) \quad \text{lorsque } n \rightarrow \infty \quad (2.5)$$

pour une fonction de distribution G non-dégénérée, alors G est de la forme :

$$G(z) = \begin{cases} \exp\{-[1 + \xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\} & \text{si } \xi \neq 0 \\ \exp\{-\exp(\frac{z-\mu}{\sigma})\} & \text{si } \xi = 0 \end{cases} \quad (2.6)$$

définit pour $\{z \in \mathbb{R} \mid 1 + \xi(z - \mu)/\sigma > 0\}$, avec $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^$, et $\xi \in \mathbb{R}$.*

Sous cette forme, G regroupe trois familles de lois de probabilités : Gumbell ($\xi = 0$), Fréchet

($\xi > 0$), et Weibull ($\xi < 0$), offrant donc une flexibilité de modélisation accrue. On nomme cette famille par son acronyme anglais : GEV (*Generalized Extreme Value*).

Généralement, on appelle respectivement les paramètres μ , σ et ξ les paramètres de localisation, d'échelle et de forme. La figure 2.3 montre le rôle du paramètre de forme ξ , contrôlant l'allure et le poids de la queue de G .

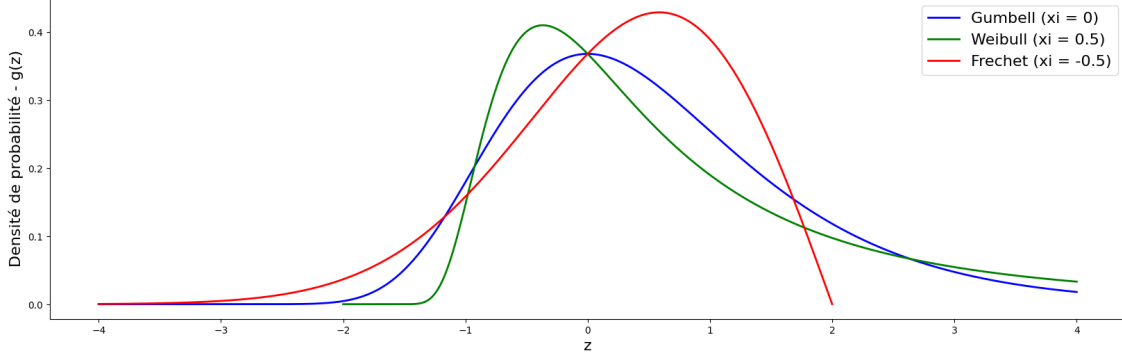


FIGURE 2.3 Densités de probabilité pour différentes valeurs de ξ correspondant aux 3 familles incluent dans la GEV, avec $\mu = 0$, et $\sigma = 1$.

En pratique, les constantes de normalisation $\{a_n > 0\}$ et $\{b_n\}$, inconnues, ne posent pas de difficultés car elles peuvent être absorbées par G . Pour n suffisamment grand, et sous l'hypothèse que :

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq z\right) \approx G(z) \quad \text{alors} \quad \mathbb{P}(M_n \leq z) \approx G\left(\frac{z - b_n}{a_n}\right) = G^*(z) \quad (2.7)$$

avec G^* appartenant également à la famille GEV. Il n'est donc pas nécessaire de considérer $\{a_n > 0\}$ et $\{b_n\}$ en pratique.

Finalement, une considération importante est à noter par rapport aux valeurs de ξ . Les propriétés asymptotiques nécessaires pour l'estimation par maximum de vraisemblance ne sont pas satisfaites pour $\xi \leq -0.5$, notamment car l'information de Fisher n'est plus définie ou tend vers l'infini [11].

2.2.2 Approche par dépassement de seuil

La modélisation des valeurs extrêmes par la GEV a l'inconvénient de restreindre fortement le nombre de points de données utilisés, alors que d'autres valeurs extrêmes peuvent être disponibles, sans garantir que les valeurs sélectionnées soient elles mêmes extrêmes. Contrairement à la modélisation des maxima par blocs, la modélisation par dépassement de seuil permet de

potentiellement tirer profit de plus grandes séries de données.

Théorème 2 (Théorème de Pickands-Balkema-De Haan) *Soit X_1, \dots, X_n une suite de variables aléatoires indépendantes et identiquement distribuées selon F . Sous les mêmes conditions que théorème 1, c'est-à-dire $\lim_{n \rightarrow \infty} F^n(z) = G(z; \mu, \sigma, \xi)$, alors pour un seuil u suffisamment grand, la fonction de distribution de la variable $Y = X - u$, conditionnellement à $X > u$, est donnée par :*

$$\lim_{u \rightarrow z+} \mathbb{P}(Y = X - u \leq y \mid X > u) = H(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\tilde{\sigma}}\right) & \text{si } \xi = 0 \end{cases} \quad (2.8)$$

défini pour $\{y \in \mathbb{R} \mid 1 + \xi y/\tilde{\sigma} > 0\}$ et $\tilde{\sigma} = \sigma + \xi(u - \mu) > 0$, avec $z+$ la limite supérieure du domaine de X .

La famille des distributions définie par l'équation (2.8) se nomme la famille de Pareto généralisée (*Generalized Pareto Distribution* en anglais, GPD).

2.2.3 Extension de la loi de Pareto généralisée

L'extension de la loi de Pareto généralisée vise à définir une distribution capable de modéliser l'ensemble de la gamme des précipitations non nulles (faibles, modérées et extrêmes) tout en préservant les propriétés asymptotiques de la théorie des valeurs extrêmes dans la queue de la distribution [1].

Une propriété intéressante de la théorie des valeurs extrêmes est sa capacité à également modéliser les valeurs de faibles intensités. Prenons comme exemple les séries de précipitations, bornées inférieurement par 0, et inversons leur signe pour construire la variable $Y = -X$. Les plus grandes valeurs de Y peuvent alors être modélisées par une distribution de Pareto généralisée, avec un paramètre de forme ξ négatif, correspondant au cas de Weibull pour la loi des valeurs extrêmes généralisée (voir figure 2.3).

Fixons un seuil v définissant la frontière avec les faibles valeurs de X . On note $\xi = -1/\kappa$ avec $\kappa > 0$. Suivant l'hypothèse précédente, on peut trouver la forme de la distribution des faibles valeurs de X [1] :

$$\mathbb{P}(X \leq x \mid X \leq v) = \mathbb{P}(Y > -x \mid Y > -v) \approx \bar{H}_{-1/\kappa}\left(\frac{-x+v}{\sigma}\right) = \left(1 - \frac{1}{\kappa} \frac{-x+v}{\sigma}\right)^\kappa,$$

la condition $\bar{H}_{-1/\kappa}(0) = 0$, découlant de la contrainte imposée par la borne supérieure de Y sur la fonction de survie $\bar{H}_{-1/\kappa}$, implique aussi que $v = \kappa\sigma$, car $\bar{H}_{-1/\kappa}(0) = (1 - \frac{v}{\kappa\sigma})^\kappa =$

$0 \implies v = \kappa\sigma$. Par conséquent, les faibles valeurs de X peuvent être approximativement décrites par une loi de puissance, et il est possible de modéliser l'ensemble de la gamme de précipitations, à l'exclusion de 0, par une loi de la forme :

$$\mathbb{P}(X \leq x) \approx \begin{cases} 1 - \text{cst} \times \bar{H}_\xi\left(\frac{x}{\sigma}\right) & \text{pour un « grand » } x, \\ 1 - \text{cst} \times x^\kappa & \text{pour un « petit » } x. \end{cases}$$

La distinction entre les « petites » et les « grandes » valeurs implique le choix d'un seuil, ce qui limite considérablement la flexibilité de la méthode. Or, celle-ci vise à modéliser l'ensemble de la distribution des précipitations de manière simple et efficace du point de vue numérique. Afin d'introduire le cœur de leur approche, les auteurs utilisent comme base une méthode permettant d'étendre les capacités de modélisation de la loi de Pareto généralisée [12]. Sachant qu'il est possible d'échantillonner une GPD par $\sigma H_\xi^{-1}(u)$ avec $u \sim \mathcal{U}(0, 1)$, on peut augmenter la flexibilité de l'échantillonnage en transformant u par l'inverse d'une fonction K partageant les propriétés d'une fonction de répartition définie sur $[0, 1]$. Ainsi, on a :

$$X = \sigma H_\xi^{-1}\{K^{-1}(u)\}. \quad (2.9)$$

Afin de garantir que la queue de la distribution de X adopte bien un comportement conforme à une GPD avec ξ comme paramètre de forme, et que les valeurs proches de zéro suivent une loi de puissance, les auteurs imposent trois contraintes que doit respecter la fonction K . Quatre familles de fonctions paramétriques satisfaisant ces contraintes ont été proposées :

1. $K(v) = v^\kappa$ avec $\kappa > 0$;
2. $K(v) = pv^{\kappa_1} + (1-p)v^{\kappa_2}$ avec $\kappa_2 \geq \kappa_1 > 0$ et $p \in [0, 1]$;
3. $K(v) = 1 - Q_\delta\{(1-v)^\delta\}$ avec $\delta > 0$ et Q_δ la fonction de répartition d'une variable aléatoire $\sim \beta(1/\delta, 2)$;
4. $K(v) = [1 - Q_\delta\{(1-v)^\delta\}]^{\kappa/2}$, avec $\kappa, \delta > 0$ et Q_δ défini précédemment.

Par soucis de simplicité, et au vu de sa performance jugée satisfaisante pour décrire les précipitations par plusieurs travaux, nous nous limiterons à l'utilisation du modèle (1) dans ce mémoire [1, 13, 14]. La figure 2.4 illustre son comportement selon plusieurs valeurs de κ , et propose une comparaison avec la loi Gamma, particulièrement pertinente pour les grandes valeurs.

Il est important de noter que des méthodes plus complexes, étendant les capacités de l'extension de la loi de Pareto généralisée pour les précipitations, ont été proposées, mais nous laissons leur utilisation à des travaux futurs [15].

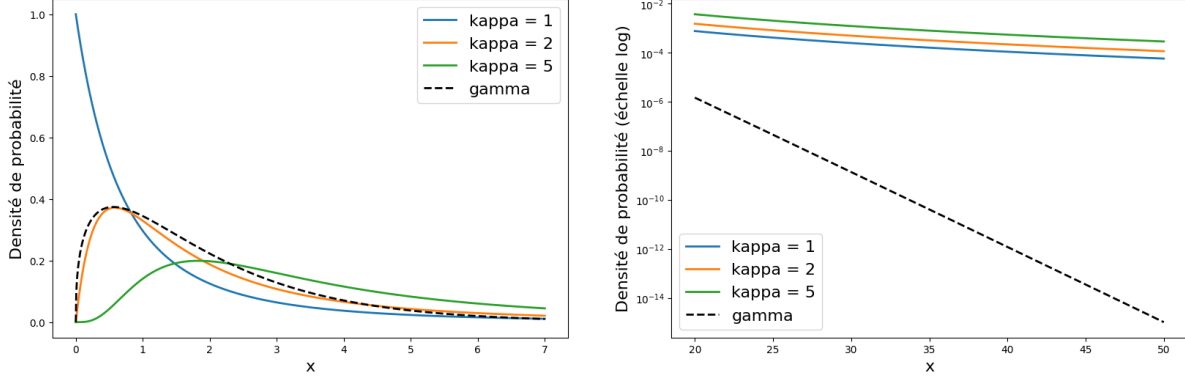


FIGURE 2.4 Densités du modèle (1) pour des « petites » et « moyennes » valeurs (gauche) et des « grandes » valeurs (droite), en fonction de κ . La forme de la GPD est retrouvée lorsque $\kappa = 1$. La distribution Gamma avec comme paramètres (1.4, 1.4) est affichée en comparaison. Figure inspirée de [1].

2.3 Relation entre les précipitations et les températures

2.3.1 Clausius-Clapeyron

La relation de Clausius-Clapeyron décrit le lien entre la pression de vapeur saturante d'un corps et sa température. Introduite entre les années 1824 et 1850 par plusieurs travaux, cette relation fondamentale de la thermodynamique joue un rôle important en sciences atmosphériques, notamment dans l'étude des précipitations, car elle quantifie la capacité de l'air à contenir de la vapeur d'eau selon sa température. Dans le cas d'un changement de phase entre liquide et vapeur à saturation, la relation de Clausius-Clapeyron s'écrit :

$$\frac{\partial e_s}{\partial T} = \frac{L_v e_s}{R_v T^2} \quad (2.10)$$

avec e_s la pression de vapeur saturante (en hPa), T la température (en Kelvin), L_v la chaleur latente de vaporisation (en $J \cdot K^{-1}$), et R_v la constante des gaz parfait spécifique au corps à l'étude (en $J \cdot kg^{-1} \cdot K^{-1}$).

Sous l'hypothèse que $L_v(T)$ soit constante, la formule d'Auguste-Roche-Magnus offre une solution approximative de l'équation de Clausius-Clapeyron pour la vapeur d'eau dans l'atmosphère :

$$e_s(T) \approx 6.1094 \exp\left(\frac{17.625T}{243.04 + T}\right) \quad (2.11)$$

avec e_s la pression saturante de la vapeur d'eau (en hPa), et T la température (en °Celsius).

Cette solution a la particularité d'être sous une forme exponentielle, et sa dérivée montrant

une évolution de la pression saturante de manière exponentielle en fonction de la température, suivant un taux entre 6 et 7%. Ce résultat représente la base physique de l'hypothèse de transférabilité de la formule de Clausius-Clapeyron aux précipitations, supposant que les extrêmes de précipitations augmenteraient en fonction des températures à un rythme similaire.

2.3.2 Estimation par méthode de regroupement

La relation entre les précipitations extrêmes et la température a été fréquemment étudiée à l'aide de la méthode dite de « regroupement » (ou *binning method* en anglais). Cette approche a été largement utilisée dans les premières tentatives visant à valider empiriquement l'hypothèse de l'amplification des précipitations extrêmes selon la relation de Clausius-Clapeyron [16, 17].

Les faibles intensités de précipitations (généralement inférieures à $0.1\text{--}2 \text{ mm} \cdot \text{h}^{-1}$) sont d'abord exclues de la série de données. Les valeurs restantes sont ensuite regroupées en intervalles définis selon la température associée. La taille de ces intervalles peut être déterminée soit en fonction d'un nombre fixe d'observations, soit en fonction d'une largeur donnée sur l'échelle des températures (*e.g.*, 2°C). Pour chaque intervalle, on calcule la température moyenne ainsi qu'un quantile élevé q (tel que le 99ème niveau de quantile) des précipitations. Une régression linéaire est ensuite ajustée sur les données transformées logarithmiquement, en modélisant la relation entre le quantile de précipitations P_q et la température moyenne \bar{T} [18] :

$$\log(P_q) = \alpha \bar{T} + \beta. \quad (2.12)$$

Le taux, exprimé en pourcentage, d'augmentation des précipitations par degré Celsius est obtenu à partir du coefficient α :

$$\frac{\partial P_q}{\partial T} = 100(e^\alpha - 1) \quad (2.13)$$

Cette méthode présente l'avantage d'être à la fois simple à mettre en œuvre et facile à interpréter. Toutefois, plusieurs limitations doivent être prises en compte. D'une part, le choix du seuil d'exclusion initial peut influencer significativement les taux d'augmentation obtenus. D'autre part, une inversion de la relation entre température et précipitations est fréquemment observée au-delà d'un certain seuil thermique, situé typiquement entre 20 et 30°C , ce qui s'explique par la diminution de l'humidité disponible à haute température. Deux stratégies sont généralement proposées pour remédier à cette non-linéarité : (1) substituer la température de l'air par le point de rosée comme variable explicative, ou (2) intégrer

explicitement un point de bascule dans la régression linéaire [19].

Enfin, une méthode alternative, fréquemment utilisée et conceptuellement proche, consiste à appliquer une régression de quantile directement sur l'ensemble des données, sans recourir au regroupement préalable [20, 21].

2.3.3 Estimation par la loi des valeurs extrêmes généralisée

Plusieurs travaux ont néanmoins émis des critiques envers l'approche de regroupement [22, 23]. Notamment, les variations saisonnières influençant simultanément les températures et les précipitations, introduisent un facteur de confusion qui remet en question l'existence d'un lien causal direct entre ces deux variables [24]. La pertinence de la méthode de regroupement pour mesurer l'impact des changements climatiques sur les précipitations extrêmes est donc remise en question, le lien entre températures et précipitations journalières (pouvant être expliqué par des variations saisonnières) étant sensiblement différent de celui lié à un réchauffement global.

Pour pallier à ce problème, on peut estimer le taux d'augmentation des précipitations en fonction des anomalies de températures en changements climatiques à l'aide d'une loi GEV non-stationnaire [24]. Les maxima annuels de précipitations sont sélectionnés et normalisés, en les divisant par la médiane de la série. Une loi GEV est ensuite ajustée sur ces données, en utilisant comme covariable les anomalies annuelles (ou limitées à la saison présentant la plus forte probabilité d'occurrence de précipitations extrêmes (*e.g.*, l'été)) de températures moyennes. La paramétrisation de la dépendance des paramètres de la loi GEV aux anomalies de température dépend des implémentations. Certains utilisent [25] :

$$\mu(T) = \mu_0 + \mu_1 T, \quad \sigma(T) = \exp(\sigma_0 + \sigma_1 T), \quad \xi(T) = \xi. \quad (2.14)$$

Le taux d'augmentation entre les précipitations extrêmes et les températures (en anglais *Temperature-Precipitation Scaling Rate*, ou TPSR) s'obtient en calculant le quantile q obtenu pour les anomalies de températures T_1 et T_2 , menant à Q_1 et Q_2 . Tout quantile de niveau q en fonction de T_i peut être obtenu par la formule :

$$Q(q; \mu(T_i), \sigma(T_i), \xi(T_i)) = \begin{cases} \mu(T_i) - \sigma(T_i) \log[-\log(q)] & \text{si } \xi(T_i) = 0 \text{ et } q \in (0, 1); \\ \mu(T_i) + \frac{\sigma(T_i)}{\xi(T_i)} \{[-\log(q)]^{-\xi(T_i)} - 1\} & \text{si } \xi(T_i) > 0 \text{ et } q \in [0, 1), \\ & \text{ou } \xi(T_i) < 0 \text{ et } q \in (0, 1]. \end{cases} \quad (2.15)$$

Finalement, le TPSR α (en $\%/^{\circ}C$) est donné comme la solution de la formule :

$$Q_2 = Q_1(1 + 0.001\alpha)^{T_2 - T_1} \quad (2.16)$$

En général on choisira $\Delta T = T_2 - T_1 = 1$, car la valeur de ΔT n'a que peu d'impact sur le résultat [26].

CHAPITRE 3 REVUE DE LITTÉRATURE

Ce chapitre vise à fournir une synthèse de la littérature concernant les principaux thèmes et enjeux abordés dans ce mémoire. Dans un premier temps, nous proposerons un bref résumé de l’histoire de la modélisation numérique du climat, de ses avancées majeures, et ses défis actuels. Nous couvrirons ensuite les méthodes standards de réduction d’échelle, avant de nous intéresser à l’utilisation de l’apprentissage profond dans ce contexte, en portant une attention particulière aux approches dites « informées par la physique ».

3.1 Modélisation numérique du climat

Le premier modèle numérique du climat est proposé en 1955, simulant l’atmosphère sur 2 niveaux verticaux, sur une durée de 31 jours [27]. Ces travaux initiaux, combinés au progrès des puissances de calcul, et à la quantité grandissante de données récoltées, mèneront à l’apparition des premiers modèles climatiques globaux (ou *Climate Global Models* en anglais, GCMs) durant les décennies suivantes [28].

Les GCMs sont des modèles numériques comprenant d’une part, un « cœur » dynamique, permettant de résoudre les équations fondamentales de l’atmosphère (*e.g.*, Navier-Stokes), et d’autre part, une partie physique qui modélise les échanges entre les grands « corps » de la planète : océans, continents, atmosphère. Certains processus physiques se produisant à une échelle plus fine que la résolution des GCMs sont représentés par des paramétrisations empiriques, sources d’incertitude qui motivent l’augmentation de la résolution spatiale et temporelle des modèles [29, 30]. En intégrant progressivement les interactions entre climat et biosphère, et en modélisant explicitement de plus en plus de phénomènes (biogéochimie marine, aérosols, etc.), les GCMs ont évolué vers des modèles maintenant appelés Modèles du Système Terre (ou *Earth System Model* en anglais, ESMs) [28, 31].

En parallèle de cette évolution, les applications de la modélisation climatique numérique se sont considérablement élargies. L’avènement des simulations climatiques globales au cours des dernières décennies a permis des avancées majeures dans la compréhension des causes et des conséquences des changements climatiques [32, 33]. Les projections climatiques établies selon différents scénarios d’émissions de gaz à effet de serre (GES) ont largement contribué à orienter les politiques nationales et internationales, ainsi qu’à anticiper l’évolution des événements météorologiques extrêmes, tels que les inondations associées à un niveau de retour de 100 ans, ou encore à estimer les besoins futurs en eau des zones urbaines et du secteur

agricole [29, 34–36]. Malgré les améliorations récentes, la complexité de calcul actuelle des ESMs reste prohibitive, ce qui rend la production de simulations à des résolutions spatiales inférieures à 50 kilomètres (km) irréalisable sur le plan opérationnel [37].

Étant donné que de nombreux processus physiques ne peuvent être résolus qu’à une résolution spatiale plus fine, les simulations à grande échelle ne disposent pas des détails physiques nécessaires pour prendre en compte la dynamique à petite échelle. Cependant, il existe un besoin aigu de simulations climatiques à haute résolution pour évaluer les impacts des changements climatiques, dans le cadre d’études d’impact climatique régionales ou locales [38, 39], d’évaluations des risques liés aux phénomènes météorologiques extrêmes [40], de la modélisation hydrologique [41], ou de la modélisation des incendies de forêt [42]. Pour répondre à ce besoin, des techniques de réduction d’échelle ont été développées afin d’accroître la résolution des simulations climatiques, le plus souvent en restreignant la couverture sur une région donnée. On distingue deux grandes approches : la réduction d’échelle *dynamique* ou *statistique*.

3.2 Réduction d’échelle dynamique et statistique

La réduction d’échelle dite *dynamique* consiste à réaliser une simulation à l’aide d’un modèle climatique régional (ou *Regional Climate Model* en anglais, RCM) à haute résolution sur un domaine spatial restreint, en le contraignant aux frontières par les sorties d’un modèle global à plus faible résolution. À l’inverse, les approches *statistiques* cherchent à établir des relations empiriques entre les données à haute résolution (*e.g.*, observations historiques, données satellitaires, simulations à haute résolution) et les simulations climatiques à basse résolution, afin d’estimer les variables climatiques locales à partir des champs globaux.

Les modèles dynamiques de réduction d’échelle intègrent des processus, dynamiques et biochimiques propres à chaque région simulée. Les conditions aux frontières latérales et supérieures sont fixées par les simulations de modèles globaux à chaque pas de temps, assurant la cohérence avec la circulation atmosphérique à large échelle [43]. Ces modèles permettent d’obtenir des simulations aux résolutions spatiales allant jusqu’à 10 km [44]. Pour référence, le RCM canadien : CRCM5 (5th Canadian Regional Climate Model), permet d’obtenir des simulations climatiques à une résolution de $0.11^\circ \approx 12$ km sur le domaine de l’Amérique du Nord, incluant 685×668 cellules, sur 56 niveaux verticaux et 17 niveaux de surface, pour un pas de temps de 5 minutes [45, 46].

L’augmentation de la résolution spatiale et temporelle des RCM rend possible la modélisation d’un plus grand nombre de phénomènes et d’interactions à fine échelle, réduisant les biais des

simulations globales [47]. Toutefois, leur coût numérique élevé limite encore la disponibilité d'ensemble de simulations climatiques à long terme, pourtant essentielles pour modéliser la variabilité naturelle du climat et évaluer une diversité de scénarios d'émissions de GES [48].

Contrairement à la méthode *dynamique*, la réduction d'échelle *statistique* ne résout pas explicitement de processus physiques. Cette famille de méthodes vise à établir des relations statistiques entre les variables atmosphériques à faible résolution issues des GCMs (ou parfois des RCMs), et des données à haute résolution (observées ou issues de simulations dans le cadre d'émulateurs). Il existe différents cadres méthodologiques séparant ces méthodes [49] :

- **Perfect Prognosis (PP)** : Ces méthodes consistent à calibrer un modèle statistique entre des prédictors à grande échelle (issus idéalement de réanalyses, c'est-à-dire des simulations alignées sur les observations) et des observations locales. Le modèle ainsi formé, reposant en général sur de la régression, est ensuite appliqué aux sorties de GCMs, corrigées, ou jugées sans biais, pour générer des projections locales. La qualité des prédictors est cruciale pour la réussite des approches PP.
- **Model Output Statistics (MOS)** : Cette approche consiste à corriger la distribution des sorties des GCMs pour les aligner avec celles des observations locales, en effectuant une correction statistique, typiquement *via* des méthodes de quantile mapping, ou de correction d'échelle. Les méthodes MOS correspondent généralement à de la correction de biais.
- **Weather Generators (WG)** : Ces modèles reposent sur des approches stochastiques pouvant s'inscrire tant dans le cadre du PP que du MOS. Les WG modélisent les propriétés marginales et la structure temporelle des variables météorologiques, tout en étant parfois capables de reproduire la dépendance spatiale entre les prédictors.

Les méthodes de réduction d'échelle statistiques sont couramment utilisées dans la littérature, en partie grâce à leur simplicité d'implémentation et aux faibles coûts de calcul associés. Par exemple, certains travaux ont évalué l'impact des réchauffements climatiques sur les inondations en appliquant des méthodes relevant de chacun des cadres de réduction d'échelle susmentionnés pour affiner la résolution d'un ensemble de simulations issues de GCMs [50].

Un autre cadre méthodologique couramment utilisé en apprentissage profond, et adopté dans ce mémoire, se nomme le **Perfect Upscaling (PU)**. Ce cadre, proche de celui du PP, repose sur l'hypothèse que la relation entre les prédictors (simulations à basse résolution, et potentielles covariables aux résolutions variables) et les prédictants (simulations, ou observations à haute résolution) est exactement l'inverse d'un processus d'agrégation (*e.g.*, une moyenne ou une accumulation). Il ignore ainsi le problème du biais des GCM.

3.3 Apprentissage profond pour la réduction d'échelle

Ces dix dernières années, l'apprentissage profond a graduellement gagné en popularité dans le domaine de la modélisation climatique [51]. Ces méthodes ont démontré leur efficacité dans des applications variées, notamment pour la simulation du climat global [52], la réalisation de prévisions météorologiques, aux performances comparables à celles des modèles numériques traditionnels [53], la paramétrisation des modèles numériques de climat afin d'améliorer la modélisation des processus de fine échelle [54], ainsi que la réduction d'échelle de simulations climatiques [55].

Les premières applications de l'apprentissage profond à la réduction d'échelle de simulations climatiques se sont inspirées des progrès réalisés en super-résolution, une tâche de vision par ordinateur visant à reconstruire une image haute résolution à partir d'une version dégradée. Problème inverse mal-posé, la super-résolution cherche à restaurer des détails spatiaux fins, perdus lors de la dégradation. Ces travaux ont jeté les bases des premières tentatives d'adaptation de ces modèles à la réduction d'échelle de simulations climatiques. Dans la continuité de la section 2.1, nous présenterons séparément les travaux ayant recouru à des approches déterministes, de ceux reposant sur des méthodes génératives.

3.3.1 Méthodes déterministes

Les modèles déterministes ont été les premiers à être utilisés pour effectuer de la réduction d'échelle de variables climatiques. Le potentiel des réseaux de neurones convolutifs profonds pour cette tâche a été démontré en adaptant un modèle issu de la littérature de la super-résolution pour la réduction d'échelle spatiale de précipitations journalières, augmentant la résolution des simulations de 1° à 0.125° , et obtenant des performances supérieures aux approches de correction de biais et désagrégation spatiale (BCSD), ou de régression automatique [56–58]. Les résultats soulignent également l'importance de l'adaptation des architectures issues de la super-résolution (traitant principalement des images RGB) pour prendre en compte les caractéristiques des variables climatiques.

Par la suite, de nombreux travaux ont implémenté et comparé des CNNs aux méthodes statistiques classiques. Certains se sont focalisés sur la modélisation des précipitations extrêmes [59], d'autres ont exploré des architectures hybrides combinant CNN et perceptrons multicouches (MLP) [60], ou ont utilisé des CNNs pour estimer les paramètres de distributions probabilistes censées représenter les données de précipitations [61]. Des améliorations du modèle issu de [58] ont également été proposées, améliorant sensiblement les performances tout en réduisant le coût numérique [62]. Des modèles de type UNet ont aussi été utilisés

pour effectuer de la réduction d'échelle de variables climatiques [63–65].

D'autres modèles déterministes plus lourds et complexes (en termes de paramètres et d'architecture) ont également été utilisés avec succès, mais ces approches ne sont pas explorées dans ce mémoire. Un modèle de type *Transformers* pré-entraîné sur de très larges jeux de données hétérogènes, a permis ensuite d'obtenir des performances remarquables sur un certain nombre de tâches, incluant la réduction d'échelle [66]. D'autres modèles *Transformers* (SwinIR et Uformer) ont été implémentés pour de la réduction d'échelle et de la correction de biais de température et vitesse de vent [67]. Le modèle Mamba, de la famille des *State Space Model*, a également été appliqué à la réduction d'échelle de précipitations contraint par la topographie [68]. Enfin, d'autres travaux ont proposé une approche basée sur les opérateurs de Fourier permettant d'effectuer de la réduction d'échelle à une résolution arbitraire [69].

La fonction de perte utilisée pour l'entraînement des modèles a également fait l'objet de plusieurs études, notamment dans le but d'améliorer la représentation des événements extrêmes. Par exemple, il a été démontré qu'un modèle de type UNet, entraîné à l'aide d'une combinaison d'une perte quadratique à pondération exponentielle et d'une métrique de similarité issue de la vision par ordinateur (SSIM), est capable de modéliser l'ensemble de la distribution des précipitations à haute résolution, y compris la queue et les valeurs extrêmes [70]. Une fonction de perte asymétrique a également été conçue pour pénaliser plus fortement les sous-estimations, de manière proportionnelle à l'intensité de la valeur cible, ce qui permet une meilleure capture des valeurs extrêmes [65]. Enfin, certains travaux avancent que la perte de Charbonnier [71] $\mathcal{L}_{\text{charb}}$ offre une alternative plus robuste aux fonctions de perte standards \mathcal{L}_1 et \mathcal{L}_2 [67] :

$$\mathcal{L}_{\text{charb}}(\theta) = \frac{1}{N} \sum_{i=1}^N \sqrt{(\hat{y}_i - y_i)^2 + \epsilon^2} \quad (3.1)$$

$$\mathcal{L}_1(\theta) = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (3.2)$$

$$\mathcal{L}_2(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (3.3)$$

où N est le nombre d'échantillons, y_i la valeur cible, $\hat{y}_i = f_{\theta}(x_i)$ la prédiction du modèle avec paramètres θ , et ϵ un petit terme positif (*e.g.*, 10^{-3}) assurant la différentiabilité.

3.3.2 Méthodes génératives

De par leur nature probabiliste, les méthodes génératives offrent deux avantages principaux par rapport aux approches déterministes : (1) en modélisant la distribution conditionnelle

d’une variable, les modèles de cette famille permettent de mieux modéliser la variabilité naturelle liée au climat, et (2) il est possible de quantifier les incertitudes de leurs prédictions, ce qui est important pour la réduction d’échelle d’évènements météorologiques extrêmes.

Afin de tirer profit des avantages des méthodes génératives, tout en conservant les capacités de l’architecture du UNet, une méthode permettant d’adapter ce modèle à un contexte probabiliste dans le cadre de la segmentation d’images médicales a été développée [72]. Les auteurs proposent d’apprendre un espace latent représentant les paramètres d’une loi gaussienne, et d’échantillonner cette loi pour enrichir les prédictions du UNet. Cette approche permet de modéliser la variabilité inhérente aux solutions, liée à la nature mal posée du problème inverse. Dans le cadre d’un projet collaboratif non mentionné dans ce mémoire, nous appliquons avec succès cette méthode à la réduction d’échelle de précipitations, et améliorons la conservation des hautes fréquences des données, ainsi que la prédiction des évènements extrêmes.

Bien qu’elles ne soient pas abordées en détail dans ce mémoire, plusieurs approches génératives ont été appliquées avec succès à la réduction d’échelle climatique. En particulier, plusieurs études récentes ont mis en évidence les performances supérieures des modèles de diffusion [73] pour la reconstruction de champs à haute résolution [74–76]. Un réseau antagoniste génératif (ou *Generative Adversarial Network* en anglais, GAN) de type CycleGAN [77, 78], capable de reproduire fidèlement les fines structures spatiales des précipitations tout en corrigeant les biais systématiques présents dans les sorties des ESM a également été récemment proposé [79].

3.4 Apprentissage profond informé par la physique

La nature de la tâche de réduction d’échelle, combinée aux contraintes physiques inhérentes de la modélisation climatique, constituent une double motivation pour le développement de méthodes dites « informées par la physique ».

Premièrement, la réduction d’échelle (dans le domaine climatique) constitue un problème inverse mal posé pour deux raisons : (1) la solution n’est pas nécessairement unique, en raison de la perte d’information entre la haute à la basse résolution, et (2) la solution ne dépend pas de manière continue des données d’entrée, dans la mesure où de faibles perturbations peuvent provoquer des variations significatives de la solution [80]. Pour remédier à ces limitations, il est nécessaire d’introduire des *a priori* sur le système, dans le but de réduire la sensibilité aux perturbations et de contraindre l’espace des solutions admissibles. Deuxièmement, les prédictions issues des méthodes d’apprentissage profond peuvent présenter des incohérences

physiques [81]. L'introduction d'*a priori*, qu'ils soient simples ou sophistiqués, vise à améliorer le réalisme des résultats, à renforcer la robustesse face aux situations hors distribution d'apprentissage, ainsi qu'à accroître la confiance des utilisateurs et des utilisatrices dans les modèles.

Parmi les approches dites « informées par la physique », on distingue généralement deux grandes catégories. La première regroupe les méthodes intégrant des contraintes, dures ou douces, à la fonction de coût afin d'imposer le respect, total ou partiel, de lois physiques ou de principes de conservation. La seconde catégorie repose sur l'incorporation d'*a priori* directement au sein de l'architecture du modèle, en s'appuyant sur des connaissances du système (également appelées « connaissances du domaine » ou « domain knowledge » en anglais) pour guider sa conception. Une revue de littérature liste les différentes approches visant à intégrer des contraintes physiques ainsi que des *a priori* issues des sciences du climat dans les méthodes d'apprentissage profond appliquées à la modélisation climatique [82].

Au sein des approches par contraintes, on différencie généralement deux types : les contraintes dures et les contraintes douces. Une contrainte est dite dure lorsqu'elle doit être strictement satisfaite par toute solution admissible, tandis qu'une contrainte douce est introduite sous forme de pénalité dans la fonction de perte, pondérée par un coefficient, ce qui offre davantage de flexibilité mais moins de garantie. Dans cette optique, des autrices et auteurs proposent plusieurs formulations de contraintes dures pour imposer la conservation de la masse dans un cadre de réduction d'échelle appliqué aux colonnes de vapeur d'eau, et comparent les performances de plusieurs modèles d'apprentissage profond sans contraintes, ou avec contraintes dures ou douces [83]. Leurs conclusions indiquent que l'introduction de contraintes permet de réduire les erreurs dans les régions à forte complexité spatiale, telles que les zones côtières ou montagneuses, contribuant également à une amélioration notable de la qualité visuelle des prédictions. D'autres travaux exploitent la relation de Clausius-Clapeyron pour intégrer des contraintes dures, ou douces respectivement dans des contextes de post-traitement et de réduction d'échelle de variables atmosphériques [84, 85]. Enfin, une autre approche par contrainte dure consiste à imposer, en post-traitement, une contrainte d'équilibre entre les prédictions du GAN à haute résolution et les valeurs agrégées des champs d'entrée, afin de renforcer la cohérence physique et d'améliorer la robustesse des prédictions en changements climatiques [70].

La seconde catégorie regroupe une famille de méthodes plus large, qui proposent d'intégrer des *a priori* à travers la conception des architectures d'apprentissage profond, ou dans les étapes du traitement de données. Par exemple, une approche issue de la littérature permettant d'améliorer la réduction d'échelle d'événements de précipitations à très haute résolution

(250 mètres), adapte une méthode de super-résolution pour vidéos, en remplaçant le champ de « flux optiques », utilisé pour transmettre l’information entre trames, par le champ d’advection du vent [86]. Cette substitution exploite les connaissances sur la forte dépendance entre la précipitation et la dynamique du vent, permettant ainsi de générer des champs à très fine échelle cohérents spatialement. Une autre approche propose d’intégrer un modèle orographique (permettant de mieux modéliser la structure des précipitations dans les régions à orographie complexe) au sein d’un GAN, et de conditionner l’espace latent du modèle génératif à l’aide d’un ensemble de processus gaussiens pour obtenir un point de départ statistiquement et physiquement enrichi [87]. Certains proposent de modéliser les processus climatiques sous la forme d’une équation aux dérivées partielles d’advection, et prédisent l’évolution du champ de vitesse à l’aide de réseaux de neurones [88]. Cette approche permet d’intégrer explicitement la dynamique physique du système climatique dans les prédictions. Enfin, d’autres auteurs et autrices introduisent une série de transformations inspirées des sciences atmosphériques, permettant d’effectuer différentes tâches de modélisation climatique dans un espace invariant en réchauffements climatiques [89].

CHAPITRE 4 DONNÉES ET ANALYSE EXPLORATOIRE

L’objectif de ce chapitre est de présenter en détail le jeu de données utilisé dans ce mémoire, et de proposer une analyse statistique exploratoire des variables d’intérêt.

4.1 ClimEx : grand ensemble de simulations en réchauffements climatiques

Le projet ClimEx (ou CRCM5-LE) fournit un ensemble de 50 simulations climatiques régionales indépendantes, à une résolution spatiale de $0.11^\circ \approx 12$ km. Ces simulations couvrent deux domaines géographiques : le nord-est de l’Amérique du Nord, et l’Europe, sur une période allant de 1950 à 2099, à une résolution temporelle de 5 minutes [90]. Afin d’obtenir ces simulations à haute résolution, une réduction d’échelle dynamique a été effectuée, en utilisant un ensemble de 50 simulations globales indépendantes générées par l’ESM canadien CanESM2, à une résolution spatiale de $2.8^\circ \approx 310$ km. Celles-ci ont servi de forçage au RCM canadien CRCM5 (version 5), développé à l’Université du Québec à Montréal (UQAM). Le caractère indépendant des membres de l’ensemble repose sur l’introduction de perturbations initiales appliquées à la simulation globale du CanESM2 : 5 perturbations en 1850, suivies de 10 en 1950. Ces perturbations permettent de générer 50 trajectoires globales indépendantes à partir de 1955, capturant la variabilité naturelle du climat. Le forçage radiatif appliqué à partir de 2006 correspond au scénario d’émission de GES RCP8.5. En résultat, une cinquantaine de variables par simulation ont été archivées, avec des pas de temps allant de l’horaire au journalier. Certaines variables sont disponibles en libre accès au pas de temps journalier via la plateforme PAVICS.

Les travaux présentés dans ce mémoire s’appuient sur un sous-ensemble du jeu de données ClimEx. Plus précisément, nous utilisons les sorties d’un seul membre de l’ensemble (membre *kdf*), en laissant l’étude de l’apprentissage et de la généralisation inter-membres à des travaux ultérieurs. Les accumulations journalières de précipitations (**pr**), exprimées en $\text{kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, sont extraites, ainsi que les statistiques journalières de température : moyennes (**tas**), minima (**tasmin**) et maxima (**tasmax**), exprimées en kelvins (K). L’intégration de covariables climatiques supplémentaires pour la réduction d’échelle des précipitations est laissée à des travaux futurs. Par ailleurs, pour des raisons de contraintes numériques, le domaine spatial initial, comprenant 280×280 cellules, est restreint à une région de 128×128 cellules, couvrant le sud du Québec, une partie des provinces maritimes canadiennes, et du nord-est des États-Unis (voir Figure 4.1 (a)).

4.2 Analyse statistique exploratoire

Afin d'identifier des tendances pouvant être liées aux changements climatiques, on distingue deux périodes : **A** et **B** (1955 – 2030 et 2025 – 2099), couvrant 75 ans de données journalières chacune, avec une superposition de 5 ans. On peut considérer la période **A** comme une période avec des changements climatiques faibles, et la période **B** avec des changements climatiques forts, selon le scénario RCP8.5. La carte de la différence absolue de la médiane des températures moyennes journalières, présentée à la Figure 4.1 (a), met en évidence le réchauffement général des températures sur tout le domaine ClimEx, et particulièrement dans le Nord. Elle souligne également l'hétérogénéité de la réponse du domaine choisi pour ce mémoire en ce qui concerne l'évolution des températures en changements climatiques. La différence la plus importante du sous-domaine est de 5.07°C , la plus petite de 2.40°C , et la médiane de 3.52° , comparées à 5.51°C , 2.11°C , et 3.37°C pour le domaine entier, suggérant que le sous-domaine choisi est représentatif du domaine ClimEx quant à la réponse aux changements climatiques.

Les cartes (b:e) de la figure 4.1 se restreignent au sous-domaine de 128×128 cellules. Les cartes (b) et (c) présentent respectivement la médiane et le 95ème quantile des précipitations non-nulles pour la période **A**. On remarque que les régions avec une médiane élevée correspondent aux régions montagneuses proches de la côte : les monts Chics-Chocs en Gaspésie, le Plateau du Cap-Breton en Nouvelle-Écosse, et les monts Long-Range en Terre-Neuve-et-Labrador. En raison de leur orographie complexe, ces régions posent généralement plus de difficultés lors de la réduction d'échelle, car les fines dynamiques liées à cette orographie ne sont pas pris en compte par les GCMs. Les fortes précipitations observées dans les régions côtières (carte (c)) s'expliquent par la convergence fréquente de systèmes dépressionnaires en provenance de l'Atlantique Nord, dont l'intensité est amplifiée par l'apport d'humidité du Gulf Stream.

Les cartes (d) et (e) illustrent les variations relatives, entre périodes **B** et **A**, de la médiane et du 95ème quantile des précipitations journalières non-nulles. On observe des motifs spatiaux marqués pour la médiane, avec une forte augmentation sur les zones continentales, atteignant jusqu'à 49.2%, une quasi-stabilité le long des côtes, ainsi qu'une diminution notable de la médiane des précipitations sur l'océan Atlantique et les Grands Lacs (en bas à gauche sur la carte (d)), jusqu'à -18.7% . Cette hétérogénéité suggère que l'évolution des précipitations faibles à modérées dépend fortement de facteurs externes, tels que l'orographie ou les conditions atmosphériques régionales. En revanche, les précipitations extrêmes (représentées par le 95ème quantile) sont en hausse sur l'ensemble du domaine, sans motif spatial particulier. L'augmentation varie de 5.14% à 27.6% , avec une moyenne de 15.99% . Toute réserve gardée, cette tendance généralisée à la hausse, en parallèle de l'élévation des températures

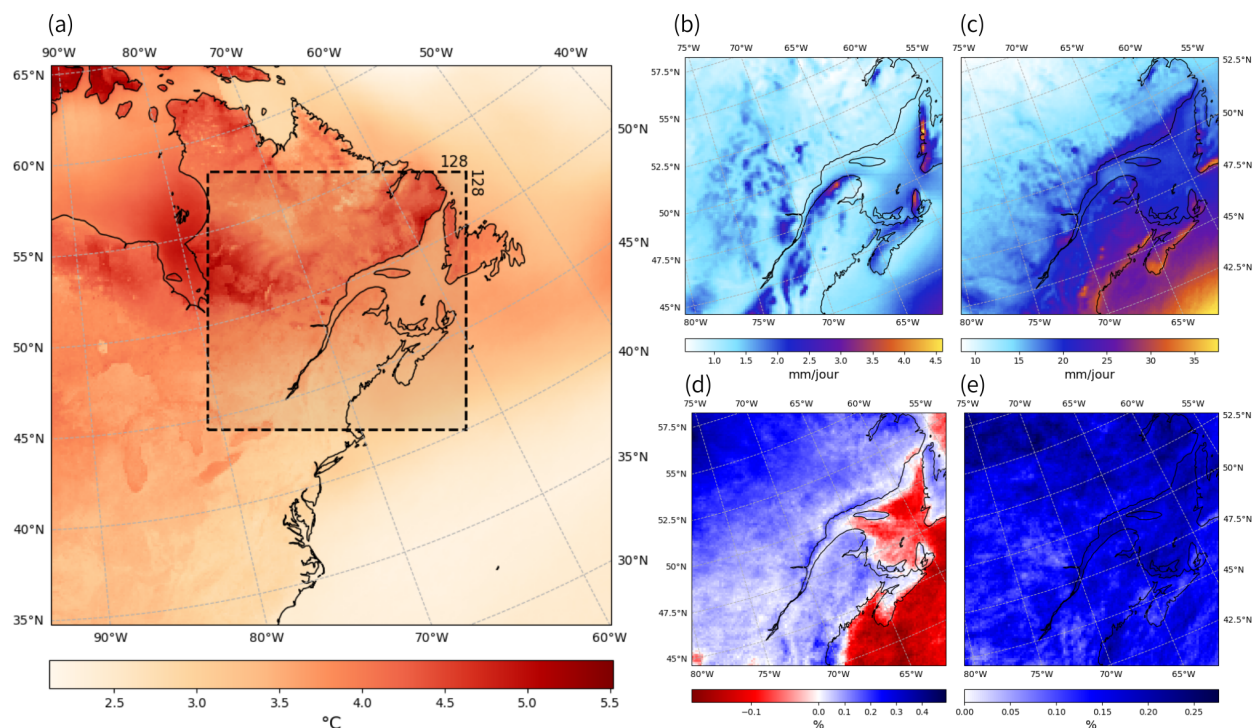


FIGURE 4.1 Soient une période **A** allant de 1955 à 2030, et une période **B** chevauchant de cinq ans allant de 2025 à 2099. (a) Différence absolue, entre périodes **B** et **A**, de la médiane des températures moyennes journalières sur le domaine de ClimEx, et en pointillés noirs : le sous-domaine utilisé dans ce mémoire. (b) Médiane des précipitations journalières non-nulles sur la période **A**. (c) 95ème quantile des précipitations journalières non-nulles sur la période **A**. (d) Différence relative de la médiane des précipitations non-nulles. (e) Différence relative du 95ème quantile des précipitations non-nulles.

moyennes (cf. carte (a)), semble cohérente avec la relation théorique de Clausius-Clapeyron, présentée à la section 2.3.

Comme énoncé à la section 3.2, les travaux de ce mémoire s'appuient sur le cadre méthodologique du *Perfect Upscaling*. On suppose donc que les champs de précipitations à haute résolution correspondent à l'inverse d'une agrégation par moyenne. On considère les données issues de ClimEx comme les données à haute résolution, et on effectue un moyennage spatial par blocs de 256 cellules (16×16) pour obtenir les données à basse résolution. Un exemple aléatoire du résultat de cette opération est présenté sur la partie droite de la figure 4.2. On observe logiquement une perte significative d'information due à l'agrégation spatiale. En l'occurrence, les motifs fins associés aux précipitations disparaissent en grande partie, ce qui rend difficile l'identification de structures météorologiques précises, ainsi que la localisation d'événements de précipitations extrêmes, souvent concentrés sur de petites zones, ou sur des bandes.

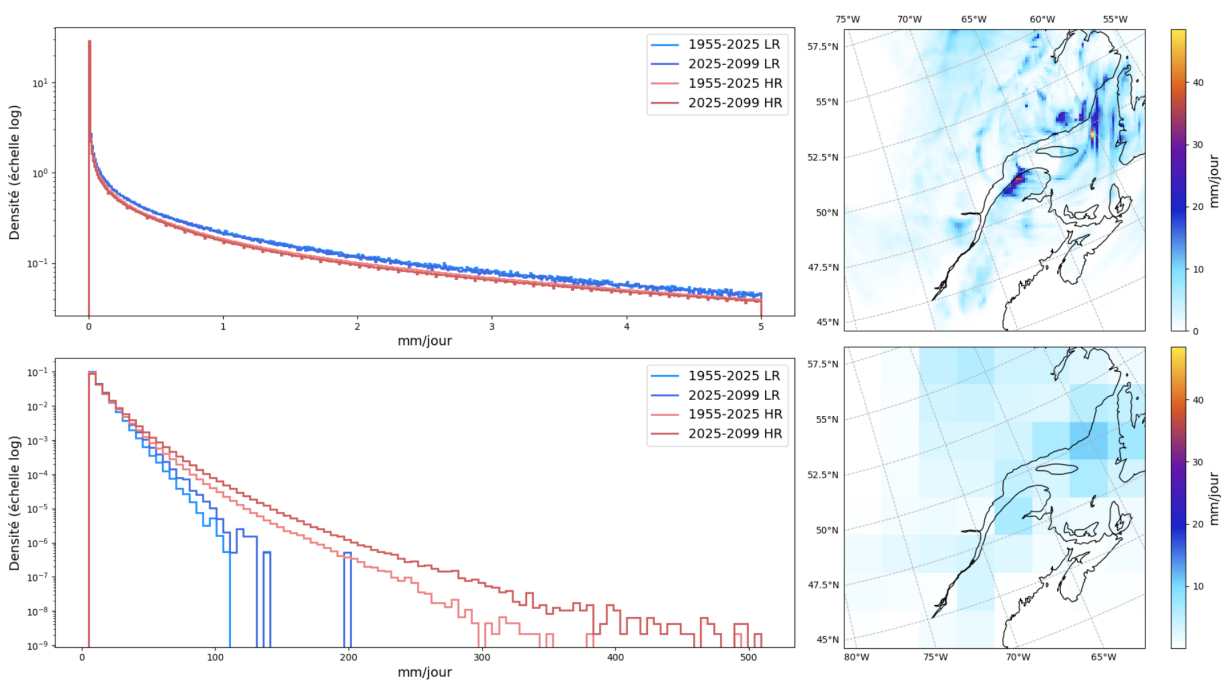


FIGURE 4.2 Afin de simuler des observations à basse résolution (LR), les précipitations journalières haute résolution (HR) issues de ClimEx sont agrégées en moyennant des blocs de 256 cellules (16×16). (**À gauche**) Les histogrammes des précipitations journalières, séparés selon les intensités ($0 - 5\text{mm/jour}$ en haut, et $\geq 5\text{mm/jour}$ en bas) comparant les périodes **A** et **B**, en haute et basse résolution. (**À droite**) Un champ de précipitations journalières sélectionné aléatoirement (en haut), et sa version agrégée (en bas).

L'analyse des histogrammes de précipitations (voir figure 4.2), calculés sur l'agrégation de

toutes les cellules du domaine, met en évidence les effets de l'agrégation spatiale : à basse résolution, la densité des faibles précipitations augmente, en raison de la dilution des cellules sans pluie dans les zones voisines. En revanche, les précipitations modérées deviennent nettement sous-représentées, conséquence directe du lissage spatial. Enfin, la majeure partie de la queue de la distribution, correspondant aux fortes précipitations à haute résolution, est fortement atténuée, voire totalement absente à basse résolution. Lorsqu'on sépare les données par période, on observe clairement que la queue de la distribution des précipitations pour la période **B** est généralement plus lourde, ce qui témoigne d'une tendance d'augmentation marquée des précipitations extrêmes en changements climatiques.

La capacité à restituer les détails spatiaux fins propres aux événements de précipitations, à reconstruire fidèlement la queue de leur distribution, et à intégrer l'amplification des valeurs extrêmes liée aux changements climatiques constituent les principaux défis que les travaux présentés dans ce mémoire cherchent à relever.

CHAPITRE 5 APPRENTISSAGE PROFOND ET RÉDUCTION D'ÉCHELLE DES PRÉCIPITATIONS EN CHANGEMENTS CLIMATIQUES

Dans ce chapitre, nous détaillons l'architecture du modèle d'apprentissage profond utilisé pour effectuer la réduction d'échelle des champs de précipitations. Nous présentons ensuite une première évaluation de ses performances, et nous analysons ses limites afin de motiver les améliorations méthodologiques proposées au chapitre suivant.

5.1 UNet pour la réduction d'échelle

Cette section vise à présenter l'architecture originale du modèle UNet utilisé dans ce mémoire, et présenter la stratégie employée pour effectuer la réduction d'échelle des champs de précipitations avec ce modèle.

5.1.1 Présentation de l'architecture

Comme indiqué précédemment, nous utilisons comme point de départ une architecture issue de travaux antérieurs [2]. L'architecture suit une structure classique encodeur/décodeur, dont chaque niveau est composé de plusieurs blocs résiduels, avec des sauts de connexions propres, et une intégration de l'information temporelle afin de guider le UNet. Nous présentons ici les détails de cette architecture, illustrées à la figure 5.1.

L'encodeur est composé de quatre niveaux, chacun doublant le nombre de canaux (dimension des « caractéristiques », comme les trois canaux RGB pour une image en couleur) du précédent, et réduisant de moitié leurs dimensions spatiales. Ainsi, une entrée de 128×128 cellules est transformée en cartes de caractéristiques de 16×16 cellules. Chaque niveau de l'encodeur comprend trois blocs résiduels : le premier réalise une agrégation spatiale par moyenne, le deuxième double le nombre de canaux, et le troisième conserve les dimensions. Le décodeur comporte lui quatre blocs résiduels par niveau : le premier effectuant le sur-échantillonnage, et les autres réduisant le nombre de canaux tout en intégrant les informations issues des sauts de connexions des blocs correspondants de l'encodeur.

La structure originale des blocs résiduels est illustrée à la Figure 5.1, et peut inclure selon le contexte, un opérateur d'agrégation ou de sur-échantillonnage par plus proche voisin. Aucun gain de performance significatif n'ayant été observé lors de tests avec des méthodes d'interpolation bilinéaire ou cubique, le sur-échantillonnage par plus proche voisin a été retenu pour sa simplicité et l'absence de biais introduit. La fonction d'activation utilisée est

$\text{SiLU}(x) = x \text{sig}(x)$, fréquemment employée en raison de sa dérivabilité en tout point, son gradient lisse comparable à celui de la sigmoïde, et son comportement non-linéaire proche de celui de la fonction $\text{ReLU}(x) = \max(0, x)$ [91]. La normalisation par groupe permet de stabiliser l’entraînement, tout en étant plus robuste que la normalisation par lot pour les lots de petites tailles [92]. La normalisation par groupe centre les valeurs au sein d’un groupe (fixé à 32 ici) de cartes de caractéristiques, et en réduit la variance à l’unité. Ces cartes correspondent aux différentes sorties (équivalent aux canaux) produites par les filtres d’une couche de convolution.

Une couche de *dropout* est introduite dans chaque bloc résiduel, afin de fixer à zéro aléatoirement $100 \times p \%$ des valeurs des cartes caractéristiques à l’entraînement. Ce mécanisme permet de limiter le sur-apprentissage, en empêchant le modèle de « mémoriser » une solution [93]. La probabilité p est un hyper-paramètre que nous gardons fixé à $p = 0.1$ tel que dans l’architecture d’origine.

Les informations temporelles liées à l’entrée X sont intégrées à plusieurs reprises dans le UNet afin de pouvoir prendre en compte les différences saisonnières, et l’évolution dans le futur. Nous utilisons une représentation permettant de modéliser à la fois le cycle des jours dans le mois, des mois dans l’année, ainsi que la progression des années. Pour chaque date (m/j/a), nous calculons la variable :

$$t = \left(\frac{a - a_{\min}}{a_{\max} - a_{\min}} \right) \cos \left(\frac{2\pi j}{31} \right) \sin \left(\frac{2\pi m}{12} \right), \quad (5.1)$$

avec $j \in [1, 31]$, $m \in [1, 12]$ et $a \in [1955, 2099]$ respectivement le jour, le mois, et l’année du champ de précipitations. a_{\max} et a_{\min} correspondent aux années maximum et minimum de la période d’entraînement. La variable t est ensuite transformée par un perceptron multicouche (ou *Multilayer Perceptron* en anglais, MLP) comportant une couche cachée de dimension 128, et produisant deux sorties : un paramètre d’échelle $\gamma \in \mathbb{R}$, et un paramètre de phase $\beta \in \mathbb{R}$. Ces deux paramètres sont utilisés pour appliquer une transformation affine aux cartes de caractéristiques x tel qu’illustré sur la figure 5.1 :

$$y = \gamma x + (1 + \beta). \quad (5.2)$$

Finalement, les paramètres θ du modèle sont initialisés aléatoirement par échantillonnage d’une distribution uniforme de « Kaiming » [94] :

$$\theta_l \sim \mathcal{U}(-b, b) \quad \text{avec} \quad b = \sqrt{\frac{9}{n_{l-1}}}, \quad (5.3)$$

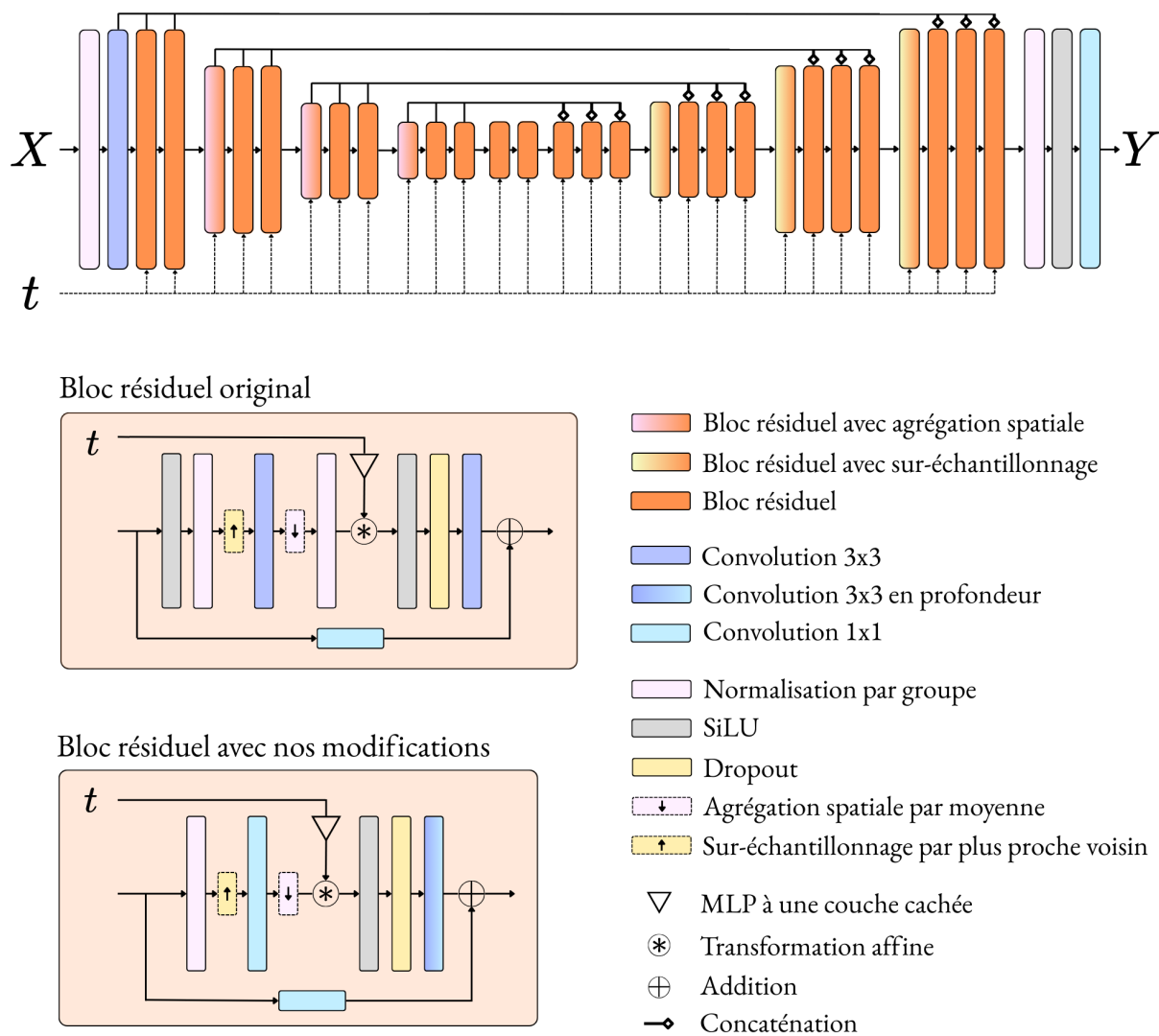


FIGURE 5.1 Présentation de l'architecture du modèle UNet utilisé, et des modifications effectuées sur le bloc résiduel original [2].

où n_{l-1} correspond au nombre de canaux en entrée de la couche l .

5.1.2 Stratégie, *pré* et *post* traitement des données

Dans le cadre du *Perfect Upscaling*, les champs de précipitations à basse résolution sont générés artificiellement par moyennage spatial des données journalières haute résolution issues de ClimEx (voir section 4). Afin d'évaluer de manière pratique les capacités de l'apprentissage profond pour la réduction d'échelle, nous adoptons un facteur d'agrégation de 16, sensiblement plus élevé que dans la majorité des travaux existants, tout en conservant un niveau d'information suffisant sur notre domaine restreint. Chaque champ basse résolution comprend ainsi 8×8 cellules, résultant de l'agrégation de blocs de 256 cellules haute résolution.

L'architecture du UNet nécessite que l'entrée et la sortie aient des dimensions spatiales similaires, afin d'assurer la continuité des sauts de connexions. Ainsi, il est nécessaire de sur-échantillonner les champs de précipitations à basse résolution avant de les fournir en entrée du modèle. Une autre approche pouvant être envisagée consiste à étendre le décodeur afin d'obtenir une sortie de dimensions plus élevée que l'entrée, mais nos expérimentations n'étant pas concluantes, nous avons décidé de nous concentrer sur la première approche [95]. Plusieurs travaux utilisent l'interpolation bilinéaire ou bicubique pour effectuer ce sur-échantillonnage. Nous avons fait le choix d'effectuer un sur-échantillonnage par plus proche voisin, afin de ne pas introduire de biais à cet étape, et ne remarquant pas d'impact sur les performances sur les expérimentations préliminaires.

Soient $x, y \in \mathbb{R}^{H \times W}$ deux champs de précipitations de haute résolution totalisant $H \times W$ cellules, respectivement obtenus par interpolation depuis la basse résolution, et issus directement des données haute résolution. Une première approche consiste à entraîner un UNet f_θ , paramétré par θ , pour prédire y à partir de x et de covariables éventuelles v , soit $f_\theta(x, v) = y$. Une alternative consiste à modéliser uniquement le résidu $r = y - x$, en apprenant $f_\theta(x, v) = r$. Cette stratégie permet de focaliser l'apprentissage sur les détails absents de l'interpolation, plutôt que sur la totalité du champ, dont certaines composantes ont déjà été perdues par l'encodeur. Nous retenons cette approche dans la suite, ayant observé des performances systématiquement supérieures, en accord avec des travaux antérieurs [2, 74].

Les champs journaliers de précipitations et de températures moyennes entre 1955 et 2020 sont utilisés pour constituer l'ensemble d'entraînement. Afin d'optimiser l'architecture du UNet, cet ensemble est scindé en un sous-ensemble d'entraînement couvrant la période 1955–2010, et un ensemble de validation allant de 2010 à 2020. L'ensemble de test, quant à lui, s'étend de 2020 à 2099, de manière à évaluer les performances du modèle dans un contexte de changements climatiques marqués, et pour analyser l'évolution de ses capacités à mesure que

ses effets s'intensifient.

L'optimisation des paramètres est effectuée à l'aide de l'algorithme **Adam** [96], avec une taille de mini-lot fixée à $N = 32$, sauf indication contraire. Le modèle UNet est entraîné en minimisant l'erreur quadratique moyenne (MSE) entre ses prédictions et les valeurs cibles. Soient $R \in \mathbb{R}^{N \times H \times W}$ les champs de résidus cibles, $X \in \mathbb{R}^{N \times H \times W}$ les champs de précipitations interpolés à partir des données basse résolution, $V \in \mathbb{R}^{N \times H \times W}$ les champs de covariables, et f_θ le UNet paramétré par θ . La fonction de perte utilisée pour l'entraînement est alors donnée par :

$$\mathcal{L}(\theta) = \frac{1}{N H W} \sum_{i=1}^N \|f_\theta(X_i, V_i) - R_i\|_2^2. \quad (5.4)$$

Pré-traitement

Comme l'illustrent les histogrammes de la figure 4.2, les précipitations sont fortement concentrées autour de zéro, avec une queue de distribution marquée. Cette structure motive l'application d'une transformation des valeurs de précipitations pour deux raisons principales : (1) élargir la plage des faibles précipitations, souvent sous-représentées, et (2) réduire la proportion de prédictions négatives, qui traduisent une incohérence physique et génèrent une surabondance de zéros après troncation. Pour répondre à ces objectifs, nous comparons deux transformations, notées T_1 et T_2 , conçues pour relâcher la contrainte de positivité sur l'espace transformé, et données par les équations suivantes [70] :

$$T_1(x, \epsilon) = \log(e^{x+\epsilon} - 1), \quad (5.5)$$

$$T_2(x, \epsilon) = \log(1 + x) - \log(\epsilon), \quad (5.6)$$

avec $\epsilon = 10^{-5}$ (correspondant à la valeur permettant d'obtenir le moins d'incohérence lors de nos expérimentations préliminaires). Les résultats pour la réduction d'échelle de précipitations présentés au tableau 5.1 montrent clairement leur efficacité pour réduire, tant la proportion de prédictions négatives, que l'amplitude de ces incohérences. Les deux transformations permettent également d'améliorer l'erreur moyenne absolue (ou *Mean Absolute Error* en anglais, MAE) de manière comparable.

Pour la suite nous conservons la transformation T_1 . Les données sont normalisées par la méthode **min/max** sur l'axe temporel. Soit $\tilde{X} \in \mathbb{R}^{N \times H \times W}$ une séquence de N champs de précipitations transformés (par T_1), le résultat de la normalisation est obtenu par : $\tilde{X}_{\text{norm}} = \frac{\tilde{X} - \min_n \tilde{X}_n}{\max_n \tilde{X}_n - \min_n \tilde{X}_n}$. Afin d'être conforme au cadre de la réduction d'échelle de simulations climatiques, nous utilisons les minima et maxima des champs de précipitations basse-résolution pour normaliser les champs haute-résolution : $\tilde{Y}_{\text{norm}} = \frac{\tilde{Y} - \min_n \tilde{X}_n}{\max_n \tilde{X}_n - \min_n \tilde{X}_n}$. Les résidus normalisés

TABLEAU 5.1 MAE (en mm/jour), proportions et la norme ℓ_2 des valeurs négatives, pour aucune transformation, T_1 ou T_2 . UNet entraîné sur la période 1960 – 2000, et évalué sur 2000 – 2010.

	MAE	Neg. %	ℓ_2 Neg.
Sans transformation	1.1236	12.88	228.73
Avec T_1	1.0866	2.20	18.3
Avec T_2	1.0856	9.47	27.92

sont donc égaux à :

$$R = \frac{\tilde{Y}_{\text{norm}} - \tilde{X}_{\text{norm}}}{\max_n \tilde{X}_n - \min_n \tilde{X}_n}. \quad (5.7)$$

Post-traitement

Une fois les opérations de prétraitement (résidus, normalisation, transformation) inversées, une étape supplémentaire de post-traitement est introduite. En effet, le modèle n’est soumis à aucune contrainte explicite garantissant la cohérence entre la valeur d’une cellule à basse résolution et la distribution des valeurs prédites dans les cellules haute résolution qui lui correspondent. Ce déséquilibre, bien que parfois discret à l’entraînement, peut devenir problématique hors distribution. Pour y remédier, nous imposons une contrainte sur les prédictions du UNet, après entraînement, assurant cette cohérence entre échelles [79] :

$$\hat{y}_i^* = \left(\frac{\sum_i^{N_{\text{grid}}} x_i}{\sum_i^{N_{\text{grid}}} \hat{y}_i} \right) \hat{y}_i, \quad (5.8)$$

où \hat{y}_i désigne une cellule de précipitations à haute résolution prédite par le UNet, x_i la valeur correspondante à basse résolution, et N_{grid} le nombre de cellules haute résolution agrégées dans chaque cellule basse résolution.

Le tableau 5.2 présente la MAE du UNet avec et sans application de la contrainte de post-traitement (5.8), ainsi que la MAE calculée au voisinage de certains niveaux de quantile (0.5, 0.9, 0.95 et 0.99). On observe qu’en dépit d’une légère hausse de l’erreur globale due principalement à une dégradation des performances sur les précipitations moyennes (*i.e.*, autour de la médiane), l’application de la contrainte permet une nette amélioration des prédictions sur les valeurs extrêmes (au-delà du 90ème quantile). Par ailleurs, l’histogramme des prédictions (voir figure 5.2) met en évidence un élargissement de la gamme des précipitations prédites, traduisant une meilleure représentation des fortes intensités.

TABLEAU 5.2 UNet entraîné sur la période 1955 – 2010 et évalué sur la période 2010 – 2020, sur 50 époques. Comparaison des performance avec ou sans la contrainte de post-traitement (5.8). MAE générale (en mm/jour), et calculée au voisinage de certains quantiles (0.5, 0.90, 0.95, et 0.99).

	MAE	MAE-q (0.5)	MAE-q (0.9)	MAE-q (0.95)	MAE-q (0.99)
sans contrainte	1.1754	0.7824	3.6363	6.3347	16.1048
avec contrainte	1.1928	0.8594	3.6044	6.0203	14.6546

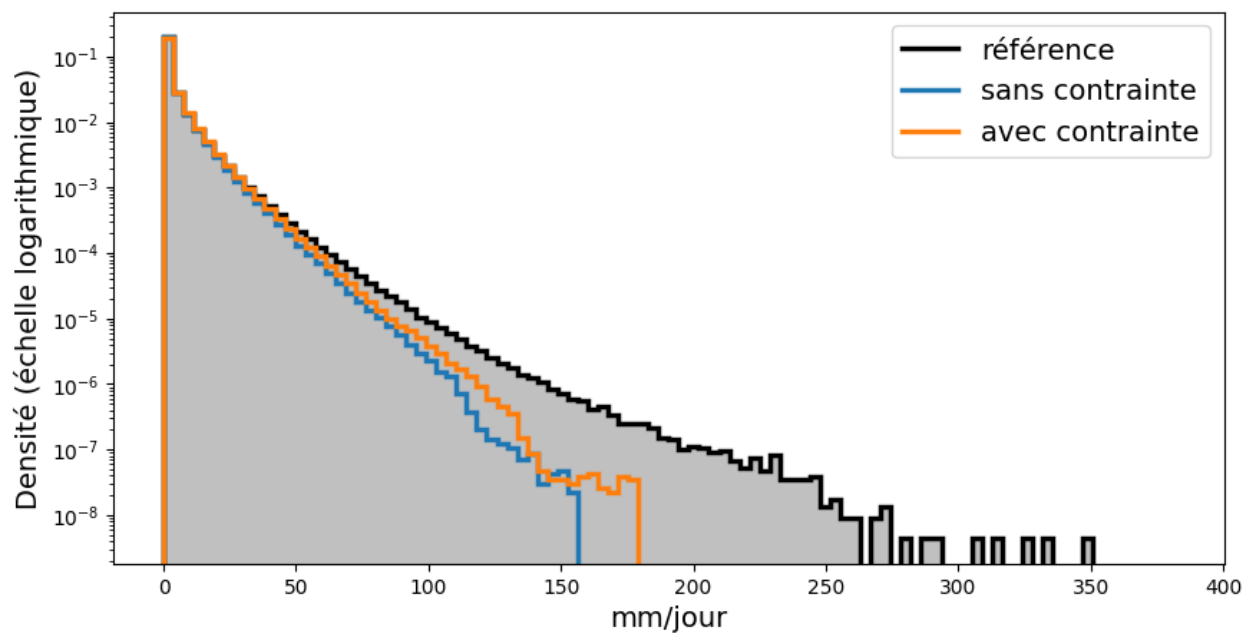


FIGURE 5.2 Histogrammes des précipitations prédites par le UNet avec ou sans la contrainte (5.8) sur la période d'évaluation (2010 – 2020).

5.2 Améliorations du UNet

Dans l’objectif d’améliorer les performances du UNet, nous intégrons plusieurs recommandations issues de la littérature, visant à renforcer les modèles convolutifs profonds [97]. En cohérence avec les auteur(rice)s de ces recommandations, nous effectuons les modifications séquentiellement, et conservons uniquement les modifications permettant d’améliorer la MAE. Il est important de souligner que notre objectif principal n’est pas ici d’identifier les choix architecturaux optimaux du UNet, ce qui impliquerait de comparer plusieurs combinaisons de modifications, mais plutôt de nous appuyer sur la littérature pour tenter d’améliorer ses performances.

La première modification consiste à ajuster le nombre de blocs résiduels par niveau, de manière à concentrer la capacité du modèle sur les niveaux pouvant être plus influents (voir Figure 5.1). Alors que l’encodeur original suit une séquence (3–3–3–3), nous expérimentons deux répartitions : (2–2–4–4), et (4–4–2–2). Le tableau 5.3 montre une amélioration marginale des performances, malgré une augmentation du nombre de paramètres pour une des séquences. **Pour la suite, nous conservons le nombre initial de blocs par niveaux.**

TABLEAU 5.3 MAE pour différentes séquences de blocs par niveau de l’encodeur (décodeur symétrique avec un bloc supplémentaire par niveau). UNet entraîné sur 50 époques avec 64 échantillons par mini-lot.

Séquence de blocs résiduels	MAE	Évol. rel. MAE	Nb. paramètres
(3 – 3 – 3 – 3)	1.1975	NA	5.709×10^6
(2 – 2 – 4 – 4)	1.1947	–0.23%	6.740×10^6
(4 – 4 – 2 – 2)	1.1953	–0.18%	4.679×10^6

Le nombre de canaux au premier niveau de l’encodeur est initialement fixé à 32 afin de respecter nos contraintes en terme de mémoire vive GPU ≈ 24 Gb (*Graphic Processor Unit*). Pour optimiser l’utilisation des paramètres, nous modifions les deux premières couches de convolution de chaque niveau. La première, chargée de faire varier le nombre de canaux (dans l’encodeur ou le décodeur), utilise des filtres de taille 1×1 au lieu de 3×3 , concentrant l’apprentissage sur la dimension des canaux uniquement. La seconde, qui conserve le nombre de canaux, est remplacée par une convolution groupée en profondeur, c’est-à-dire une convolution appliquée indépendamment à chaque canal, sans interaction entre eux. Ces ajustements réduisent significativement le nombre de paramètres, ce qui permet de doubler le nombre de canaux initiaux tout en maintenant un budget mémoire constant. Le tableau 5.4 met en évidence l’amélioration nette de l’erreur absolue moyenne (MAE) obtenue grâce à ces modifications.

Pour la suite, nous conservons cette modification.

TABLEAU 5.4 MAE pour différents types de convolutions et nombre de canaux initiaux. UNet entraîné sur 50 époques.

Types de convolution	Nb. canaux init.	MAE	Évol. rel. MAE	Nb. param
(3×3) , (3×3)	32	1.1928	NA	5.709×10^6
(1×1) , (3×3) en profondeur	64	1.1782	-1.22%	4.609×10^6

Grâce à la réduction du nombre de paramètres du modèle, il est suggéré qu’augmenter la taille des filtres des couches de convolutions (3×3) à (7×7) permettrait d’améliorer les performances en agrandissant le champ réceptif. Le tableau 5.5 indique une forte dégradation dans notre cas.

Pour la suite, nous conservons la taille de filtre (3×3) .

TABLEAU 5.5 MAE pour différentes tailles de filtres de convolution. UNet entraîné sur 50 époques.

Taille de filtre	MAE	Évol. rel. MAE	Nb. paramètres
(3×3)	1.1782	NA	4.609×10^6
(7×7)	1.1883	+0.86%	4.796×10^6

Les dernières modifications effectuées consistent à : (1) retirer la première couche SiLU du bloc résiduel, et (2) retirer la deuxième couche de normalisation du bloc résiduel. Ces recommandations sont inspirées de l’architecture des modèles de type *Transformers*. Le tableau 5.6 indique que ces deux modifications permettent de légèrement améliorer les performances du UNet. Notons également que l’erreur diminue pour tous les quantiles évalués, indiquant un réel gain de performances.

Pour la suite, nous retirons ces couches des bloc résiduels.

TABLEAU 5.6 MAE avec certaines couches du bloc résiduel retirées : (1) seulement SiLU, et (2) SiLU et normalisation par groupe (GN). UNet entraîné sur 50 époques.

Couches retirées	MAE	Évol. rel. MAE	Nb. paramètres
Aucune	1.1782	NA	4.609×10^6
SiLU	1.1741	-0.35%	4.609×10^6
SiLU + GN	1.1732	-0.42%	4.600×10^6

Les modifications apportées à l’architecture originale permettent de retrouver une MAE similaire à celle obtenue sans la contrainte de post-traitement (5.8), soit une réduction totale de 1.64% de la MAE, tout en conservant les améliorations pour les hauts quantiles de précipitations.

5.3 Évaluation en changements climatiques

Pour motiver la pertinence de l’objectif de ce projet, qui vise à renforcer la robustesse en changements climatiques des modèles convolutifs profonds pour la réduction d’échelle des précipitations, il est d’abord nécessaire de dresser le constat de la sensibilité du UNet aux changements climatiques.

Nous entraînons le UNet pendant 50 époques sur la période 1955–2020, soit un total de 23 725 champs de précipitations journalières. L’évaluation de ses performances est ensuite réalisée sur un jeu de test couvrant la période 2020 – 2099. Il convient de souligner que, sur cette période future, la température moyenne du domaine connaît une hausse significative, doublant entre les décennies 2020 – 2029 (5,37 °C) et 2090 – 2099 (11,04 °C). Cette période d’évaluation est donc représentative d’un climat fortement impacté par les changements climatiques.

La figure 5.3 présente (en haut) la MAE annuelle du UNet sur la période d’évaluation (en mm/jour), et (en bas) l’évolution relative à l’erreur moyenne sur 2020 – 2025 de la MAE évaluée aux voisinages de plusieurs quantiles. On remarque d’abord une nette augmentation de l’erreur alors que les changements climatiques s’intensifient, passant de ≈ 1.2 mm/jour entre 2020 et 2030 à plus de ≈ 1.4 mm/jour sur la dernière décennie, correspondant à une augmentation de 16.67%. Si on ajuste une relation linéaire aux erreurs annuelles, on obtient une tendance linéaire de l’erreur de 0.0031, c’est-à-dire + 0.15 mm/jour d’erreur moyenne supplémentaire tous les 50 ans. C’est significatif, car cet ordre de grandeur correspond à peu près au premier quartile des précipitations sur le domaine de test.

L’évaluation de la MAE dans le voisinage de différents quantiles permet d’identifier les intensités de précipitations qui contribuent le plus à la dégradation des performances du modèle. Le second graphique de la figure 5.3 montre que cette dégradation est d’autant plus marquée que le quantile considéré est élevé : l’évolution relative de la MAE atteint environ 25% pour le quantile de niveau $q = 0.99$, contre seulement 5 à 8% pour le quantile de niveau $q = 0.5$ en fin de période. Cette observation appuie l’hypothèse selon laquelle le UNet ne parvient pas à modéliser de manière satisfaisante la relation de dépendance entre l’augmentation des températures et l’intensification des précipitations extrêmes.

On sépare la période de test en deux sous-périodes : 2020 – 2059 et 2060 – 2099 couvrant chacune la moitié de la période de test. L’évolution relative de la MAE entre la première et la seconde sous-période, présentée à la figure 5.4a, est positive sur à peu près l’entièreté du domaine géographique. On peut observer une corrélation au nord-ouest du domaine, entre les augmentations marquées de la MAE et de la température moyenne (figure 5.4b). Notons que la partie du domaine recouverte par l’océan ne peut pas être analysée sous la même perspective,

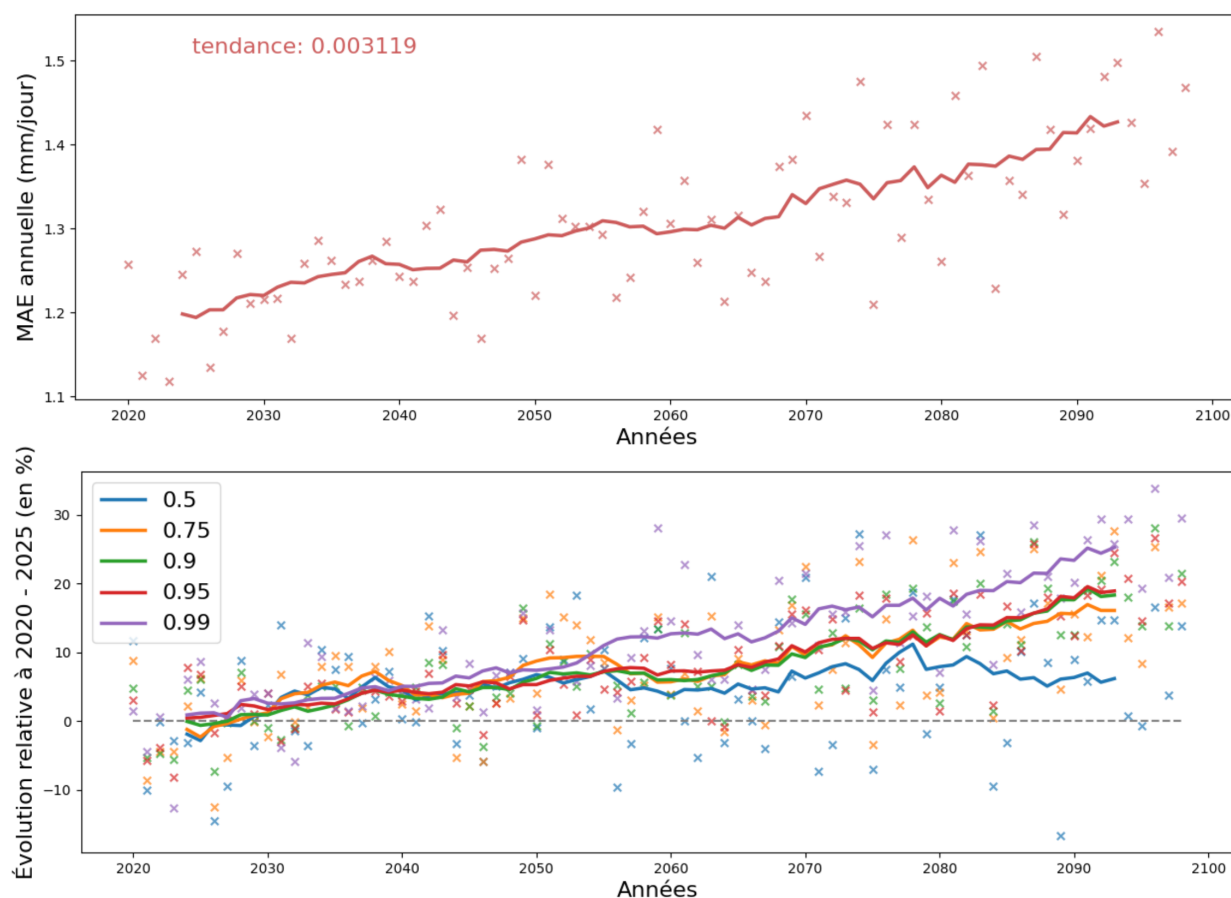
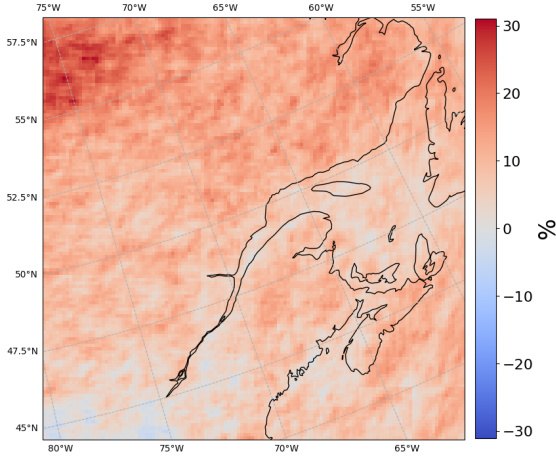
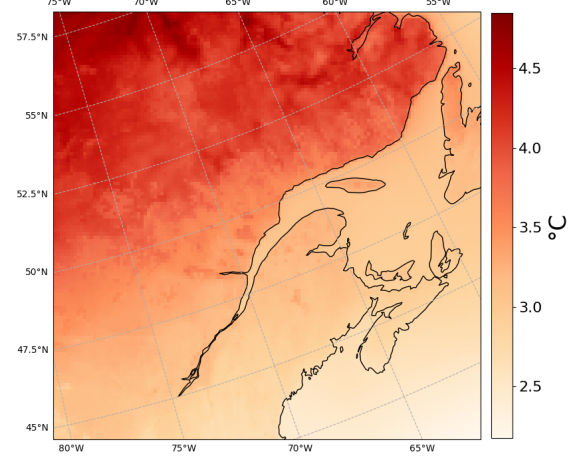


FIGURE 5.3 (**Haut**) MAE moyenne annuelle (en mm/jour) du UNet sur la période de test. (**Bas**) Évolution relative (en %) de la MAE annuelle, comparée à la période 2020–2025, calculée au voisinage des niveaux de quantile 0.5, 0.75, 0.90, 0.95 et 0.99. Courbes obtenues par moyenne mobile sur 10 valeurs.

car les précipitations sont soumises à des phénomènes différents (*e.g.*, dépressions océaniques, Gulf Stream). Ce constat renforce notre hypothèse selon laquelle le UNet n'intègre pas bien le mécanisme d'augmentation des précipitations sous élévation des températures.



(a) Évolution relative de la MAE.



(b) Évolution absolue de la température.

FIGURE 5.4 Comparaison de l'évolution de l'erreur absolue moyenne (MAE) des prédictions de précipitations haute résolution du UNet entre les périodes 2020 – 2059 et 2060 – 2099, par rapport l'évolution des températures sur ces mêmes périodes.

5.3.1 Estimation des TPSRs par la loi des valeurs extrêmes généralisée

Afin de mieux comprendre la relation liant les performances du UNet en changements climatiques à l'évolution des températures, nous appliquons la méthode décrite en section 2.3.3 pour estimer les taux de progression des maxima annuels de précipitations en fonction des anomalies de température (TPSRs), à la fois sur les données de référence et sur les prédictions issues du UNet. Nous adoptons la méthodologie suivante, détaillée en plusieurs étapes :

Nous sélectionnons les maxima annuels de précipitations par cellule sur la période d'estimation. Les séries de maxima sont ensuite normalisées (division par la médiane sur la période). Parallèlement, les anomalies saisonnières de température, calculées pour la période de mai à septembre de chaque année, sont extraites afin d'analyser la sensibilité des précipitations extrêmes à l'évolution des températures [25]. Pour chaque cellule, une loi GEV est ajustée aux maxima de précipitations, avec des paramètres qui sont fonction des anomalies de température.

Afin de réduire les incertitudes liées à l'échantillonnage lors de l'estimation des paramètres de la loi GEV, nous appliquons un regroupement spatial consistant à agréger les données des

cellules voisines disposées en grille 3×3 autour de chaque cellule analysée [26]. Les paramètres de chaque cellule seront donc estimés en utilisant les données de ses neuf cellules voisines. Pour éviter d'obtenir des paramètres et des taux incohérents, nous ne considérons pas les cellules sur les bords du domaine.

Pour déterminer la forme fonctionnelle de la dépendance des paramètres de la loi GEV aux anomalies de température T , nous considérons les modèles suivants :

— **Paramètre de localisation** $\mu(T)$: trois modèles :

$$\begin{aligned} M_{\mu,1} : \mu(T) &= \mu_0, \\ M_{\mu,2} : \mu(T) &= \mu_0 + \mu_1 T, \\ M_{\mu,3} : \mu(T) &= \exp(\mu_0 + \mu_1 T). \end{aligned}$$

— **Paramètre d'échelle** $\sigma(T)$: deux modèles :

$$\begin{aligned} M_{\sigma,1} : \sigma(T) &= \sigma_0, \\ M_{\sigma,2} : \sigma(T) &= \exp(\sigma_0 + \sigma_1 T). \end{aligned}$$

— **Paramètre de forme** : indépendant de T car complexe à estimer :

$$\xi(T) = \xi.$$

Les paramètres sont estimés par maximisation de la log-vraisemblance, en utilisant l'algorithme d'optimisation **L-BFGS-B** (ou en anglais *Limited-memory BFGS for Bounds constraints*) pour contrôler les limites du paramètres de forme : $\xi \in (-0.5, 0.5)$, et assurer la positivité du paramètre d'échelle : $\sigma(T) > 0$ [98]. Soient y_1, \dots, y_n les maxima de précipitations normalisés pour une cellule donnée, t_1, \dots, t_n les anomalies de températures, et les paramètres $\mu(T)$, $\sigma(T)$ et ξ , la log-vraisemblance de la GEV est donnée par :

$$\ell(\mu, \sigma, \xi) = \begin{cases} -\sum_{i=1}^n \log \sigma(t_i) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \log \left(1 + \xi \frac{y_i - \mu(t_i)}{\sigma(t_i)}\right) \\ -\sum_{i=1}^n \left(1 + \xi \frac{y_i - \mu(t_i)}{\sigma(t_i)}\right)^{-1/\xi} & \text{si } \xi \neq 0, \\ -\sum_{i=1}^n \log \sigma(t_i) - \sum_{i=1}^n \frac{y_i - \mu(t_i)}{\sigma(t_i)} - \sum_{i=1}^n \exp\left(\frac{y_i - \mu(t_i)}{\sigma(t_i)}\right) & \text{si } \xi = 0. \end{cases} \quad (5.9)$$

définie sur le support $\forall i : 1 + \xi \frac{y_i - \mu(t_i)}{\sigma(t_i)} > 0$. On utilise la librairie **JAX** pour calculer automatiquement la jacobienne de $\ell(\mu, \sigma, \xi)$ et paralléliser l'optimisation des cellules. Le tableau 5.7 indique la log-vraisemblance négative (NLL) moyenne pour chaque combinaison des modèles de $\mu(T)$ et $\sigma(T)$. Les résultats confirment la pertinence de l'intégration des anomalies

de températures, la combinaison $M_0 + M_{\sigma,0}$ ayant la plus forte NLL. Nous conservons la combinaison $M_{\mu,3} + M_{\sigma,2}$.

TABLEAU 5.7 Log-vraisemblance négative moyenne des lois de GEV du domaine sur la période 2020 – 2099 pour toutes les combinaisons de modèles des paramètres $\mu(T)$ et $\sigma(T)$.

Combinaison de modèles	NLL moyenne
$M_{\mu,1} + M_{\sigma,1}$	289.91
$M_{\mu,2} + M_{\sigma,1}$	281.67
$M_{\mu,3} + M_{\sigma,1}$	281.56
$M_{\mu,1} + M_{\sigma,2}$	285.99
$M_{\mu,2} + M_{\sigma,2}$	275.67
$M_{\mu,3} + M_{\sigma,2}$	275.53

Une fois les paramètres estimés pour chaque cellule, nous calculons les TPSRs à partir de la loi GEV ajustée. Pour ce faire, on calcule la moyenne des anomalies de températures $T^{(1)} = (1/n) \sum_{i=1}^n t_i$, et $T^{(2)} = T_1 + 1$, et les quantiles de niveaux q associés, $z_q^{(1)}$ et $z_q^{(2)}$, par la fonction de quantile de la GEV :

$$z_q(T) = \begin{cases} \mu(T) + \frac{\sigma(T)}{\xi} [(-\log q)^{-\xi} - 1] & \text{si } \xi \neq 0, \\ \mu(T) - \sigma(T) \log(-\log q) & \text{si } \xi = 0. \end{cases} \quad (5.10)$$

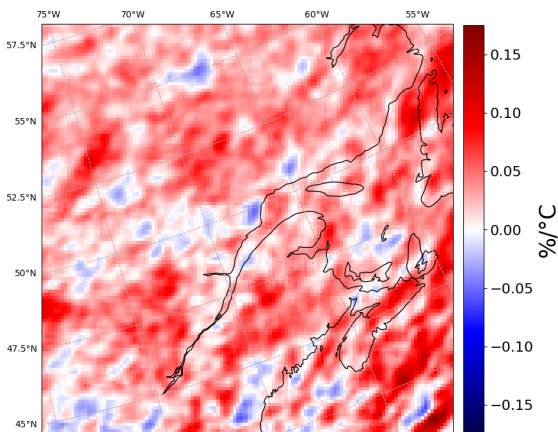
Finalement, on calcule pour chaque cellule le TPSR α (en $\%/^\circ C$) par la formule :

$$z_q^{(2)} = z_q^{(1)} (1 + 0.01\alpha)^{T^{(2)} - T^{(1)}}. \quad (5.11)$$

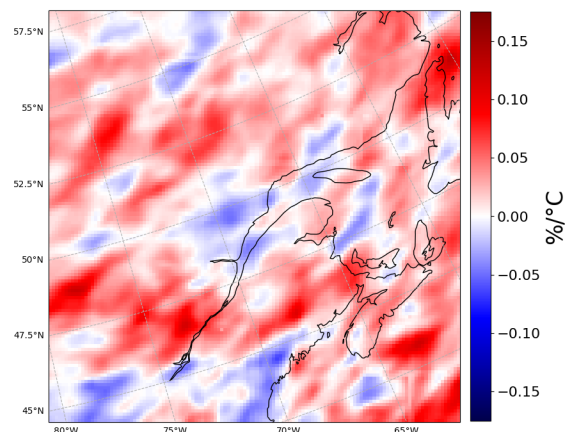
Les log-vraisemblances négatives par cellule de la combinaison $M_{\mu,4} + M_{\sigma,2}$, pour les données de référence, et les prédictions du UNet, sont disponibles à l'annexe A. Les valeurs par cellule de chaque paramètre, pour les données de référence et les prédictions du UNet sont également incluses.

Résultats

La figure 5.5 présente les TPSRs évalués au niveau de quantile $q = 0.99$ pour les données de référence (figure 5.5a) et les prédictions du UNet (figure 5.5b) pour la période de test 2020 – 2099. Il est important de noter que l'évaluation des TPSRs comporte de fortes incertitudes entre membres de simulations climatiques. Le taux moyen de $3.2\%/^\circ C$ observé sur les données de référence ici appartient à l'intervalle des taux observés par d'autres travaux à travers tous les membres de ClimEx [25].



(a) TPSRs au 99ème quantile des données de référence sur la période de test 2020 – 2099.



(b) TPSRs au 99ème quantile des prédictions du UNet sur la période de test.

FIGURE 5.5 Comparaison des TPSRs pour le niveau de quantile $q = 0.99$ estimés sur la période de test 2020 – 2099 à partir des données de référence, et des prédictions du UNet.

On observe que la proportion de taux négatifs semble plus élevée lorsque estimés sur les prédictions du UNet, et de plus forte intensité. Cette observation est partagée pour les taux de tous les quantiles estimés, comme indiqués par le tableau 5.8, et la figure 5.6, présentant respectivement la moyenne, et les quartiles des taux de chaque quantile. Néanmoins, on note tout de même que la distribution spatiale des TPSRs est à peu près similaire entre les données de référence et les prédictions du UNet.

Les taux estimés à partir des prédictions du UNet étant systématiquement inférieurs à ceux obtenus sur les données de référence, cela suggère que les précipitations extrêmes prédites progressent moins rapidement avec l’augmentation des températures que dans la réalité. On en déduit que les difficultés du UNet à représenter correctement les extrêmes sont amplifiées dans un contexte de changements climatiques.

TABLEAU 5.8 Moyennes des TPSRs estimés à plusieurs niveaux de quantiles pour les données de référence, et les prédictions du UNet, pour la période de test 2020 – 2099.

Quantiles	Référence	UNet
0.5	2.24%	1.49%
0.75	2.61%	1.69%
0.90	2.87%	1.83%
0.95	3.00%	1.91%
0.99	3.20%	1.92%

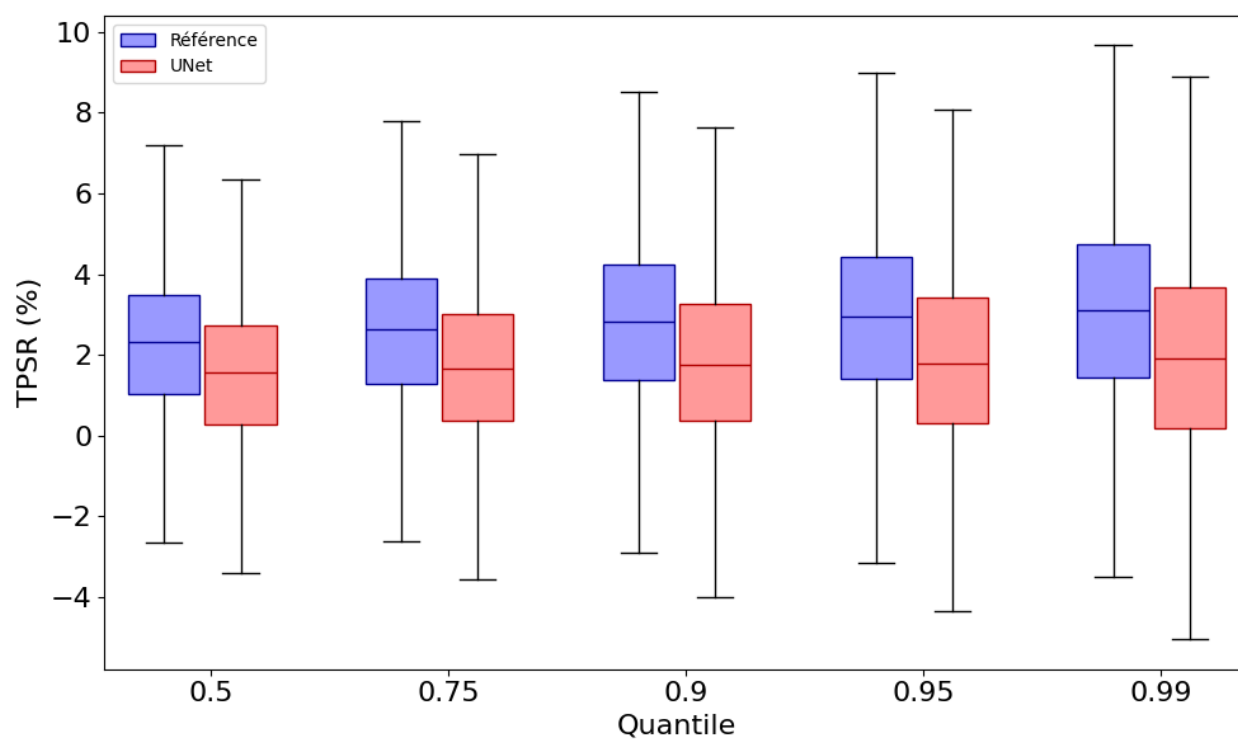


FIGURE 5.6 Boîtes à moustaches des TPSRs estimés à plusieurs quantiles pour les données de référence, et les prédictions du UNet, pour la période de test 2020 – 2099.

CHAPITRE 6 INTÉGRATION DE LA RELATION TEMPÉRATURE-PRÉCIPITATIONS POUR LA RÉDUCTION D'ÉCHELLE

Au chapitre précédent, nous avons formulé l'hypothèse que la dégradation des performances du UNet en changements climatiques résulte en partie de son incapacité à reproduire correctement les taux d'augmentation des précipitations extrêmes en fonction de l'évolution des températures. Dans ce chapitre, nous explorons plusieurs approches visant à mieux intégrer cette relation lors de l'entraînement du modèle d'apprentissage profond.

6.1 Contrainte douce de Clausius-Clapeyron par approche de regroupement

Comme indiqué à la section 2.3.2, deux méthodes permettent d'évaluer si les extrêmes de précipitations sont cohérents avec la dépendance à la température décrite par la relation de Clausius-Clapeyron. Dans cette section, nous nous concentrons sur la première approche, dite « par regroupement », et présentons une tentative d'intégration de cette méthode au sein de l'apprentissage du modèle UNet, dans le but de mieux intégrer la relation entre température et précipitations.

6.1.1 Intégration de la relation précipitations-températures par contrainte douce

Contrairement à la pression de vapeur saturante, dont la dépendance à la température est exactement décrite par la relation de Clausius-Clapeyron, les précipitations extrêmes n'obéissent pas strictement à cette loi. La relation Clausius-Clapeyron pour les précipitations constitue donc une approximation statistique observée dans certaines conditions, et non une contrainte déterministe. Ainsi, il n'est pas pertinent d'imposer au UNet une contrainte physique exacte, ou « dure ». Néanmoins, afin de mieux intégrer la relation de dépendance existante entre les extrêmes de précipitations et la température, on peut chercher à « guider » le UNet vers un équilibre, *via* une pénalité sur sa fonction de perte, c'est-à-dire une contrainte « douce ».

On note $\mathcal{C} : Y \in \mathbb{R}^{N \times 2 \times H \times W} \rightarrow c \in \mathbb{R}^{H \times W}$ la fonction qui permet de calculer un champ de taux c (en $\%/^{\circ}\text{C}$) d'évolution des extrêmes de précipitations en fonction de la température pour une séquence de champs de précipitations et de températures journalières Y . Dans le cadre de l'entraînement du UNet par mini-lots, on considère un mini-lot de N champs de précipitations interpolés à partir de données de basse-résolution $x \in \mathbb{R}^{N \times h \times w}$, les champs de résidus cibles à haute résolution $r \in \mathbb{R}^{N \times H \times W}$, les champs de températures à haute résolution

$T \in \mathbb{R}^{N \times H \times W}$, le UNet f_θ paramétré par θ , et $\gamma \geq 0$ la pondération de la pénalité. La fonction de perte avec contrainte douce par approche de regroupement est donnée par :

$$\mathcal{L}_\theta(x, T, r) = \mathcal{L}_{\text{reco}}(x, T, r, \theta) + \gamma \mathcal{L}_{CC}(x, T, r, \theta), \quad (6.1)$$

$$\text{avec } \mathcal{L}_{\text{reco}}(x, T, r, \theta) = \frac{1}{N H W} \sum_{n=1}^N \sum_{i=1}^H \sum_{j=1}^W (f_\theta(x_n, T_n)_{i,j} - y_{n,i,j})^2, \quad (6.2)$$

$$\text{et } \mathcal{L}_{CC}(x, T, r, \theta) = \frac{1}{H W} \sum_{i=1}^H \sum_{j=1}^W \{\mathcal{C}(f_\theta(x, T)_{.,i,j} + r_{.,i,j}, T_{.,i,j}) - \mathcal{C}(x_{.,i,j} + r_{.,i,j}, T_{.,i,j})\}^2. \quad (6.3)$$

La fonction \mathcal{C} regroupe les différentes étapes de l'approche dite « par regroupement ». Étant donné une suite de N précipitations journalières $y_{1:N}$ (en mm/jour) et températures $t_{1:N}$ (en °C) :

1. **Filtrage des très faibles intensités** : Conserver uniquement les journées dont les précipitations dépassent un seuil s :

$$\mathcal{S} = \{n^* \in \{1, \dots, N\} \mid y_{n^*} > s\}. \quad (6.4)$$

2. **Tri et regroupement** : Ordonner les indices $n^* \in \mathcal{S}$ par ordre croissant des températures t_l , puis diviser la liste obtenue en D intervalles $\{G_1, \dots, G_D\}$ de capacité constante m :

$$G_d = \{n_d^{*(1)}, \dots, n_d^{*(m)}\}, \quad \text{où } m \approx \frac{|\mathcal{S}|}{D}. \quad (6.5)$$

3. **Calcul des statistiques par groupe** : Pour chaque groupe $d \in \{1, \dots, D\}$:

$$\bar{t}_d = \frac{1}{m} \sum_{n^* \in G_d} t_{n^*}, \quad y_d^q = \text{quantile}_q(\{y_{n^*} : n^* \in G_d\}). \quad (6.6)$$

4. **Régression pour estimer la dépendance** : On suppose une relation log-linéaire :

$$\log y_d^q = \alpha + \beta \bar{t}_d + \varepsilon_d. \quad (6.7)$$

L'estimateur des moindres carrés pour β est donné par :

$$\hat{\beta} = (T_D^T T_D)^{-1} T_D^T Q_y, \quad (6.8)$$

où

$$T_D = [\bar{t}_1, \dots, \bar{t}_D], \quad Q_y = [\log y_1^q, \dots, \log y_D^q]. \quad (6.9)$$

5. **Calcul du taux :** Finalement, on calcule le taux c par la formule suivante :

$$c = \exp(\hat{\beta}) - 1. \quad (6.10)$$

N’ayant pas observé de forte sensibilité au nombre D d’intervalles, nous fixons $D = 10$ pour obtenir des capacités d’intervalles suffisamment élevées. N’ayant pas non plus observés de forte sensibilité aux niveaux de quantile parmi $q \in \{0.90, 0.95, 0.99\}$, nous fixons $q = 0.90$ afin de réduire les erreurs d’estimation lorsque m , le nombre de données par intervalle, est petit. Afin de s’assurer que chaque mini-lot de prédictions du UNet comporte assez de valeurs supérieurs au seuil s pour remplir $D = 10$ intervalles, on entraîne d’abord le modèle sur $\mathcal{L}_{\text{reco}}$ pendant deux époques. Les résultats et l’impact des hyper-paramètres s et γ sont présentés à la section 7.2.1.

6.1.2 Limites de l’approche par regroupement

La fonction \mathcal{C} , utilisée pour estimer la dépendance entre températures et précipitations extrêmes, est définie à partir d’un regroupement nécessitant un nombre suffisant d’observations pour produire des taux stables. Dans notre contexte, l’entrée de \mathcal{C} est constituée des couples (y_n, t_n) appartenant au mini-lot d’apprentissage de N couples $\{(y_n, t_n)\}_{n=1}^N$.

Or, pour obtenir une estimation robuste des taux issus de cette procédure, la taille de mini-lot N doit être relativement grande. Cette contrainte implique des mini-lots volumineux, augmentant significativement la mémoire vive requise lors de l’entraînement du réseau et limitant l’application de la méthode pour des réseaux de neurones comme le UNet.

Pour illustrer cette limite, nous avons restreint le domaine spatial à 32×32 cellules afin de libérer de la mémoire et d’augmenter la taille des mini-lots d’apprentissage. La figure 6.1 présente l’évolution de \mathcal{L}_{CC} en fonction de la taille de mini-lots, et montre clairement la nécessité de pouvoir fixer $N > 512$ pour que les erreurs convergent vers zéro. Malheureusement, les ressources à notre disposition (déjà significatives) ne nous permettent d’augmenter la taille des mini-lots que jusqu’à $N = 150$ pour le domaine complet.

Une piste non explorée dans ce mémoire consisterait à entraîner le modèle UNet sur des « patches » spatiaux superposés du domaine. Cette approche permettrait de réduire le poids des dimensions spatiales des mini-lots, et donc d’en augmenter le nombre d’échantillons [99].

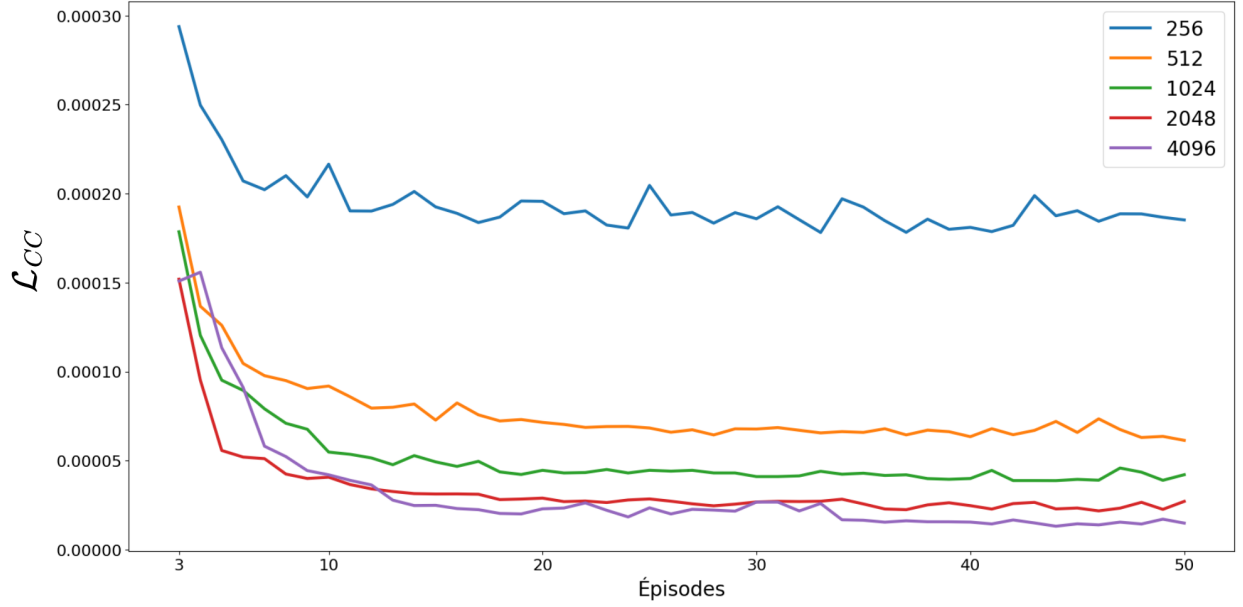


FIGURE 6.1 \mathcal{L}_{CC} pour la période 1955 – 2020 sur un domaine restreint en fonction de tailles croissantes de mini-lots (256, 512, 1024, 2048, et 4096).

6.2 Intégration d'*a priori* statistiques issus de la théorie des valeurs extrêmes

Tel qu'expliqué à la section 2.3.3, la méthode par approche de regroupement comporte l'inconvénient de ne pas pouvoir réellement mesurer le lien causal entre évolution des températures et extrêmes de précipitations, les variations saisonnières introduisant un facteur de confusion important [24].

Basées sur la théorie des valeurs extrêmes, les approches alternatives modélisent les précipitations extrêmes en capturant la non-stationnarité *via* des paramètres dépendants des anomalies annuelles ou saisonnières de température. Dans cette section, nous étudions comment exploiter ces distributions comme *a priori* statistiques afin de renforcer la robustesse du modèle d'apprentissage profond dans un contexte de changements climatiques.

6.2.1 Modélisation des précipitations par extension de la loi de Pareto généralisée

L'extension de la loi de Pareto généralisée (ou *Extended Generalized Pareto Distribution* en anglais, ExtGPD), présentée à la section 2.2.3, permet de modéliser la totalité de la gamme des précipitations (à l'exclusion des précipitations nulles), et de conserver les propriétés asymptotiques de la théorie des valeurs extrêmes [1]. Notre objectif consiste à obtenir

une distribution (dépendante des anomalies de température) précise des précipitations pour chaque cellule de notre domaine, en utilisant nos données d'entraînement, pour ensuite utiliser cette distribution lors de la réduction d'échelle par apprentissage profond.

La fonction de répartition d'une variable aléatoire suivant une loi ExtGPD est donnée par :

$$F(x) = K \left\{ H_\xi \left(\frac{x}{\sigma} \right) \right\}, \quad (6.11)$$

avec K un modèle paramétrique soumis à certaines conditions, H_ξ la fonction de répartition de la loi de Pareto généralisée (voir eq. 2.8) avec $\xi \geq 0$, et $\sigma > 0$ le paramètre d'échelle. Nous utilisons le modèle proposé par les auteurs $K(v) = v^\kappa$ avec $\kappa > 0$. Par conséquent, la fonction de densité s'écrit :

$$f(x) = \frac{\partial}{\partial x} F(x) = \frac{\partial}{\partial x} H_\xi \left(\frac{x}{\sigma} \right)^\kappa = \kappa H_\xi \left(\frac{x}{\sigma} \right)^{\kappa-1} h_\xi \left(\frac{x}{\sigma} \right) \sigma^{-1}, \quad (6.12)$$

avec H_ξ et h_ξ respectivement les fonctions de répartition et de densité de la loi de Pareto généralisée :

$$H_\xi(z) = \begin{cases} 1 - (1 + \xi z)^{-1/\xi} & \text{si } \xi \neq 0, \\ 1 - \exp(-z) & \text{si } \xi = 0, \end{cases} \quad (6.13)$$

$$h_\xi(z) = \begin{cases} (1 + \xi z)^{-(\xi+1)/\xi} & \text{si } \xi \neq 0, \\ \exp(-z) & \text{si } \xi = 0. \end{cases} \quad (6.14)$$

Nous devons estimer trois paramètres pour chaque cellule du domaine : $\kappa > 0$, $\sigma > 0$, et $\xi \geq 0$ pour ajuster l'ExtGPD aux données de précipitations. Nous effectuons l'estimation des paramètres par maximisation de la log-vraisemblance censurée à gauche, tel que suggéré par les auteurs. L'effet de bruine (ou *drizzle effect* en anglais), responsable d'une surreprésentation des faibles valeurs de précipitations, est pris en compte par l'application d'un seuil de 1 mm/jour [100]. Étant donné notre choix $K(v) = v^\kappa$, F la fonction de répartition de la loi ExtGPD, et f la fonction de densité, la log-vraisemblance des paramètres de l'ExtGPD, pour le seuil $C = 1\text{mm/jour}$, et $y_1, \dots, y_n \stackrel{\text{i.i.d.}}{\sim} F$ une série de précipitations non-nulles, s'écrit :

$$\ell_C(\kappa, \sigma, \xi) = \log \left\{ \prod_{i:y_i < C} F(C) \prod_{i:y_i \geq C} f(y_i) \right\} \quad (6.15)$$

$$= \kappa \sum_{i:y_i < C} \log \left\{ H_\xi \left(\frac{C}{\sigma} \right) \right\} + \sum_{i:y_i \geq C} \log \kappa + (\kappa - 1) \sum_{i:y_i \geq C} \log \left\{ H_\xi \left(\frac{y_i}{\sigma} \right) \right\} \quad (6.16)$$

$$+ \sum_{i:y_i \geq C} \log \left\{ h_\xi \left(\frac{y_i}{\sigma} \right) \right\} - \sum_{i:y_i \geq C} \log \sigma. \quad (6.17)$$

Ainsi, pour $\xi = 0$, on obtient :

$$\ell_C(\kappa, \sigma, \xi = 0) = \kappa \sum_{i:y_i < C} \log \left\{ 1 - \exp \left(-\frac{C}{\sigma} \right) \right\} + (\kappa - 1) \sum_{i:y_i \geq C} \log \left\{ 1 - \exp \left(-\frac{y_i}{\sigma} \right) \right\} \quad (6.18)$$

$$+ \sum_{i:y_i \geq C} \log \kappa - \sum_{i:y_i \geq C} \frac{y_i}{\sigma} - \sum_{i:y_i \geq C} \log \sigma, \quad (6.19)$$

et pour $\xi > 0$:

$$\ell_C(\kappa, \sigma, \xi > 0) = \kappa \sum_{i:y_i < C} \log \left\{ 1 - \left(1 + \xi \frac{C}{\sigma} \right)^{-1/\xi} \right\} + (\kappa - 1) \sum_{i:y_i \geq C} \log \left\{ 1 - \left(1 + \xi \frac{y_i}{\sigma} \right)^{-1/\xi} \right\} \quad (6.20)$$

$$+ \sum_{i:y_i \geq C} \log \kappa - (\xi + 1)/\xi \sum_{i:y_i \geq C} \log \left\{ 1 + \xi \frac{y_i}{\sigma} \right\} - \sum_{i:y_i \geq C} \log \sigma. \quad (6.21)$$

Pour estimer les paramètres de la loi en maximisant la log-vraisemblance définie par l'équation (6.17), nous utilisons l'algorithme d'optimisation **L-BFGS**. Afin de garantir le respect des contraintes sur les paramètres, nous appliquons des transformations appropriées de ces derniers. Cette approche s'est révélée plus stable que l'utilisation directe de l'algorithme **L-BFGS-B**, qui impose des bornes mais a montré des instabilités pour l'ajustement de cette loi.

Dans le but d'inclure la non-stationnarité liée à l'évolution des températures, nous choisissons de faire dépendre les paramètres σ et κ aux anomalies annuelles de température. Les paramètres κ, σ et ξ , pour les anomalies de température T , sont définis lors de l'optimisation par les transformations suivantes :

$$\kappa(T) = \exp(\alpha_0 + \alpha_1 T) \quad (6.22)$$

$$\sigma(T) = \exp(\beta_0 + \beta_1 T) \quad (6.23)$$

$$\xi(T) = 0.5 / \{1 + \exp(-\nu)\} \quad (6.24)$$

avec $\alpha_0, \alpha_1, \beta_0, \beta_1, \nu \in \mathbb{R}$. Afin d'éviter que les paramètres κ et σ ne divergent vers des valeurs irréalistes lors de l'optimisation, nous ajoutons un terme de pénalisation de 10^6 à la log-vraisemblance négative pour chaque cas où $\kappa(T) > 2$, ou $\sigma(T) > 30$. Ces seuils ont été définis empiriquement à partir des bornes maximales observées lors d'estimations précédentes avec contraintes explicites. Pour rappel, nous fixons $C = 1$ mm/jour.

La log-vraisemblance négative moyenne obtenue dans le cas stationnaire ($\alpha_1 = \beta_1 = 0$) est de 43 784.86, contre 43 778.69 dans le cas non-stationnaire. La figure 6.2 présente la distribution

spatiale de la log-vraisemblance négative dans le cas non-stationnaire. On observe des valeurs élevées dans les régions à forte orographie : on distingue notamment les Laurentides, les monts Chics-Chocs, le plateau du Cap-Breton et les monts Long-Range. La NLL est également élevée dans le sud-est du domaine, en zone océanique, caractérisée par de fortes précipitations comme le montre la figure 4.1(c).

Les figures 6.3 et 6.4 présentent la distribution spatiale des paramètres obtenus dans le cas non-stationnaire. On observe à la figure 6.3a que le paramètre stationnaire $\kappa(T)$ est plus élevé sur la zone continentale, en particulier dans les régions montagneuses. Cela suggère des distributions plus étalées, avec une densité réduite autour des faibles intensités de précipitations. La figure 6.3b montre quant à elle que le paramètre stationnaire $\sigma(T)$ distingue nettement les zones continentales et océaniques, cette dernière étant associée à des valeurs plus élevées du paramètre d'échelle.

Concernant les composantes non stationnaires, représentées aux figures 6.3c et 6.3d, on observe une dynamique opposée sur les zones maritimes, avec une augmentation du paramètre d'échelle $\sigma(T)$ et une diminution de $\kappa(T)$. Par ailleurs, le coefficient β_1 est globalement positif sur l'ensemble du domaine, indiquant une réponse croissante des précipitations extrêmes avec l'augmentation des températures. Enfin, la figure 6.4 révèle que le paramètre de forme ξ est plus élevé sur les zones continentales, traduisant des queues de distribution plus lourdes.

Quatre cellules, représentatives de la diversité du domaine géographique, sont identifiées par les étoiles 1 à 4 de la figure 6.2. Afin de s'assurer de la viabilité de la loi ExtGPD pour décrire les précipitations journalières, et évaluer la pertinence de la prise en compte des températures, nous calculons les diagrammes quantile-quantile de chacune de ces cellules. La figure 6.5 montre les diagrammes quantile-quantile pour chaque cellule dans le cas stationnaire ($\alpha_1 = \beta_1 = 0$), ainsi que les histogrammes des quantiles obtenus. La figure 6.6 présente les mêmes éléments dans le cas non-stationnaire. Afin d'évaluer la pertinence de la prise en compte de l'évolution des températures, les comparaisons sont effectuées sur les données de la période 2020 – 2099, tandis que les paramètres ont été estimés sur la période historique 1955 – 2020.

Ces diagrammes quantile-quantile, ainsi que les histogrammes correspondants, illustrent la pertinence de la prise en compte de l'évolution des anomalies de températures. En effet, dans le cas stationnaire ($\alpha_1 = \beta_1 = 0$), présenté à la figure 6.5, on observe une sous-estimation systématique des quantiles élevés (sur la période 2020–2099) par la loi ajustée sur la période 1955–2020. À l'inverse, la figure 6.6 montre des diagrammes quantile-quantile quasi parfaits et des histogrammes très bien alignés. Ainsi, l'inclusion des anomalies de températures permet d'adapter plus fidèlement la distribution des précipitations non nulles dans un contexte de

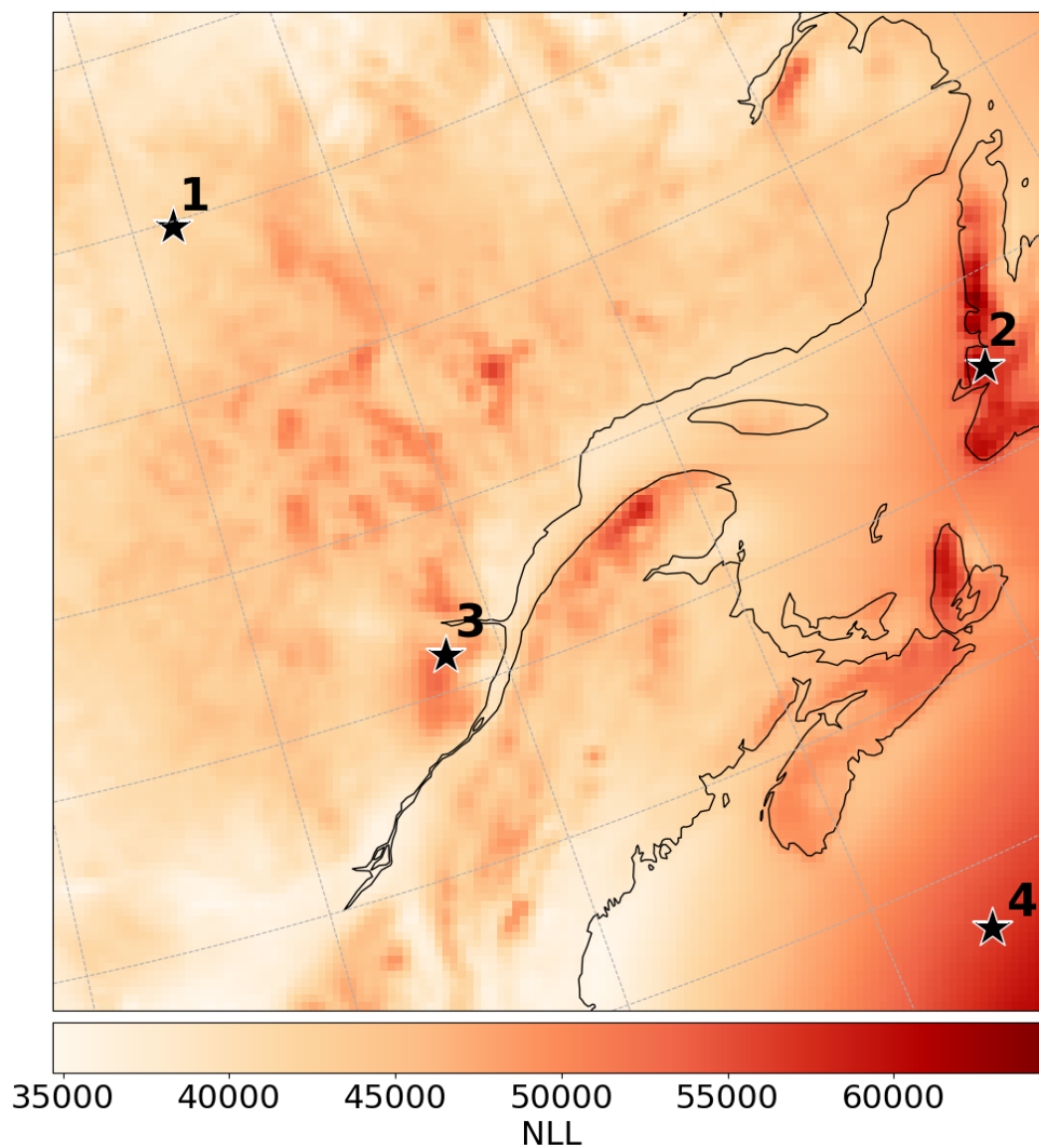
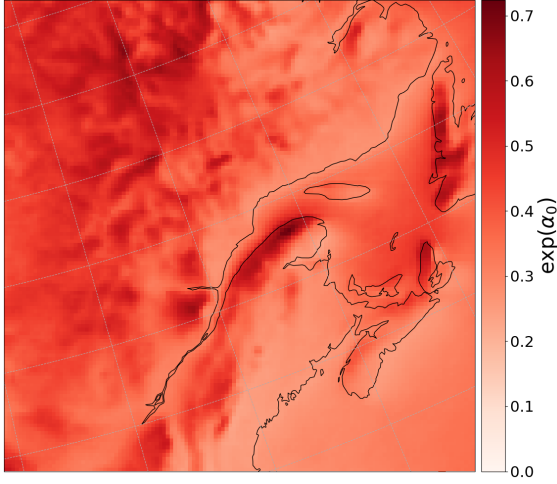
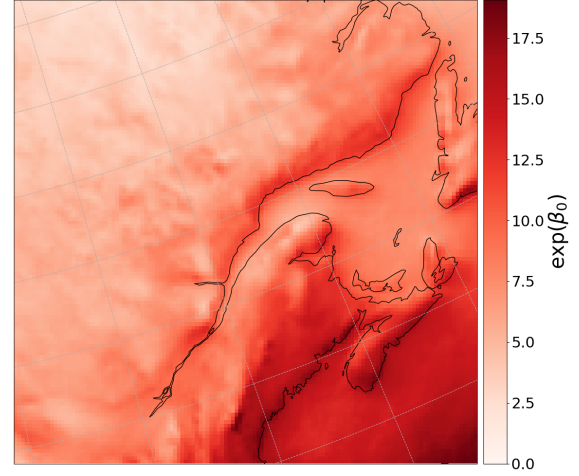


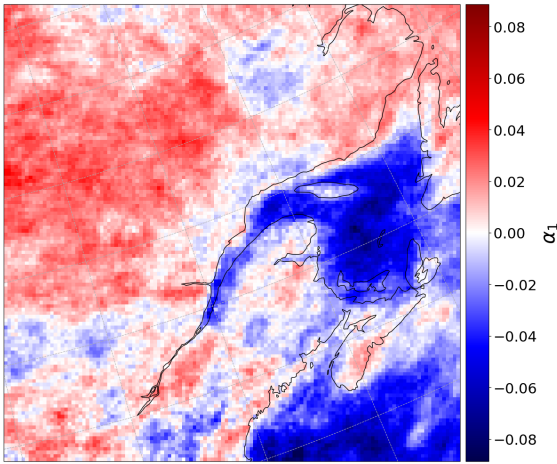
FIGURE 6.2 Log-vraisemblance négative obtenue après optimisation des paramètres définis en (6.22-6.24) de la loi ExtGPD, ajustée indépendamment sur chaque cellule du domaine à partir des données de référence couvrant la période 1955–2020. Les étoiles numérotées 1 à 4 identifient les cellules analysées aux figures 6.5 et 6.6.



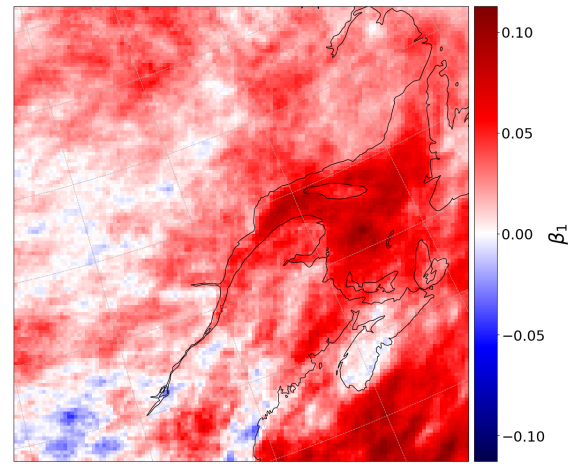
(a) Valeurs du paramètre $\exp(\alpha_0)$ correspondant à la partie stationnaire du paramètre $\kappa(T)$.



(b) Valeurs de $\exp(\beta_0)$ correspondant à la partie stationnaire du paramètre $\sigma(T)$.



(c) Valeurs de α_1 correspondant à la partie non-stationnaire du paramètre $\kappa(T)$.



(d) Valeurs de β_1 correspondant à la partie non-stationnaire du paramètre $\sigma(T)$.

FIGURE 6.3 Valeurs des paramètres $\exp(\alpha_0)$, α_1 , $\exp(\beta_0)$, β_1 pour chaque cellule du domaine après maximisation de la log-vraisemblance (6.17).

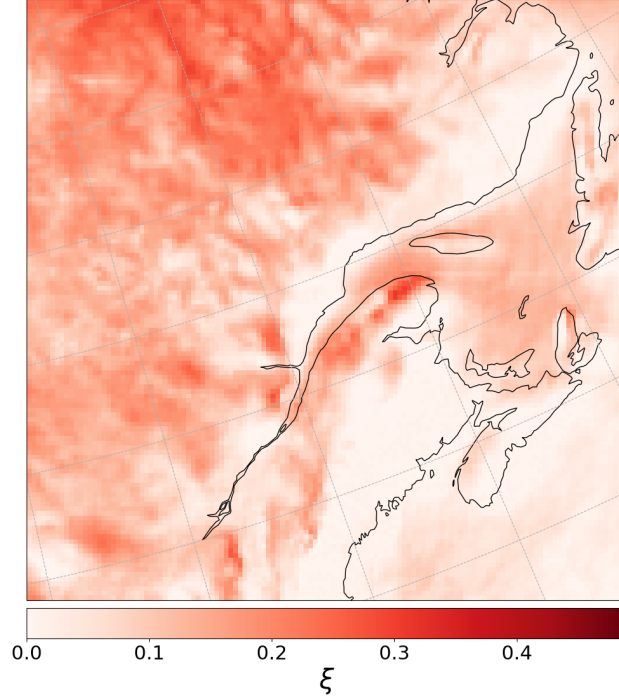


FIGURE 6.4 Valeurs du paramètre ξ après application de la transformation (6.24).

changements climatiques.

6.2.2 Réduction d'échelle par prédiction des niveaux de quantile

Contrairement à l'approche classique de réduction d'échelle, qui consiste à entraîner un modèle d'apprentissage profond afin de prédire directement la valeur de précipitation à haute résolution (ou, dans notre cas, le résidu par rapport à la basse résolution), nous proposons une méthode fondée sur la structure statistique des précipitations haute résolution. Plus précisément, pour chaque cellule, nous exploitons la loi ExtGPD ajustée précédemment.

Soient x et y deux champs de précipitations respectivement à basse et haute résolution, associés à un champ d'anomalies saisonnières de températures tp . Pour chaque cellule (i, j) de la grille haute résolution, le modèle de réduction d'échelle a pour objectif de prédire le niveau de quantile $\hat{p}_{ij}(x) \in (0, 1)$ tel que \hat{y}_{ij} , la valeur de précipitation correspondante, s'obtienne par inversion de la fonction de répartition F_{ij} , dépendante des anomalies de températures, estimée à partir de la loi ExtGPD ajustée pour cette cellule :

$$\hat{y}_{ij} = F_{ij}^{-1}(\hat{p}_{ij}(x), tp_{i,j}). \quad (6.25)$$

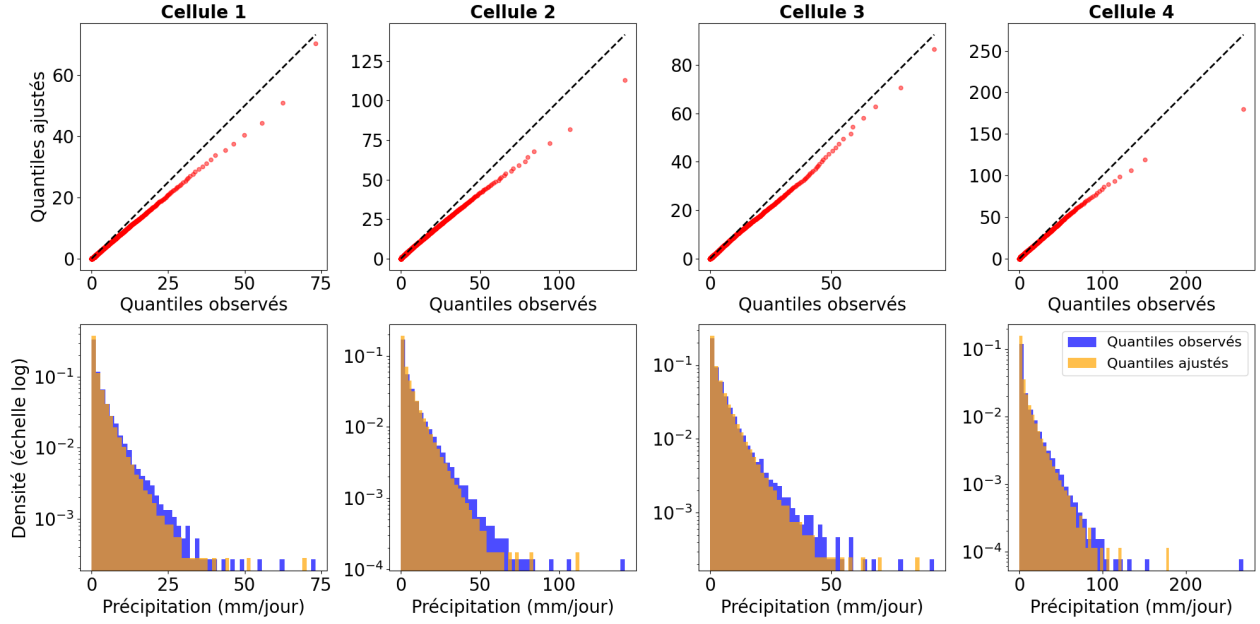


FIGURE 6.5 (**Première ligne**) Diagrammes quantile-quantile des données de la période 2020 – 2099 provenant des cellules 1 à 4 (voir figure 6.2), pour les modèles stationnaires ($\alpha_1 = \beta_1 = 0$) ajustés sur la période 1955 – 2020. (**Deuxième ligne**) Histogramme des quantiles obtenus.

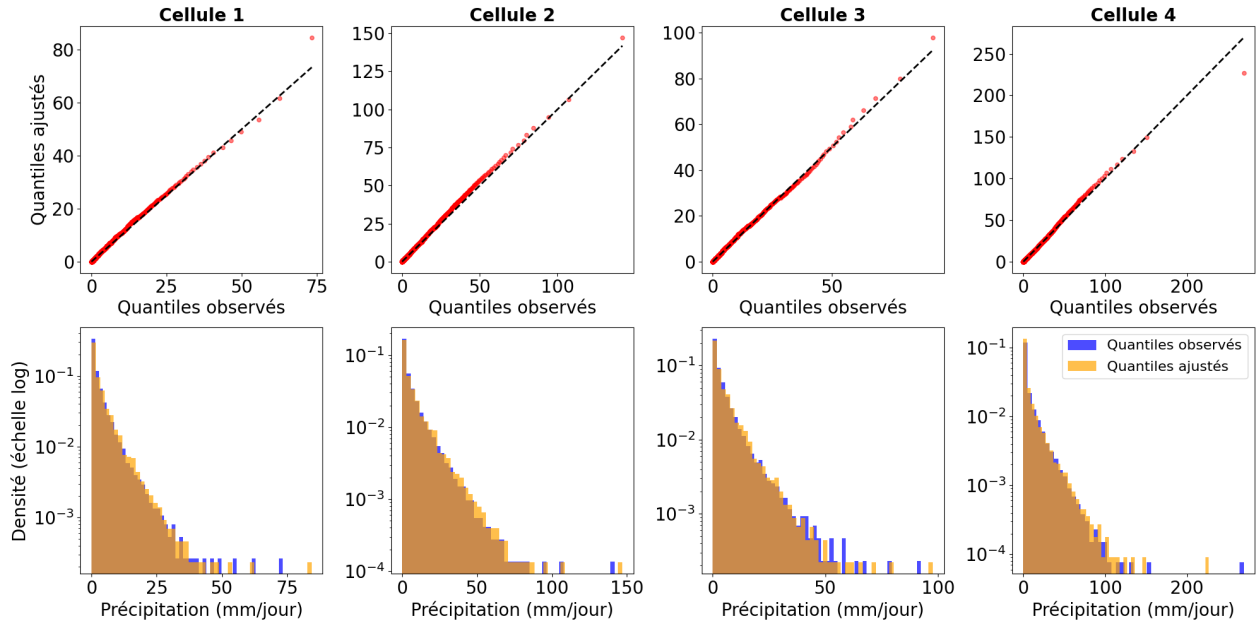


FIGURE 6.6 (**Première ligne**) Diagrammes quantile-quantile des données de la période 2020 – 2099 provenant des cellules 1 à 4 (voir figure 6.2), pour les modèles non-stationnaires ajustés sur la période 1955 – 2020. (**Deuxième ligne**) Histogramme des quantiles obtenus.

Cette approche permet d'intégrer les propriétés de la théorie des valeurs extrêmes pour améliorer la représentation des précipitations de forte intensité. De plus, en faisant dépendre la fonction de répartition $F_{i,j}$ aux anomalies de températures annuelles, on renforce la robustesse du modèle en changements climatiques. L'inverse de la fonction de répartition de la loi ExtGPD peut se calculer par la formule :

$$F^{-1}(p) = \begin{cases} \frac{\sigma}{\xi} \left[(1 - p^{1/\kappa})^{-\xi} - 1 \right] & \text{si } \xi > 0, \\ -\sigma \log(1 - p^{1/\kappa}) & \text{si } \xi = 0. \end{cases} \quad (6.26)$$

L'ensemble du champ haute résolution prédit \hat{y} est alors reconstruit en appliquant cette opération à chaque cellule de la grille \mathcal{G}_{HR} :

$$\hat{y} = \left\{ \hat{y}_{ij} = F_{ij}^{-1}(\hat{p}_{ij}(x), tp_{i,j}) \mid (i, j) \in \mathcal{G}_{HR} \right\}. \quad (6.27)$$

Les niveaux de quantile associés aux lois ExtGPD ajustées n'étant pas uniformément répartis (ils sont en grande majorité concentrés vers les faibles valeurs, proches de zéro), nous choisissons d'entraîner le modèle f_θ à prédire une transformation logarithmique du niveau de quantile. Plus précisément, si $p_{i,j} = F(y_{i,j}, tp_{i,j})$ désigne le niveau de quantile correspondant à la valeur cible $y_{i,j}$, selon la fonction de répartition F de la loi ExtGPD dépendante de $tp_{i,j}$, alors le modèle est entraîné à approximer :

$$f_\theta(x)_{i,j} = \log(1 - p_{i,j}) \implies \hat{p}_{i,j} = 1 - \exp(-f_\theta(x)_{i,j}) \quad (6.28)$$

où x désigne l'entrée interpolée de la basse résolution, $f_\theta(x)_{i,j}$ la sortie du modèle pour la cellule (i, j) , et $\hat{p}_{i,j}$ le niveau de quantile prédit.

Soient $Y \in \mathbb{R}^{N \times H \times W}$ un mini-lot de N champs de précipitations journalières de référence à haute résolution, $X \in \mathbb{R}^{N \times H \times W}$ un mini-lot correspondant interpolé à partir de données basse résolution, et $T \in \mathbb{R}^{N \times H \times W}$ le mini-lot des anomalies de températures haute résolution, les champs de paramètres $\kappa, \sigma_0, \sigma_1, \xi \in \mathbb{R}^{H \times W}$ ajustés au préalable pour chaque cellule de la grille, et $F(\cdot, T_{i,j})$ la fonction de répartition de la loi ExtGPD dépendante par les anomalies de température, pour la cellule (i, j) .

Nous entraînons alors le UNet en minimisant la fonction de perte suivante :

$$\mathcal{L}_{\text{ExtGPD}_1}(\theta) = \frac{1}{N H W} \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W \left(f_\theta(X_n)_{h,w} - \log(1 - F(Y_{n,h,w}, T_{n,h,w})) \right)^2, \quad (6.29)$$

6.2.3 Apprentissage par alignement des niveaux de quantile

Une autre approche consiste à utiliser la prédiction de précipitations $\hat{Y}_{n,i,j} = f_\theta(X_n)_{i,j}$ produite par le UNet, puis à projeter cette valeur dans l'espace des niveaux de quantiles en appliquant la fonction de répartition $F_{i,j}(\cdot, T_{n,i,j})$ associée à la cellule (i, j) . On compare ensuite le niveau de quantile prédit au niveau de quantile cible $p_{n,i,j} = F(Y_{n,i,j}, T_{n,i,j})$, sans appliquer la transformation logarithmique détaillée précédemment, car la valeur de précipitations prédite n'est pas directement dépendante du niveau de quantile :

$$\mathcal{L}_{\text{ExtGPD}_2}(\theta) = \frac{1}{NHW} \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W \left(F(f_\theta(X_n)_{h,w}, T_{n,h,w}) - F(Y_{n,h,w}, T_{n,h,w}) \right)^2. \quad (6.30)$$

CHAPITRE 7 ÉVALUATION

Ce chapitre vise à présenter et analyser les résultats de chacune des méthodes proposées au chapitre précédent, en comparaison à ceux du UNet présenté au chapitre 5, entraîné sur la perte quadratique des résidus (5.4).

7.1 Métriques

L’objectif principal de la réduction d’échelle est de reproduire les structures fines des champs de précipitations à haute résolution à partir de champs dégradés qui ne capturent que les structures de grande échelle. Le modèle de réduction d’échelle doit également être en mesure de généraliser dans un contexte de changements climatiques, alors que les températures moyennes augmentent.

Atteindre cet objectif nécessite la satisfaction simultanée de plusieurs critères de performance :

1. **Restitution des intensités locales** : le modèle doit être capable de reproduire fidèlement les valeurs de précipitations à haute résolution, ainsi que l’étendue des évènements de précipitations.
2. **Reconstruction de la distribution statistique** : il est essentiel que le modèle puisse reconstituer la distribution des précipitations à haute résolution, et en particulier de bien pouvoir représenter les valeurs extrêmes. Cela suppose une bonne estimation des queues de distribution.
3. **Fidélité fréquentielle** : les champs reconstruits doivent contenir la bonne quantité d’information à chaque échelle spatiale. Cela suppose une reconstruction précise des structures de précipitation sur un intervalle de fréquences spatiales, avec le moins de distorsion possible.

Afin de comparer la capacité des modèles à satisfaire le critère (1), nous mesurons l’erreur absolue moyenne (MAE) (exprimée en mm/jour) entre les champs de précipitations générés par la réduction d’échelle et les données de référence. La prédiction des précipitations extrêmes à haute résolution représentent un défi particulier pour les modèles. Pour évaluer plus finement les performances dans les régions extrêmes de la distribution, nous calculons aussi la MAE aux voisinages de plusieurs niveaux de quantiles, en suivant la méthodologie suivante :

1. Les quantiles de niveaux $q \pm 0.025$ sont estimés sur l'ensemble des données de référence.
2. Nous identifions les points de données de l'ensemble de référence dont les valeurs de précipitation sont comprises entre ces deux quantiles. Cela définit une plage locale autour du quantile de niveau q .
3. La MAE est ensuite calculée uniquement sur cet ensemble restreint de points.

Nous calculons également la MAE pour chaque année de la période de test. Cette évaluation permet d'analyser l'évolution des erreurs au fil du temps, et de comparer la capacité de généralisation aux changements climatiques entre modèles.

Pour évaluer le respect du critère (2), nous comparons, en agrégeant toutes les cellules du domaine, les histogrammes des précipitations sous une échelle logarithmique pour se concentrer sur la queue de la distribution.

Finalement, pour évaluer le respect du critère (3), nous analysons la fidélité fréquentielle des champs de précipitations reconstitués à travers la densité spectrale de puissance moyennée radialement (ou *Radial-Averaged Power Spectral Density* en anglais, RA-PSD). Cette métrique permet de quantifier la distribution de la variance spatiale du champ en fonction des fréquences spatiales, et fournit ainsi une mesure directe de la capacité des modèles à restituer les structures présentes à différentes échelles.

Pour un champ de précipitations $Y \in \mathbb{R}_+^{H \times W}$, la RA-PSD est obtenue à partir de la transformée de Fourier discrète bidimensionnelle :

$$\hat{Y}(k_h, k_w) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} Y_{h,w} e^{-2\pi i \left(\frac{k_h h}{H} + \frac{k_w w}{W} \right)} \quad (7.1)$$

$$\text{PSD}(k_h, k_w) = \frac{1}{HW} \left| \hat{Y}(k_h, k_w) \right|^2 \quad (7.2)$$

La densité spectrale de puissance (PSD) bidimensionnelle est ensuite moyennée radialement, c'est-à-dire par intervalle d'anneaux successifs de rayon $k = \sqrt{k_h^2 + k_w^2}$, afin d'obtenir une PSD unidimensionnelle notée $\text{PSD}(k)$. Les vecteurs de PSD unidimensionnels ainsi obtenus pour chaque champ de précipitations sont ensuite moyennés, fréquence par fréquence, pour produire une PSD moyenne représentative de l'ensemble des données pour chaque modèle.

Pour évaluer la qualité de reconstruction fréquentielle des modèles, nous analysons la RA-PSD relative entre les champs générés et les données de référence, définie par :

$$\text{PSD}_{\text{rel}}(k) = \frac{\text{PSD}_{\text{modèle}}(k)}{\text{PSD}_{\text{référence}}(k)} \quad (7.3)$$

Une valeur de $\text{PSD}_{\text{rel}}(k) = 1$ indique une restitution parfaite de la variance à l'échelle spatiale correspondante. Des valeurs inférieures à 1 traduisent une perte de variance (sous-représentation des structures), tandis que des valeurs supérieures à 1 révèlent une surestimation.

7.2 Résultats

Chaque modèle comparé est entraîné selon la même procédure, et avec les mêmes hyperparamètres d'entraînement. La graine aléatoire est fixée à 351 pour l'ensemble des entraînements. L'algorithme d'optimisation utilisé est **Adam**, avec ses paramètres standards et un taux d'apprentissage de 10^{-3} . Les champs de précipitations journalières sont mélangés aléatoirement, et la taille du mini-lot est fixée à 32 journées, sauf pour les variantes de l'approche par regroupement, pour lesquelles elle est fixée à 150. Chaque modèle est entraîné pendant 50 époques.

Pour rappel, l'ensemble d'entraînement couvre la période 1955–2020, correspondant à 23 725 points de données par cellule, et l'ensemble de test couvre la période allant de 2020 à 2099, totalisant 28 835 points de données par cellule.

L'interpolation bicubique est utilisée comme méthode de référence (*baseline*), fournissant une borne inférieure de comparaison pour évaluer les performances des modèles. L'erreur absolue moyenne de cette méthode est présentée au tableau 7.1.

TABLEAU 7.1 MAE globale et évaluée aux voisinages de plusieurs niveaux de quantiles pour l'interpolation bicubique sur la période de test.

MAE	MAE-q (0.5)	MAE-q (0.75)	MAE-q (0.9)	MAE-q (0.95)	MAE-q (0.99)
1.570	1.181	2.446	4.050	6.983	20.949

7.2.1 Contrainte douce par approche de regroupement

Nous analysons les performances de l'approche consistant à ajouter à la fonction de perte du modèle une contrainte douce censée permettre d'aligner les prédictions du modèle sur la relation entre les températures et les extrêmes de précipitation. Cette contrainte définie par l'approche de regroupement est définie à la section 6.1.

Dans l'objectif de mieux comprendre le comportement du modèle lors de l'application de cette contrainte, nous comparons les résultats en fonction de plusieurs valeurs du paramètre $\gamma \in \{0, 1, 10, 50, 100\}$, qui module l'intensité de la pénalité appliquée à la fonction de perte

($\gamma = 0$ correspondant à la perte quadratique, et $\gamma = 100$ impliquant une pénalité à peu près du même ordre de grandeur que la fonction de perte initiale). Nous comparons également les performances des modèles en fonction du seuil de filtrage s (6.4), fixé à 0 ou 1 mm/jour.

Les tableaux 7.2 et 7.3 présentent les valeurs de MAE sur la période de test pour les différentes configurations étudiées. Aucun gain significatif de performance n'est observé suite à l'introduction de la pénalité, quel que soit le seuil s . Au contraire, pour des coefficients $\gamma > 10$, une dégradation marquée des performances est constatée. Les résultats obtenus pour les coefficients modérés ($\gamma \in 1, 10$) sont toutefois systématiquement meilleurs lorsqu'un filtrage des précipitations inférieures à $s = 1$ mm/jour est appliqué — à l'exception de la médiane. Cette amélioration peut s'expliquer par une concentration de l'apprentissage sur les fortes intensités, les faibles précipitations étant écartées. Néanmoins, cet effet restant marginal, ces résultats doivent être interprétés avec prudence et nécessitent des investigations complémentaires pour être consolidés.

TABLEAU 7.2 MAE globale et évaluée aux voisinages de plusieurs niveaux de quantiles, en fonction de la valeur du coefficient γ de la fonction de perte (6.1), pour $s = 0.0$ mm/jour (seuil de filtrage (6.4)).

γ	MAE	MAE-q (0.5)	MAE-q (0.75)	MAE-q (0.9)	MAE-q (0.95)	MAE-q (0.99)
0	1.329	0.918	2.273	4.045	6.744	17.151
1	1.331	0.931	2.290	4.044	6.710	17.123
10	1.328	0.944	2.255	3.939	6.624	17.306
50	1.375	0.987	2.314	4.032	6.813	17.773
100	1.393	1.010	2.315	4.028	6.810	18.073

TABLEAU 7.3 MAE globale et évaluée au voisinage de plusieurs niveaux de quantiles, en fonction de la valeur du coefficient γ de la fonction de perte (6.1), pour $s = 1.0$ mm/jour (seuil de filtrage (6.4)).

γ	MAE	MAE-q (0.5)	MAE-q (0.75)	MAE-q (0.9)	MAE-q (0.95)	MAE-q (0.99)
0	1.329	0.918	2.273	4.045	6.744	17.151
1	1.323	0.933	2.268	3.989	6.599	17.083
10	1.321	0.937	2.260	3.908	6.503	17.376
50	1.365	1.010	2.272	3.879	6.594	18.128
100	1.391	1.037	2.297	3.948	6.730	18.308

Pour la suite des évaluations, nous retenons pour chaque seuil s la configuration avec $\gamma = 10$, ce coefficient offrant généralement les meilleures performances. L'évolution de la MAE sur la période de test, présentée en figure 7.1, montre que les modèles pénalisés (courbes orange et violette) ne parviennent pas à atténuer la tendance croissante de l'erreur au cours du temps.

Par ailleurs, la figure 7.2 confirme les limites évoquées en section 6.1.2 : les modèles entraînés avec une pénalité ($\gamma > 0$) n'améliorent pas l'alignement des TPSRs avec ceux des données de référence, en comparaison avec l'entraînement basé uniquement sur $\mathcal{L}_{\text{reco}}$. La séparation en deux sous-périodes suggère enfin que ces taux d'évolution ne présentent pas de dynamique sensible à l'accélération des changements climatiques.

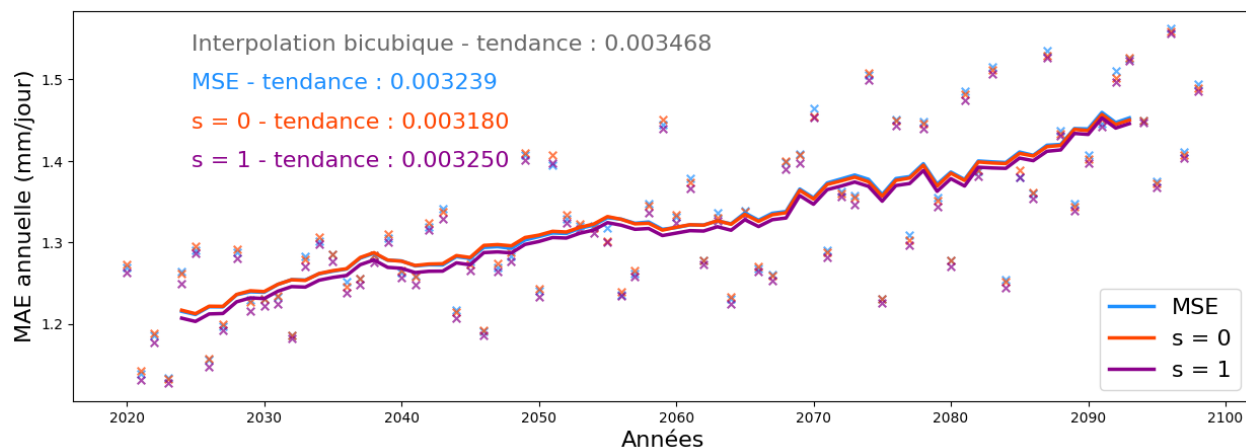


FIGURE 7.1 MAE annuelle (en mm/jour) pour la perte MSE ($\gamma = 0$), et pour $s = 0$ ou $s = 1$ mm/jour avec $\gamma = 10$. L'erreur de l'interpolation cubique n'est pas affichée pour préserver la lisibilité de la figure. Lissage par moyenne mobile sur 10 valeurs.

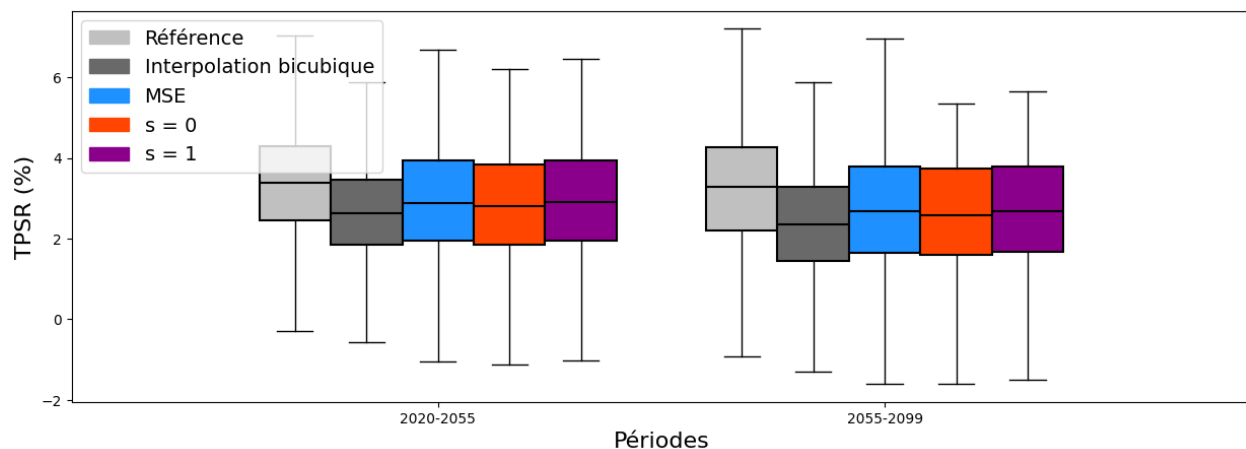


FIGURE 7.2 Boîtes à moustaches des TPSRs (en $\%/^{\circ}\text{C}$) obtenus *via* la méthode définie à la section 6.1 pour chaque modèle. On sépare l'ensemble de test en deux périodes : 2020 – 2050, et 2050 – 2099, pour évaluer l'évolution temporelle.

L'étendue des précipitations prédites par les modèles n'est pas améliorée par l'ajout de la contrainte, comme le montre la figure 7.3, qui présente les histogrammes des précipitations

pour chaque modèle sur la période de test (toutes les cellules du domaine étant agrégées). On observe en revanche une amélioration marginale de la PSD relative pour les fréquences supérieures à environ $0,047 \text{ km}^{-1}$, ce qui correspond à une échelle spatiale inférieure à quatre cellules. La chute des courbes de PSD à ces fréquences s’observe également dans les données de référence. Elle peut traduire d’une distorsion propre au modèle RCM à cette échelle, possiblement liée à une transition entre deux schémas physiques, ou encore à un effet de filtrage numérique.

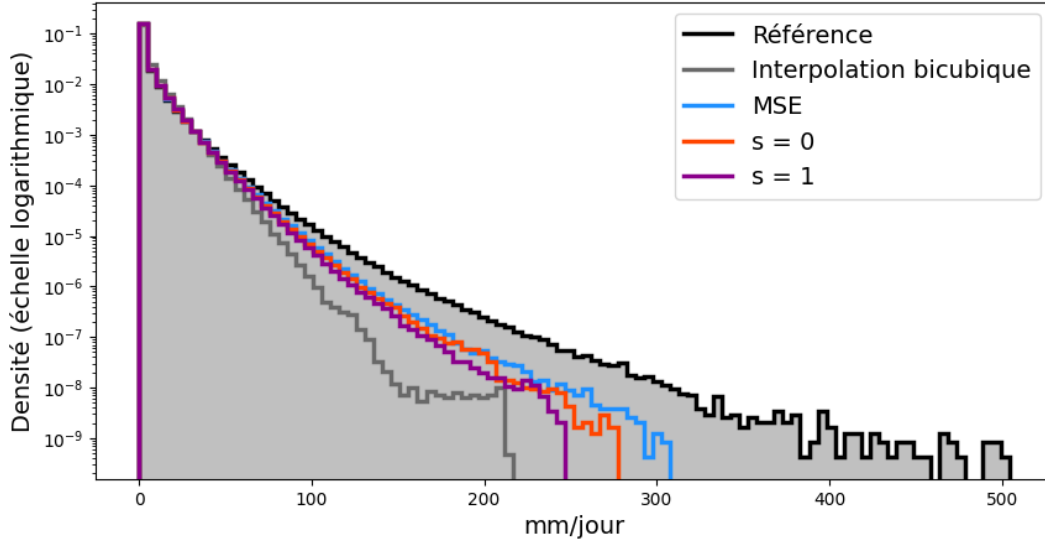


FIGURE 7.3 Histogramme des précipitations pour chaque modèle, incluant l’ensemble des cellules du domaine sur la période de l’ensemble de test 2020 – 2099

Comme discuté à la section 6.1.2, une taille de lot bien supérieure à $N = 150$ échantillons serait nécessaire pour obtenir une estimation robuste des TPSRs en utilisant la méthode par regroupement. Une telle estimation est indispensable pour que la pénalisation issue de cette méthode définisse un paysage d’optimisation cohérent avec l’objectif annoncé : aligner les taux de progression des extrêmes de précipitations en fonction des températures avec ceux observés dans les données de référence.

Cependant, les limitations en mémoire vive rendent impossible l’utilisation de tailles de lot supérieures à 256 tout en conservant un domaine suffisamment vaste pour contenir une diversité d’information.

Une solution consiste à entraîner le modèle sur un ensemble de sous-domaines superposés (ou *patches* en anglais) extraits du domaine initial. Cette approche permettrait une réduction significative du besoin de mémoire, rendant possible l’apprentissage du UNet avec la fonction de perte proposée, pour des tailles de lot dépassant 512. Nous encourageons donc la poursuite

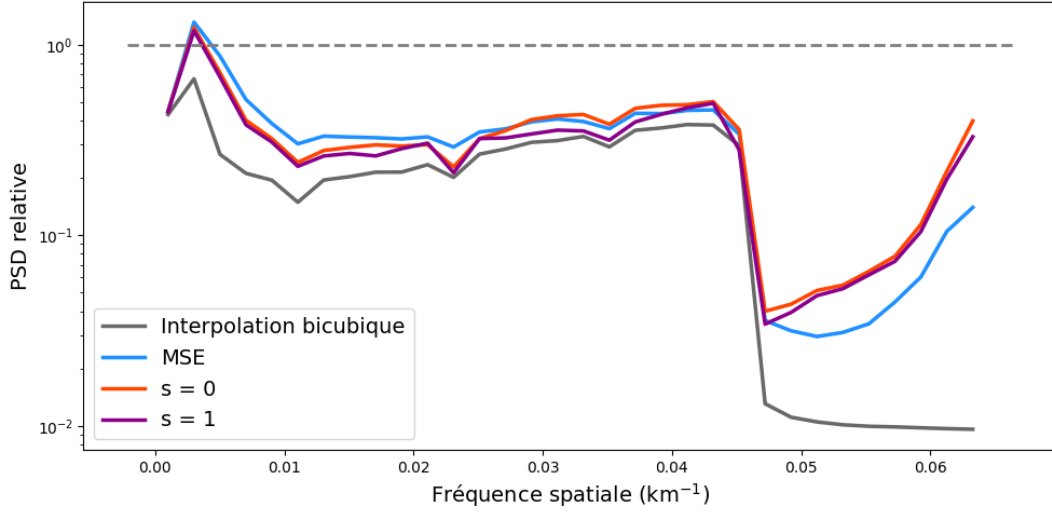


FIGURE 7.4 PSD_{rel} pour chaque modèle par rapport à la densité spectrale de puissance des données de référence, calculées suivant la méthode détaillée à la section 7.1.

de ces travaux, afin de pouvoir réellement évaluer la pertinence de cette approche pour améliorer la prédiction des extrêmes et la robustesse aux changements climatiques.

7.2.2 Intégration d'*a priori* statistiques issus de la théorie des valeurs extrêmes

Nous évaluons chacune des méthodes proposées dans deux configurations : stationnaire et non-stationnaire, correspondant respectivement à l'exclusion ou à l'inclusion des anomalies de températures dans la fonction de perte, et le calcul du quantile. Toutefois, ces anomalies ne sont jamais utilisées comme covariables explicites à l'entrée des modèles. En effet, leur inclusion aux côtés des précipitations à basse résolution et de la température moyenne journalière a systématiquement conduit à une dégradation des performances lors de nos expérimentations. Seulement un modèle parmi ceux évalués a accès aux anomalies de températures annuelles lors de l'évaluation sur la période de test : $\mathcal{L}_{\text{ExtGPD}_1}(T)$, car cette variable est nécessaire pour le calcul du quantile de précipitation pour le cas de la loi ExtGPD non-stationnaire.

Le tableau 7.4 met en évidence l'intérêt d'utiliser l'*a priori* statistique offert par la loi ExtGPD pour les niveaux de quantile $q \in [0.75, 0.95]$. Néanmoins, notons que la médiane des précipitations haute résolution est systématiquement sujette à des erreurs plus importantes en suivant ces approches. Malgré des améliorations, les progrès restent marginaux, on ne peut donc pas tirer de conclusion évidente quant à l'intérêt de ces méthodes pour améliorer la prédiction des extrêmes de précipitations.

Similairement, l'inclusion des anomalies de températures, bien que bénéfique pour modéliser

TABLEAU 7.4 MAE globale (en mm/jour) et évaluée au voisinage de plusieurs niveaux de quantiles, pour le UNet standard $\mathcal{L}_{\text{reco}}$ et les méthodes de prédictions et d'alignement sur les niveaux de quantile de la loi ExtGPD : $\mathcal{L}_{\text{ExtGPD}_1}$ (1) et $\mathcal{L}_{\text{ExtGPD}_2}$ (2), avec ou sans anomalies de températures T .

Modèle	MAE	MAE (0.5)	MAE (0.75)	MAE (0.9)	MAE (0.95)	MAE (0.99)
$\mathcal{L}_{\text{reco}}$	1.329	0.918	2.273	4.045	6.744	17.151
1	1.330	0.942	2.223	3.920	6.662	17.729
1 + T	1.336	0.943	2.257	3.993	6.734	17.629
2	1.318	0.928	2.251	3.956	6.590	17.221
2 + T	1.321	0.926	2.258	3.986	6.631	17.175

la distribution des précipitations en changements climatiques, comme noté à la section 6.2.1, semble avoir l'effet inverse de celui attendu, avec une augmentation plus rapide des erreurs par rapport aux autres modèles, ce qu'on peut observer à la figure 7.5. Néanmoins, cela ne doit pas forcément être considéré comme une dégradation des performances, car comme indiqué par l'histogramme des précipitations du modèle $\mathcal{L}_{\text{ExtGPD}_1}(T)$ à la figure 7.7, la prédiction des niveaux de quantile et l'inclusion des anomalies de températures permet d'élargir la limite des intensités obtenues en changements climatiques. Or, dans le cas où l'évènement de précipitations contenant une de ces intensités extrêmes n'est pas exactement bien situé par le UNet, le modèle sera doublement pénalisé par rapport à un autre modèle ayant largement sous-estimé les intensités de précipitations du champ.

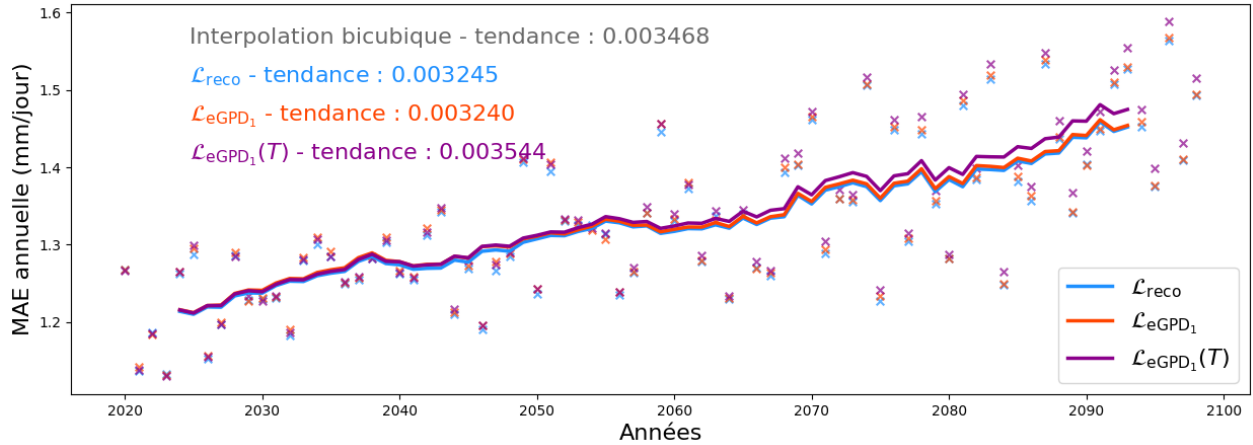


FIGURE 7.5 MAE (en mm/jour) annuelle pour le UNet standard ($\mathcal{L}_{\text{reco}}$) et les variantes stationnaire et non-stationnaire de la méthode de prédiction des niveaux de quantile (voir section 6.2.2) des lois ExtGPD ajustées précédemment : $\mathcal{L}_{\text{ExtGPD}_1}$ et $\mathcal{L}_{\text{ExtGPD}_1}(T)$. La courbe est obtenue par moyenne mobile sur une fenêtre de 10 valeurs.

La double pénalité remet donc en cause la pertinence de la seule utilisation de la MAE

pour mesurer les performances du UNet en changements climatiques, surtout concernant les extrêmes de précipitations.

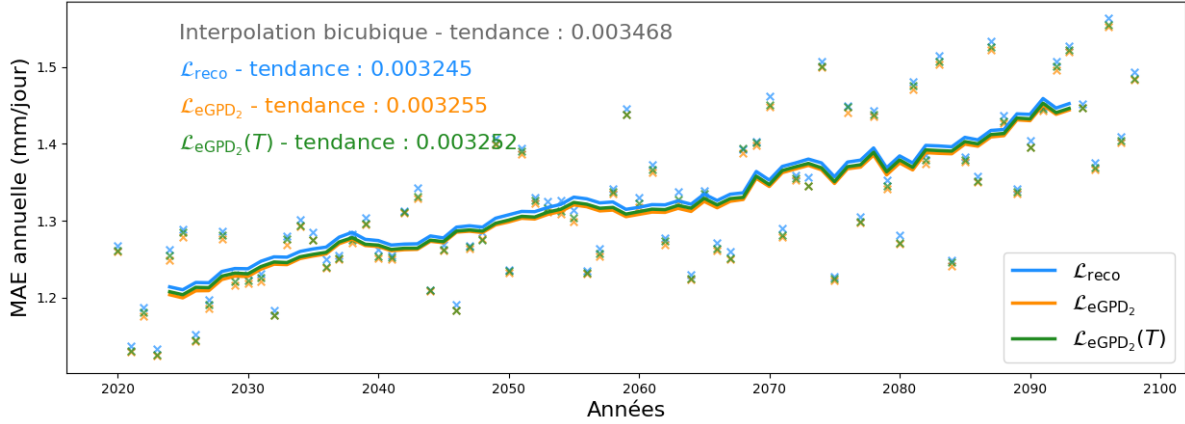


FIGURE 7.6 MAE (en mm/jour) annuelle pour le UNet standard ($\mathcal{L}_{\text{reco}}$) et les variantes stationnaire et non-stationnaire de la méthode d'alignement des niveaux de quantile (voir section 6.2.3) : $\mathcal{L}_{\text{ExtGPD}_1}$ et $\mathcal{L}_{\text{ExtGPD}_1}(T)$. La courbe est obtenue par moyenne mobile sur une fenêtre de 10 valeurs.

Lors des expérimentations menées avec les modèles de prédiction des niveaux de quantile des lois ExtGPD, nous avons observé une tendance systématique à sous-estimer les hauts quantiles, conduisant à des précipitations simulées nettement inférieures à celles observées dans les données de référence. Afin d'atténuer cet effet, la transformation logarithmique a été introduite dans la fonction de perte (6.29), ce qui a permis de corriger partiellement ce biais. L'amélioration de la représentation des précipitations extrêmes et la meilleure conservation de la variance spatiale (voir figure 7.8) suggèrent que l'intégration d'informations issues de la loi ExtGPD au sein des modèles d'apprentissage profond pour la réduction d'échelle des précipitations conserve un potentiel important.

La figure 7.9 illustre les champs de précipitations simulés pour deux journées distinctes de la période de test, en comparant les sorties des différents modèles étudiés aux données de référence à haute résolution issues de ClimEx. Cette visualisation permet de mettre en évidence la complexité de la tâche de réduction d'échelle des précipitations, surtout sans avoir recourt à de nombreuses covariables.

Les méthodes proposées n'atteignent pas pleinement les objectifs fixés : (1) améliorer la réduction d'échelle des précipitations extrêmes, et (2) renforcer la robustesse du UNet face aux changements climatiques, notamment en intégrant la relation de dépendance entre précipitations et évolution des températures. Malgré cela, les résultats obtenus restent prometteurs et encouragent à poursuivre les recherches afin de mieux exploiter le potentiel de ces approches.

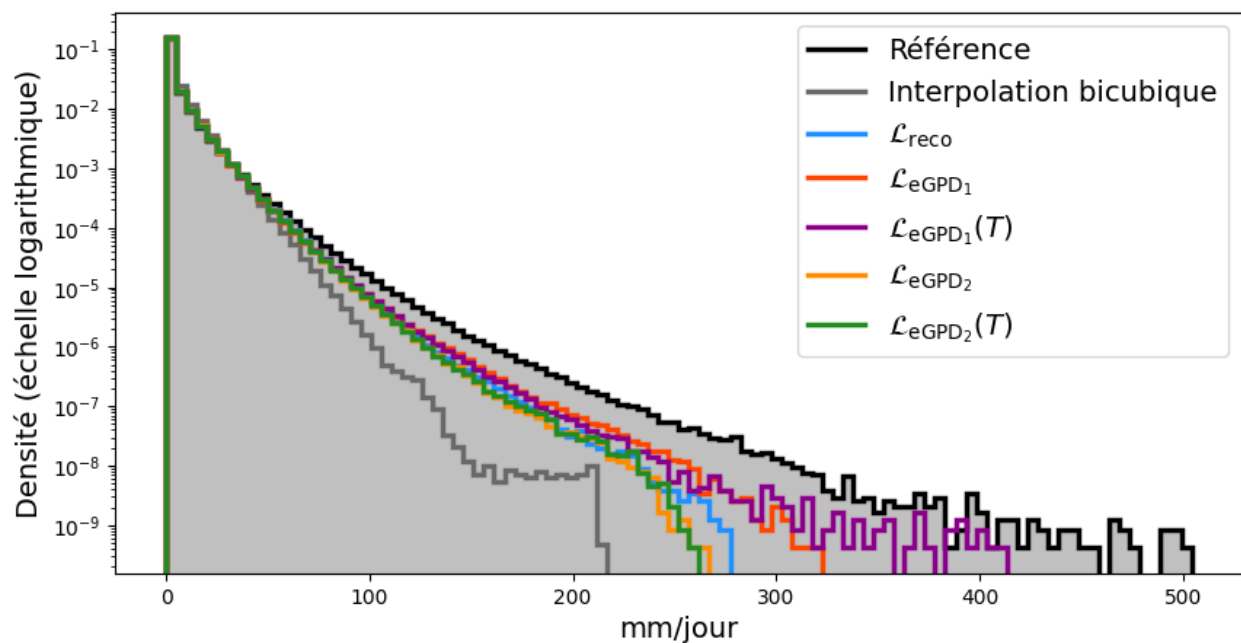


FIGURE 7.7 Histogramme des précipitations pour chaque modèle, incluant l'ensemble des cellules du domaine sur la période de l'ensemble de test 2020 – 2099.

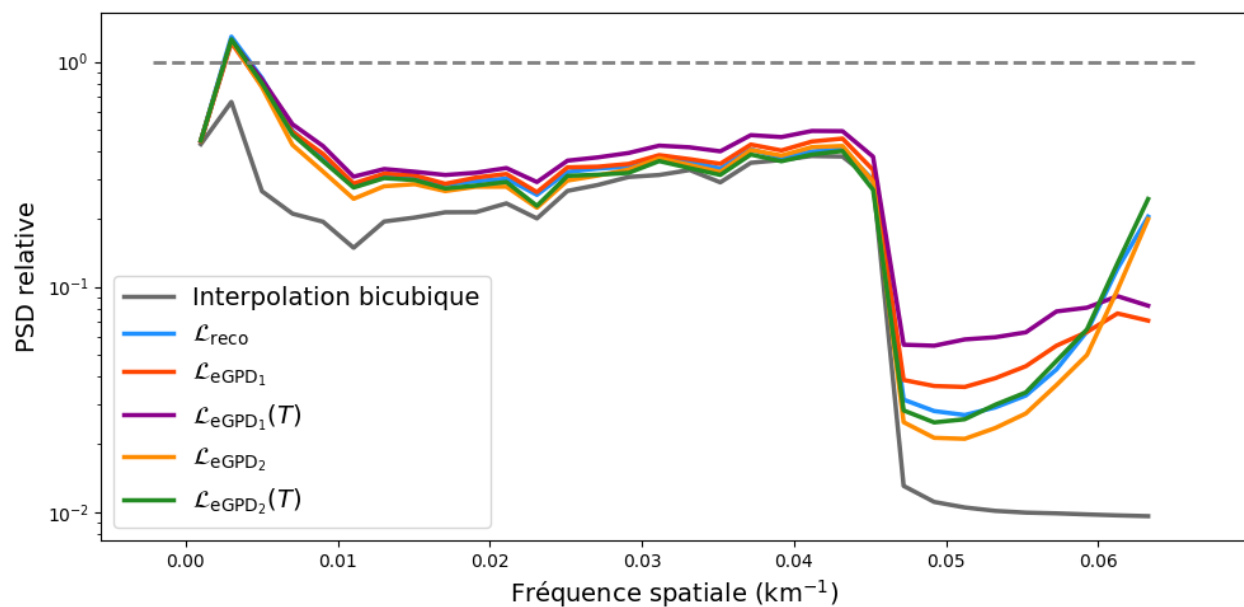
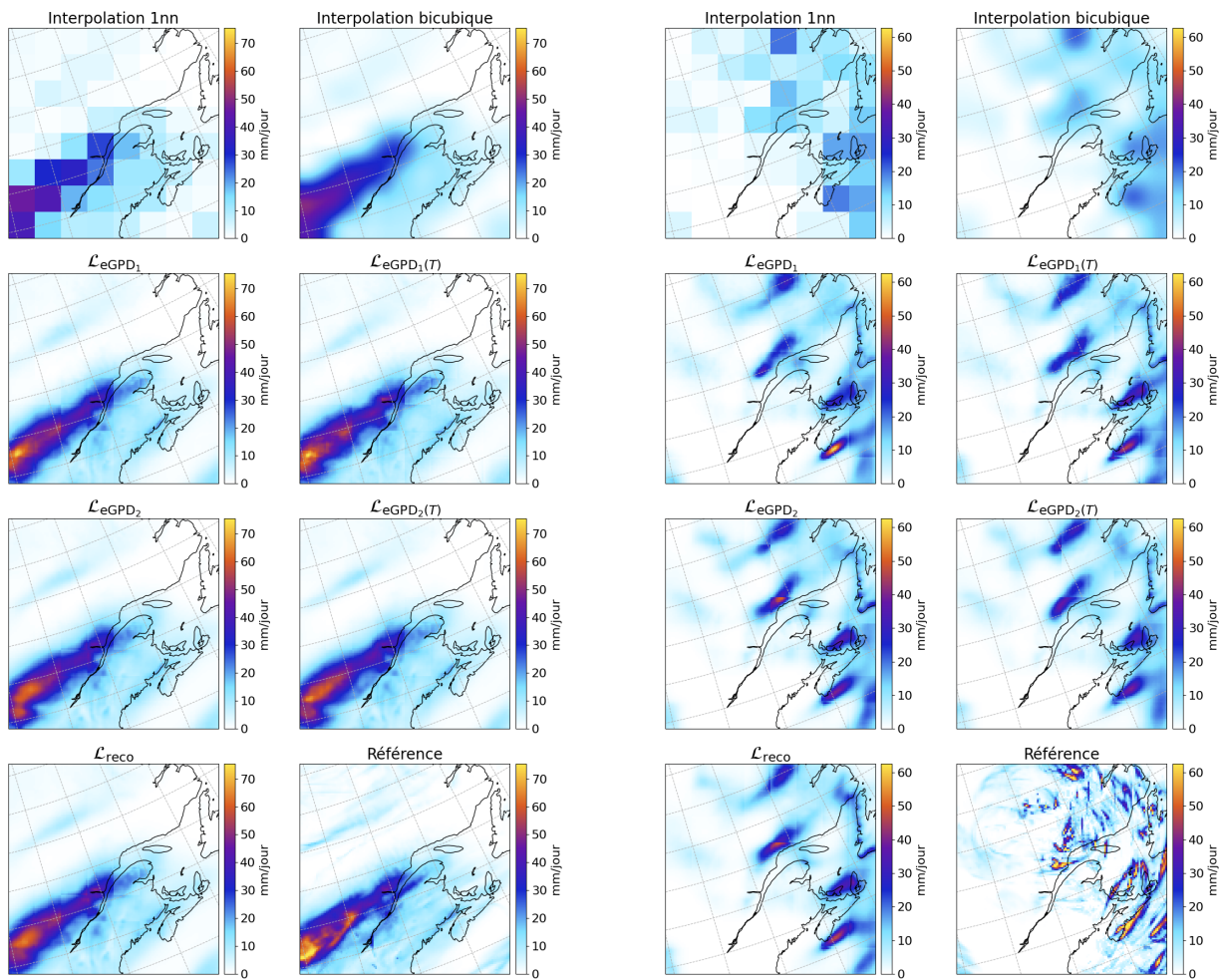


FIGURE 7.8 PSD_{rel} pour chaque modèle par rapport à la PSD des données de référence, calculées suivant la méthode détaillée à la section 7.1.



(a) Précipitations du 03 mai 2078.

(b) Précipitations du 21 juillet 2085.

FIGURE 7.9 Champs de précipitations à basse résolution (interpolation 1nn), interpolés par méthode bicubique, prédits par chacun des modèles étudiés, et à haute résolution issus de ClimEx, pour deux dates sélectionnées aléatoirement sur la période de test.

CHAPITRE 8 CONCLUSION

8.1 Synthèse des travaux

Les changements climatiques intensifient les événements extrêmes et modifient profondément la distribution de variables atmosphériques comme les précipitations, avec des impacts majeurs sur les populations et les écosystèmes. Des projections climatiques à haute résolution sont essentielles pour anticiper ces évolutions et s’y adapter. Les méthodes d’apprentissage profond représentent une opportunité importante pour produire des simulations à haute résolution à bas coût. Ce mémoire s’est concentré sur la réduction d’échelle des précipitations par apprentissage profond, en mettant en lumière certaines vulnérabilités des réseaux de neurones, notamment leur difficulté à prédire les extrêmes et leur manque de robustesse face aux changements climatiques.

Après avoir implémenté un réseau convolutif profond de type UNet pour la tâche de réduction d’échelle des champs de précipitations, incluant la définition d’une stratégie de *pré-* et *post-*traitement des données, plusieurs améliorations structurelles issues de la littérature ont été intégrées afin d’en renforcer les performances.

L’évaluation du UNet en changements climatiques, accompagnée de l’implémentation de méthodes d’estimation de TPSRs inspirées en partie de la théorie des valeurs extrêmes, a permis de formuler l’hypothèse supposant que la dégradation des performances de réduction d’échelle des précipitations en changements climatiques serait liée à la non-intégration de la relation de Clausius-Clapeyron au sein des caractéristiques apprises.

De ce constat a découlé la proposition d’incorporer au sein de l’entraînement du UNet l’information issue des TPSRs, en utilisant l’approche de regroupement. Malgré une implémentation fonctionnelle, le potentiel de la méthode n’a pas pu être réellement explorée, incitant à repenser la stratégie d’intégration de l’information obtenue par cette approche. Deux approches alternatives ont donc été développées, consistant à conditionner l’apprentissage par une fonction de répartition, issue de la théorie des valeurs extrêmes, représentant la distribution des précipitations de chaque cellule en fonction de l’évolution des anomalies de températures. Les paramètres de cette loi ont été estimés pour chaque cellule du domaine, et nous avons explicitement modélisé la dépendance aux anomalies de températures, permettant de mieux représenter les précipitations en changements climatiques.

Les résultats obtenus ne permettent pas, à ce stade, de conclure favorablement quant à l’efficacité des méthodes proposées pour pallier aux limites identifiées du UNet pour la réduction

d'échelle des précipitations. Toutefois, certains éléments laissent entrevoir un potentiel non négligeable, et suggèrent que des travaux complémentaires sont nécessaires pour mieux exploiter les approches développées.

8.2 Perspectives de recherche

Comme évoqué dans ce mémoire, une stratégie prometteuse consisterait à partitionner le domaine spatial en sous-domaines superposés, et à entraîner le UNet sur cette nouvelle structure. Une telle approche permettrait d'augmenter significativement la taille des mini-lots, condition essentielle pour garantir une estimation robuste des TPSRs *via* l'approche par regroupement. Comme le suggèrent les résultats de la section 6.1.2, cette méthode faciliterait également la convergence de la fonction de perte \mathcal{L}_{CC} définie en (6.1).

Concernant les méthodes fondées sur l'extension de la loi de Pareto généralisée, une analyse plus approfondie de l'influence des anomalies de température sur la modélisation des précipitations serait nécessaire. Il serait également pertinent d'envisager une combinaison des fonctions de perte proposées, dans le but de combiner les forces de chaque approche. Enfin, en s'inspirant des travaux en apprentissage profond appliqués aux problèmes inverses, l'introduction de couches d'échantillonnage directement issues de la loi ExtGPD au sein du UNet pourrait constituer une piste prometteuse. Cette intégration architecturale permettrait d'incorporer de manière explicite les *a priori* sur la forme des distributions de précipitations, ainsi que leur dépendance aux températures.

RÉFÉRENCES

- [1] P. Naveau, R. Huser, P. Ribereau et A. Hannart, “Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection,” *Water Resources Research*, vol. 52, n°. 4, p. 2753–2769, 2016. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR018552>
- [2] R. A. Watt et L. A. Mansfield, “Generative Diffusion-based Downscaling for Climate,” avr. 2024, arXiv :2404.17752 [physics]. [En ligne]. Disponible : <http://arxiv.org/abs/2404.17752>
- [3] Y. LeCun, Y. Bengio et G. Hinton, “Deep learning,” *Nature*, vol. 521, n°. 7553, p. 436–444, mai 2015, publisher : Nature Publishing Group. [En ligne]. Disponible : <https://www.nature.com/articles/nature14539>
- [4] I. Goodfellow, Y. Bengio et A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [5] A. Krizhevsky, I. Sutskever et G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” dans *Advances in Neural Information Processing Systems*, vol. 25. Curran Associates, Inc., 2012. [En ligne]. Disponible : https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html
- [6] O. Ronneberger, P. Fischer et T. Brox, “U-Net : Convolutional Networks for Biomedical Image Segmentation,” dans *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells et A. F. Frangi, édit. Cham : Springer International Publishing, 2015, p. 234–241.
- [7] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha et V. K. Asari, “Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation,” *CoRR*, vol. abs/1802.06955, 2018. [En ligne]. Disponible : <http://arxiv.org/abs/1802.06955>
- [8] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh et J. Liang, “Unet++ : A nested u-net architecture for medical image segmentation,” dans *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, p. 3–11.
- [9] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker et D. Rueckert, “Attention

- u-net : Learning where to look for the pancreas,” dans *Medical Imaging with Deep Learning*, 2018. [En ligne]. Disponible : <https://openreview.net/forum?id=Skft7cijM>
- [10] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. London : Springer London, 2001. [En ligne]. Disponible : <http://dx.doi.org/10.1007/978-1-4471-3675-0>
- [11] R. L. Smith, “Maximum likelihood estimation in a class of nonregular cases,” *Biometrika*, vol. 72, n^o. 1, p. 67–90, 1985. [En ligne]. Disponible : <http://www.jstor.org/stable/2336336>
- [12] I. Papastathopoulos et J. Tawn, “Extended generalised pareto models for tail estimation,” *Journal of Statistical Planning and Inference*, vol. 143, n^o. 1, p. 131–143, 2013.
- [13] A. Haruna, J. Blanchet et A.-C. Favre, “Performance-based comparison of regionalization methods to improve the at-site estimates of daily precipitation,” *Hydrology and Earth System Sciences*, vol. 26, n^o. 10, p. 2797–2811, 2022. [En ligne]. Disponible : <https://hess.copernicus.org/articles/26/2797/2022/>
- [14] J. S. Nanditha, G. Villarini et P. Naveau, “Assessing future changes in daily precipitation extremes across the contiguous United States with the extended Generalized Pareto distribution,” *Journal of Hydrology*, vol. 659, p. 133212, 2025. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0022169425005505>
- [15] P. Gamet et J. Jalbert, “A flexible extended generalized pareto distribution for tail estimation,” *Environmetrics*, vol. 33, n^o. 6, p. e2744, 2022. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2744>
- [16] G. Lenderink et E. van Meijgaard, “Increase in hourly precipitation extremes beyond expectations from temperature changes,” *Nature Geoscience*, vol. 1, n^o. 8, p. 511–514, août 2008. [En ligne]. Disponible : <https://doi.org/10.1038/ngeo262>
- [17] R. Hardwick Jones, S. Westra et A. Sharma, “Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity,” *Geophysical Research Letters*, vol. 37, n^o. 22, 2010, _eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2010GL045081>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2010GL045081>
- [18] H. Ali, H. J. Fowler, G. Lenderink, E. Lewis et D. Pritchard, “Consistent Large-Scale Response of Hourly Extreme Precipitation to Temperature Variation Over Land,” *Geophysical Research Letters*, vol. 48, n^o. 4, p. e2020GL090317, 2021, _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL090317>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL090317>
- [19] A. Papadopoulos-Zachos et C. Anagnostopoulou, “A methodological approach to improving extreme precipitation reanalysis data using the clausius-clapeyron

- relationship : A case study in a mediterranean city,” *Atmosphere*, vol. 15, n°. 10, 2024. [En ligne]. Disponible : <https://www.mdpi.com/2073-4433/15/10/1195>
- [20] Y. Moustakis, C. J. Onof et A. Paschalis, “Atmospheric convection, dynamics and topography shape the scaling pattern of hourly rainfall extremes with temperature globally,” *Communications Earth & Environment*, vol. 1, n°. 1, p. 1–9, août 2020, publisher : Nature Publishing Group. [En ligne]. Disponible : <https://www.nature.com/articles/s43247-020-0003-0>
- [21] H. Van de Vyver, B. Van Schaeybroeck, R. De Troch, R. Hamdi et P. Termonia, “Modeling the Scaling of Short-Duration Precipitation Extremes With Temperature,” *Earth and Space Science*, vol. 6, n°. 10, p. 2031–2041, 2019, _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019EA000665>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2019EA000665>
- [22] J. Bao, S. C. Sherwood, L. V. Alexander et J. P. Evans, “Future increases in extreme precipitation exceed observed scaling rates,” *Nature Climate Change*, vol. 7, n°. 2, p. 128–132, févr. 2017. [En ligne]. Disponible : <https://www.nature.com/articles/nclimate3201>
- [23] Q. Sun, F. Zwiers, X. Zhang et G. Li, “A comparison of intra-annual and long-term trend scaling of extreme precipitation with temperature in a large-ensemble regional climate simulation,” *Journal of Climate*, vol. 33, n°. 21, p. 9233 – 9245, 2020. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/clim/33/21/jcliD190920.xml>
- [24] X. Zhang, F. W. Zwiers, G. Li, H. Wan et A. J. Cannon, “Complexity in estimating past and future extreme short-duration rainfall,” *Nature Geoscience*, vol. 10, n°. 4, p. 255–259, avr. 2017. [En ligne]. Disponible : <https://www.nature.com/articles/ngeo2911>
- [25] A. Pérez Bello, A. Mailhot et D. Paquin, “The Response of Daily and Sub-Daily Extreme Precipitations to Changes in Surface and Dew-Point Temperatures,” *Journal of Geophysical Research : Atmospheres*, vol. 126, n°. 16, p. e2021JD034972, 2021, _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021JD034972>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2021JD034972>
- [26] C. Li, F. Zwiers, X. Zhang et G. Li, “How Much Information Is Required to Well Constrain Local Estimates of Future Precipitation Extremes?” *Earth’s Future*, vol. 7, n°. 1, p. 11–24, 2019, _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018EF001001>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2018EF001001>
- [27] N. A. Phillips, “The general circulation of the atmosphere : A numerical experiment,” *Quarterly Journal of the Royal Meteorological Society*, vol. 82, n°. 352, 1956. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.49708235202>

- [28] A. Dahan, “Putting the Earth System in a numerical box ? The evolution from climate modeling toward global change,” *Studies in History and Philosophy of Science Part B : Studies in History and Philosophy of Modern Physics*, vol. 41, n°. 3, p. 282–292, sept. 2010. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1355219810000511>
- [29] T. Schneider, J. Teixeira, C. S. Bretherton, F. Brient, K. G. Pressel, C. Schär et A. P. Siebesma, “Climate goals and computing the future of clouds,” *Nature Climate Change*, vol. 7, n°. 1, p. 3–5, janv. 2017, publisher : Nature Publishing Group. [En ligne]. Disponible : <https://www.nature.com/articles/nclimate3190>
- [30] V. Balaji, F. Couvreur, J. Deshayes, J. Gautrais, F. Hourdin et C. Rio, “Are general circulation models obsolete?” *Proceedings of the National Academy of Sciences*, vol. 119, n°. 47, p. e2202075119, nov. 2022, publisher : Proceedings of the National Academy of Sciences. [En ligne]. Disponible : <https://www.pnas.org/doi/abs/10.1073/pnas.2202075119>
- [31] G. M. Flato, “Earth system models : an overview,” *WIREs Climate Change*, vol. 2, n°. 6, p. 783–800, 2011. [En ligne]. Disponible : <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.148>
- [32] T. J. Crowley, “Causes of climate change over the past 1000 years,” *Science*, vol. 289, n°. 5477, p. 270–277, 2000. [En ligne]. Disponible : <https://www.science.org/doi/abs/10.1126/science.289.5477.270>
- [33] IPCC, *Climate Change 2022 : Impacts, Adaptation and Vulnerability : Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1^{er} éd. Cambridge University Press, juin 2023. [En ligne]. Disponible : <https://www.cambridge.org/core/product/identifier/9781009325844/type/book>
- [34] IPCC, *Climate Change 2022 : Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, P. R. Shukla, J. Skea, A. R. Reisinger et IPCC, édit. Geneva : Cambridge University Press, 2022. [En ligne]. Disponible : <https://www.cambridge.org/core/product/identifier/9781009157926/type/book>
- [35] S. Seneviratne, X. Zhang, M. Adnan, W. Badi, C. Dereczynski, A. Di Luca, S. Ghosh, I. Iskandar, J. Kossin, S. Lewis, F. Otto, I. Pinto, M. Satoh, S. Vicente-Serrano, M. Wehner et B. Zhou, *Weather and Climate Extreme Events in a Changing Climate*. Cambridge, United Kingdom and New York, NY, USA : Cambridge University Press, 2021, p. 1513–1766.
- [36] P. C. D. Milly, R. T. Wetherald, K. A. Dunne et T. L. Delworth, “Increasing risk of great floods in a changing climate,” *Nature*, vol. 415, n°. 6871, p.

- 514–517, janv. 2002, publisher : Nature Publishing Group. [En ligne]. Disponible : <https://www.nature.com/articles/415514a>
- [37] P. Neumann, P. Düben, P. Adamidis, P. Bauer, M. Brück, L. Kornblueh, D. Klocke, B. Stevens, N. Wedi et J. Biercamp, “Assessing the scales in numerical weather and climate predictions : will exascale be the rescue?” *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, vol. 377, n°. 2142, p. 20180148, févr. 2019, publisher : Royal Society. [En ligne]. Disponible : <https://royalsocietypublishing.org/doi/10.1098/rsta.2018.0148>
- [38] C. L. Walsh, D. Roberts, R. J. Dawson, J. W. Hall, A. Nickson et R. Hounscome, “Experiences of integrated assessment of climate impacts, adaptation and mitigation modelling in London and Durban,” *Environment & Urbanization*, vol. 25, n°. 2, p. 361–380, oct. 2013, publisher : SAGE Publications Ltd. [En ligne]. Disponible : <https://doi.org/10.1177/0956247813501121>
- [39] A. T. Ogunrinde, O. , Phillip G., A. , Akinola S., et J. T. Fasinmirin, “Evaluation of the impact of climate change on the characteristics of drought in Sahel Region of Nigeria : 1971–2060,” *African Geographical Review*, vol. 40, n°. 2, p. 192–210, avr. 2021, publisher : Routledge. [En ligne]. Disponible : <https://www.tandfonline.com/doi/full/10.1080/19376812.2020.1814826>
- [40] N. Pilguy, M. Taszarek, M. Kryza et H. E. Brooks, “Reconstruction of Violent Tornado Environments in Europe : High-Resolution Dynamical Downscaling of ERA5,” *Geophysical Research Letters*, vol. 49, n°. 11, p. e2022GL098242, 2022, _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2022GL098242>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2022GL098242>
- [41] A. Foughali, Y. Tramblay, Z. Bargaoui, J. Carreau et D. Ruelland, “Hydrological Modeling in Northern Tunisia with Regional Climate Model Outputs : Performance Evaluation and Bias-Correction in Present Climate Conditions,” *Climate*, vol. 3, n°. 3, p. 459–473, sept. 2015, number : 3 Publisher : Multidisciplinary Digital Publishing Institute. [En ligne]. Disponible : <https://www.mdpi.com/2225-1154/3/3/459>
- [42] J. Gaudreau, L. Perez et P. Drapeau, “BorealFireSim : A GIS-based cellular automata model of wildfires for the boreal forest of Quebec in a climate change paradigm,” *Ecological Informatics*, vol. 32, p. 12–27, mars 2016. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1574954115002022>
- [43] F. Giorgi et W. J. G. Jr, “Regional Dynamical Downscaling and the CORDEX Initiative,” *Annual Review of Environment and Resources*, vol. 40, n°. Volume 40, 2015, p. 467–490, nov. 2015, publisher : Annual Reviews. [En ligne]. Disponible : <https://www.annualreviews.org/content/journals/10.1146/annurev-environ-102014-021217>

- [44] F. Giorgi, “Thirty years of regional climate modeling : Where are we and where are we going next ?” *Journal of Geophysical Research : Atmospheres*, vol. 124, n°. 11, p. 5696–5723, 2019. [En ligne]. Disponible : <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD030094>
- [45] A. Martynov, R. Laprise, L. Sushama, K. Winger, L. Šeparović et B. Dugas, “Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5 : model performance evaluation,” *Climate Dynamics*, vol. 41, n°. 11, p. 2973–3005, déc. 2013. [En ligne]. Disponible : <https://doi.org/10.1007/s00382-013-1778-9>
- [46] D. Paquin, M. Giguère, C. McCray, O. Asselin, C. Gauthier, D. Matte, M.-P. Labonté et P. Bourgault, “CRCM5-CMIP6 : A dynamically-downscaled ensemble of CMIP6 simulations.” avr. 2025. [En ligne]. Disponible : <https://zenodo.org/records/15196952>
- [47] C. E. Iles, R. Vautard, J. Strachan, S. Joussaume, B. R. Eggen et C. D. Hewitt, “The benefits of increasing resolution in global and regional climate simulations for European climate extremes,” *Geoscientific Model Development*, vol. 13, n°. 11, p. 5583–5607, nov. 2020, publisher : Copernicus GmbH. [En ligne]. Disponible : <https://gmd.copernicus.org/articles/13/5583/2020/>
- [48] F. von Trentini, M. Leduc et R. Ludwig, “Assessing natural variability in RCM signals : comparison of a multi model EURO-CORDEX ensemble with a 50-member single model large ensemble,” *Climate Dynamics*, vol. 53, n°. 3, p. 1963–1979, août 2019. [En ligne]. Disponible : <https://doi.org/10.1007/s00382-019-04755-8>
- [49] D. Maraun, M. Widmann et J. M. Gutiérrez, “Statistical downscaling skill under present climate conditions : A synthesis of the VALUE perfect predictor experiment,” *International Journal of Climatology*, vol. 39, n°. 9, p. 3692–3703, 2019, _eprint : <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5877>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.5877>
- [50] H. Tabari, S. M. Paz, D. Buekenhout et P. Willems, “Comparison of statistical downscaling methods for climate change impact analysis on precipitation-driven drought,” *Hydrology and Earth System Sciences*, vol. 25, n°. 6, p. 3493–3517, 2021. [En ligne]. Disponible : <https://hess.copernicus.org/articles/25/3493/2021/>
- [51] C. O. de Burgh-Day et T. Leeuwenburg, “Machine learning for numerical weather and climate modelling : a review,” *Geoscientific Model Development*, vol. 16, n°. 22, p. 6433–6477, nov. 2023, publisher : Copernicus GmbH. [En ligne]. Disponible : <https://gmd.copernicus.org/articles/16/6433/2023/>
- [52] O. Watt-Meyer, G. Dresdner, J. McGibbon, S. K. Clark, B. Henn, J. Duncan, N. D. Brenowitz, K. Kashinath, M. S. Pritchard, B. Bonev, M. E. Peters et

- C. S. Bretherton, “ACE : A fast, skillful learned global atmospheric model for climate prediction,” déc. 2023, arXiv :2310.02074 [physics]. [En ligne]. Disponible : <http://arxiv.org/abs/2310.02074>
- [53] I. Price, A. Sanchez-Gonzalez, F. Alet, T. R. Andersson, A. El-Kadi, D. Masters, T. Ewalds, J. Stott, S. Mohamed, P. Battaglia, R. Lam et M. Willson, “Probabilistic weather forecasting with machine learning,” *Nature*, vol. 637, n°. 8044, p. 84–90, janv. 2025. [En ligne]. Disponible : <https://doi.org/10.1038/s41586-024-08252-9>
- [54] S. Rasp, M. S. Pritchard et P. Gentine, “Deep learning to represent subgrid processes in climate models,” *Proceedings of the National Academy of Sciences*, vol. 115, n°. 39, p. 9684–9689, 2018. [En ligne]. Disponible : <https://www.pnas.org/doi/abs/10.1073/pnas.1810286115>
- [55] N. Rampal, S. Hobeichi, P. B. Gibson, J. Baño-Medina, G. Abramowitz, T. Beucler, J. González-Abad, W. Chapman, P. Harder et J. M. Gutiérrez, “Enhancing regional climate downscaling through advances in machine learning,” *Artificial Intelligence for the Earth Systems*, vol. 3, n°. 2, p. 230066, 2024. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/aies/3/2/AIES-D-23-0066.1.xml>
- [56] A. W. Wood, E. P. Maurer, A. Kumar et D. P. Lettenmaier, “Long-range experimental hydrologic forecasting for the eastern united states,” *Journal of Geophysical Research : Atmospheres*, vol. 107, n°. D20, p. ACL–6, 2002.
- [57] M. Hessami, P. Gachon, T. Ouarda et A. St-Hilaire, “Automated regression-based statistical downscaling tool,” *Environmental Modelling & Software*, vol. 23, p. 813–834, 06 2008.
- [58] T. Vandal, E. Kodra, S. Ganguly, A. Michaelis, R. Nemani et A. R. Ganguly, “DeepSD : Generating High Resolution Climate Change Projections through Single Image Super-Resolution,” dans *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’17. New York, NY, USA : Association for Computing Machinery, 2017, p. 1663–1672. [En ligne]. Disponible : <https://dl.acm.org/doi/10.1145/3097983.3098004>
- [59] T. Vandal, E. Kodra et A. R. Ganguly, “Intercomparison of machine learning methods for statistical downscaling : the case of daily and extreme precipitation,” *Theoretical and Applied Climatology*, vol. 137, n°. 1, p. 557–570, juill. 2019. [En ligne]. Disponible : <https://doi.org/10.1007/s00704-018-2613-3>
- [60] B. Pan, K. Hsu, A. AghaKouchak et S. Sorooshian, “Improving Precipitation Estimation Using Convolutional Neural Network,” *Water Resources Research*, vol. 55, n°. 3, p. 2301–2321, 2019, _eprint : <https://agu->

- pubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018WR024090. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024090>
- [61] J. Baño-Medina, R. Manzananas et J. M. Gutiérrez, “Configuration and intercomparison of deep learning neural models for statistical downscaling,” *Geoscientific Model Development*, vol. 13, n°. 4, p. 2109–2124, 2020. [En ligne]. Disponible : <https://gmd.copernicus.org/articles/13/2109/2020/>
- [62] L. S. Passarella, S. Mahajan, A. Pal et M. R. Norman, “Reconstructing High Resolution ESM Data Through a Novel Fast Super Resolution Convolutional Neural Network (FSRCNN),” *Geophysical Research Letters*, vol. 49, n°. 4, p. e2021GL097571, 2022, _eprint : <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2021GL097571>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2021GL097571>
- [63] K. Höhle, M. Kern, T. Hewson et R. Westermann, “A comparative study of convolutional neural network models for wind field downscaling,” *Meteorological Applications*, vol. 27, n°. 6, p. e1961, nov. 2020. [En ligne]. Disponible : <https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.1961>
- [64] M. van der Meer, S. de Roda Husman et S. Lhermitte, “Deep learning regional climate model emulators : A comparison of two downscaling training frameworks,” *Journal of Advances in Modeling Earth Systems*, vol. 15, n°. 6, p. e2022MS003593, 2023, e2022MS003593 2022MS003593. [En ligne]. Disponible : <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022MS003593>
- [65] A. Doury, S. Somot et S. Gadat, “On the suitability of a convolutional neural network based RCM-emulator for fine spatio-temporal precipitation,” *Climate Dynamics*, juill. 2024. [En ligne]. Disponible : <https://doi.org/10.1007/s00382-024-07350-8>
- [66] T. Nguyen, J. Brandstetter, A. Kapoor, J. K. Gupta et A. Grover, “ClimaX : A foundation model for weather and climate,” dans *Proceedings of the 40th International Conference on Machine Learning*. PMLR, juill. 2023, p. 25 904–25 938, iSSN : 2640-3498. [En ligne]. Disponible : <https://proceedings.mlr.press/v202/nguyen23a.html>
- [67] X. Zhong, F. Du, L. Chen, Z. Wang et H. Li, “Investigating transformer-based models for spatial downscaling and correcting biases of near-surface temperature and wind-speed forecasts,” *Quarterly Journal of the Royal Meteorological Society*, vol. 150, n°. 758, p. 275–289, 2024, _eprint : <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4596>. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.4596>
- [68] Z. Liu, H. Chen, L. Bai, W. Li, W. Ouyang, Z. Zou et Z. Shi, “MambaDS : Near-Surface Meteorological Field Downscaling With Topography Constrained Selective State-Space

- Modeling,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, p. 1–15, 2024. [En ligne]. Disponible : <https://ieeexplore.ieee.org/document/10752514/>
- [69] Q. Yang, A. Hernandez-Garcia, P. Harder, V. Ramesh, P. Sattigeri, D. Szwarcman, C. D. Watson et D. Rolnick, “Fourier Neural Operators for Arbitrary Resolution Climate Data Downscaling,” *Journal Machine Learning Research*, n°. 25, 2024.
- [70] P. Hess et N. Boers, “Deep Learning for Improving Numerical Weather Prediction of Heavy Rainfall,” *Journal of Advances in Modeling Earth Systems*, vol. 14, n°. 3, p. e2021MS002765, mars 2022. [En ligne]. Disponible : <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002765>
- [71] P. Charbonnier, L. Blanc-Feraud, G. Aubert et M. Barlaud, “Two deterministic half-quadratic regularization algorithms for computed imaging,” dans *Proceedings of 1st International Conference on Image Processing*, vol. 2, nov. 1994, p. 168–172 vol.2. [En ligne]. Disponible : <https://ieeexplore.ieee.org/document/413553/>
- [72] S. Kohl, B. Romera-Paredes, C. Meyer, J. De Fauw, J. R. Ledsam, K. Maier-Hein, S. M. A. Eslami, D. Jimenez Rezende et O. Ronneberger, “A Probabilistic U-Net for Segmentation of Ambiguous Images,” dans *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., 2018. [En ligne]. Disponible : https://proceedings.neurips.cc/paper_files/paper/2018/hash/473447ac58e1cd7e96172575f48dca3b-Abstract.html
- [73] J. Ho, A. Jain et P. Abbeel, “Denoising Diffusion Probabilistic Models,” dans *Advances in Neural Information Processing Systems*, vol. 33. Curran Associates, Inc., 2020, p. 6840–6851. [En ligne]. Disponible : <https://proceedings.neurips.cc/paper/2020/hash/4c5bcfec8584af0d967f1ab10179ca4b-Abstract.html>
- [74] M. Mardani, N. Brenowitz, Y. Cohen, J. Pathak, C.-Y. Chen, C.-C. Liu, A. Vahdat, M. A. Nabian, T. Ge, A. Subramaniam, K. Kashinath, J. Kautz et M. Pritchard, “Residual corrective diffusion modeling for km-scale atmospheric downscaling,” *Communications Earth & Environment*, vol. 6, n°. 1, p. 1–10, févr. 2025, publisher : Nature Publishing Group. [En ligne]. Disponible : <https://www.nature.com/articles/s43247-025-02042-5>
- [75] Z. Y. Wan, R. Baptista, Y.-f. Chen, J. Anderson, A. Boral, F. Sha et L. Zepeda-Núñez, “Debias Coarsely, Sample Conditionally : Statistical Downscaling through Optimal Transport and Probabilistic Diffusion Models,” oct. 2023, arXiv :2305.15618 [cs]. [En ligne]. Disponible : <http://arxiv.org/abs/2305.15618>
- [76] H. Addison, E. Kendon, S. Ravuri, L. Aitchison et P. A. Watson, “Machine learning emulation of precipitation from km-scale regional climate simulations using

- a diffusion model,” juill. 2024, arXiv :2407.14158 [physics]. [En ligne]. Disponible : <http://arxiv.org/abs/2407.14158>
- [77] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville et Y. Bengio, “Generative Adversarial Nets,” dans *Advances in Neural Information Processing Systems*, vol. 27. Curran Associates, Inc., 2014. [En ligne]. Disponible : https://papers.nips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html
- [78] J.-Y. Zhu, T. Park, P. Isola et A. A. Efros, “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,” dans *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice : IEEE, oct. 2017, p. 2242–2251. [En ligne]. Disponible : <http://ieeexplore.ieee.org/document/8237506/>
- [79] P. Hess, M. Drücke, S. Petri, F. M. Strnad et N. Boers, “Physically constrained generative adversarial networks for improving precipitation fields from Earth system models,” *Nature Machine Intelligence*, vol. 4, n°. 10, p. 828–839, oct. 2022, publisher : Nature Publishing Group. [En ligne]. Disponible : <https://www.nature.com/articles/s42256-022-00540-1>
- [80] C. Charles, “Introduction aux problèmes inverses,” janv. 2014, note de statistique et d’informatique - Université de Liège.
- [81] J. Gonzalez-Abad, A. Hernandez-Garcia, P. Harder, D. Rolnick et J. M. Gutierrez, “Multi-Variable Hard Physical Constraints for Climate Model Downscaling,” *Proceedings of the AAAI Symposium Series*, vol. 2, n°. 1, p. 62–67, janv. 2024. [En ligne]. Disponible : <https://ojs.aaai.org/index.php/AAAI-SS/article/view/27650>
- [82] K. Kashinath, M. Mustafa, A. Albert, J.-L. Wu, C. Jiang, S. Esmailzadeh, K. Azizzadenesheli, R. Wang, A. Chattopadhyay, A. Singh, A. Manepalli, D. Chirila, R. Yu, R. Walters, B. White, H. Xiao, H. A. Tchelepi, P. Marcus, A. Anandkumar, P. Hassanzadeh et n. Prabhat, “Physics-informed machine learning : case studies for weather and climate modelling,” *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, vol. 379, n°. 2194, p. 20200093, févr. 2021, publisher : Royal Society. [En ligne]. Disponible : <https://royalsocietypublishing.org/doi/full/10.1098/rsta.2020.0093>
- [83] P. Harder, A. Hernandez-Garcia, V. Ramesh, Q. Yang, P. Sattigeri, D. Szwarcman, C. D. Watson et D. Rolnick, “Hard-constrained deep learning for climate downscaling,” *J. Mach. Learn. Res.*, vol. 24, n°. 1, p. 365 :17 534–365 :17 573, 2023.
- [84] F. Zanetta, D. Nerini, T. Beucler et M. A. Liniger, “Physics-Constrained Deep Learning Postprocessing of Temperature and Humidity,” déc. 2023, section :

- Artificial Intelligence for the Earth Systems. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/aies/2/4/AIES-D-22-0089.1.xml>
- [85] P. Jiang, Z. Yang, J. Wang, C. Huang, P. Xue, T. C. Chakraborty, X. Chen et Y. Qian, “Efficient Super-Resolution of Near-Surface Climate Modeling Using the Fourier Neural Operator,” *Journal of Advances in Modeling Earth Systems*, vol. 15, n°. 7, p. e2023MS003800, 2023. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1029/2023MS003800>
- [86] B. Teufel, F. Carmo, L. Sushama, L. Sun, M. N. Khaliq, S. Bélair, A. Shamseldin, D. N. Kumar et J. Vaze, “Physics-informed deep learning framework to model intense precipitation events at super resolution,” *Geoscience Letters*, vol. 10, n°. 1, p. 19, avr. 2023. [En ligne]. Disponible : <https://doi.org/10.1186/s40562-023-00272-z>
- [87] A. Saha et S. Ravela, “Statistical-physical adversarial learning from data and models for downscaling rainfall extremes,” *Journal of Advances in Modeling Earth Systems*, vol. 16, n°. 6, p. e2023MS003860, 2024, e2023MS003860 2023MS003860. [En ligne]. Disponible : <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2023MS003860>
- [88] Y. Verma, M. Heinonen et V. Garg, “ClimODE : Climate forecasting with physics-informed neural ODEs,” dans *The Twelfth International Conference on Learning Representations*, 2024. [En ligne]. Disponible : <https://openreview.net/forum?id=xuY33XhEGR>
- [89] T. Beucler, P. Gentine, J. Yuval, A. Gupta, L. Peng, J. Lin, S. Yu, S. Rasp, F. Ahmed, P. A. O’Gorman, J. D. Neelin, N. J. Lutsko et M. Pritchard, “Climate-invariant machine learning,” *Science Advances*, vol. 10, n°. 6, p. eadj7250, févr. 2024, publisher : American Association for the Advancement of Science. [En ligne]. Disponible : <https://www.science.org/doi/10.1126/sciadv.adj7250>
- [90] M. Leduc, A. Mailhot, A. Frigon, J.-L. Martel, R. Ludwig, G. B. Brietzke, M. Giguère, F. Brissette, R. Turcotte, M. Braun et J. Scinocca, “The climex project : A 50-member ensemble of climate change projections at 12-km resolution over europe and northeastern north america with the canadian regional climate model (crcm5),” *Journal of Applied Meteorology and Climatology*, vol. 58, n°. 4, p. 663 – 693, 2019. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/apme/58/4/jamc-d-18-0021.1.xml>
- [91] S. Elfving, E. Uchibe et K. Doya, “Sigmoid-weighted linear units for neural network function approximation in reinforcement learning,” *Neural Networks*, vol. 107, p. 3–11, 2018. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0893608017302976>

- [92] Y. Wu et K. He, “Group Normalization,” *International Journal of Computer Vision*, vol. 128, n°. 3, p. 742–755, mars 2020. [En ligne]. Disponible : <https://doi.org/10.1007/s11263-019-01198-w>
- [93] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever et R. Salakhutdinov, “Dropout : a simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, n°. 1, p. 1929–1958, janv. 2014.
- [94] K. He, X. Zhang, S. Ren et J. Sun, “Delving deep into rectifiers : Surpassing human-level performance on imagenet classification,” dans *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, p. 1026–1034.
- [95] J.-S. Giroux, S.-P. Breton et J. Carreau, “Interpolation-free deep learning for meteorological downscaling on unaligned grids across multiple domains with application to wind power,” 2024. [En ligne]. Disponible : <https://arxiv.org/abs/2410.03945>
- [96] D. P. Kingma et J. Ba, “Adam : A method for stochastic optimization,” 2017. [En ligne]. Disponible : <https://arxiv.org/abs/1412.6980>
- [97] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell et S. Xie, “A convnet for the 2020s,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.
- [98] R. H. Byrd, P. Lu, J. Nocedal et C. Zhu, “A limited memory algorithm for bound constrained optimization,” *SIAM Journal on Scientific Computing*, vol. 16, n°. 5, p. 1190–1208, 1995. [En ligne]. Disponible : <https://doi.org/10.1137/0916069>
- [99] A. Pérez, M. S. Cruz, D. S. Martín et J. M. Gutiérrez, “Transformer based super-resolution downscaling for regional reanalysis : Full domain vs tiling approaches,” oct. 2024. [En ligne]. Disponible : <http://arxiv.org/abs/2410.12728>
- [100] E. Kjellström, F. Boberg, M. Castro, J. H. Christensen, G. Nikulin et E. Sánchez, “Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models,” *Climate Research*, vol. 44, n°. 2-3, p. 135–150, déc. 2010. [En ligne]. Disponible : <https://www.int-res.com/abstracts/cr/v44/n2-3/p135-150/>

ANNEXE A ESTIMATION DES TPSRS PAR LA LOI DES VALEURS EXTRÊMES GÉNÉRALISÉE

Les figures A.1, A.2, A.3, A.4, A.5, et A.6 présentent respectivement les NLLs, et les paramètres μ_0 , μ_1 , σ_0 , σ_1 et ξ estimés à partir des données de référence et des prédictions du UNet pour la période de test 2020 – 2029.

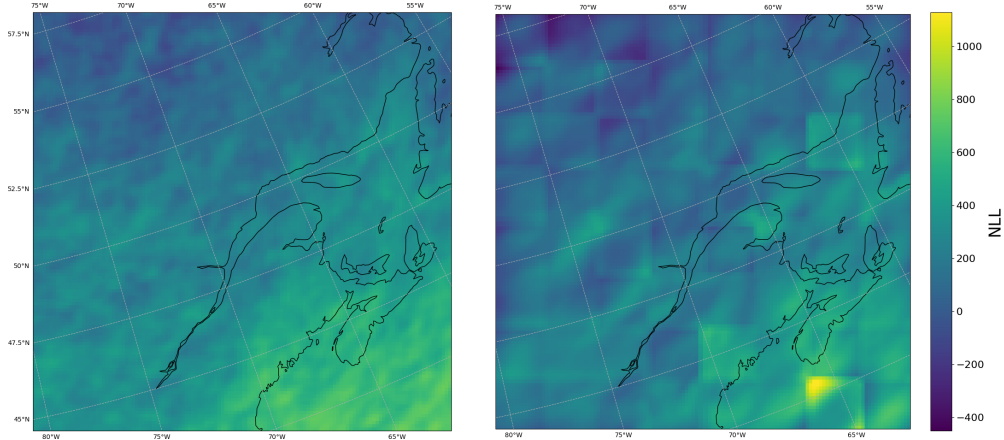


FIGURE A.1 Log-vraisemblance négative à l'issue de l'optimisation des paramètres sous la combinaison $M_{\mu,4} + M_{\sigma,2}$ pour les données de référence (**à gauche**) et les prédictions du UNet (**à droite**).

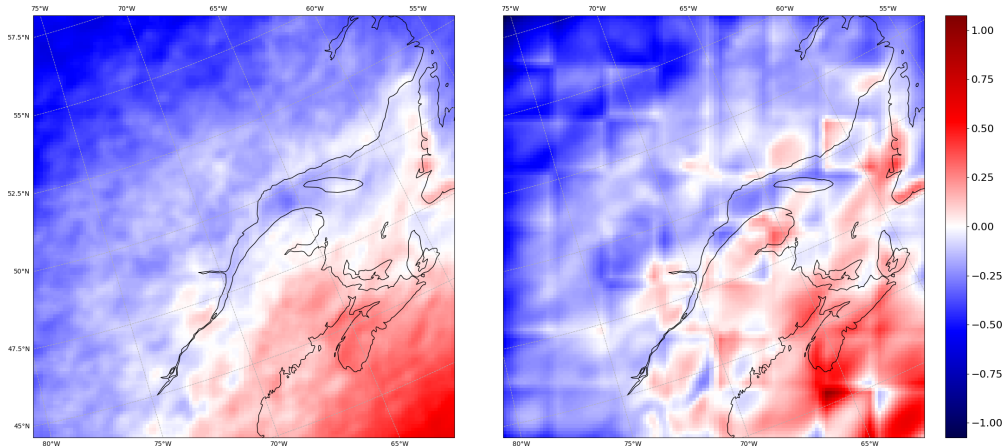


FIGURE A.2 Paramètres μ_0 de $M_{\mu,4}$ pour les données de référence (**à gauche**) et les prédictions du UNet (**à droite**).

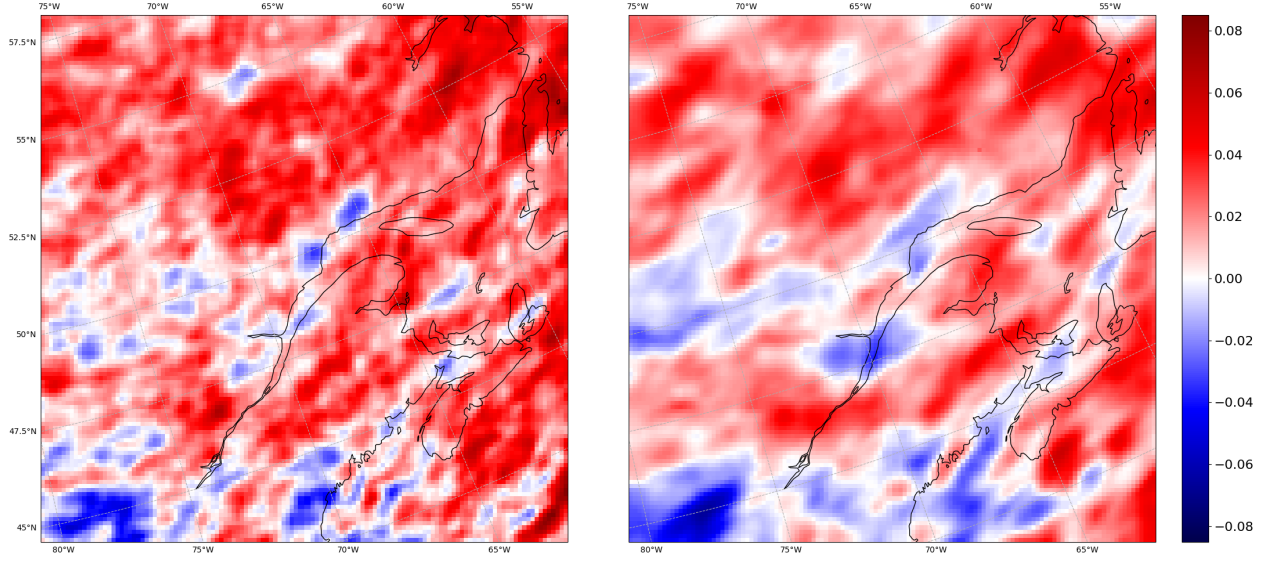


FIGURE A.3 Paramètres μ_1 de $M_{\mu,4}$ pour les données de référence (**à gauche**) et les prédictions du UNet (**à droite**).

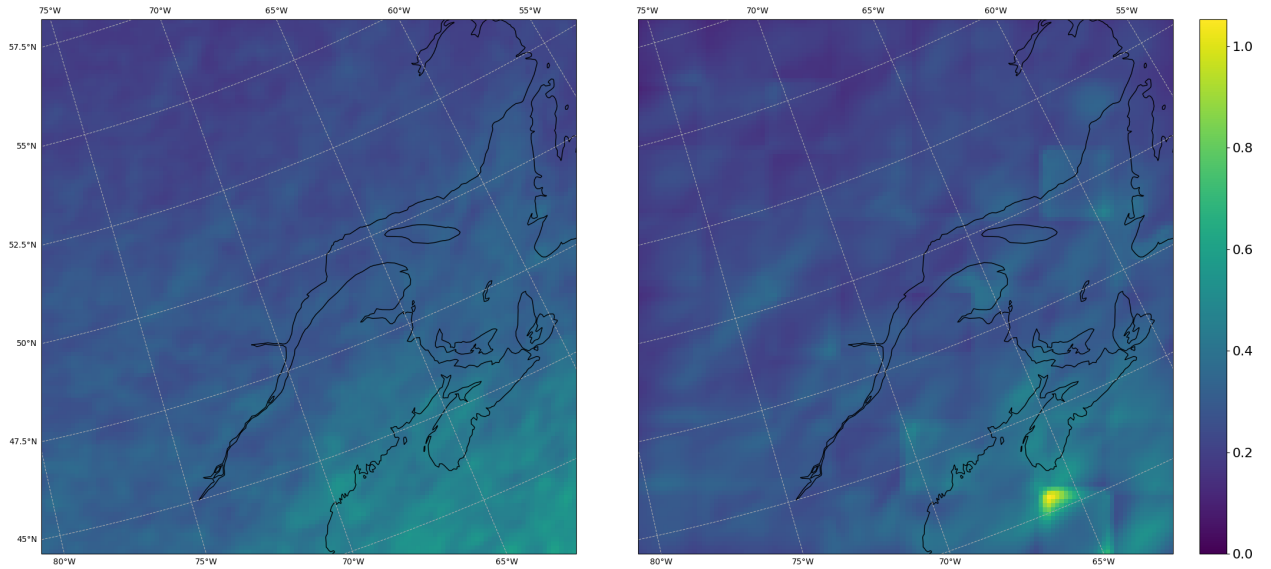


FIGURE A.4 Paramètres σ_0 de $M_{\sigma,2}$ pour les données de référence (**à gauche**) et les prédictions du UNet (**à droite**).

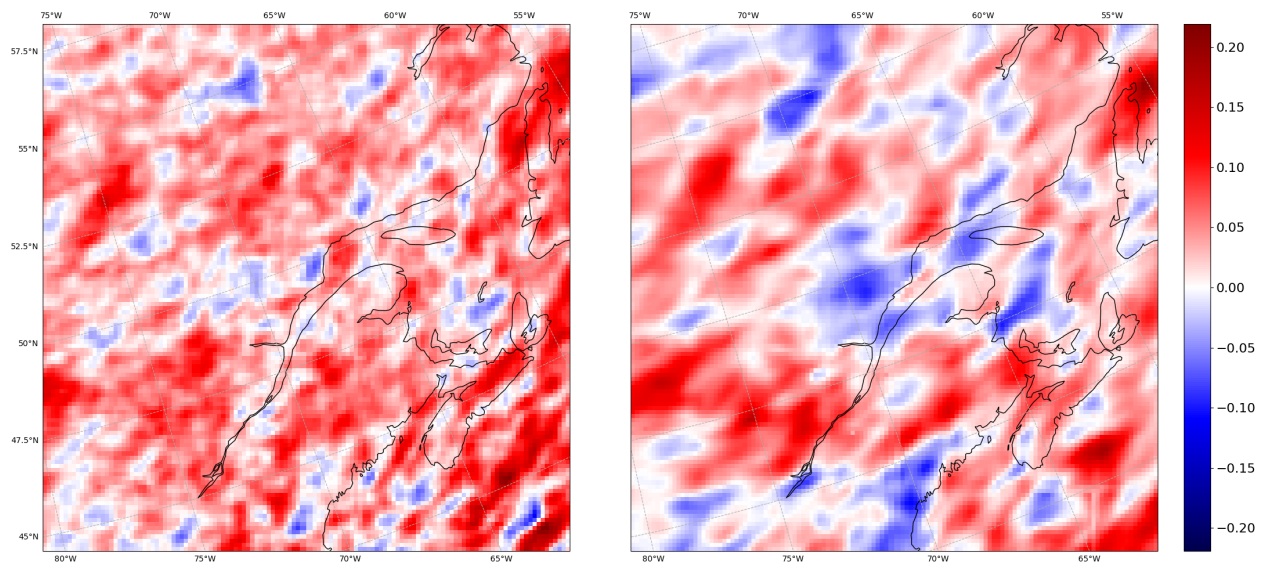


FIGURE A.5 Paramètres σ_1 de $M_{\sigma,2}$ pour les données de référence (**à gauche**) et les prédictions du UNet (**à droite**).

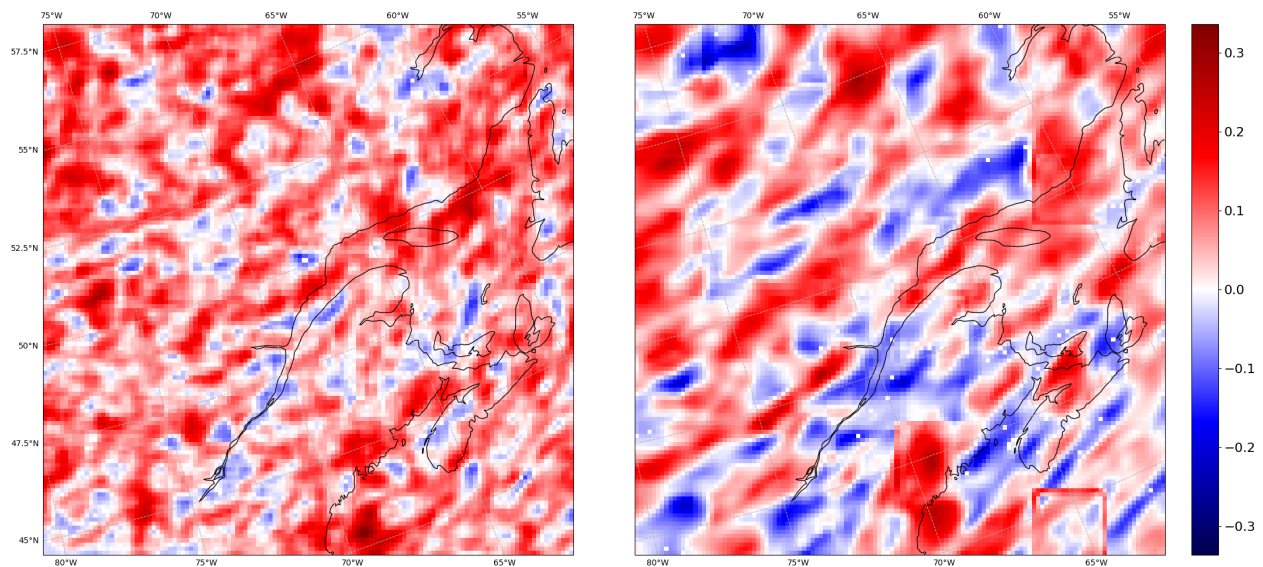


FIGURE A.6 Paramètres ξ pour les données de référence (**à gauche**) et les prédictions du UNet (**à droite**).