

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

ÉCOLE NATIONALE POLYTECHNIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



LABORATOIRE MATÉRIAUX DE GÉNIE CIVIL
ET ENVIRONNEMENT

Mémoire de projet de fin d'études

Pour l'obtention du diplôme d'ingénieur d'état en Hydraulique

Anticipation de l'impact du changement climatique sur
les sécheresses dans le nord de l'Algérie à l'aide du
downscaling basé sur l'intelligence artificielle

BOUHA Amel et OUALI Hamza

Sous la direction de Mr. BERMAD Abdelmalek

Présenté et soutenu publiquement le (8/07/2025)

Composition du jury :

Président	Mr. LEFKIR Abdelouahab	Professeur	ENSTP
Promoteur	Mr. BERMAD Abdelmalek	Professeur	ENP
Co-Promoteur	Mr. ALIANE Youcef	Doctorant	U. Annaba
Examineur	Mr. OULEBSIR Rafik	Docteur/ MCB	USTHB

ENP 2025

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

ÉCOLE NATIONALE POLYTECHNIQUE



المدرسة الوطنية المتعددة التقنيات
Ecole Nationale Polytechnique



LABORATOIRE MATÉRIAUX DE GÉNIE CIVIL
ET ENVIRONNEMENT

Mémoire de projet de fin d'études

Pour l'obtention du diplôme d'ingénieur d'état en Hydraulique

Anticipation de l'impact du changement climatique sur
les sécheresses dans le nord de l'Algérie à l'aide du
downscaling basé sur l'intelligence artificielle

BOUHA Amel et OUALI Hamza

Sous la direction de Mr. BERMAD Abdelmalek

Présenté et soutenu publiquement le (8/07/2025)

Composition du jury :

Président	Mr. LEFKIR Abdelouahab	Professeur	ENSTP
Promoteur	Mr. BERMAD Abdelmalek	Professeur	ENP
Co-Promoteur	Mr. ALIANE Youcef	Doctorant	U. Annaba
Examineur	Mr. OULEBSIR Rafik	Docteur/ MCB	USTHB

ENP 2025

ملخص :

يُعدّ الجفاف أحد أبرز التهديدات المناخية المُلحة في شمال إفريقيا، ولا سيما في الجزائر، حيث تُشكل التقلّبات المتزايدة في كميات الأمطار تحديات جسيمة للزراعة، وإدارة الموارد المائية، والاستقرار الاجتماعي والاقتصادي. وفي هذا السياق، تتناول هذه الدراسة حالات الجفاف الماضية والمتوقعة في حوض سد قدارة (الجزائر)، في ظلّ التحديات المتزايدة المرتبطة بتغير المناخ. ولمعالجة محدودية النماذج المناخية العالمية، التي غالبًا ما تفتقر إلى الدقة على المستوى المحلي، تم استخدام تقنيات تقليص النطاق (Downscaling) القائمة على خوارزميات التعلم الآلي (XGBoost و CNN)، وذلك لتحسين دقة التوقعات المناخية. ولتحريز هذا التحليل، استُخدم مؤشر شدة الجفاف المعياري (SPI) لتقييم تواتر الجفاف، وشِدته، ومدّته، وذلك للتوقعات المناخية خلال الفترة الممتدة من 2026 إلى 2100، وفقًا لثلاثة سيناريوهات مناخية مختلفة (SSP1-2.6، SSP2-4.5، SSP5-8.5). وقد أظهرت النتائج اتجاهًا واضحًا نحو تفاقم موجات الجفاف في المستقبل. وفي ضوء هذه النتائج، قُدمت توصيات عملية لتعزيز القدرة على الصمود في وجه التحديات المائية والزراعية، بما يتماشى مع أهداف التنمية المستدامة وتوصيات الهيئة الحكومية الدولية المعنية بتغير المناخ (GIEC).

الكلمات المفتاحية: الجفاف، تغيّر المناخ، CNN، XGBoost، downscaling، مؤشر SPI، التوقعات المناخية، GIEC.

Abstract

Drought constitutes one of the most urgent climate threats in North Africa, particularly in Algeria, where the increasing variability of precipitation poses serious challenges to agriculture, water management, and socio-economic stability. In this context, this study examines past and future droughts in the Keddara Dam watershed (Algeria), in the face of growing challenges related to climate change. In order to overcome the limitations of global climate models, whose resolution is often insufficient at the local scale, downscaling techniques based on machine learning (XGBoost, CNN) were used to refine climate projections. To support this analysis, the SPI index was used for climate projection to assess the frequency, intensity, and duration of droughts over the 2026—2100 period, according to three contrasting climate scenarios (SSP1—2.6, SSP2—4.5, SSP5—8.5). The results obtained highlight a clear trend towards the worsening of future drought episodes. In view of these findings, concrete recommendations are formulated to strengthen water and agricultural resilience, in line with the Sustainable Development Goals and the GIEC guidelines.

0.5cm]

Keywords : Drought - climate change - downscaling - XGBoost - CNN - SPI - climate projections - GIEC.

Résumé

La sécheresse constitue l'une des menaces climatiques les plus urgentes en Afrique du Nord, en particulier en Algérie, où la variabilité croissante des précipitations pose de sérieux défis à l'agriculture, à la gestion de l'eau et à la stabilité socio-économique. Dans ce contexte, cette étude examine les sécheresses passées et futures dans le bassin versant du barrage de Keddara (Algérie), face aux défis croissants liés au changement climatique. Afin de pallier les limites des modèles climatiques globaux, dont la résolution est souvent insuffisante à l'échelle locale, des techniques de downscaling basées sur l'apprentissage automatique (XGBoost, CNN) ont été mobilisées pour affiner les projections climatiques. Pour appuyer cette analyse, l'indice SPI a été utilisé pour la projection climatique afin d'évaluer la fréquence, l'intensité et la durée

des sécheresses sur la période 2026–2100, selon trois scénarios climatiques contrastés (SSP1–2.6, SSP2–4.5, SSP5–8.5). Les résultats obtenus mettent en évidence une tendance claire à l’aggravation des épisodes de sécheresse à venir. Face à ces constats, des recommandations concrètes sont formulées pour renforcer la résilience hydrique et agricole, en cohérence avec les objectifs du développement durable et les orientations du GIEC.

Mots clés : Sécheresse - changement climatique - downscaling - XGBoost - CNN - SPI - projections climatiques - GIEC.

Dédicace

«Chaque personne que l'on croise nous enseigne quelque chose.»

À vous toutes et tous, merci.

Hamza et Amel

Remerciements

Nous exprimons d'abord notre plus profonde gratitude à Allah, Le Tout-Puissant, pour nous avoir guidées tout au long de ce parcours, pour nous avoir accordé la santé, la force et la volonté nécessaires afin de mener à bien ce travail, depuis ses premiers pas jusqu'à son achèvement.

Nous tenons à adresser nos plus sincères remerciements à notre encadrant, Monsieur Abdelmalek BERMAD, pour sa confiance, ses conseils avisés et son accompagnement constant tout au long de cette aventure scientifique. Son soutien, à la fois humain et professionnel, a été une source de motivation et d'enrichissement continu. Il a su nous orienter avec rigueur et bienveillance, et nous lui en sommes profondément reconnaissantes.

Nos remerciements chaleureux vont également à Monsieur Youcef ALIANE, pour le temps qu'il nous a consacré, son assistance précieuse ainsi que pour les outils méthodologiques essentiels qu'il nous a transmis, sans lesquels ce travail n'aurait pu aboutir.

Nous exprimons également notre profonde reconnaissance à Monsieur Rafik OULEBSIR, pour son soutien précieux dans l'orientation en programmation. Sa disponibilité et ses conseils techniques ont grandement facilité la mise en œuvre des méthodes utilisées dans cette étude.

Nous remercions sincèrement les membres du jury d'avoir pris le temps et fait preuve de patience pour examiner ce travail. Leurs remarques, critiques constructives et suggestions ne manqueront pas de contribuer à son amélioration et à notre progression personnelle.

Nos pensées reconnaissantes vont à nos familles respectives, pour leur soutien inconditionnel, leur patience et leur présence indéfectible durant les moments de doute et de difficulté. Leur appui moral et affectif a été d'un immense réconfort.

À nos ami(e)s et à toutes les personnes qui nous ont soutenues et encouragées tout au long de cette aventure académique, votre soutien moral et vos discussions enrichissantes ont été d'une valeur inestimable pour nous. En particulier, nous tenons à remercier Mademoiselle Aya CHENNIT et Monsieur Mohamed Yanis BENNABI, sans lesquels nous n'aurions pas pu aller jusqu'au bout.

Nous remercions également toutes les personnes qui s'intéressent à ce travail, avec l'espoir qu'elles y trouveront des éléments utiles pour leurs propres recherches ou projets.

Enfin, à toutes celles et ceux qui ont, de près ou de loin, contribué à l'aboutissement de ce travail, nous exprimons notre respect, notre gratitude et nos remerciements les plus sincères.

Le meilleur reste à venir.

Table des matières

Liste des tableaux

Table des figures

Liste des acronymes

Introduction générale	14
1 GÉNÉRALITÉS SUR LES SÉCHERESSES	17
1.1 Introduction	17
1.2 Changement Climatique	17
1.2.1 Définition du changement climatique	17
1.2.2 Impact du changement climatique	18
1.3 La sécheresse	18
1.3.1 Définition	18
1.3.2 Caractéristiques de la sécheresse	19
1.3.3 Types de sécheresse	19
1.4 Impacts de la sécheresse à différentes échelles	21
1.4.1 Sécheresse dans le monde	21
1.4.2 Les sécheresses en Algérie	23
1.5 Les indices de sécheresse	24
1.5.1 Palmer Drought Severity Index (PDSI) (GUTTMAN N. B., 1998)	25
1.5.2 Standardized Precipitation Index (SPI) (Gutmann, NB 1998, Giddings, L et al., 2005)	25
1.6 Conclusion	27
2 BASES CONCEPTUELLES DU DOWNSCALING STATISTIQUE	28

2.1	Introduction	28
2.2	Les données de réanalyse	28
2.3	Les modèles climatiques globaux (GCMs)	29
2.3.1	Description des modèles climatiques globaux	29
2.3.2	Définition et rôle des composantes des variantes RANs dans CMIP6 . . .	31
2.4	Méthodes de désagrégation spatiale du climat (Downscaling) :	31
2.4.1	Principe du Downscaling	31
2.4.2	Statistical Downscaling	32
2.5	Le Quantile Mapping (QM)	33
2.6	Conclusion	35
3	APPRENTISSAGE AUTOMATIQUE (MACHINE LEARNING)	36
3.1	Introduction	36
3.2	Définitions	37
3.2.1	Intelligence Artificielle (Artificial Intelligence)	37
3.2.2	Apprentissage Automatique (Machine Learning)	37
3.2.3	Apprentissage Profond (Deep Learning)	37
3.3	Classification des types d'apprentissage automatique	37
3.3.1	Extreme Gradient Boosting – Xgboost	38
3.3.2	Convolutional Neural Networks (CNN)	41
3.4	Critères d'évaluation	45
3.4.1	Critere Numerique	45
3.4.2	Critere Graphique	47
3.5	Conclusion	47
4	MATÉRIELS ET MÉTHODES	49
4.1	Introduction	49
4.2	Région d'Étude	49
4.3	Méthodologie	50
4.4	Données utilisées pour le downscaling statistique	52
4.4.1	Données de référence locales et calcul du SPI :	53

4.4.2	Données à l'échelle globale : Sélection des modèles climatiques globaux (GCM)	57
4.5	Le Quantile Mapping (QM) comme méthode d'ajustement post-downscaling . .	61
4.6	Techniques de Downscaling Statistique	62
4.6.1	Approche XGBoost (Extreme Gradient Boosting)	62
4.6.2	Approche par Réseaux de Neurones Convolutifs (CNN)	64
4.7	Conclusion	67
5	RÉSULTATS ET DISCUSSION	68
5.1	Introduction	68
5.2	Évolution des sécheresses au barrage de Keddara (1951–2013)	68
5.2.1	Interprétation des résultats calculés	69
5.2.2	Validation des calculs du SPI à partir des constats historique	71
5.3	Évaluation comparative des modèles	72
5.3.1	Critères Numériques	72
5.3.2	Critères Graphiques	76
5.3.3	Résultats de l'évaluation	81
5.4	Conclusion	82
6	PROJECTIONS CLIMATIQUES FUTURES	83
6.1	Introduction	83
6.2	Évaluation du Quantile Mapping	84
6.3	Description des scénarios climatiques (SSP)	86
6.4	Projections futures de la sécheresse au barrage de Keddara	86
6.4.1	Scénario SSP1–2.6	87
6.4.2	Scénario SSP2–4.5	88
6.4.3	Scénario SSP5–8.5	89
6.5	Identification des dangers révélés par les résultats obtenus	90
6.5.1	Une vulnérabilité hydrique croissante	91
6.5.2	Impacts sectoriels – Le cas critique de l'agriculture	91
6.5.3	Analyse sectorielle des impacts de la sécheresse à Keddara et dans les régions semi-arides d'Algérie	92

6.6	Plan National Sécheresse (Algérie, 2019)	93
6.6.1	Définition du Plan National de la Sécheresse (PNS)	93
6.6.2	Objectifs principaux du PNS	93
6.6.3	Recommandation institutionnelle	93
6.6.4	Missions du groupe thématique « sécheresse »	94
6.7	Révision du Plan National de Sécheresse	94
6.8	Perspectives pour l'Algérie à l'horizon 2030	96
6.8.1	Promouvoir l'agriculture de conservation des sols et L'agroforesterie . . .	96
6.8.2	Mobiliser la finance pour l'adaptation agricole	97
6.8.3	Renforcer le relais social pour la résilience hydrique	97
6.9	Conclusion	97
CONCLUSION GÉNÉRALE		99
Bibliographie		101

Liste des tableaux

3.1	Intervalles d'appréciation des critères de validation pour variable précipitation .	47
3.2	Intervalles d'appréciation des critères de validation pour la variable SPI.	47
4.1	Tableau récapitulatif des meilleures distributions statistiques ajustées aux séries de précipitations (in situ et réanalyses)	56
4.2	Caractéristiques des modèles climatiques globaux (GCM) et des réalisations sélectionnées pour l'étude	59
4.3	Liste des variables prédictives issues des GCM et leur influence sur les précipitations	60
4.4	Hyperparamètres du Modèle	63
4.5	Architecture du modèle CNN utilisé	65
5.1	Performances du modèle CNN_sr sur la précipitation	72
5.2	Performances du modèle Xgboost_sur la précipitation	72
5.3	Performances du modèle CNN sur l'indice SPI-1	73
5.4	Performances du modèle XGBoost sur l'indice SPI-1	73
5.5	Performances du modèle CNN sur l'indice SPI-3	74
5.6	Performances du modèle XGBoost sur l'indice SPI-3	74
5.7	Performances du modèle CNN sur l'indice SPI-6	75
5.8	Performances du modèle XGBoost sur l'indice SPI-6	75
5.9	Performances du modèle CNN sur l'indice SPI-12	76
5.10	Performances du modèle XGBoost sur l'indice SPI-12	76
6.1	Balance alimentaire céréalière (blé) en Algérie [1].	91

Table des figures

1.1	Types de sécheresse, manifestations dans le temps et leurs impacts [2]	21
2.1	Representation of GCMs and RCMs. Source : Ambrizzi et al. (2018)[3]	30
2.2	Illustration du passage de l'échelle grossière à l'échelle fine	33
2.3	Quantile Mapping : une valeur simulée, correspondant à un certain quantile de la distribution simulée, est remplacée par le quantile de la distribution observée ayant la même probabilité cumulée.	34
3.1	Hierarchie des sous-domaines de l'intelligence artificielle [4]	36
3.2	Classification des types d'apprentissage automatique (Machine Learning) [5] . .	38
3.3	Schéma des arbres Xgboost [6]	39
3.4	Conv1D : Convolution on time dimension	43
3.5	Illustration du Max Pooling avec un filtre 2×2 et un pas de 2×2 appliqué à une matrice d'entrée [7]	43
3.6	Représentation schématique de l'architecture du modèle.	44
4.1	Localisation géographique et délimitation du lac de retenue du barrage de Ked-dara [8]	50
4.2	Shéma de la méthodologie	52
4.3	Calcul du SPI multiéchelle à partir des données in situ et des données de réanalyse	57
4.4	Interface du portail ESGF-IPSL pour la recherche et le téléchargement des simulations climatiques CMIP6	58
4.5	Schéma récapitulatif de l'algorithme XGBoost	64
4.6	schéma récapitulatif du modele CNN	66
5.1	Évolution Temporelle SPI-1 pour la période [1951–2013]	69
5.2	Évolution Temporelle SPI-3 pour la période [1951–2013]	69
5.3	Évolution Temporelle SPI-6 pour la période [1951–2013]	70

5.4	Évolution Temporelle SPI-12 pour la période [1951–2013]	70
5.5	QQPLOT de quelques exemples de GCMS de la variable précipitation	76
5.6	QQPLOT de quelques exemples de GCMS de la variable précipitation	77
5.7	QQPLOT de quelques exemples de GCMS de l'indice SPI	78
5.8	QQPLOT de quelques exemples de GCMS de l'indice SPI	79
5.9	Évolution temporelle de quelques exemples de GCMs pour la précipitation . . .	79
5.10	Évolution temporelle de quelques exemples de GCMs pour l'indice SPI	80
5.11	Évolution temporelle de quelques exemples de GCMs pour l'indice SPI	81
6.1	Distribution comparative des données brutes (GCM), downscalées (CNN), et ajustées (QM)	85
6.2	Évolution temporelle pour les différents SPI selon le scénario SSP1-2.6	88
6.3	Évolution Temporelle pour les différents SPI selon le scénario ssp245	89
6.4	Évolution Temporelle pour les différents SPI selon le scénario ssp585	90
6.5	Structure organisationnelle du groupe (sécheresse)	94
6.6	Principales composantes pour une action résilient	94

Liste des acronymes

- **GIEC** : Groupe d'experts intergouvernemental sur l'évolution du climat
- **OMM** : Organisation météorologique mondiale
- **ISO** : Organisation internationale de normalisation
- **ONU** : Organisation des Nations Unies
- **FAO** : Food and Agriculture Organization
- **ANRH** : Agence Nationale des Ressources Hydrauliques (Algérie)
- **GCM** : Global Climate Model / Modèle Climatique Global
- **RCM** : Regional Climate Model
- **CMIP6** : Coupled Model Intercomparison Project Phase 6
- **PDSI** : Palmer Drought Severity Index
- **SPI** : Standardized Precipitation Index
- **AED** : Atmospheric Evaporative Demand
- **SD** : Statistical Downscaling
- **QM** : Quantile Mapping
- **CNN** : Convolutional Neural Network
- **RNN** : Recurrent Neural Network
- **LSTM** : Long Short-Term Memory
- **GBM** : Gradient Boosting Machine
- **SSP** : Shared Socioeconomic Pathways
- **PNS** : Plan National de la Sécheresse
- **PNC** : Plan National Climat
- **CNC** : Comité National Climat
- **ODD** : Objectifs de Développement Durable

Introduction générale

La sécheresse est l'un des phénomènes climatiques les plus dévastateurs affectant les régions arides et semi-arides du globe, tant par sa fréquence croissante que par ses impacts multiformes sur les ressources en eau, l'agriculture, les écosystèmes, l'économie et les conditions de vie des populations.

En Algérie, pays dont plus de 85 % du territoire est soumis à un climat aride à très aride, la sécheresse constitue un risque majeur, avec des répercussions directes sur la sécurité hydrique et alimentaire. Ce phénomène est aujourd'hui exacerbé par les effets du changement climatique mondial, qui amplifie la variabilité des précipitations et l'intensité des vagues de chaleur.

Au cours des dernières décennies, la région a connu une diminution sensible des précipitations et une élévation progressive des températures moyennes, entraînant une intensification des épisodes de sécheresse. Cette tendance, largement documentée par les études nationales et internationales, affecte le fonctionnement des infrastructures hydrauliques (notamment les barrages), la recharge des nappes souterraines, les rendements agricoles, la biodiversité, ainsi que l'accès à l'eau potable pour les populations. Face à ces défis, il devient crucial de mieux comprendre la dynamique passée des sécheresses et d'anticiper leur évolution future.

Le sixième rapport d'évaluation du **GIEC** (2023a) met clairement en évidence l'aggravation attendue des événements climatiques extrêmes, en particulier les sécheresses, dans un contexte de réchauffement global. Alors que le seuil de $+1,5^{\circ}\text{C}$, fixé par l'Accord de Paris comme limite de sécurité, semblait devoir être franchi vers 2040 selon les projections antérieures (GIEC, 2019), les nouvelles estimations indiquent un dépassement probable entre 2030 et 2035. Un rapport publié le 28 mai 2025 par l'**Organisation Météorologique Mondiale (OMM)** estime même qu'il existe une probabilité de 70 % que la température moyenne mondiale dépasse temporairement ce seuil dès la période 2025–2029, et 86 % de chance qu'au moins une année durant cette période dépasse ce seuil. De plus, bien que plus faible, un risque de 1 % existe pour qu'au moins une année dépasse les $+2^{\circ}\text{C}$ d'ici 2029. Il est important de noter que ce dépassement temporaire ne signifie pas nécessairement que l'objectif des $+1,5^{\circ}\text{C}$ de l'Accord de Paris est perdu, mais il constitue un signe précurseur d'un franchissement durable du seuil de sécurité climatique si des mesures urgentes ne sont pas prises.

En Afrique du Nord, et plus particulièrement au Maghreb, les projections à l'horizon 2050 indiquent une intensification des sécheresses, avec des conséquences sévères sur l'agriculture, les ressources hydriques, les écosystèmes et la sécurité alimentaire. Des baisses importantes des rendements agricoles sont anticipées en cas de réchauffement de $+3^{\circ}\text{C}$: -17% pour le blé, -5% pour le maïs, -15% pour le sorgho, et jusqu'à -40% pour le millet. Le Maghreb figure ainsi parmi les régions les plus vulnérables aux effets du changement climatique, quels que soient les scénarios de forçage radiatif envisagés (GIEC, 2023b).

En Algérie, cette vulnérabilité est particulièrement marquée. Seuls 13% du territoire, localisés dans la bande nord, bénéficient de précipitations suffisantes pour répondre aux besoins

domestiques, agricoles et industriels. La façade littorale, relativement arrosée, subit une pression croissante liée à la densité de population et aux activités économiques. Les hautes plaines et les steppes, essentielles pour la production céréalière et pastorale, souffrent d'un déficit hydrique chronique et d'une désertification avancée. Le Sahara, qui couvre plus de 85% du territoire national, dépend quant à lui exclusivement des ressources souterraines et est de plus en plus exposé à des vagues de chaleur extrême. À l'horizon 2100, les projections annoncent une hausse des températures moyennes comprise entre +2°C et +3°C, accompagnée d'une baisse pouvant atteindre 40% des précipitations annuelles, accentuant ainsi les déséquilibres déjà existants entre l'offre et la demande en eau.

Face à ces signaux d'alerte, les stratégies d'adaptation doivent reposer sur une approche intégrée, territorialisée et anticipative, visant à renforcer la résilience des territoires.. Le GIEC recommande notamment la mise en place de systèmes d'alerte précoce, le renforcement de la résilience des systèmes agricoles, la généralisation de l'irrigation raisonnée, ainsi que le recours à des solutions fondées sur la nature. Il insiste également sur les risques de maladaptation, comme l'irrigation intensive dans des zones à forte probabilité de sécheresse, qui peut conduire à une surexploitation des nappes et à des effets contre-productifs.

Problématique centrale

Quelles sont les vulnérabilités spécifiques liées à la sécheresse dans la région du Centre-Nord algérien, et quelles mesures d'adaptation peuvent être envisagées pour y faire face efficacement selon l'intensité des risques projetés ?

Objectifs et démarche

C'est dans ce contexte que s'inscrit le présent travail, qui vise à analyser en profondeur les sécheresses passées et futures en Algérie, à partir d'un cas d'étude représentatif : le bassin versant du barrage de Keddara, situé dans la région nord-centre du pays . Ce mémoire est scindé en six chapitres dans lesquels nous avons exposé l'ensemble du travail réalisé :

Après une brève introduction générale où on a mis l'accent sur la présentation du contexte climatique, des enjeux liés à la sécheresse et des objectifs du travail.

Le chapitre 1 constitue le socle théorique de ce mémoire. Il explore en profondeur la notion de sécheresse, en en définissant les types, les mécanismes et les impacts, tout en mettant en évidence son aggravation sous l'effet du changement climatique. Cette contextualisation est essentielle pour appréhender la complexité du phénomène dans un cadre scientifique rigoureux.

Le chapitre 2 prolonge cette base en présentant les outils scientifiques et les données climatologiques mobilisés pour l'étude. Il expose les spécificités des modèles climatiques globaux (GCMs) et des données de réanalyse, tout en introduisant les méthodes de downscaling statistique, avec un accent particulier sur la méthode du Quantile Mapping, reconnue pour son efficacité dans l'amélioration des projections climatiques à l'échelle locale.

Dans la continuité, le chapitre 3 s'attache à poser les fondements du recours à l'intelligence artificielle, en particulier au Machine Learning, dans le traitement des données climatiques. Il justifie le choix des algorithmes XGBoost et CNN, en mettant en lumière l'importance de l'optimisation des hyperparamètres, des fonctions de perte et des métriques d'évaluation pour garantir des modèles robustes et précis.

S'appuyant sur ces éléments, le chapitre 4 élabore une méthodologie intégrée combinant downscaling statistique et apprentissage automatique pour améliorer la précision des projections locales. Il teste notamment l'efficacité de l'ajustement par Quantile Mapping en phase post-

downscaling, afin de rapprocher les sorties des modèles des observations réelles et d'identifier l'approche la plus fiable.

Le chapitre 5 met en application cette méthodologie à travers l'analyse des sécheresses historiques, en calculant l'indice SPI à différentes échelles temporelles et en comparant les résultats avec des événements connus. Ce chapitre évalue également les performances des différentes approches de downscaling afin de retenir la combinaison la plus performante pour les projections futures.

Enfin, le chapitre 6 est consacré à l'application régionale de l'approche développée. Il analyse les projections climatiques pour la région de Keddara, en identifiant les tendances futures en matière de sécheresse et les risques hydriques associés à l'horizon 2100. Ce dernier chapitre propose également une stratégie d'adaptation réaliste pour 2030, tenant compte des vulnérabilités locales et des enjeux socio-environnementaux du territoire.

Enfin, une **conclusion générale**, dans laquelle on a mis l'accent sur l'importance des résultats obtenus pour l'analyse des sécheresses futures dans la région de Keddara, fondée sur les données climatiques projetées, qui confirment que même dans les scénarios les plus optimistes, l'Algérie restera exposée à des épisodes extrêmes de déficit hydrique. Ces résultats renforcent l'urgence d'agir dès aujourd'hui pour anticiper, adapter et atténuer les impacts à venir.

Chapitre 1

GÉNÉRALITÉS SUR LES SÉCHERESSES

1.1 Introduction

Dans un contexte de changement climatique global, les épisodes de sécheresse suscitent une attention croissante en raison de leur intensification et de leurs multiples impacts sur les écosystèmes, les ressources naturelles et les sociétés humaines. Affectant aussi bien les régions arides que les zones traditionnellement tempérées, ces événements se distinguent par leur complexité, tant dans leurs manifestations que dans leurs mécanismes de déclenchement et de propagation.

Comprendre les sécheresses exige une approche intégrée fondée sur une compréhension rigoureuse des fondements scientifiques, des dynamiques climatiques générales. En tant que phénomène complexe et multiforme, la sécheresse ne peut être appréhendée sans une définition claire de ses composantes, une connaissance approfondie des processus physiques sous-jacents, ainsi qu'une maîtrise des indicateurs permettant de la quantifier.

L'ensemble des éléments présentés dans ce chapitre constitue une base essentielle pour analyser les sécheresses dans leur diversité, en identifiant leurs types, leurs causes, leurs indicateurs associés, ainsi que leurs impacts potentiels sur les systèmes naturels et anthropiques. Cette compréhension est indispensable pour développer des outils de surveillance efficaces, orienter les politiques de gestion des risques climatiques et renforcer la résilience des territoires face à un phénomène appelé à devenir de plus en plus fréquent et sévère.

1.2 Changement Climatique

1.2.1 Définition du changement climatique

Selon les Nations Unies [9] et l'Organisation internationale de normalisation (ISO)[10], les changements climatiques désignent les variations à long terme de la température et des modèles météorologiques.

D'après la définition donnée par le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) dans son sixième rapport, le changement climatique correspond à :

« Une variation de l'état du climat, qu'on peut déceler par des modifications de la

moyenne et/ou de la variabilité de ses propriétés, et qui persiste pendant une longue période, généralement durant des décennies ou plus. » [11]

Donc comme une définition synthétique : Le changement climatique désigne une modification durable des paramètres climatiques globaux ou régionaux de la Terre, tels que la température, les précipitations, ou la fréquence et l'intensité des événements extrêmes, qui persiste pendant plusieurs décennies ou plus.

Ce dérèglement rapide et global peut être attribué à des causes naturelles, telles que les cycles solaires ou les éruptions volcaniques massives.

Cependant, depuis le début de l'ère industrielle au XIX^e siècle, les activités humaines (anthropiques) en sont devenues la principale cause, affectant profondément les écosystèmes, les sociétés humaines et la santé.

1.2.2 Impact du changement climatique

- **Températures plus élevées** : La hausse des gaz à effet de serre entraîne une élévation rapide des températures mondiales, accentuant les vagues de chaleur, les incendies de forêt et les maladies liées à la chaleur, en particulier dans l'Arctique où le réchauffement est deux fois plus rapide que la moyenne.
- **Tempêtes plus violentes** : Le réchauffement amplifie l'évaporation et l'humidité atmosphérique, provoquant des tempêtes plus fréquentes, plus intenses et plus destructrices, notamment les cyclones tropicaux.
- **Augmentation des sécheresses** : Le changement climatique aggrave les pénuries d'eau et les sécheresses agricoles, écologiques et hydrologiques, avec des impacts sévères sur les écosystèmes et les terres agricoles.
- **Réchauffement et montée des océans** : Les océans se réchauffent, se dilatent et montent, en partie à cause de la fonte des glaces, tout en devenant plus acides, menaçant la biodiversité marine et les zones côtières.
- **Pénuries de denrées alimentaires** : Les changements climatiques réduisent la productivité agricole, affectent la pêche et l'élevage, et aggravent la malnutrition dans les régions les plus vulnérables.
- **Accroissement des risques sanitaires** : Le changement climatique aggrave la pollution de l'air et de l'eau, favorise la propagation des maladies infectieuses, accroît la mortalité liée aux catastrophes naturelles (canicules, inondations, incendies), et met sous pression les systèmes de santé publique. Il est responsable, directement ou indirectement, de millions de décès chaque année.
- **Pauvreté et déplacements de populations** : Les catastrophes climatiques forcent chaque année des millions de personnes à se déplacer, tout en piégeant les plus vulnérables dans la pauvreté et l'insécurité.

1.3 La sécheresse

1.3.1 Définition

La sécheresse est un phénomène naturel complexe, multidimensionnel et contextuel, qui résiste à une définition unique en raison de sa grande variabilité spatiale, temporelle et selon les secteurs d'activité concernés.

De manière générale, elle peut être définie comme une réduction significative et prolongée des précipitations sur une période étendue (souvent une saison ou davantage) entraînant des déficits hydriques affectant les systèmes environnementaux, agricoles, hydrologiques et socio-économiques [12]. Contrairement à une simple anomalie météorologique passagère, la sécheresse s'inscrit dans le fonctionnement global du système Terre, révélant les interactions entre climat, usage des ressources et vulnérabilités humaines.

En raison de cette nature multifactorielle et de sa complexité, les scientifiques et les décideurs recourent généralement à deux types de définitions complémentaires :

- **Les définitions conceptuelles** : elles permettent de cerner la notion générale de sécheresse et constituent un cadre essentiel pour orienter les politiques publiques et la compréhension du phénomène.
- **Les définitions opérationnelles** : elles s'appuient sur des données mesurables (telles que les précipitations ou l'humidité du sol) pour détecter, surveiller et gérer les épisodes de sécheresse dans des contextes spécifiques, comme l'agriculture ou l'approvisionnement en eau.

1.3.2 Caractéristiques de la sécheresse

La sécheresse est fondamentalement décrite selon plusieurs dimensions clés : la sévérité, la durée et l'étendue spatiale. Ces dimensions permettent de quantifier l'ampleur d'un épisode de sécheresse ainsi que son impact sur les systèmes naturels et humains. Cette approche a été proposée par Salas (1993) dans son étude sur les caractéristiques des sécheresses hydrologiques [13] :

- **Durée** : Elle indique combien de temps la sécheresse persiste, allant de quelques jours à plusieurs années selon les conditions locales.
- **Magnitude** : Elle mesure la quantité totale d'eau manquante (déficit cumulé) par rapport à un seuil durant toute la période de sécheresse. Elle constitue un indicateur clé pour évaluer la sévérité d'une sécheresse, au-delà de sa simple durée ou intensité.
- **Intensité** : Elle exprime le déficit moyen en eau par unité de temps, indiquant la vitesse et la gravité de l'assèchement.
- **Sévérité** : Elle désigne soit l'ampleur du déficit hydrique, soit les impacts socioenvironnementaux causés, selon le contexte.
- **Étendue géographique** : Elle représente la superficie touchée par la sécheresse, qui peut s'élargir ou se réduire avec le temps.
- **Fréquence** : Elle correspond au temps moyen séparant deux sécheresses de même intensité ou plus, utile pour l'analyse du risque.

1.3.3 Types de sécheresse

La sécheresse englobe divers processus du système Terre, impliquant notamment des composantes météorologiques (anomalies de précipitations), hydrologiques (baisse des débits et des niveaux des réservoirs), agricoles (déficits en humidité du sol) et écologiques ou biophysiques (stress de la végétation). Les répercussions de ces déficits s'étendent aux sphères socio-économiques, affectant les moyens de subsistance, la sécurité alimentaire, l'approvisionnement en eau et la production d'énergie [14]. Chaque type de sécheresse se distingue par des processus physiques spécifiques ainsi que par les impacts environnementaux et socio-économiques qui lui sont propres [15].

a) Sécheresse météorologique :

La sécheresse météorologique est définie comme une période prolongée de précipitations inférieures à la normale, par rapport aux moyennes climatologiques de long terme. Elle constitue généralement la première étape dans l'évolution d'un épisode de sécheresse, et agit comme un déclencheur des autres types de sécheresse [16].

b) Sécheresse agricole :

La sécheresse agricole est principalement liée à **un déficit en humidité du sol**, qui affecte de manière significative la croissance des plantes et les rendements agricoles [17].

La sécheresse agricole se manifeste lorsqu'un déficit de précipitations (typique de la sécheresse météorologique) s'accompagne d'une augmentation de la demande évaporative de l'atmosphère (AED). Cette demande accrue est généralement liée à des températures élevées, qui intensifient les flux d'évapotranspiration, tant à partir du sol nu que de la couverture végétale.

Il en résulte un déséquilibre hydrique dans la zone racinaire, compromettant la croissance des cultures et les rendements agricoles [18] [19].

Les impacts socio-économiques de la sécheresse agricole peuvent être graves et persistants, se prolongeant bien au-delà de la fin de l'épisode de sécheresse météorologique. Ces impacts s'accumulent progressivement et peuvent compromettre la sécurité alimentaire, l'emploi rural et les moyens de subsistance [20].

Cette situation est particulièrement critique dans les pays méditerranéens, où l'agriculture pluviale prédomine. Dans des pays comme l'Algérie, le Maroc et la Tunisie, elle représente plus de 90 % des terres cultivées, ce qui rend ces régions particulièrement vulnérables aux stress liés à la sécheresse [21].

c) Sécheresse Hydrologique :

La sécheresse hydrologique se caractérise par une période prolongée de déficit en disponibilité en eau dans les systèmes d'eau de surface (tels que les cours d'eau, rivières, lacs) et les nappes souterraines.

Elle se manifeste par un écoulement faible des rivières, une baisse du niveau des réservoirs ou une diminution des nappes phréatiques, dans des conditions considérées comme naturelles, c'est-à-dire en l'absence d'interventions anthropiques telles que l'exploitation des barrages, les prélèvements pour l'irrigation ou les transferts interbassins [22].

Bien que le déficit de précipitations constitue généralement le facteur déclencheur, l'apparition et l'intensité des sécheresses hydrologiques sont retardées par rapport aux sécheresses météorologiques, en raison des processus complexes de stockage et de transport au sein du cycle hydrologique. De ce fait, toute sécheresse météorologique ne se traduit pas nécessairement par une sécheresse hydrologique [23].

Le développement progressif de la sécheresse et sa nature multifactorielle sont illustrés à la Figure I-1.

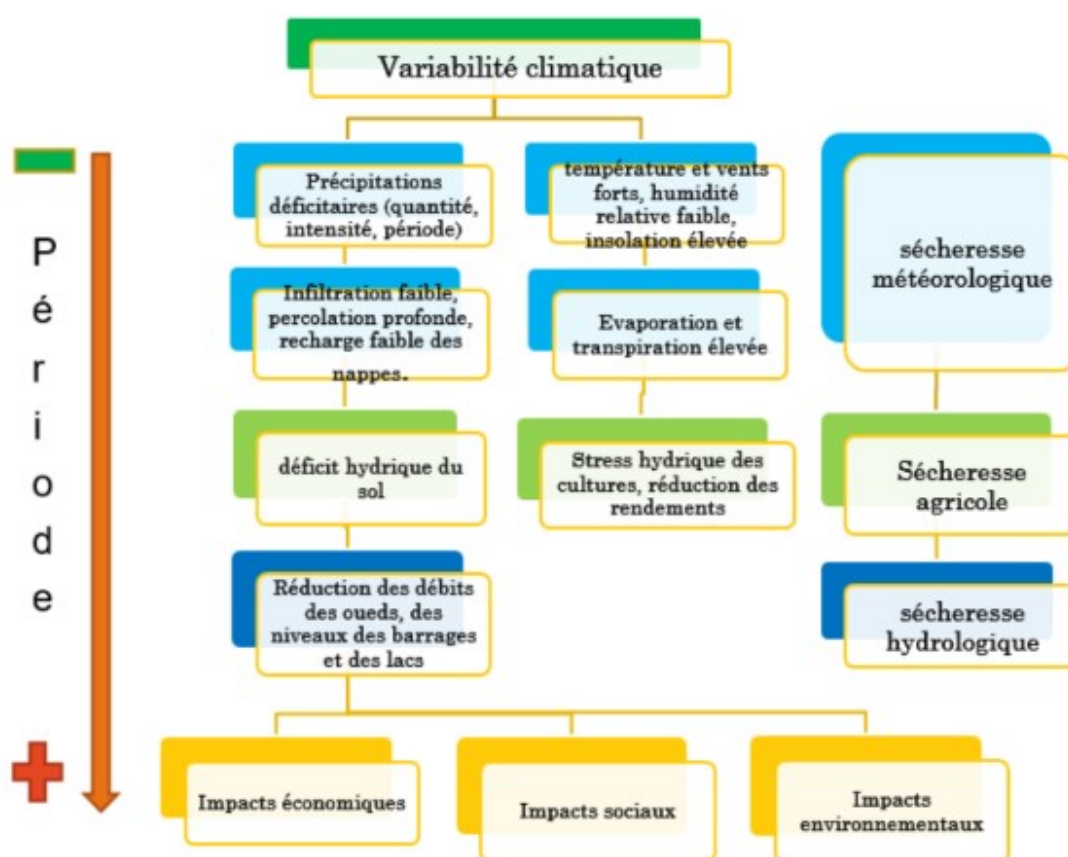


FIGURE 1.1 – Types de sécheresse, manifestations dans le temps et leurs impacts [2]

Ce schéma conceptuel met en évidence la manière dont **un déficit en précipitations**, souvent initié par la **variabilité climatique naturelle**, peut déclencher **une cascade d'effets**. Plus précisément, il entraîne une diminution de **l'humidité du sol** et du **ruissellement de surface**, ce qui conduit ensuite à une réduction **des débits des cours d'eau**. Ces altérations hydrologiques ont des répercussions importantes sur **les secteurs socio-économiques**, en particulier dans les régions dépendantes d'activités sensibles à l'eau, telles que **l'agriculture**, **la production d'énergie** et **l'approvisionnement en eau potable** [24].

1.4 Impacts de la sécheresse à différentes échelles

1.4.1 Sécheresse dans le monde

En Amérique du Nord, les Grandes Plaines des États-Unis ont été gravement touchées par une sécheresse majeure dans les années 1950, provoquant des perturbations profondes dans le secteur agricole. En réponse à cette crise, les autorités ont mis en œuvre la Soil Conservation Act, une politique axée sur la préservation des sols et la gestion durable des ressources en eau.

Malgré ces efforts, la région reste exposée à des sécheresses prolongées, engendrant pertes économiques, déclin des rendements agricoles et parfois des mouvements migratoires.

Plus récemment, entre 2012 et 2016, la Californie a connu l'une des pires sécheresses de son histoire, avec des pertes économiques estimées à près de 10 milliards de dollars. Cette crise a mis en évidence la vulnérabilité des systèmes hydriques et la nécessité de renforcer les politiques

de gestion intégrée des ressources en eau.

En Amérique du Sud, dans les années 1980, la région du Nord-Est du Brésil a été frappée par une sécheresse particulièrement sévère, affectant gravement une agriculture fortement dépendante des précipitations. Cette crise a provoqué d'importantes pertes de production, accentuant les inégalités sociales et entraînant des déplacements de populations rurales vers les zones urbaines.

En Afrique, au début des années 1970, la région du Sahel (notamment le Niger, le Mali et le Burkina Faso) a subi l'une des crises humanitaires les plus graves du XX^e siècle, causée par une sécheresse prolongée. Elle a engendré de graves famines, la perte massive de cheptel, des migrations forcées à grande échelle, et la mort de plus de 100 000 personnes. En réaction, les Nations Unies et plusieurs ONG ont mis en place des programmes de surveillance, d'alerte précoce et d'aide alimentaire.

Plus récemment, la Corne de l'Afrique (Éthiopie, Kenya, Somalie) a connu une sécheresse d'une intensité extrême, aggravée par des conflits armés et une faible résilience locale. En Somalie, cette crise a provoqué une famine généralisée, la perte de millions de têtes de bétail, et la mort de plus de 250 000 personnes, dont une majorité d'enfants.

En Asie, plusieurs épisodes marquants ont été documentés. Une sécheresse prolongée de 1998 à 2002 a affecté la Chine, l'Inde et le Pakistan, entraînant de graves pertes agricoles, une insécurité alimentaire accrue et de graves pénuries d'eau. En réponse, la Chine a investi dans de grands projets d'irrigation et des campagnes de conservation de l'eau, l'Inde a renforcé ses infrastructures hydrauliques et introduit des cultures résistantes à la sécheresse, tandis que le Pakistan a lancé des programmes communautaires de gestion intégrée de l'eau.

En Asie centrale, le cas de la mer d'Aral illustre une forme de sécheresse hydrologique d'origine anthropique. Depuis les années 1960, le détournement massif des eaux des fleuves Amou-Daria et Syr-Daria pour l'irrigation du coton a conduit à une perte de 90 % du volume de la mer, l'effondrement des pêcheries locales, une salinisation massive des sols, et des problèmes de santé publique, en raison de tempêtes de poussière toxiques chargées de sels et de polluants.

En Europe, en 2022, l'Espagne a été confrontée à l'une de ses pires sécheresses des dernières décennies, dans un contexte de canicules répétées, de précipitations exceptionnellement faibles et de températures record. Les répercussions ont été multiples et marquées, notamment sur l'agriculture avec de fortes baisses de rendement, notamment pour les céréales, l'olivier et les cultures maraîchères, l'approvisionnement en eau potable avec des tensions sur les réservoirs et des restrictions d'usage dans plusieurs régions, ainsi que la production d'énergie avec une baisse de la production hydroélectrique due au faible niveau des barrages.

Cette sécheresse a mis en évidence la vulnérabilité croissante du Sud de l'Europe face au changement climatique et a relancé le débat sur la gestion durable des ressources hydriques en zone méditerranéenne.

Tableau I-1 : Études de cas d'épisodes de sécheresse majeurs [9]

Région / Période	Aspect de la sécheresse	Impacts hydrologiques	Impacts environnementaux	Impacts socio-économiques
Asie – Chine (2010–2011)	La pire sécheresse depuis 60 ans dans 8 provinces du nord, affectant la zone de production de blé.	Le lac Hubei réduit à 1/8 de sa surface et 20% de sa profondeur.	7,73 millions d'hectares de blé d'hiver touchés, ensemencement de nuages (3 mm de neige).	3 234 habitants relogés, 15 milliards USD d'aides aux agriculteurs, 1 milliard USD pour 1 350 puits.
Europe (2022)	La pire en 500 ans. Restrictions d'eau, sécheresse record en Espagne, Italie, Portugal, Angleterre.	Rhin < 40 cm, Po - 30%, Danube et Espagne niveaux bas, réservoirs UK à 65%.	Récolte de maïs - 18,5%, sorgho toxique, réduction eau bétail (50%) et irrigation (80%), 860 000 ha brûlés.	Transport +25–35%, céréales -14%, légumes -9,7%, fruits -19%, hydroélectricité -20%.
Afrique – Corne de l'Afrique (2020–2023)	Sécheresse extrême en Somalie, Éthiopie, Kenya. Saison des pluies 2022 la plus sèche depuis 40 ans.	2,4 millions de têtes de bétail mortes au Kenya.	1 210 animaux sauvages morts (éléphants, zèbres, buffles...).	43 000 morts en Somalie, 3,1 millions de déplacés, conflits fonciers, 150–200 000 morts de faim en Éthiopie.
Amérique du Nord – USA (2011–2017)	Sécheresse historique en Californie, précipitations < 34% en 2013, 13 mois consécutifs.	9 réservoirs < 50%, lac Mead à son plus bas (1074,71 pieds).	95% des saumons chinook morts, 102 millions d'arbres morts (62 M en 2016).	50% de réduction eau agricole, 430 000 acres en jachère, hydroélectricité -50%, restrictions eau 25% (2015).
Amérique du Sud – Brésil (2014–2017)	Sécheresse la plus grave en 100 ans dans le sud-est (São Paulo, Rio, Minas Gerais).	Réservoir de Cantareira à 6% (février 2015).	1/3 de la récolte de café perdue, soja -17%.	Rationnement prolongé, prix du café +50%, pertes hydroélectriques : 15,8 milliards de réais.
Océanie – Australie (2017–2019)	Années très sèches dans le sud-est, températures supérieures à la moyenne 36 mois consécutifs.	Débit minimal dans le bassin Murray-Darling (décembre 2019), nappes phréatiques très basses en 2018.	3 M ha de terres protégées, 1,5 M ha de forêts, 820 000 ha agricoles brûlés.	24,3 M ha brûlés, 3 500 maisons détruites, 34 morts, pertes : 53 Mds USD (sécheresse) et 10 Mds USD (feux).

1.4.2 Les sécheresses en Algérie

L'Algérie, pays à climat aride au Sud et semi-aride au Nord, est fortement exposée aux sécheresses. Ces dernières affectent surtout le Nord du pays, où sont concentrées l'essentiel des terres agricoles, des ressources hydriques et de la population.

Les sécheresses y sont récurrentes, parfois longues et sévères, et ont des répercussions profondes sur l'agriculture, l'approvisionnement en eau, l'environnement et les conditions socio-économiques.

Selon le rapport de la FAO (document sur l'eau n°45)[25], les sécheresses dans le Nord de l'Algérie se traduisent par une diminution des volumes dans les barrages ; le tarissement des sources rurales d'eau potable ; l'augmentation de la salinité dans les retenues et les nappes superficielles ; la baisse des rendements agricoles, notamment des céréales et des fourrages.

Les archives gouvernementales rapportent une sécheresse particulièrement dévastatrice entre 1945 et 1947, qui a causé des pertes massives dans l'élevage, notamment la disparition de près

de 90 % du cheptel dans la région d'Aïn Sefra (Sud d'Oran), avec environ 900 000 moutons perdus et près de 3 000 décès par famine sur une population estimée à 80 000 habitants.

Par la suite, l'année 1966 a été marquée par des précipitations exceptionnellement faibles, entraînant une mauvaise germination des cultures et des pertes agricoles importantes.

Entre 1922 et 1992, plusieurs périodes prolongées de sécheresse ont été recensées dans le Nord du pays, notamment entre 1910–1920, 1939–1948, et 1973–1992, caractérisées par des impacts notables tels que : la baisse des apports aux barrages, le tarissement des nappes et puits peu profonds, et des crises alimentaires, incendies de forêts et troubles sociaux.

Au cours des quatre dernières décennies, plusieurs épisodes de sécheresse extrême ou sévère ont affecté les régions Nord du pays, compromettant à la fois l'agriculture, les ressources en eau et les conditions de vie des populations locales [26]. Trois événements marquants peuvent être cités : Les années 1980, avec une période prolongée de déficit hydrique, provoquant des pertes agricoles et pastorales majeures, ainsi que des difficultés d'approvisionnement en eau domestique et d'irrigation ;

Les années 1990, marquées par une chute significative de la production agricole, notamment en agriculture pluviale, et une raréfaction préoccupante des ressources en eau souterraine et de surface ;

L'année 2003, particulièrement chaude, durant laquelle le Nord de l'Algérie, comme d'autres régions européennes, a subi une sécheresse intense, couplée à une canicule, ayant entraîné des pénuries d'eau critiques, une baisse importante des rendements agricoles, une flambée des prix alimentaires, et des tensions sur l'approvisionnement en eau pour les usages domestiques et agricoles. Face à la multiplication des sécheresses, l'Algérie doit renforcer la gestion intégrée des ressources en eau, le suivi hydrométéorologique, l'adaptation agricole (cultures résistantes, irrigation efficiente), et les plans de résilience et d'alerte précoce.

1.5 Les indices de sécheresse

Indicateurs : Ce sont des variables physiques mesurables utilisées pour décrire les conditions de sécheresse, telles que les précipitations, la température, le débit des rivières, les niveaux des nappes phréatiques et des réservoirs, l'humidité du sol ou encore l'enneigement.

Indices : Ce sont des valeurs numériques calculées à partir d'indicateurs climatiques permettant d'estimer la sévérité d'une sécheresse sur une période donnée. Ils traduisent de manière synthétique l'état de sécheresse dans le paysage.

Les indices les plus utilisés pour la caractérisation du régime pluviométrique, et qui ont fait l'objet de plusieurs études, notamment dans le suivi de la sécheresse, sont :

- L'indice de sécheresse de Palmer, (Palmer Drought Index, PDI), et
- L'indice pluviométrique standardisé, (Standardized Precipitation Index, SPI).

Ces indices sont aussi considérés comme indicateurs de changement ou de tendance dans le régime pluviométrique.

1.5.1 Palmer Drought Severity Index (PDSI) (GUTTMAN N. B., 1998)

L'indice de sécheresse de Palmer a été conçu pour quantifier une sécheresse météorologique. Cet indice a été élaboré par W.C. Palmer en 1965 et est appelé l'indice de sévérité de sécheresse de Palmer, (The Palmer Drought Severity Index : PDSI).

L'indice est obtenu sur la base d'un calcul hydrologique pour chaque mois sur une longue série d'années. Il exige donc à la fois les observations des précipitations et de la température qui est la variable utilisée dans l'estimation de l'évapotranspiration et tient compte aussi de l'humidité du sol.

Le PDSI a été développé pour être standardisé afin qu'il soit utilisé à tous les endroits et toutes les époques. Cependant, certaines études montrent que la standardisation, en réalité, n'est pas valide. La variation spatiale et temporelle de l'indice de Palmer peut donc être trompeuse et conduire à des conclusions erronées.

Pendant de nombreuses années, le PDSI est resté le seul indice opérationnel de sécheresse, et il demeure encore aujourd'hui largement utilisé à travers le monde.

1.5.2 Standardized Precipitation Index (SPI) (Gutmann, NB 1998, Giddings, L et al., 2005)

L'indice de précipitation standardisé (Standardized Precipitation Index, SPI) est reconnu comme l'indicateur de sécheresse le plus fiable et le plus largement applicable, en particulier pour l'évaluation des sécheresses à moyen et long terme dans les régions arides et semi-arides. Il a été mis en œuvre par McKee en 1993 pour améliorer la détection de la sécheresse et la capacité de surveillance.

Contrairement à l'indice de Palmer qui est basé sur le bilan d'eau uniquement à l'échelle du mois, impliquant les précipitations, évapotranspiration, ruissellement et humidité du sol, le SPI a été développé afin de quantifier un déficit de précipitation à différentes échelles de temps qui semblent être importantes à l'hydrologue.

Cet indice s'est imposé comme un outil essentiel pour la surveillance des sécheresses météorologiques. En 2009, l'Organisation météorologique mondiale (OMM) a officiellement recommandé le SPI comme indice principal pour l'évaluation des conditions de sécheresse météorologique, en raison de sa simplicité d'utilisation et de son efficacité dans des contextes climatiques variés (OMM, 2009).

Cette adoption généralisée s'explique par le fait que le SPI repose exclusivement sur les données de précipitations, ce qui le rend simple à calculer et applicable à différentes échelles temporelles. Cette flexibilité renforce considérablement son utilité dans les analyses de risque de sécheresse (Handbook of Drought Indicators and Indices, 2016) [27]

Le SPI correspond au rapport de l'écart à la moyenne sur l'écart type des hauteurs pluviométriques annuelles.

$$\text{SPI} = \frac{x_i - \bar{x}}{\sigma_i} \quad (1.1)$$

où :

- SPI : Indice pluviométrique de l'année ;
- x_i : Hauteur de pluie précipitée au cours de l'année ;
- \bar{x} : Moyenne interannuelle des hauteurs de pluie sur la période observée ;
- σ_i : Écart-type des hauteurs de pluie observées.

Le SPI a la vocation à indiquer à lui seul si l'année peut être qualifiée d'excédentaire (SPI>0) ou de déficitaire (SPI<0). McKee a suggéré une classification du SPI de 2 à -2.

Plusieurs études ont évalué les indices de sécheresse à l'aide d'analyses fondées sur des critères décisionnels, en prenant en compte des dimensions telles que la robustesse, la facilité de traitement, la transparence, la sophistication, l'extensibilité et la dimensionnalité. Ces évaluations montrent que le SPI obtient le score global le plus élevé, en raison de sa forte robustesse, un critère jugé déterminant car il conditionne l'efficacité de l'indice dans une large gamme de contextes climatiques et météorologiques.

Dans la région méditerranéenne, le SPI a été identifié comme l'indice le plus performant pour la surveillance de la sécheresse (Spinoni et al., 2016 ; Vicente-Serrano et al., 2014 ; Trambly et al., 2020) . Ce constat se vérifie particulièrement en Algérie, où le SPI affiche une performance supérieure dans la caractérisation des épisodes de sécheresse et constitue un outil d'aide à la décision pour la gestion des ressources en eau.

le SPI peut être calculé sur différentes périodes d'accumulation des précipitations (généralement de 1 à 48 mois), les valeurs obtenues permettent d'estimer divers impacts liés à la sécheresse :

- **SPI-1 à SPI-3** : Lorsqu'il est calculé sur de courtes périodes (1 à 3 mois), le SPI peut servir d'indicateur de la diminution de l'humidité du sol, ce qui reflète principalement une sécheresse agricole.
- **SPI-3 à SPI-12** : Lorsqu'il est calculé sur des périodes moyennes (environ 3 à 12 mois), le SPI peut indiquer une baisse des débits fluviaux et du stockage dans les réservoirs, correspondant à une sécheresse hydrologique.
- **SPI-12 à SPI-48** : Lorsqu'il est calculé sur de longues périodes (12 à 48 mois), le SPI peut signaler une diminution de la recharge des réservoirs et des nappes phréatiques, traduisant une sécheresse hydrologique prolongée.

Tableau I-2 : Classification de la sécheresse selon le SPI d'après l'EDO (European Drought Observatory) et le GDO (Global Drought Observatory)

Anomalie	Plage de valeurs SPI	Régime de précipitation
Positive	$> 2,0$	Extrêmement humide
	$1,5 < \text{SPI} \leq 2,0$	Très humide
	$1,0 < \text{SPI} \leq 1,5$	Modérément humide
Près de la normale	$-1,0 < \text{SPI} \leq 1,0$	Précipitations normales
Négative	$-1,5 < \text{SPI} \leq -1,0$	Modérément sec
	$-2,0 < \text{SPI} \leq -1,5$	Très sec
	$\leq -2,0$	Extrêmement sec

1.6 Conclusion

En somme, la sécheresse, bien qu'étant un phénomène naturel, est de plus en plus amplifiée par le changement climatique, ce qui en fait un défi majeur pour la gestion durable des ressources en eau, la sécurité alimentaire et la stabilité des systèmes humains et écologiques.

Sa complexité, sa variabilité dans le temps et l'espace, ainsi que la diversité de ses formes (météorologique, agricole, hydrologique, socio-économique) exigent une approche multidimensionnelle, combinant :

- Des connaissances scientifiques approfondies,
- Des outils de surveillance et de diagnostic,
- Et des stratégies d'adaptation à différentes échelles.

La maîtrise des indicateurs climatiques, tels que le Precipitation Index (SPI), associée à une analyse fine des dynamiques atmosphériques, constitue un levier essentiel pour :

- Anticiper les risques,
- Réduire la vulnérabilité,
- Et renforcer la résilience des territoires face à des sécheresses de plus en plus fréquentes, longues et intenses.

Ainsi, une compréhension intégrée et territorialisée de la sécheresse s'avère indispensable pour orienter efficacement les politiques d'atténuation et d'adaptation, dans un contexte mondial marqué par des mutations climatiques rapides et profondes.

Chapitre 2

BASES CONCEPTUELLES DU DOWNSCALING STATISTIQUE

2.1 Introduction

Les modèles climatiques globaux (GCM, Global Climate Models) représentent des instruments essentiels pour analyser les processus climatiques à grande échelle et projeter leur évolution future à l'échelle planétaire. Cependant, malgré leur apport considérable, ces modèles présentent des limites notables, en particulier en ce qui concerne leur résolution spatiale et temporelle, ce qui entrave leur capacité à reproduire fidèlement les phénomènes climatiques à échelle fine [28].

L'un des défis majeurs liés à l'utilisation des simulations issues des GCM pour des applications à l'échelle régionale ou locale réside dans la présence de biais systématiques, ainsi que dans leur résolution relativement grossière, souvent de l'ordre de 100 km.

Cette faible résolution limite considérablement leur pertinence pour les études d'impact climatique à échelle locale, qui nécessitent des données plus détaillées et spatialement précises. Dans ce contexte, le recours à des techniques de downscaling, s'appuyant notamment sur des données à échelle régionale telles que celles issues des réanalyses climatiques, permet de raffiner les sorties des GCM.

Ce chapitre est consacré à la présentation des approches de downscaling, en détaillant les principes, les méthodes et les outils mobilisés pour améliorer la représentativité spatiale des projections climatiques à des échelles adaptées aux besoins locaux.

2.2 Les données de réanalyse

La réanalyse est une méthode scientifique avancée qui vise à reconstituer l'état de l'atmosphère, des océans et de la surface terrestre pour chaque jour des dernières décennies, avec un niveau de précision et de cohérence optimal.

Elle repose sur l'assimilation de données, un processus qui intègre des milliers d'observations historiques (provenant de stations météorologiques, satellites, ballons-sondes, navires, etc.) dans un modèle numérique de l'atmosphère, basé sur les lois fondamentales de la physique.

La réanalyse s'inspire de la prévision numérique du temps, utilisée quotidiennement par les centres météorologiques. En prévision classique, une estimation de l'état de l'atmosphère (appelée "analyse") est produite toutes les 6 ou 12 heures, en combinant une ancienne prévision avec de nouvelles observations.

Contrairement à une simple base de données météorologiques, une réanalyse intègre simultanément un grand nombre de variables :

- Atmosphériques : température, vent, pression, humidité, etc.
- Océaniques : hauteur des vagues, température de surface de la mer, courants, etc.
- Terrestres : humidité du sol, couverture neigeuse, etc.
- Chimie atmosphérique : ozone, monoxyde de carbone, aérosols, etc.

Ce traitement global permet d'analyser de manière cohérente les interactions entre les différentes composantes du système climatique, en tenant compte de leurs dynamiques physiques [29].

Tableau II-1 : Les sources des données de réanalyse

Réanalyse	Période couverte	Type	Résolution spatiale	Méthode d'assimilation des données	Agences responsables
NCEP-R1	1948 – présent	Spectral	T62 (~210 km)	3D-Var	National Centers for Environmental Prediction (NCEP) – NCAR (USA)
NCEP-R2	1979 – présent	Spectral	T62 (~210 km)	3D-Var	NCEP – Département de l'Énergie des États-Unis (DOE) (USA)
CFSR	1979 – présent	Spectral	T382 (~38 km)	3D-Var	NCEP (USA)
20CR	1871 – 2008	Spectral	T62 (~210 km)	Filtre de Kalman, Ensemble	NOAA/ESRL – CIRES (USA)
MERRA	1979 – présent	Finite volume	0.5° × 0.66° (~65 km)	3D-Var / IAU	NASA GMAO (USA)
ERA-Interim	1979 – présent	Spectral	T255 (~79 km)	4D-Var	ECMWF (Centre européen pour les prévisions à moyen terme) (Europe)
JRA-25	1979 – présent	Spectral	T106 (~120 km)	3D-Var	JMA (Agence météorologique japonaise) (Japon)

2.3 Les modèles climatiques globaux (GCMs)

2.3.1 Description des modèles climatiques globaux

Les modèles de circulation générale (General Circulation Models – GCMs), également appelés modèles climatiques globaux, sont des outils informatiques sophistiqués conçus pour simuler le système climatique de la Terre dans sa complexité. Ces modèles intègrent des représentations mathématiques des processus physiques et dynamiques ayant lieu dans l'atmosphère, les océans, la surface terrestre et la cryosphère, dans le but de comprendre et de prévoir le comportement du climat à différentes échelles spatiales et temporelles [30], [31].

Les modèles climatiques constituent des instruments essentiels pour la compréhension des régimes météorologiques et climatiques. Ils simulent les interactions complexes entre les composants du système climatique, y compris les échanges d'énergie, d'humidité et de gaz à effet de serre. Grâce à ces capacités, ils permettent la production de projections climatiques dans divers scénarios d'émissions de gaz à effet de serre, offrant ainsi des éléments de réponse aux grandes incertitudes liées à l'évolution future du climat (GIEC, 2021).

Ces modèles jouent un rôle central dans les études climatiques, notamment pour l'évaluation des impacts du changement climatique sur les ressources en eau, l'agriculture, la santé et les écosystèmes. Ils sont également incontournables dans l'élaboration de politiques publiques climatiques et dans la mise en œuvre de stratégies d'adaptation et d'atténuation [32]. Leur développement repose sur une coopération internationale entre centres de recherche et institutions météorologiques, renforçant ainsi leur fiabilité scientifique et leur utilité décisionnelle [23].

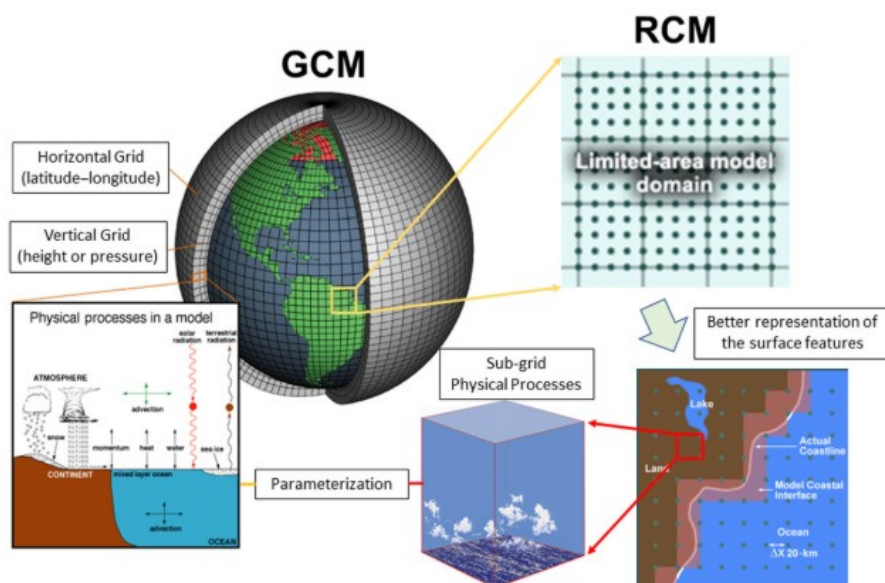


FIGURE 2.1 – Representation of GCMs and RCMs. Source : Ambrizzi et al. (2018)[3]

Dans le cadre du projet CMIP6 (Coupled Model Intercomparison Project Phase 6), une étiquette de variante (RANs) identifie une simulation spécifique d'une expérience de modèle climatique. Étant donné que plusieurs simulations peuvent être réalisées avec le même modèle mais avec de légères différences (par exemple, des conditions initiales ou des configurations du modèle), cette étiquette permet de distinguer les différentes simulations.

Les **RANs** sont composées de quatre indices :

- **r** : Indice de réalisation (*Realization index*) : différentes réalisations stochastiques (c'est-à-dire des simulations débutant à partir de conditions initiales légèrement différentes).
- **i** : Indice d'initialisation (*Initialization index*) : différentes procédures d'initialisation (par exemple, modifications dans les méthodes d'assimilation de données).
- **p** : Indice de physique (*Physics index*) : différentes versions de la physique du modèle ou de ses paramétrisations.
- **f** : Indice de forçage (*Forcing index*) : différences dans les jeux de données de forçage appliqués (par exemple, modifications de l'occupation des sols ou des émissions).

Exemple : **r1i1p1f1** correspond à la première réalisation, première initialisation, première configuration physique, et premier jeu de forçages appliqués.

2.3.2 Définition et rôle des composantes des variantes RANs dans CMIP6

- **Réalisation** : désigne différentes simulations utilisant le même modèle et le même protocole expérimental, mais avec des conditions initiales légèrement différentes.
(Exemple : r1 et r2 représentent deux membres d'un ensemble initialisés à partir de points différents dans une simulation de contrôle, pour observer comment le bruit aléatoire météorologique influence les régimes de précipitations.)
- **Initialisation** : indique des variations dans les procédures d'initialisation du modèle, telles que l'utilisation de jeux de données d'observation différents ou de méthodes alternatives.
(Exemple : i1 et i2 pourraient correspondre à des simulations basées sur différentes bases de données de température océanique.)
- **Configuration physique** : fait référence à différentes versions ou configurations des paramétrisations physiques du modèle, afin d'explorer l'impact de ces changements sur les résultats climatiques, notamment pour des processus comme la formation des nuages ou le rayonnement.
(Exemple : p1 peut utiliser la configuration physique standard, tandis que p2 intègre des interactions aérosols actualisées.)
- **Jeu de forçage** : identifie le jeu de forçages externes appliqués à la simulation, permettant de distinguer les simulations selon les intrants climatiques utilisés, comme les concentrations de gaz à effet de serre ou les changements d'occupation des sols.
(Exemple : f1 peut signifier que les forçages historiques standards ont été utilisés (CO₂, aérosols volcaniques, activité solaire, utilisation des terres), tandis que f2 pourrait représenter une configuration sans aérosols volcaniques.)

2.4 Méthodes de désagrégation spatiale du climat (Downscaling) :

2.4.1 Principe du Downscaling

Le downscaling (ou désagrégation spatiale) désigne l'ensemble des méthodes permettant de traduire les sorties climatiques à grande échelle des GCM vers des échelles plus fines, adaptées aux besoins régionaux ou locaux.

La valeur ajoutée du downscaling peut être définie comme la capacité de la méthode à enrichir le signal climatique issu du GCM, en intégrant des éléments à fine échelle, tels qu'une meilleure représentation du relief, les contrastes terre-mer, ou encore la simulation explicite des dynamiques atmosphériques locales. En parallèle, le downscaling permet de mieux comprendre et anticiper les réactions des systèmes naturels et anthropiques face aux changements climatiques à échelle locale, ce qui en fait un outil central dans les études d'Impacts, d'Adaptation et de Vulnérabilité (IAV) [33].

Deux grandes approches de downscaling sont couramment utilisées :

- Le downscaling statistique, qui repose sur l'établissement de relations empiriques entre des variables climatiques à grande échelle (prédicteurs) et des variables locales (prédictees), afin de générer des scénarios climatiques régionalisés [34] ;

- Le downscaling dynamique, qui consiste à imbriquer un modèle climatique régional (RCM) à haute résolution dans un GCM, permettant de simuler explicitement les processus physiques à des échelles plus fines [35].

2.4.2 Statistical Downscaling

Le downscaling statistique (ou désagrégation statistique) repose sur l'identification de relations statistiques entre un vecteur de prédicteurs à grande échelle, généralement composé de plusieurs variables atmosphériques (telles que la température, la pression, le vent ou l'humidité à différents niveaux), ainsi que de la variable cible à downscaler, et la variable climatique régionale d'intérêt, observée localement. Ce vecteur multivarié permet de représenter les mécanismes climatiques de grande échelle influençant les conditions locales .

Ces relations sont dérivées à partir de données historiques observées, puis appliquées à des projections futures issues des modèles climatiques globaux (GCMs), afin d'estimer les conditions locales nécessaires aux études d'impact [36].

L'hypothèse centrale de cette approche repose sur la stationnarité statistique, c'est-à-dire que la relation entre les prédicteurs et les prédites observée dans le climat passé restera valable dans un climat futur [37] ; [38]. Or, cette hypothèse peut être mise à mal dans un contexte de changement climatique marqué par des altérations profondes du système climatique terrestre.

Le processus de modélisation statistique est généralement itératif : il implique la sélection et l'évaluation de multiples prédicteurs et configurations jusqu'à l'obtention d'une relation optimale [37]. Plusieurs études ont mis en évidence que la qualité des résultats dépend fortement de plusieurs facteurs : le choix des variables prédictives, le type de méthode statistique utilisée, la durée de la période de calibration, les caractéristiques des stations météorologiques, ainsi que la qualité des observations disponibles [38] ; [39].

L'un des principaux avantages du downscaling statistique réside dans ses faibles exigences en ressources de calcul, sa facilité d'implémentation, et sa souplesse d'intégration dans les études d'impact, contrairement au downscaling dynamique beaucoup plus coûteux et complexe [38] ; [40] ; [41].

Cependant, le SD présente plusieurs limites importantes. Premièrement, il ne prend pas directement en compte les processus physiques complexes (tels que les rétroactions atmosphériques ou les interactions océan-atmosphère) qui régissent le climat régional [Dayon et al., 2015 ; Lanzante et al., 2018]. Deuxièmement, il peut négliger la variabilité naturelle et la non-linéarité du système climatique, ce qui limite sa capacité à représenter fidèlement l'évolution des extrêmes climatiques et des régimes de circulation [39] ; [42].

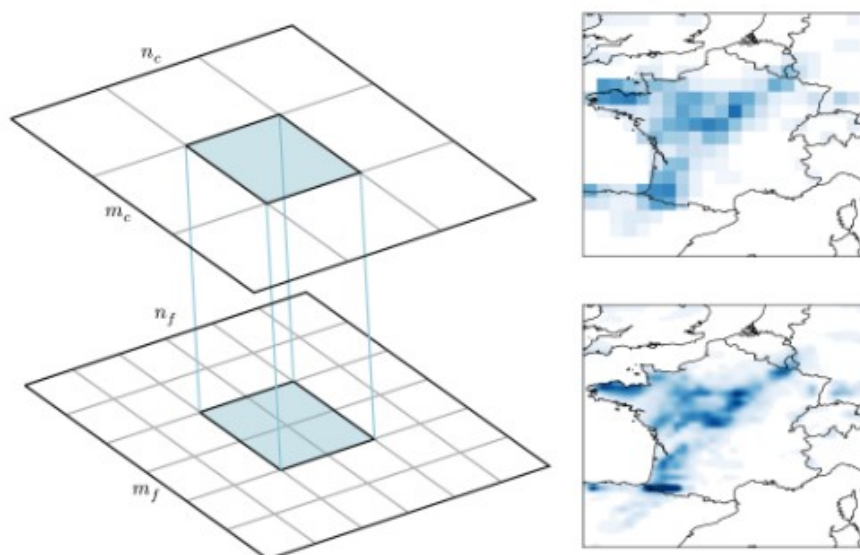


FIGURE 2.2 – Illustration du passage de l'échelle grossière à l'échelle fine

On distingue plusieurs méthodes pour l'élaboration de relations statistiques entre les valeurs à l'échelle grossière (globale) et celles à l'échelle locale. Parmi celles-ci, on peut citer :

- **Apprentissage automatique**

L'intégration des techniques d'apprentissage automatique (machine learning) dans le processus de désagrégation statistique connaît un essor considérable. Ces approches se distinguent par leur capacité à modéliser des relations complexes et non linéaires entre les variables atmosphériques à grande échelle et les caractéristiques climatiques locales. Elles permettent ainsi de dépasser les limites des méthodes statistiques empiriques classiques. Leur aptitude à capturer la non-linéarité intrinsèque des processus climatiques en fait des outils particulièrement prometteurs pour l'étude des impacts du changement climatique à l'échelle régionale.

- **Méthodes d'apprentissage profond**

Les approches d'apprentissage profond (deep learning) se sont récemment imposées comme des outils très performants dans le cadre du downscaling statistique. Leur efficacité repose sur leur capacité à modéliser des relations hautement non linéaires et à capturer des dépendances spatio-temporelles complexes au sein des données climatiques. Ces méthodes offrent ainsi un fort potentiel pour améliorer la résolution des projections climatiques régionales.

2.5 Le Quantile Mapping (QM)

Il s'agit d'une méthode de transformation des distributions fondée sur les quantiles, largement utilisée dans les études de correction de biais en climatologie [43].

Dans sa version la plus répandue, cette méthode remplace chaque quantile de la distribution simulée par le même quantile de la distribution observée. En d'autres termes, chaque valeur issue du modèle climatique est transformée de façon à conserver son rang dans la distribution,

tout en étant projetée sur la distribution observée correspondante. Cette opération permet une correction systématique et cohérente des biais, tant sur la moyenne que sur la forme de la distribution.

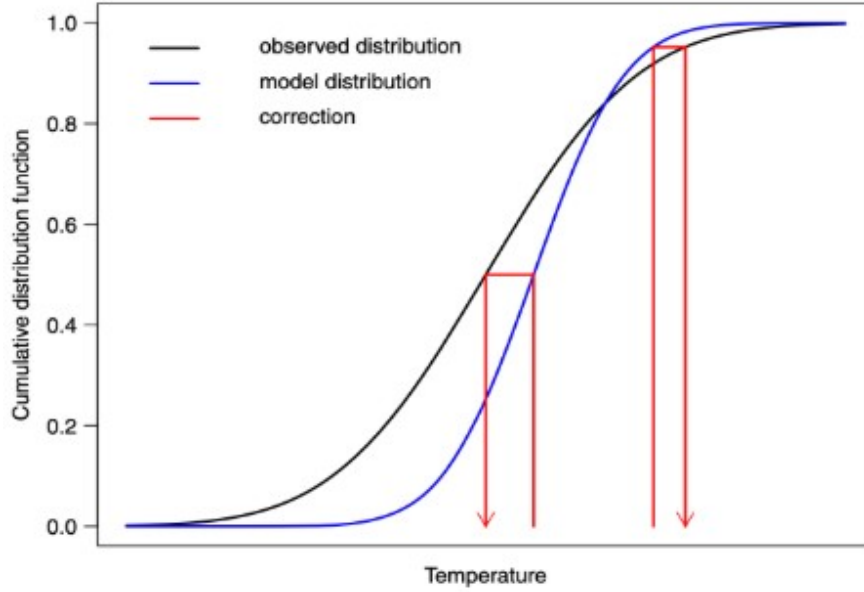


FIGURE 2.3 – Quantile Mapping : une valeur simulée, correspondant à un certain quantile de la distribution simulée, est remplacée par le quantile de la distribution observée ayant la même probabilité cumulée.

Le Quantile Mapping peut s'écrire de manière générale comme suit :[44]

$$y_{QM}(t) = F_{\text{obs,historique}}^{-1} [F_{\text{mod,historique}}(x_{\text{mod}}(t))] \quad (2.1)$$

- $x_{\text{mod}}(t)$: variable simulée (par un GCM ou RCM) à l'instant t
- $F_{\text{mod,historique}}$: fonction de répartition cumulée (CDF) de $x_{\text{mod}}(t)$, calculée sur la période historique
- $F_{\text{obs,historique}}^{-1}$: fonction quantile (inverse de la CDF) des observations historiques
- $y_{QM}(t)$: valeur corrigée de biais à l'instant t

En pratique, on projette la valeur simulée dans l'espace de probabilité via la CDF du modèle, puis on la projette à nouveau dans l'espace des observations via la CDF inverse des observations.

Cas d'application hors période historique :

Lorsqu'on applique la correction sur une période projetée (par exemple un scénario futur), une adaptation de l'équation 1 est utilisée :

$$y_{QM}(t) = F_{\text{obs,historique}}^{-1} \left[F_{\text{mod,historique}} \left(\left(\frac{x_{\text{mod,historique}}}{x_{\text{mod,p}}(t)}, x_{\text{mod,p}}(t) \right) \right) \right] \cdot \frac{\overline{x_{\text{mod,p}}(t)}}{x_{\text{mod,historique}}} \quad (2.2)$$

- $x_{\text{mod,p}}(t)$: valeur simulée pendant la période projetée p
- $\overline{x_{\text{mod,historique}}}$: moyenne climatologique simulée sur la période historique
- $\overline{x_{\text{mod,p}}(t)}$: moyenne glissante simulée dans la période projetée autour de t

Cette version ajuste dynamiquement la distribution du futur en tenant compte de l'évolution du climat simulé dans le temps, en normalisant les données du futur par rapport aux statistiques historiques.

2.6 Conclusion

Ce chapitre a posé les fondements essentiels à la compréhension et à l'usage des outils climatologiques permettant l'analyse du climat à différentes échelles. Les modèles climatiques globaux (GCMs) y ont été présentés comme des instruments incontournables pour simuler le climat passé, présent et futur. Malgré leur puissance, ces modèles présentent certaines limitations, notamment en termes de résolution spatiale et de biais systématiques.

Pour pallier ces insuffisances, les données de réanalyse ont été introduites. Produites à partir de modèles assimilant des observations, elles fournissent des jeux de données cohérents, complets et spatialement explicites, servant à la fois de référence et de base pour des techniques de raffinement spatial.

Dans ce contexte, les techniques de downscaling, et en particulier les approches statistiques, ont été détaillées. Celles-ci permettent de transposer les sorties des GCMs à des échelles locales plus fines, indispensables pour les études d'impact. Une attention particulière a été portée aux méthodes récentes issues de l'intelligence artificielle, telles que les modèles d'apprentissage automatique et d'apprentissage profond, dont l'explication fera l'objet d'une analyse approfondie dans le chapitre suivant.

Enfin, la méthode de Quantile Mapping a été présentée comme une étape importante dans le processus de correction des biais. Elle permet d'ajuster les distributions statistiques des données simulées pour qu'elles soient cohérentes avec celles des observations ou des données de réanalyse, améliorant ainsi la fiabilité des projections climatiques locales.

Chapitre 3

APPRENTISSAGE AUTOMATIQUE (MACHINE LEARNING)

3.1 Introduction

Au cours des dernières décennies, l'intelligence artificielle a profondément transformé notre manière de traiter et d'analyser l'information. Parmi ses branches les plus influentes, le Machine Learning (ou apprentissage automatique) occupe une place de choix. Il s'agit d'un domaine de recherche et d'application qui permet aux machines d'apprendre automatiquement à partir de données et de s'améliorer sans intervention humaine explicite [45].

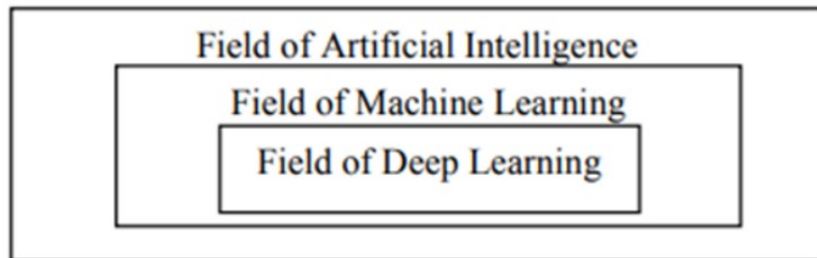


FIGURE 3.1 – Hiérarchie des sous-domaines de l'intelligence artificielle [4]

Le développement d'algorithmes d'apprentissage automatique a ouvert la voie à une grande diversité d'applications : diagnostic médical assisté par ordinateur, reconnaissance d'images et de voix, moteurs de recommandation ou encore prévisions économiques. L'efficacité de ces systèmes repose essentiellement sur le choix et la performance des algorithmes utilisés, chacun présentant des avantages et des limites selon la nature du problème et le type de données traitées.

Cette étude s'intéresse plus particulièrement à deux familles représentatives de l'apprentissage automatique : d'une part, les algorithmes de type boosting, avec notamment XGBoost, reconnu pour ses performances exceptionnelles sur des données structurées ; d'autre part, les réseaux de neurones convolutifs (CNN), emblématiques du deep learning, qui peuvent être considérés comme une sous-catégorie des réseaux de neurones supervisés, selon la classification présentée dans la figure 3-1.

3.2 Définitions

3.2.1 Intelligence Artificielle (Artificial Intelligence)

L'intelligence artificielle désigne la science et l'ingénierie de systèmes informatiques capables de percevoir leur environnement, d'apprendre, de raisonner et d'agir de manière autonome pour atteindre des buts précis, notamment en s'adaptant efficacement dans des contextes où les connaissances et ressources sont limitées [46].

3.2.2 Apprentissage Automatique (Machine Learning)

Le Machine Learning (apprentissage automatique) est une sous-discipline de l'intelligence artificielle qui se concentre sur le développement d'algorithmes capables d'apprendre des données et de s'améliorer sans être explicitement programmés pour chaque tâche spécifique. Cette approche diffère des méthodes symboliques classiques par l'usage massif des données comme source de connaissance, permettant ainsi l'automatisation de la détection de motifs complexes et de la prise de décision [47].

3.2.3 Apprentissage Profond (Deep Learning)

Le deep learning, ou apprentissage profond, est un sous-domaine du machine learning qui consiste à entraîner des réseaux de neurones artificiels multi-couches (ou profonds) capables d'apprendre des représentations hiérarchiques à partir de données brutes, sans avoir besoin d'extraction manuelle de caractéristiques.

Il exploite des réseaux neuronaux profonds pour traiter de grands volumes de données et offre des performances souvent supérieures à celles des techniques classiques de classification [48].

3.3 Classification des types d'apprentissage automatique

L'apprentissage automatique (Machine Learning) englobe un large éventail de méthodes, regroupées en plusieurs catégories principales selon la nature des données disponibles et la manière dont l'apprentissage est réalisé. Parmi ces grandes familles, on trouve notamment l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage semi-supervisé, l'apprentissage par renforcement ainsi que d'autres approches hybrides ou spécifiques. Chacune de ces catégories se distingue par ses principes de fonctionnement et ses applications typiques. La figure suivante illustre de manière synthétique cette classification.

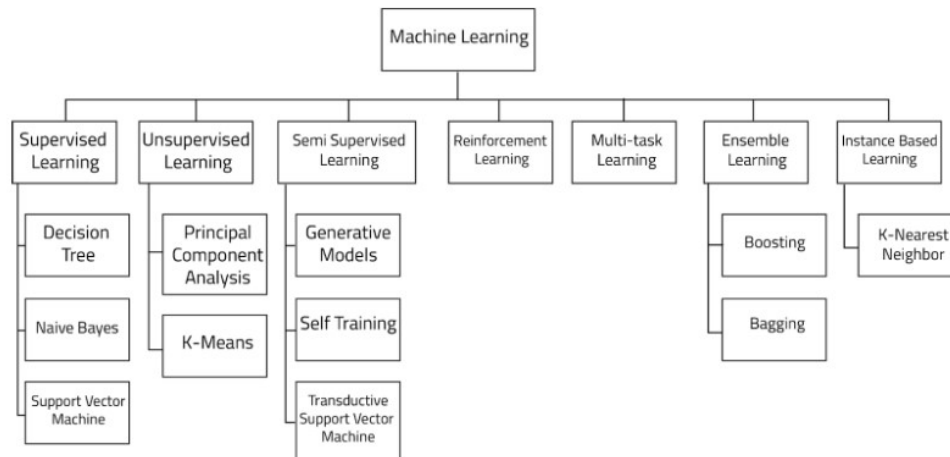


FIGURE 3.2 – Classification des types d'apprentissage automatique (Machine Learning) [5]

3.3.1 Extreme Gradient Boosting – Xgboost

L'algorithme XGBoost (Extreme Gradient Boosting) est une technique d'apprentissage supervisée introduite par Chen et Guestrin (2016). Il repose sur le principe du boosting de gradient, où plusieurs arbres de décision faibles sont entraînés de manière séquentielle, chacun corrigeant les erreurs du précédent.

Contrairement aux variantes classiques du Gradient Boosting, XGBoost se distingue par une implémentation hautement optimisée, à la fois parallélisable, scalable, et dotée de mécanismes de régularisation (L1 et L2). Ces caractéristiques permettent de :

- Améliorer la performance prédictive sur des ensembles de données volumineux ou complexes,
- Réduire le risque de surapprentissage (overfitting),
- Et accélérer considérablement l'entraînement du modèle.

XGBoost est aujourd'hui largement utilisé dans de nombreuses disciplines, y compris les sciences climatiques, pour la modélisation non linéaire, la réduction de biais dans les projections, et l'analyse des facteurs explicatifs.

Principe fondamental :

XGBoost est une implémentation optimisée de l'algorithme Gradient Boosting Machines (GBM). Le principe de base consiste à :

- Construire une séquence d'arbres de décision (souvent peu profonds),
- Chaque nouvel arbre corrige les erreurs commises par la somme des arbres précédents,
- L'ensemble final est une somme pondérée d'arbres faibles mais complémentaires.

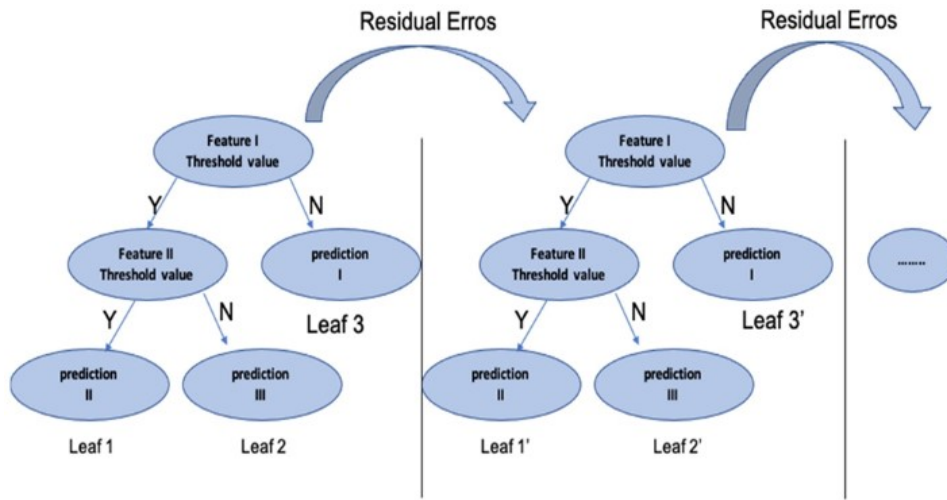


FIGURE 3.3 – Schéma des arbres Xgboost [6]

À chaque itération, l'algorithme ajoute un nouvel arbre $f_t(\mathbf{x})$ à un modèle existant de prédiction $\hat{y}^{(t)}$ pour corriger les erreurs précédentes. La fonction objective globale est définie comme suit :

$$L^{(t)} = \sum_i l(y_i, \hat{y}_i^{(t)}) + \sum \Omega(f_k) \quad (3.1)$$

où :

- l est la fonction de perte (par ex. erreur quadratique, pseudo-Huber),
- $\Omega(f)$ est un terme de régularisation contrôlant la complexité des arbres, défini par :

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda |w|^2 \quad (3.2)$$

Avec :

- T le nombre de feuilles,
- et w les poids associés aux feuilles.

Selon l'étude de Chen Guestrin (2016), cette formulation permet à XGBoost d'atteindre une efficacité computationnelle sans précédent tout en conservant des propriétés statistiques favorables à la généralisation.

Fonction de perte pseudo-Huber :

Dans ce travail, nous avons retenu la fonction de perte pseudo-Huber, particulièrement adaptée aux données présentant des valeurs extrêmes ou du bruit. Contrairement à la perte quadratique (sensible aux outliers), la pseudo-Huber combine la douceur du MSE avec la robustesse du MAE. Sa formulation est la suivante :

$$L_\delta(a) = \delta^2 \left(\sqrt{1 + \left(\frac{a}{\delta}\right)^2} - 1 \right) \quad (3.3)$$

Où δ : paramètre contrôlant la transition entre comportement quadratique et linéaire.

Cette fonction a été démontrée comme efficace pour les données climatologiques où des fluctuations extrêmes (ex. : sécheresse sévère, précipitations anormales) peuvent fausser les prédictions si elles ne sont pas correctement traitées [49].

Optimisation par recherche bayésienne (Optuna) :

L'ajustement des hyperparamètres de XGBoost est essentiel pour obtenir un compromis optimal entre biais et variance. Nous avons utilisé Optuna [50], une bibliothèque d'optimisation automatique par recherche bayésienne, qui construit un modèle probabiliste de la fonction objectif pour guider la recherche des meilleurs paramètres, notamment :

- **Le nombre d'estimateurs (n_estimators)** : correspond au nombre total d'arbres de décision faibles (appelés « estimateurs »). Chaque arbre corrige les erreurs de prédiction des précédents, permettant ainsi un apprentissage progressif. Un nombre trop faible d'estimateurs peut mener à un sous-apprentissage (underfitting), tandis qu'un nombre trop élevé peut favoriser le surapprentissage (overfitting).
- **La profondeur maximale des arbres (max_depth)** : plus la profondeur est grande, plus le modèle peut capturer des relations complexes entre les variables explicatives et la variable cible, augmentant ainsi sa capacité prédictive. Cependant, une profondeur excessive accroît également la complexité du modèle, ce qui peut conduire à un surapprentissage.
- **Le taux d'apprentissage (learning_rate)** : hyperparamètre fondamental qui contrôle l'impact de chaque nouvel arbre sur la prédiction finale dans le processus de boosting. Un faible taux d'apprentissage limite la contribution de chaque arbre, favorisant une convergence plus stable et une meilleure généralisation, mais nécessitant davantage d'itérations. À l'inverse, un taux trop élevé accélère l'apprentissage, au prix d'une plus grande instabilité et d'un risque accru de surapprentissage.
- **Le sous-échantillonnage (subsample)** : correspond à la fraction de l'ensemble de données utilisée pour entraîner chaque arbre dans le cadre du boosting. Cette technique introduit une forme de régularisation en réduisant la corrélation entre les arbres, ce qui peut améliorer la généralisation du modèle et réduire le risque de surapprentissage. Une valeur trop basse peut nuire à la stabilité du modèle.
- **Le coefficient de régularisation L1 (alpha)** : Cette forme de pénalisation favorise la parcimonie du modèle en poussant certains poids vers zéro, ce qui peut améliorer la capacité de généralisation et faciliter l'interprétation en sélectionnant implicitement les variables les plus pertinentes. L'ajout de régularisation L1 est particulièrement utile pour limiter le surapprentissage dans des contextes à forte dimensionnalité.

Avantages de XGBoost dans le contexte climatique :

Dans le cadre de cette étude dédiée à la prédiction d'un indice climatique tel que le Standardized Precipitation Index (SPI), l'algorithme XGBoost présente plusieurs atouts majeurs :

1. Modélisation des relations non linéaires

XGBoost est capable de capturer des relations complexes et non linéaires entre les variables climatiques (précipitations, température, humidité, etc.), ce qui le rend particulièrement adapté à la modélisation de phénomènes atmosphériques souvent non linéaires et interdépendants.

2. Robustesse au bruit et aux données manquantes

Grâce à sa structure d'apprentissage par arbres, XGBoost peut gérer efficacement des jeux de données partiellement incomplets ou bruités, courants dans les bases climatiques (stations météorologiques, séries de réanalyse, données satellitaires).

3. Évaluation de l'importance des variables explicatives

L'algorithme fournit automatiquement des mesures d'importance des variables (feature importance), permettant d'identifier les facteurs climatiques les plus influents dans la variation de l'indice SPI et de guider les analyses d'impact ou les stratégies d'adaptation.

3.3.2 Convolutional Neural Networks (CNN)

Les réseaux de neurones convolutifs (CNN) sont devenus un pilier incontournable de l'apprentissage profond, en particulier dans le domaine de la vision par ordinateur. Leur architecture, inspirée du cortex visuel humain, permet l'extraction automatique de caractéristiques hiérarchiques, des plus simples (bords, textures) aux plus complexes (formes, objets), directement à partir des données brutes.

Fonctionnement général :

Un CNN est constitué de couches spécialisées :

- Convolutionnelles, qui détectent les motifs locaux dans les images ;
- ReLU, qui introduisent la non-linéarité ;
- Pooling, qui réduisent la dimension spatiale tout en conservant les caractéristiques essentielles ;
- Et parfois fully connected, qui synthétisent les informations pour la prédiction finale.

Évolution des architectures CNN :

- **LeNet-5 (1998) – Yann LeCun** : premier CNN largement reconnu, conçu pour la reconnaissance des chiffres manuscrits (base MNIST). Il posait les bases de l'alternance convolution/pooling.
- **AlexNet (2012) – Krizhevsky et al.** : vainqueur du concours ImageNet, il a marqué une rupture majeure avec :
 - o l'usage de GPU pour l'entraînement accéléré,
 - o la fonction ReLU comme activation non linéaire,
 - o et la technique de *dropout* pour lutter contre le surapprentissage.
- **VGGNet (2014) – Simonyan et Zisserman** : architecture profonde mais simplifiée, utilisant uniquement des filtres 3×3, montrant que la profondeur améliore les performances.
- **GoogLeNet / Inception (2014) – Szegedy et al.** : introduction des modules multi-échelles (Inception modules) afin de réduire le coût computationnel tout en conservant la richesse des représentations.
- **ResNet (2015) – He et al.** : grande avancée grâce aux connexions résiduelles (skip connections), permettant d'entraîner des réseaux très profonds (>100 couches) sans dégradation des performances.

Ces innovations successives ont considérablement amélioré la capacité des CNN à apprendre des représentations riches et hiérarchiques pour des tâches complexes de vision par ordinateur.

Applications courantes :

Les CNN sont utilisés dans :

- La classification d'images (ex. reconnaissance d'objets),
- La détection d'objets (bounding boxes),
- La segmentation sémantique (pixels classés individuellement),
- Et même, plus récemment, dans des domaines hors image, comme la modélisation climatique, la sismologie, ou la santé environnementale.

Principe fondamental :

Les CNN sont basés sur l'idée d'appliquer des filtres convolutifs aux données d'entrée pour extraire des caractéristiques locales. Chaque filtre apprend à reconnaître un motif spécifique présent dans les données (par exemple, une tendance, une variation cyclique ou une discontinuité). Dans leur application classique en vision par ordinateur, ces réseaux détectent des bords, textures et objets. Toutefois, comme le montrent [51].

Le cœur de l'architecture CNN est la couche de convolution (1D dans le cas des séries temporelles), suivie généralement de couches de sous-échantillonnage (comme MaxPooling) et de couches entièrement connectées. Ces composants permettent aux CNN d'apprendre des représentations hiérarchiques de plus en plus abstraites à mesure que l'on progresse dans la profondeur du réseau.

CNN pour les séries temporelles : motivations et avantages :

Dans le contexte de la prévision de séries temporelles, les CNN offrent plusieurs avantages :

- Efficacité computationnelle : Contrairement aux réseaux récurrents (RNN, LSTM), les CNN permettent un parallélisme complet lors de l'entraînement, ce qui réduit considérablement les temps de calcul [52].
- Détection automatique de motifs temporels : Les couches convolutives sont capables d'extraire automatiquement des motifs temporels à court terme, sans recourir à une ingénierie manuelle des caractéristiques [53].
- Moindre dépendance à la longueur de séquence : Alors que les RNN souffrent d'un gradient qui s'atténue avec des séquences longues, les CNN permettent de capturer les dépendances sur plusieurs pas de temps grâce à l'empilement de couches convolutives ou à l'utilisation de dilated convolutions [51].

Architecture du modèle :

L'architecture du CNN présentée dans ce travail reprend les principes classiques avec certaines adaptations spécifiques aux données temporelles. Elle est composée de :

- **Couche convolutive 1D** avec activation non linéaire LeakyReLU, qui introduit une flexibilité supplémentaire par rapport à ReLU classique [54]. Son rôle est d'extraire des motifs locaux dans le temps en glissant chaque filtre convolutif (kernel) sur la séquence temporelle (lag), effectuant une multiplication élément par élément entre le filtre et la sous-séquence, puis une sommation.

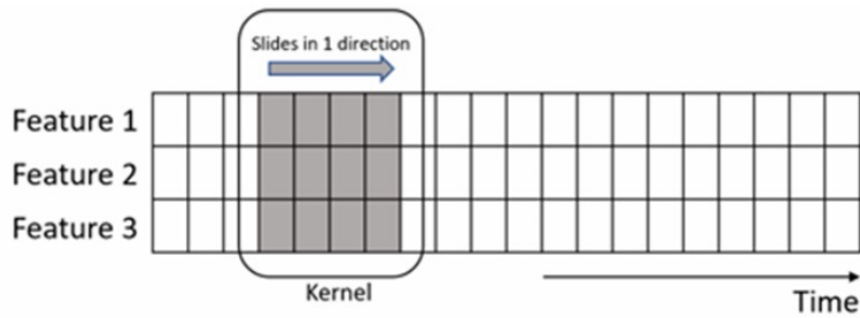


FIGURE 3.4 – Conv1D : Convolver on time dimension

- **Couche de MaxPooling**, qui réduit la dimensionnalité tout en conservant l'information la plus saillante, facilitant la généralisation du modèle. Cela permet d'éviter la redondance dans l'information tout en renforçant les signaux forts.

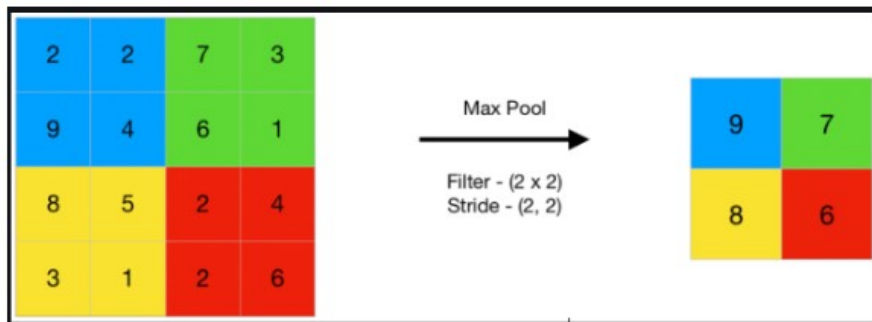


FIGURE 3.5 – Illustration du Max Pooling avec un filtre 2×2 et un pas de 2×2 appliqué à une matrice d'entrée [7]

- **Couche d'aplatissement**, qui transforme la sortie 3D (échantillons, pas de temps, filtres) en un vecteur 1D utilisable par les couches denses.
- **Couche dense avec activation ELU**, Cette couche apprend des combinaisons complexes à partir des caractéristiques extraites, avec de bonnes performances en vitesse de convergence et en robustesse face aux valeurs extrêmes [55].
- **Couche de régularisation par Dropout**, permettant de prévenir le surapprentissage (overfitting) [56].
- **Couche de sortie linéaire**, adéquate pour la régression.

Cette structure a été utilisée avec succès dans plusieurs études appliquées à la prévision de séries environnementales, notamment en hydrologie [53], et en climatologie [57].

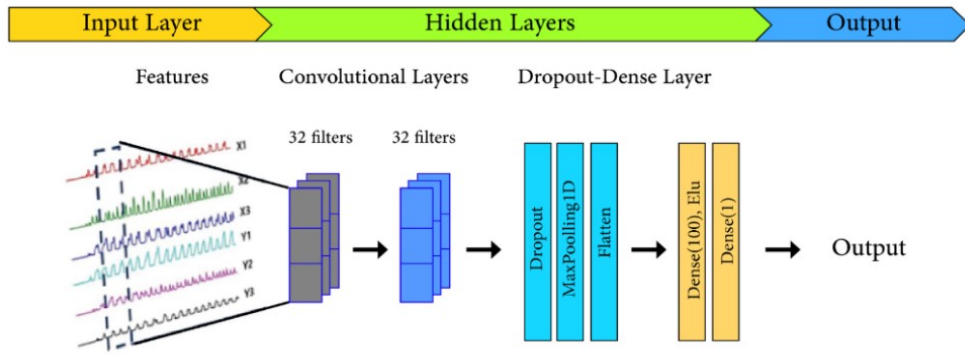


FIGURE 3.6 – Représentation schématique de l'architecture du modèle.

Fonction de perte :

La fonction de perte est un élément fondamental dans l'entraînement des réseaux de neurones. Elle mesure l'écart entre les valeurs réelles et les valeurs prédites par le modèle, et guide l'optimisation des poids.

Dans le cadre des tâches de régression, la fonction de perte la plus couramment utilisée est la MSE (Mean Squared Error) ou erreur quadratique moyenne, sa formule est donnée par :

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.4)$$

Avec :

- y_i : la valeur réelle (observée)
- \hat{y}_i : la valeur prédite par le modèle
- n : le nombre total d'exemples dans l'ensemble

Optimiseur Adam : Adaptive Moment Estimation :

L'entraînement d'un réseau de neurones consiste à ajuster les poids internes du modèle pour minimiser la fonction de perte. Pour ce faire, on utilise un algorithme d'optimisation. Dans ce travail, l'optimiseur Adam, introduit par [58], est utilisé, car il est particulièrement efficace pour les modèles profonds comme les CNN.

Adam combine les avantages de deux méthodes classiques :

- Momentum, qui accélère l'apprentissage en utilisant la moyenne des gradients passés.
- RMSProp, qui adapte le taux d'apprentissage à chaque paramètre.

Le fonctionnement de l'optimiseur Adam repose sur deux estimations :

1. Le premier moment (m) : une moyenne des gradients (représente la direction).

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (3.5)$$

2. Le second moment (v) : une moyenne des carrés des gradients (représente la vitesse).

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (3.6)$$

Avec :

$$\beta_1 \approx 0.9 \quad (\text{Momentum}) \quad \beta_2 \approx 0.999 \quad (\text{Stabilité})$$

À chaque itération t :

- On met à jour m et v en les combinant avec le gradient courant.
- On corrige les biais dus aux moyennes exponentielles.

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (3.7)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (3.8)$$

- Enfin, on met à jour les poids du modèle avec une formule adaptative :

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \quad (3.9)$$

où :

- θ_t : poids du modèle à l'itération t ,
- η : taux d'apprentissage,
- \hat{m}_t et \hat{v}_t : versions corrigées de m et v ,
- ε : constante pour éviter la division par zéro.

Avantages d'Adam :

- Convergence rapide et stable.
- Adapté aux données bruyantes ou rares (très utile en hydrologie et climatologie).
- Ne nécessite pas de réglage manuel du taux d'apprentissage pour chaque paramètre.
- Combine les bénéfices du Momentum et de l'adaptation locale du taux d'apprentissage.

3.4 Critères d'évaluation

3.4.1 Critere Numerique

L'évaluation des performances des modèles de régression repose sur l'utilisation de critères numériques qui permettent de quantifier l'écart entre les prédictions du modèle et les valeurs observées. Ces métriques, telles que l'erreur quadratique moyenne (MSE), l'erreur absolue moyenne (MAE), le coefficient de détermination (R^2) ou encore l'erreur en pourcentage symétrique ($sMAPE$), jouent un rôle fondamental dans la sélection, la comparaison et l'optimisation des modèles. Leur choix dépend des objectifs de l'analyse, de la nature des données et de la sensibilité souhaitée aux erreurs extrêmes. Une compréhension approfondie de ces critères est donc essentielle pour garantir une évaluation rigoureuse et pertinente des modèles prédictifs.

Dans ce qui suit, nous présentons une description détaillée des principales métriques utilisées dans notre étude pour évaluer les performances des modèles retenus.

Erreur Quadratique Moyenne (MSE) :

L'erreur quadratique moyenne (Mean Squared Error) mesure la moyenne des carrés des écarts entre les valeurs prédites et les valeurs réelles. Elle est très sensible aux grandes erreurs, ce qui en fait une métrique utile lorsque les grosses déviations sont pénalisées [59].

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.10)$$

Erreur Absolue Moyenne (MAE) : La Mean Absolute Error (MAE) représente la moyenne des écarts absolus entre les valeurs prédites et les valeurs observées. Moins sensible que la MSE aux valeurs extrêmes, elle est souvent utilisée pour une évaluation plus robuste [59].

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.11)$$

Le coefficient de détermination, noté R^2 :

Le R^2 , ou coefficient de détermination, mesure la proportion de la variance totale des données expliquée par le modèle. Il s'agit d'un indicateur global de la qualité de l'ajustement [59].

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (3.12)$$

Racine de l'Erreur Quadratique Moyenne (RMSE) :

La Root Mean Squared Error (RMSE) est la racine carrée de la MSE. Elle a l'avantage d'être exprimée dans la même unité que la variable cible, facilitant l'interprétation des résultats [59].

$$\text{RMSE} = \sqrt{\text{MSE}} \quad (3.13)$$

Coefficient de Détermination Ajusté (R^2 ajusté) :

Le R^2 ajusté corrige le R^2 standard en tenant compte du nombre de variables explicatives. Il est particulièrement pertinent pour comparer des modèles ayant des dimensions différentes.

$$R^2_{\text{adj}} = 1 - (1 - R^2) \cdot \frac{n - 1}{n - p - 1} \quad (3.14)$$

Où n , est le nombre d'observations, et p le nombre de prédictors [60].

Erreur Absolue en Pourcentage Symétrique (SMAPE) :

La Symmetric Mean Absolute Percentage Error (SMAPE) est une métrique en pourcentage qui prend en compte la symétrie des erreurs relatives, évitant notamment la division par zéro [59].

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^n \frac{2|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|} \quad (3.15)$$

Les métriques d'évaluation utilisées dans cette étude sont résumées dans le tableau ci-dessous, qui illustre les niveaux de performance associés à chaque indicateur. Ces seuils permettent d'interpréter les résultats des modèles de manière objective et comparative.

TABLEAU 3.1 – Intervalles d'appréciation des critères de validation pour variable précipitation

Paramètre	Très bon	Bon	Satisfaisant	Insatisfaisant
MSE	0–10	10–50	50–100	> 100
MAE	0–2	2–5	5–10	> 10
R^2	> 0,80	0,65–0,80	0,50–0,65	< 0,50
RMSE	0–3	3–6	6–10	> 10
R^2 ajusté	> 0,80	0,65–0,80	0,50–0,65	< 0,50
SMAPE	0–10 %	10–20 %	20–30 %	> 30 %

TABLEAU 3.2 – Intervalles d'appréciation des critères de validation pour la variable SPI.

Paramètre	Très bon	Bon	Satisfaisant	Insatisfaisant
R^2	> 0,80	0,65–0,80	0,50–0,65	< 0,50
$R^2_{\text{ajusté}}$	> 0,80	0,65–0,80	0,50–0,65	< 0,50
SMAPE	0–10 %	10–20 %	20–30 %	> 30 %

3.4.2 Critere Graphique

En complément des critères d'évaluation numériques tels que le MSE, le MAE ou le R^2 , des méthodes d'évaluation graphique ont également été mobilisées afin d'obtenir une analyse plus approfondie des performances des modèles prédictifs. Ces représentations visuelles permettent de diagnostiquer plus finement le comportement des modèles, notamment en ce qui concerne la distribution des résidus, les biais éventuels ou les erreurs systématiques.

À cet effet, le *QQ-plot* (Quantile-Quantile plot) a été utilisé pour vérifier l'hypothèse de normalité des erreurs, ce qui est essentiel pour valider certaines hypothèses statistiques sous-jacentes à la régression. D'autres visualisations telles que les comparaisons entre les prédictions et les observations réelles au fil du temps ont également été exploitées. Ces outils permettent d'identifier visuellement des phénomènes comme l'overfitting, l'underfitting, ou encore des relations non capturées par le modèle. L'intégration de ces évaluations graphiques renforce ainsi la robustesse et la fiabilité de l'analyse des résultats.

3.5 Conclusion

Ce chapitre a permis d'explorer les fondements théoriques, les catégories principales et les techniques avancées du Machine Learning, en mettant l'accent sur deux approches phares : XGBoost et les réseaux de neurones convolutifs (CNN). D'une part, XGBoost s'est imposé comme un algorithme de choix pour les données structurées en raison de sa capacité à modéliser des relations complexes, à intégrer la régularisation et à limiter le surapprentissage, tout

en bénéficiant d'une optimisation computationnelle poussée. D'autre part, les CNN, originellement conçus pour la vision par ordinateur, ont démontré leur efficacité dans l'analyse des séries temporelles grâce à leur aptitude à extraire automatiquement des motifs locaux et hiérarchiques, tout en offrant des performances robustes et une convergence rapide, notamment via l'optimiseur Adam.

Chapitre 4

MATÉRIELS ET MÉTHODES

4.1 Introduction

Dans le cadre des études d'impact du changement climatique, la nécessité de disposer de données climatiques fiables et spatialement détaillées s'impose avec force. Or, les modèles climatiques globaux (GCMs), bien qu'indispensables pour simuler le climat à l'échelle planétaire, présentent des limites majeures pour les applications à échelle régionale ou locale, notamment en raison de leur faible résolution spatiale et de leurs biais systématiques.

Afin de surmonter ces limitations et de produire des projections climatiques plus adaptées aux réalités locales, ce chapitre s'attache à décrire la stratégie méthodologique adoptée pour affiner les sorties brutes des GCMs à l'aide du downscaling statistique. Cette approche s'appuie sur des méthodes d'apprentissage automatique, combinant données de réanalyse (ERA5) et simulations issues des GCMs, pour restituer avec plus de finesse les signaux climatiques pertinents.

Deux techniques ont été mobilisées dans ce cadre : le réseau de neurones convolutif (CNN), reconnu pour sa capacité à modéliser des dépendances spatio-temporelles complexes, et le modèle XGBoost, apprécié pour sa robustesse et ses performances en régression. Ce chapitre détaille la démarche méthodologique mise en œuvre pour identifier la variable climatique cible la plus pertinente à downscaler, sélectionner le modèle le plus adapté à la projection future, et évaluer l'efficacité de l'ajustement statistique appliqué aux sorties issues du downscaling.

4.2 Région d'Étude

Le Nord-Centre de l'Algérie, composé de dix wilayas, représente seulement 1,49% de la superficie nationale. Pourtant, cette région s'impose comme la plus urbanisée, la plus industrialisée et la plus densément peuplée du pays. Elle concentre près d'un tiers de la population algérienne, passée de 7,7 millions en 1987 à plus de 10 millions en 2008. Cette zone abrite la capitale Alger ainsi que plusieurs grandes agglomérations urbaines, tout en intégrant une plaine agricole stratégique : la Mitidja. Considérée comme l'un des poumons agricoles du pays, la plaine de la Mitidja se distingue par la richesse de ses sols et son activité maraîchère et arboricole.

Par ailleurs, le Nord-Centre constitue un pôle industriel majeur, regroupant une part significative des infrastructures économiques et des unités de production nationales. Dans ce contexte

de forte pression démographique et de besoins croissants en ressources hydriques, le barrage de Keddara, situé dans la wilaya de Boumerdès, joue un rôle essentiel. Avec une capacité de stockage de 145,6 millions de mètres cubes (hm^3) et une localisation stratégique dans la chaîne côtière de l'Atlas tellien, ce barrage figure parmi les principaux ouvrages hydrauliques assurant l'approvisionnement en eau potable et l'irrigation dans la région.

C'est dans cette perspective que cette étude se concentre sur le barrage de Keddara, qui constitue une référence majeure pour l'analyse de la gestion et de la mobilisation des ressources en eau dans le Nord-Centre de l'Algérie.

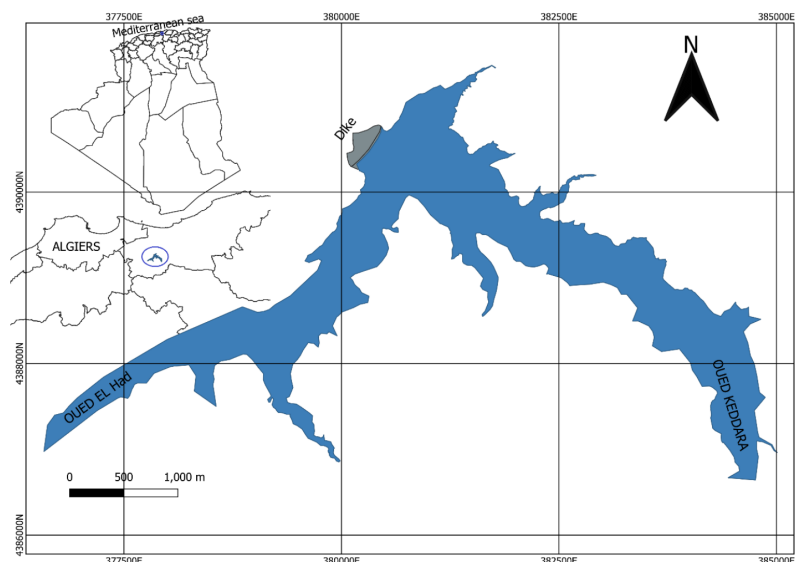


FIGURE 4.1 – Localisation géographique et délimitation du lac de retenue du barrage de Keddara [8]

4.3 Méthodologie

Afin de garantir une projection climatique à la fois fiable et opérationnelle, en appui à la prise de décision et à la mise en œuvre de mesures d'adaptation, une attention particulière a été portée à l'étape du downscaling statistique.

Contrairement à une approche unique et directe, cette étude repose sur une exploration systématique de plusieurs combinaisons méthodologiques, dans le but d'identifier la démarche la plus robuste et la plus représentative du signal climatique local. Cette stratégie s'inscrit dans une logique d'optimisation intégrant les choix des données d'entrée, les cibles climatiques à affiner, l'application du quantile mapping comme méthode d'ajustement statistique, ainsi que les algorithmes de downscaling testés.

Les données mobilisées sont regroupées en deux grandes catégories :

- (i) les données de réanalyse (telles que ERA5), considérées comme une référence à haute résolution spatiale ;
- (ii) les données issues des modèles climatiques globaux (GCMs), caractérisées par une résolution plus grossière et nécessitant un raffinement spatial pour une utilisation à l'échelle locale.

Deux types de cibles climatiques ont été considérés dans les scénarios de downscaling :

- les précipitations mensuelles ;
- l'indice standardisé de précipitation (SPI), calculé aux échelles temporelles de 1, 3, 6 et 12 mois (SPI-1, SPI-3, SPI-6, SPI-12), permettant de représenter différents types de sécheresse.

Deux approches d'apprentissage automatique ont été mobilisées pour le downscaling :

- un réseau de neurones convolutif (CNN), adapté à la capture de motifs spatio-temporels complexes ;
- le modèle XGBoost (eXtreme Gradient Boosting), reconnu pour ses performances en classification et en régression sur des ensembles de données hétérogènes et bruitées.

Par ailleurs, l'effet de l'ajustement statistique par quantile mapping a été appliqué afin d'améliorer la distribution statistique du climat futur downscalé, selon l'algorithme et le paramètre cible déterminés à l'issue de l'évaluation des performances des modèles.

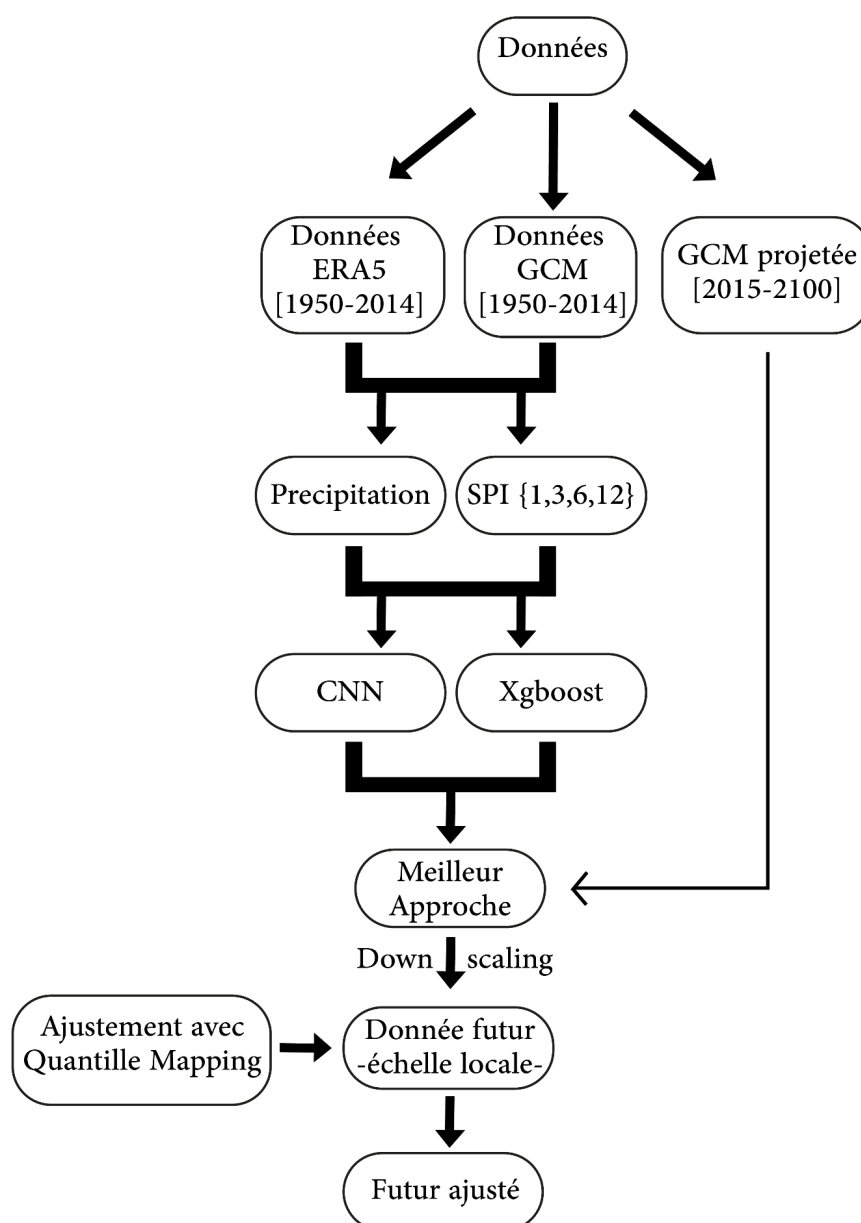


FIGURE 4.2 – Shéma de la méthodologie

4.4 Données utilisées pour le downscaling statistique

La mise en œuvre du downscaling statistique repose sur la combinaison de deux principales sources de données climatiques, permettant de faire le lien entre l'échelle globale des modèles climatiques et l'échelle locale, plus pertinente pour l'analyse des impacts.

Données de réanalyse :

Les données de réanalyse constituent une référence à haute résolution spatiale. Elles représentent les conditions climatiques observées à l'échelle régionale en intégrant des observations multiples dans un modèle atmosphérique global. Ces données fournissent une base fiable et cohérente pour caractériser le climat régional et sont utilisées comme variable explicative dans

les modèles de downscaling statistique.

Données des modèles climatiques globaux (GCM) :

Les modèles climatiques globaux (GCM) fournissent des projections climatiques à large échelle, généralement à une résolution spatiale relativement grossière. Ces projections nécessitent un ajustement via le downscaling statistique pour être exploitables à l'échelle locale, notamment pour étudier la variabilité spatio-temporelle des précipitations et les épisodes de sécheresse.

Ces deux types de données sont indispensables pour développer, entraîner et valider les modèles de downscaling statistique, assurant ainsi une meilleure précision dans la représentation des phénomènes climatiques locaux à partir des projections globales.

4.4.1 Données de référence locales et calcul du SPI :

1- Réanalyse ERA5 – Source de référence :

Pour représenter les conditions climatiques locales, l'étude s'est appuyée sur les données de réanalyse ERA5, produites par le Centre Européen pour les Prévisions Météorologiques à Moyen Terme (ECMWF) dans le cadre du programme Copernicus Climate Change Service (C3S). ERA5 est la cinquième génération de réanalyse climatique, succédant à ERA-Interim, et constitue aujourd'hui la réanalyse globale la plus complète.

La base *ERA5_mensuelles*, utilisée dans cette étude [61], fournit une reconstitution cohérente des conditions atmosphériques, océaniques et de surface depuis 1940, avec une résolution spatiale élevée de $0,25^\circ \times 0,25^\circ$ (~ 31 km). Elle est mise à jour quotidiennement, avec une latence d'environ cinq jours, et accessible au format NetCDF via le *Climate Data Store* (programme Copernicus). Son exploitation est facilitée par des bibliothèques Python spécialisées telles que *Xarray* et *CD0*, couramment utilisées pour les traitements climatologiques sous Python.

2- Extraction des données au voisinage des stations :

Pour garantir une comparabilité spatiale rigoureuse entre les données de réanalyse et les observations in situ, les valeurs de précipitations issues d'ERA5 ont été extraites aux emplacements les plus proches des stations hydrométriques considérées.

Cette opération s'est appuyée sur le calcul de la distance de Haversine, via la fonction *geodesic()* de la bibliothèque Python *Geopy*, permettant d'identifier automatiquement le point de grille ERA5 le plus proche de chaque station, et ainsi de préserver les caractéristiques climatiques locales.

3- Calcul de l'indice SPI :

L'indice de précipitation standardisé (SPI) a été utilisé comme indicateur clé pour quantifier les anomalies pluviométriques à différentes échelles temporelles (SPI-1, SPI-3, SPI-6, SPI-12).

Le SPI permet d'évaluer :

- les sécheresses météorologiques (court terme, SPI-1) ;
- les sécheresses agricoles (moyen terme, SPI-3/SPI-6) ;
- les sécheresses hydrologiques (long terme, SPI-12).

Son calcul repose sur l'ajustement d'une loi Gamma aux précipitations agrégées, suivi d'une transformation en loi normale centrée réduite, ce qui facilite l'analyse statistique et la comparaison inter-sites.

La première étape du calcul de l'indice SPI consiste à choisir l'échelle temporelle appropriée selon le type de sécheresse à diagnostiquer (SPI-1 pour les sécheresses météorologiques, SPI-3 à SPI-6 pour les sécheresses agricoles, SPI-12 pour les sécheresses hydrologiques). Une fois les précipitations agrégées selon cette échelle, il est nécessaire d'ajuster une loi de probabilité à la distribution empirique de ces données.

Bien que plusieurs distributions soient envisageables (log-normale, normale, Weibull), la loi Gamma à deux paramètres est la plus fréquemment utilisée en raison de sa flexibilité à modéliser les distributions asymétriques, caractéristiques des séries pluviométriques. Ce choix est largement validé dans la littérature scientifique (McKee et al., 1993 ; Guttman, 1999 ; Stagge et al., 2015 ; Oladipo et al., 2022 ; Aghelpour et al., 2020).

La fonction de densité de probabilité (PDF) de la loi Gamma est donnée par l'expression suivante, pour les valeurs réelles strictement positives ($x > 0$) et s'exprime comme suit :

$$f(x, \alpha, \beta) = \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad x > 0, \alpha > 0, \beta > 0 \quad (4.1)$$

où :

- α : paramètre de forme (shape)
- β : paramètre d'échelle (scale)
- $\Gamma(\alpha)$: fonction Gamma d'Euler, définie par :

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \quad (4.2)$$

Une fois les paramètres α et β estimés, la fonction de répartition cumulative (CDF) est calculée, puis transformée en valeurs de la loi normale centrée réduite (Z) pour produire l'indice SPI.

Cependant, les séries de précipitations contiennent fréquemment des valeurs nulles, notamment aux courtes échelles temporelles (SPI-1 ou SPI-3). Étant donné que la loi Gamma n'est définie que pour les valeurs strictement positives, il est nécessaire d'ajuster la distribution afin de tenir compte de la probabilité d'occurrence de zéros dans les données.

Dans ce cadre, une distribution composée est définie. La fonction de répartition cumulative ajustée, notée $H(x)$, est exprimée comme suit :

$$H(x) = \begin{cases} q, & \text{si } x = 0 \\ q + (1 - q) \cdot G(x), & \text{si } x > 0 \end{cases} \quad (4.3)$$

où :

- q est la probabilité empirique d'observer une précipitation nulle, soit $q = m/n$, avec m le nombre de zéros dans la série et n le nombre total d'observations ;
- $G(x)$ est la fonction de répartition de la loi Gamma ajustée à la partie non nulle de la série.

Cette formulation permet de modéliser de manière cohérente la distribution empirique des précipitations en intégrant les jours sans pluie, tout en conservant les propriétés statistiques nécessaires à la transformation en indice normalisé SPI.

4- Transformation finale et calcul de l'indice SPI :

Une fois la fonction de répartition cumulative ajustée $H(x)$ obtenue, celle-ci est transformée en variable normale centrée réduite (moyenne nulle, écart-type égal à 1) à l'aide de l'inverse de la fonction de répartition de la loi normale standard. Le résultat de cette transformation constitue la valeur du SPI, exprimée en unités d'écart-type, et permettant une interprétation statistique standardisée de l'anomalie pluviométrique.

Sur la base des séries temporelles extraites, l'indice SPI a été calculé à différentes échelles temporelles :

- **SPI-1** : échelle mensuelle ;
- **SPI-3** et **SPI-6** : échelles saisonnières ;
- **SPI-12** : échelle annuelle.

Ces indices ont été évalués à partir :

- des données de réanalyse ERA5,
- ainsi que des données in situ issues des stations de l'Agence Nationale des Ressources Hydrauliques (ANRH), pour la période d'observation disponible.

5- Vérification de la loi de distribution optimale :

Avant de procéder au calcul de l'indice SPI, il est essentiel de vérifier si la loi Gamma constitue le meilleur ajustement pour les données de précipitations étudiées. En effet, bien que couramment utilisée, cette distribution n'est pas universellement la plus adaptée, et un mauvais ajustement peut compromettre la fiabilité de l'indice.

Pour cela, la bibliothèque `fitter` a été utilisée dans un objectif de comparaison systématique de plusieurs lois de probabilité (Gamma, log-normale, Weibull, etc.) sur la base des données empiriques.

Cette étape préliminaire vise à modéliser avec précision la variabilité statistique des précipitations, condition indispensable à un calcul fiable du SPI :

- identifier la loi la plus représentative de la distribution des précipitations,
- garantir la robustesse du calcul du SPI,
- et minimiser les erreurs d'ajustement, qui peuvent avoir un impact significatif sur la détection des anomalies sèches.

Prétraitement et estimation des paramètres

Les séries de précipitations sont tout d'abord converties en vecteurs numériques compatibles avec les fonctions de la bibliothèque `fitter`.

L'algorithme procède ensuite, pour chaque distribution candidate (Gamma, Log Normale, Weibull, Normale, etc.), à une estimation des paramètres par la méthode du maximum de vraisemblance, en s'appuyant sur les fonctions de `scipy.stats`.

Critères de sélection

Une fois les paramètres estimés, `fitter` évalue la qualité de l'ajustement à l'aide de plusieurs critères statistiques, notamment :

- la somme des carrés des erreurs (SSE) ;
- le critère d'information d'Akaike (AIC) ;
- et le critère bayésien d'information (BIC).

Ces critères permettent de mesurer l'adéquation entre les données observées et le modèle théorique, tout en intégrant la complexité du modèle afin d'éviter les phénomènes de surajustement.

Dans cette étude, le critère d'information d'Akaike (AIC) a été retenu comme critère principal de comparaison. Il offre un compromis équilibré entre la précision de l'ajustement et la parcimonie du modèle, en pénalisant les distributions comportant un nombre excessif de paramètres.

6-Sélection finale et intégration au calcul du SPI :

La distribution optimale est identifiée à l'aide de la méthode `get_best()` de la bibliothèque `Fitter`, en spécifiant le critère de sélection (ici, l'AIC). Cette méthode retourne :

- la distribution présentant la valeur d'AIC la plus faible,
- ainsi que ses paramètres estimés.

La distribution retenue est ensuite utilisée comme base pour l'estimation des probabilités dans le cadre du calcul de l'indice SPI, en remplacement éventuel de la loi Gamma si celle-ci n'est pas optimale.

TABLEAU 4.1 – Tableau récapitulatif des meilleures distributions statistiques ajustées aux séries de précipitations (in situ et réanalyses)

Station	Colonne	Meilleure distribution
Keddara	insitu_pr	Gamma
Keddara	reanalys_pr	Gamma

7- Validation des données de réanalyse :

Les figures suivantes présentent la comparaison entre les indices SPI calculés à partir des données ERA 5 et ceux issus des observations in situ pour la station d'étude .

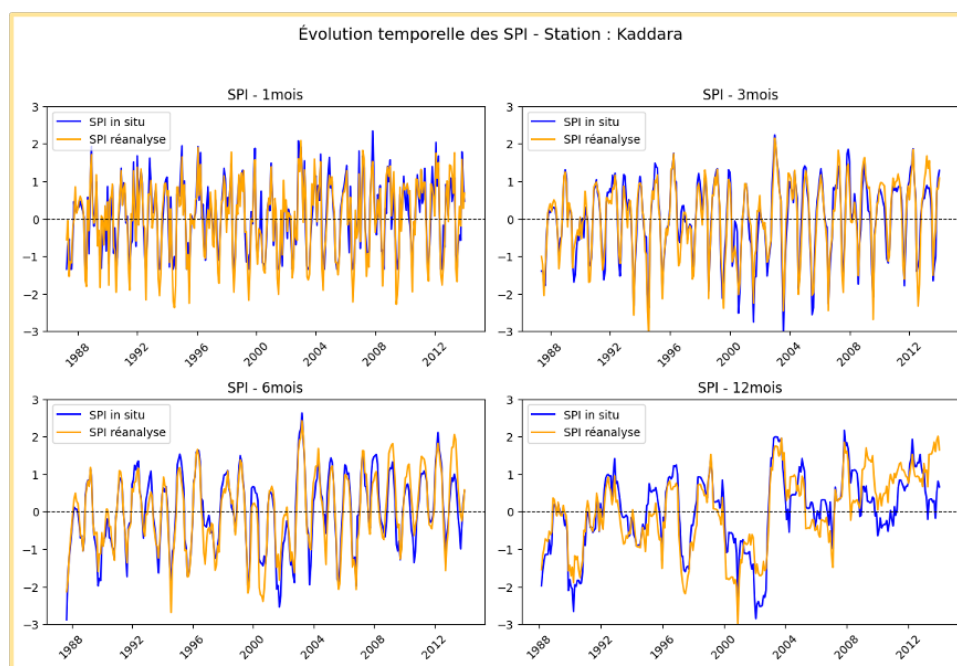


FIGURE 4.3 – Calcul du SPI multiéchelle à partir des données in situ et des données de réanalyse

Il apparaît que les données de réanalyse reproduisent de manière satisfaisante les dynamiques historiques observées, aussi bien en termes de tendances générales que de variabilité interannuelle. La concordance entre les deux sources de données (observations in situ et données issues d'ERA5) se révèle particulièrement nette sur l'ensemble des séries temporelles analysées.

Cette cohérence est observée non seulement pour les valeurs moyennes, mais également pour les extrêmes pluviométriques (pics de sécheresse ou épisodes pluvieux intenses), ainsi que pour les périodes intermédiaires, témoignant de la fiabilité des jeux de données ERA5 dans le contexte de l'étude.

4.4.2 Données à l'échelle globale : Sélection des modèles climatiques globaux (GCM)

Dans le cadre de cette étude, une sélection rigoureuse de modèles climatiques globaux (GCM) a été effectuée, en tenant compte à la fois de la pertinence scientifique, de la qualité des données disponibles et de leur compatibilité avec le cas d'étude.

La sélection s'est basée sur des modèles issus du projet CMIP6 (Coupled Model Intercomparison Project, Phase 6), reconnu comme référence internationale pour l'étude du climat à l'échelle planétaire. Les modèles ont été choisis parmi ceux proposés par plusieurs institutions scientifiques de renom, selon leur couverture temporelle, leur résolution, la disponibilité de variables nécessaires, et la diversité de leurs Réalisations Aléatoires Numériques (RAN).

Les simulations climatiques associées ont été téléchargées via le portail ESGF (Earth System Grid Federation), à travers l'interface proposée par l'Institut Pierre-Simon Laplace (IPSL), accessible à l'adresse suivante : <https://esgf-node.ipsl.upmc.fr/projects/cmip6/>

Cette base permet l'accès à une large gamme de simulations historiques et futures, dans différents scénarios d'émissions (SSP126, SSP245, SSP370, SSP585), indispensables pour explorer

les incertitudes liées au changement climatique.



FIGURE 4.4 – Interface du portail ESGF-IPSL pour la recherche et le téléchargement des simulations climatiques CMIP6

Ce site constitue l'un des nœuds principaux du réseau mondial ESGF (Earth System Grid Federation), qui fournit un accès ouvert, structuré et standardisé aux données climatiques produites par une large variété de modèles issus d'institutions internationales reconnues. Grâce à cette infrastructure, les chercheurs peuvent télécharger des ensembles de données homogènes, conformes aux protocoles du projet CMIP6, ce qui facilite à la fois la comparabilité intermodèles, la reproductibilité des résultats et la transparence des analyses climatiques.

Dans le cadre de cette étude, nous avons retenu plusieurs modèles GCM ainsi que leurs Réalisations Aléatoires Numériques (RAN), selon des critères de représentativité géographique, de disponibilité temporelle et de pertinence vis-à-vis des objectifs scientifiques. La sélection a privilégié les modèles offrant :

- une couverture temporelle continue (périodes historique et de projection),
- une résolution spatiale raisonnable,
- et une accessibilité aux variables climatiques clés (notamment les précipitations mensuelles).

Le tableau ci-dessous présente les modèles climatiques retenus, accompagnés de leurs principales caractéristiques :

TABLEAU 4.2 – Caractéristiques des modèles climatiques globaux (GCM) et des réalisations sélectionnées pour l'étude

Modèle GCM	Institution	Pays	Type de modèle	Résolution nominale [km]	RAN
GISS-E2-1-G	NASA Goddard Institute for Space Studies (GISS)	États-Unis	Atmosphère-Océan couplé	250	r26i1p1
CNRM-CM6-1	CNRM & CERFACS	France	Atmosphère-Océan couplé	250	r18i1p1
					r16i1p1
					r12i1p1
					r10i1p1
					r11i1p1
					r12i1p2
					r6i1p1
					r20i1p1
					r17i1p1
IPSL-CM6A-LR	Institut Pierre-Simon Laplace (IPSL)	France	Atmosphère-Océan couplé	250	r14i1p1 r19i1p1 r18i1p1 r26i1p1 r15i1p1
CAS-ESM2-0	Chinese Academy of Sciences (CAS)	Chine	Système Terre couplé	100	r1i1p1
CNRM-CM6-1-HR	CNRM & CERFACS	France	Atmosphère-Océan couplé (haute résolution)	50	r1i1p1
CNRM-ESM2-1	CNRM & CERFACS	France	Système Terre couplé avec cycle biogéochimique	250	r7i1p1
MRI-ESM2-0	Meteorological Research Institute (MRI)	Japon	Système Terre couplé	100	r7i1p1
MIROC-ES2H	JAMSTEC, AORI & NIES	Japon	Système Terre couplé à haute résolution biogéochimique	250	r1i1p1

- **RAN** : Réalisation Aléatoire Numérique ;
- **SSP** : Shared Socioeconomic Pathway (scénario d'émissions).

Variables climatiques sélectionnées :

Dans le cadre de cette étude, les données issues de plusieurs modèles climatiques globaux

(GCM) ont été mobilisées non seulement pour la variable principale à affiner (la précipitation (pr)), mais également pour un ensemble élargi de variables explicatives, sélectionnées en fonction de leur pertinence climatique et de leur disponibilité.

En effet, les approches de downscaling statistique nécessitent la construction d'un vecteur prédictif multidimensionnel, capable de capturer les interactions complexes entre les variables climatiques à large échelle et les phénomènes météorologiques locaux. Cette démarche vise à améliorer la précision et la robustesse des projections régionales en tenant compte des facteurs physiques influençant la dynamique des précipitations.

Les variables climatiques retenues ont été choisies sur la base de leur influence directe ou indirecte sur les précipitations, comme le suggèrent de nombreux travaux dans la littérature scientifique (Benestad et al., 2008 ; Maraun et al., 2010 ; Gutmann et al., 2014).

Ces variables ont été extraites à l'échelle mensuelle, sur la base de leur cohérence temporelle avec les observations et de leur disponibilité dans l'ensemble des modèles GCM retenus. Elles constituent les entrées du modèle de downscaling appliqué à la prédiction de la variable cible (précipitation ou SPI).

TABLEAU 4.3 – Liste des variables prédictives issues des GCM et leur influence sur les précipitations

Nom de la variable	Description	Unité	Lien avec les précipitations (influence climatique)
pr	Précipitations totales mensuelles	mm/mois	C'est la variable cible à affiner (prédicand) dans le modèle de downscaling.
evspsbl	Évaporation réelle à la surface (évapotranspiration)	mm/mois	Affecte le bilan hydrique : une évaporation élevée réduit l'eau disponible pour les précipitations. Elle est aussi liée aux pertes d'humidité du sol.
mrso	Humidité du sol (soil moisture)	kg/m ² ou % de saturation	Reflète la capacité du sol à retenir l'eau. Un sol saturé augmente le risque de précipitations convectives, tandis qu'un sol sec limite l'humidité disponible pour l'atmosphère.
psl	Pression atmosphérique au niveau de la mer	hPa (ou Pa)	Indique les systèmes de haute ou basse pression : les basses pressions sont généralement associées à des conditions favorables aux précipitations.
rsds	Rayonnement solaire incident en surface	W/m ²	Contrôle le réchauffement de la surface, ce qui influence les mouvements verticaux (convection) et donc le développement des nuages et pluies.

Suite du tableau 4.3

Nom de la variable	Description	Unité	Lien avec les précipitations (influence climatique)
sfcWind	Vitesse du vent en surface	m/s	Influence le transport d'humidité atmosphérique : des vents forts peuvent amener ou éloigner l'humidité des bassins versants.
tas	Température moyenne de l'air à 2 mètres	°C	Conditionne l'évaporation et la teneur en vapeur d'eau de l'air, deux éléments essentiels à la formation des précipitations.
tasmax	Température maximale quotidienne	°C	Impacte l'évapotranspiration et peut intensifier l'activité convective lors des épisodes chauds.
tasmin	Température minimale quotidienne	°C	Agit sur les processus de condensation nocturne, et peut influencer le déclenchement de certains régimes pluviométriques.

L'ensemble de ces variables constitue ainsi le vecteur prédictif utilisé pour l'entraînement des modèles de downscaling statistique. En combinant des informations issues de différents champs climatiques (thermiques, dynamiques, hydriques et radiatifs), ce vecteur permet de mieux représenter les processus physiques sous-jacents à la variabilité spatiale et temporelle des précipitations.

Cette approche multidimensionnelle renforce la capacité des modèles à capturer les relations non linéaires, les effets d'interaction entre variables, ainsi que les signatures climatiques régionales qui influencent fortement le comportement des précipitations à l'échelle locale. Elle constitue ainsi un levier essentiel pour améliorer la précision des projections climatiques régionales et renforcer leur utilité pour l'analyse d'impacts et l'aide à la décision.

4.5 Le Quantile Mapping (QM) comme méthode d'ajustement post-downscaling

Afin d'améliorer la représentativité statistique des données climatiques projetées à l'échelle locale, cette étude a recours à la méthode du Quantile Mapping (QM), appliquée ici comme un ajustement a posteriori des sorties du downscaling statistique. Bien que cette technique soit couramment utilisée comme outil de correction des biais systématiques présents dans les simulations brutes des modèles climatiques globaux (GCM), son cadre statistique souple permet également de l'utiliser dans un objectif d'optimisation des distributions projetées, après l'étape de raffinement spatial.

Contrairement à l'approche traditionnelle où le QM intervient en amont du downscaling, cette étude adopte une stratégie inversée : l'ajustement est appliqué après la projection locale, dans le but de réaligner la distribution statistique des données climatiques futures downscalées avec celles issues des références historiques (réanalyses ou GCMs).

L'implémentation technique s'appuie sur la fonction `quantile_correction()` de la bibliothèque Python `bias_correction`. Trois séries temporelles sont exploitées.

- **obs_data** : données d'observation de référence à haute résolution spatiale, issues de la réanalyse ERA5 ;
- **mod_data** : données historiques simulées et downscalées à partir des GCM ;
- **sce_data** : données futures issues des GCM, après downscaling statistique.

Le paramètre `modified=False` active la version classique du Quantile Mapping, dans laquelle la tendance climatique projetée n'est pas explicitement conservée (forme non paramétrique). Cette configuration est jugée adéquate dans ce contexte où l'objectif principal est d'obtenir des projections localement réalistes et statistiquement fiables.

Principe de fonctionnement :

Pour chaque valeur issue de *sce_data*, *leQM* :

- **obs_data** : données d'observation de référence à haute résolution spatiale, issues de la réanalyse ERA5 ;
- **mod_data** : données historiques simulées et downscalées à partir des GCM ;
- **sce_data** : données futures issues des GCM, après downscaling statistique.

Ce processus permet d'ajuster les distributions statistiques futures sur la structure empirique observée, tout en conservant la variabilité intra-annuelle et les caractéristiques extrêmes.

4.6 Techniques de Downscaling Statistique

Afin de mieux représenter les caractéristiques climatiques locales à partir des sorties de modèles climatiques globaux (GCM), deux approches issues de l'intelligence artificielle ont été mises en œuvre dans cette étude :

- XGBoost, une méthode d'apprentissage automatique supervisée basée sur le gradient boosting ;
- Les réseaux de neurones convolutifs (CNN), relevant de l'apprentissage profond (deep learning), adaptés aux structures spatio-temporelles complexes.

Ces techniques ont été sélectionnées pour leur capacité à modéliser des relations non linéaires complexes entre les prédicteurs à grande échelle (variables explicatives issues des GCM) et la variable cible locale (précipitations ou SPI). Cette approche comparative vise à évaluer dans un cadre rigoureux l'efficacité respective de modèles classiques de machine learning et de modèles profonds pour le raffinement spatial (downscaling) des projections climatiques.

4.6.1 Approche XGBoost (Extreme Gradient Boosting)

Développé par Chen et Guestrin (2016), XGBoost est un algorithme basé sur l'assemblage d'arbres de décision faiblement performants (boosting), avec régularisation intégrée, optimisation parallèle, et mécanismes de contrôle du surapprentissage (overfitting). Il est reconnu pour ses performances exceptionnelles dans des contextes à forte hétérogénéité et bruit des données.

Dans cette étude, XGBoost a été utilisé pour prédire :

- la précipitation mensuelle (**rea_pr** [mm/mois]) dans l'approche 1 ;
- l'indice SPI (**SPI_rea_pr**) dans l'approche 2, calculé à partir d'un ensemble de variables prédictives dérivées, incluant notamment des variables retardées (*lagged features*) pour capturer la dépendance temporelle.

Étapes de mise en œuvre :

a. Prétraitement des données :

Pour chaque fichier climatique (correspondant à un GCM donné), les étapes suivantes ont été exécutées :

- lecture et structuration des données au format temporel (indexation par date) ;
- suppression des colonnes vides ou inutilisables ;
- création de variables explicatives retardées, telles que :
 - `pr_t-1` : précipitation du mois précédent ;
 - `SPI_t-1` : indice SPI du mois précédent.

b. Découpage des ensembles de données

Les données ont été divisées en trois sous-ensembles temporellement cohérents : Train + Validation : 80 % du jeu de données ; Validation seule : 10 % du total Train+Validation (utilisée pour le tuning) ; Test : 20 % du total, réservé à l'évaluation finale des performances.

c. Optimisation des hyperparamètres

L'optimisation des hyperparamètres a été réalisée grâce à la bibliothèque Optuna, qui permet une exploration intelligente de l'espace de recherche en minimisant l'erreur quadratique moyenne (MSE) sur l'ensemble de validation.

Les hyperparamètres optimisés incluent, entre autres :

- `n_estimators` : nombre d'arbres ;
- `max_depth` : profondeur maximale des arbres ;
- `learning_rate` : taux d'apprentissage ;
- `subsample` : fraction des échantillons utilisés pour chaque arbre ;
- `reg_alpha` : terme de régularisation L1.

TABLEAU 4.4 – Hyperparamètres du Modèle

Hyperparamètre	Intervalle de recherche	Description
<code>n_estimators</code>	100 à 500	Nombre total d'arbres
<code>max_depth</code>	1 à 20	Profondeur maximale des arbres
<code>learning_rate</code>	0.001 à 0.01	Taux d'apprentissage pour la mise à jour des poids
<code>subsample</code>	0.5 à 1.0	Fraction d'échantillons utilisés par arbre
<code>alpha</code>	0.0 à 1.0	Poids de la régularisation L1 (Lasso)

La fonction objectif évalue chaque combinaison d'hyperparamètres sur 50 itérations ($n_{\text{trials}} = 50$).

Entraînement final et évaluation :

Après optimisation, le meilleur jeu d'hyperparamètres est utilisé pour réentraîner le modèle sur l'ensemble **Train + Validation**.

Les performances sont mesurées sur **Train + Validation** ainsi que sur l'ensemble **Test**, à l'aide des métriques suivantes :

- MSE : Mean Squared Error ;

- MAE : Mean Absolute Error ;
- R^2 : coefficient de détermination ;
- RMSE : Root Mean Squared Error ;
- R^2 ajusté : ajusté pour le nombre de paramètres ;
- sMAPE : Symmetric Mean Absolute Percentage Error.

Des *QQ plots* sont générés afin de comparer les valeurs prédites et observées sur les ensembles **Train** + **Validation** et **Test**.

Ci-dessous un schéma récapitulatif de l'algorithme utilisé :

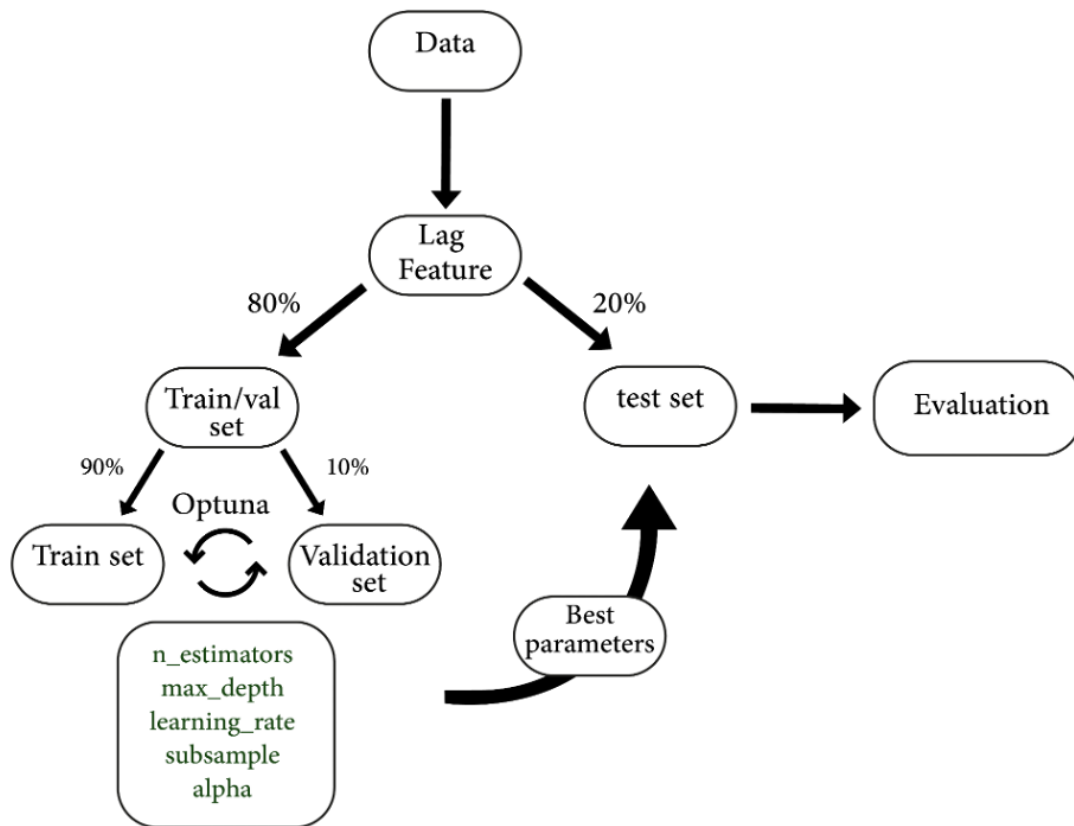


FIGURE 4.5 – Schéma récapitulatif de l'algorithme XGBoost

4.6.2 Approche par Réseaux de Neurones Convolutifs (CNN)

Les Convolutional Neural Networks (CNN) sont des architectures de réseaux de neurones artificiels particulièrement adaptées au traitement de données structurées en grille, comme les images ou les séries temporelles. Inspirés du fonctionnement du cortex visuel humain, ils sont capables d'extraire automatiquement des motifs hiérarchiques dans les données d'entrée.

Dans le cadre de cette étude, une architecture 1D CNN a été utilisée pour modéliser deux variables cibles :

- Les précipitations mensuelles (**rea_pr** [mm/mois]) dans l'approche 1 ;
- L'indice SPI (**SPI_rea_pr**) dans l'approche 2.

Ces modèles s'appuient sur un ensemble de variables explicatives dérivées des GCMs, incluant des

décalages temporels (lag) pour tenir compte de la dynamique historique du climat.

Étapes de mise en œuvre :

a. Prétraitement des données

Pour chaque fichier de données (chaque GCM), les opérations suivantes ont été effectuées :

- lecture des données et indexation temporelle via la colonne **Date** ;
- suppression des colonnes vides afin d'assurer une entrée propre au réseau ;
- séparation entre les variables explicatives et la variable cible ;
- normalisation indépendante des entrées et de la sortie à l'aide de `MinMaxScaler()`, ce qui permet de stabiliser et d'accélérer la convergence du modèle ;
- création de variables *laggées* avec une fenêtre temporelle de 3 pas ($\text{lag} = 3$), afin de capter les dépendances temporelles dans les séries climatiques.

b. Division des ensembles de données

Les données normalisées ont été séparées en :

- **Ensemble Train + Validation** : 67 % des données totales ;
- **Ensemble Test** : 33 % des données totales.

Une validation interne a été réalisée via un fractionnement de 30 % de l'ensemble d'entraînement, utilisé lors de l'entraînement du modèle CNN.

c. Architecture du modèle CNN

L'architecture du CNN 1D utilisée est décrite comme suit :

TABLEAU 4.5 – Architecture du modèle CNN utilisé

Composant	Détails techniques
Entrée	Séquence temporelle de 3 pas ($\text{lag} = 3$), chaque pas contenant n variables explicatives
Convolution 1D	32 filtres, fonction d'activation LeakyReLU
MaxPooling 1D	Réduction de la dimension temporelle, extraction des motifs dominants
Flatten	Passage d'un tenseur 3D à une entrée 2D pour la couche dense
Dense (cachée)	100 neurones, activation ELU, Dropout pour la régularisation
Dense (sortie)	1 neurone linéaire (prédiction de la variable cible)
Optimiseur	Adam, fonction de perte : Mean Squared Error (MSE)

d. Entraînement du modèle

- Nombre d'époques : 150
- Taille de batch : 100
- Validation split : 0.3 (interne à l'ensemble d'entraînement)
- Ces paramètres ont été choisis pour favoriser la stabilité de l'entraînement tout en réduisant le risque de surapprentissage (overfitting).

e. Évaluation des performances

Une fois l'entraînement terminé :

- les prédictions effectuées sur les ensembles **Train** et **Test** sont reconverties à leur échelle d'origine à l'aide de la méthode `inverse_transform()` du scaler ;
- les performances du modèle CNN sont évaluées à l'aide des mêmes métriques que celles utilisées pour XGBoost (RMSE, MAE, R^2) ;
- des QQ-plots sont générés afin de comparer visuellement la distribution des valeurs prédites à celle des valeurs observées, sur les ensembles **Train + Validation** et **Test** ;
- les deux modèles (XGBoost et CNN) sont récapitulés dans le schéma synthétique ci-dessous.

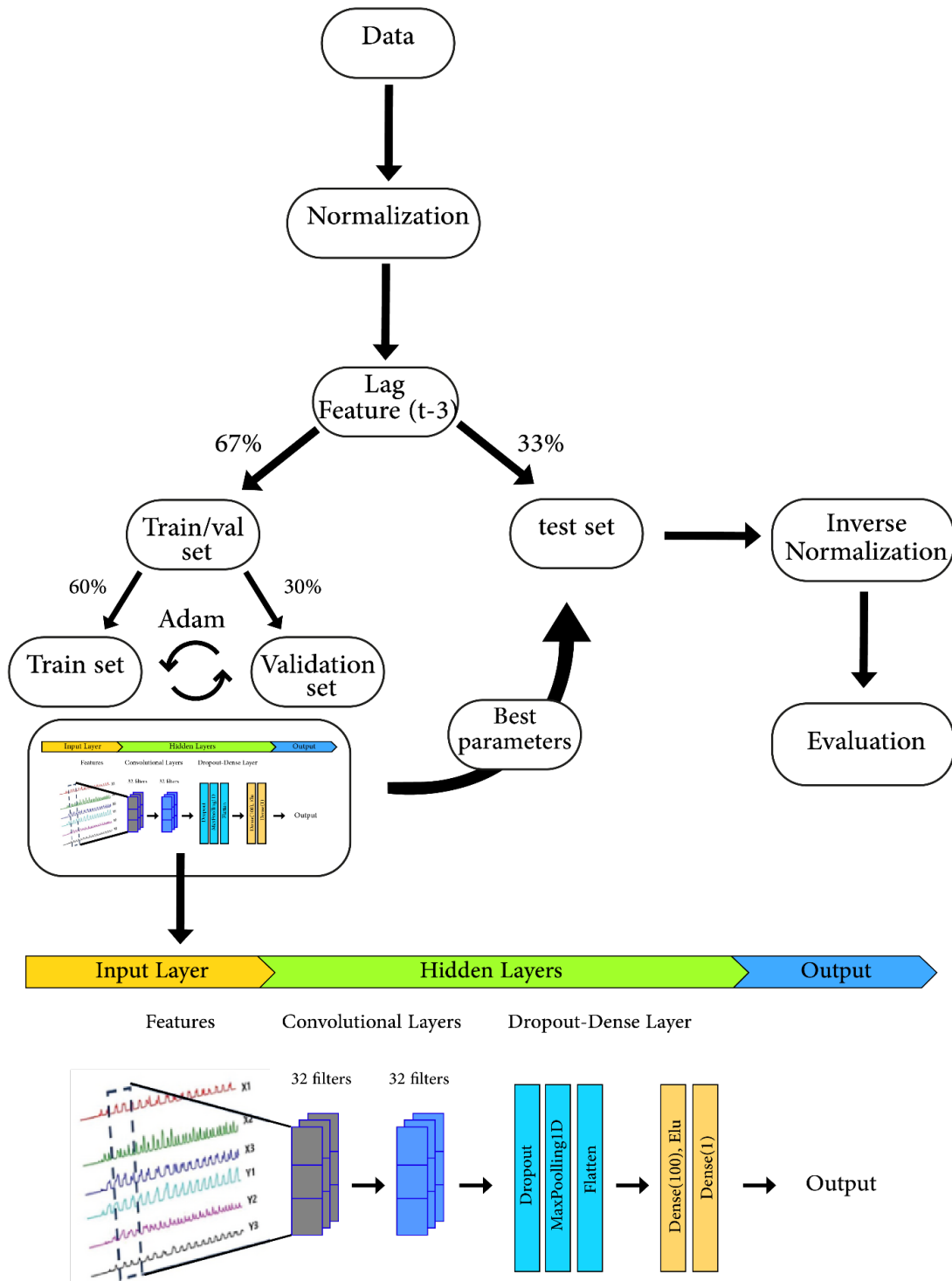


FIGURE 4.6 – schéma récapitulatif du modèle CNN

4.7 Conclusion

Ce chapitre a présenté la stratégie méthodologique adoptée pour affiner les projections climatiques brutes issues des modèles climatiques globaux (GCMs), en vue d'une meilleure représentation du signal climatique à l'échelle locale. Face aux limites intrinsèques des GCMs, notamment leur résolution spatiale grossière et leurs biais systématiques, un downscaling statistique a été mis en œuvre, mobilisant deux approches d'apprentissage automatique : le réseau de neurones convolutif (CNN) et le modèle XGBoost. Ces modèles ont permis de générer des données climatiques raffinées à partir de plusieurs variables explicatives issues à la fois des GCMs et de la réanalyse ERA5.

L'objectif de cette démarche a été double : d'une part, identifier la variable climatique cible la plus pertinente à downscaler (précipitations mensuelles ou indice SPI) ; d'autre part, sélectionner le modèle le plus performant sur la base de critères statistiques rigoureux, afin de garantir des projections fiables pour l'étude du climat futur.

En complément, une étape d'ajustement statistique a posteriori a été introduite à l'aide de la méthode du quantile mapping, dans le but d'optimiser la distribution statistique des données projetées après le downscaling. Bien que traditionnellement utilisée pour la correction des biais des données climatiques brutes, cette méthode a ici été détournée de son usage classique pour jouer un rôle d'ajustement statistique, en s'appuyant sur la structure des données historiques issues de la réanalyse et des GCMs. Elle permet ainsi d'améliorer la cohérence des projections futures en alignant leur distribution sur les caractéristiques locales observées.

Les résultats de cette démarche seront présentés dans le chapitre suivant, afin d'identifier l'approche la plus fiable pour le downscaling des projections climatiques à échelle locale.

Chapitre 5

RÉSULTATS ET DISCUSSION

5.1 Introduction

Ce chapitre présente une analyse approfondie des résultats obtenus à partir du calcul de l'indice SPI (Standardized Precipitation Index) appliqué au bassin versant du barrage de Keddara. L'objectif principal est d'évaluer la pertinence de cet indice pour caractériser les conditions de sécheresse sur la période historique, à partir des données de réanalyse ERA5 couvrant la période 1951–2013. Cette analyse permet également de confronter les résultats du SPI aux observations empiriques disponibles dans les archives nationales et les rapports institutionnels, afin de valider la fiabilité des indices climatiques utilisés dans le contexte algérien.

Un second objectif de ce chapitre est d'identifier l'approche de downscaling statistique la plus performante pour la projection future des sécheresses. Il s'agit notamment de comparer les performances des différents modèles testés (CNN et XGBoost) et de déterminer quelle variable cible — les précipitations ou l'indice SPI — se prête le mieux à un désagrégement statistique fiable en contexte de changement climatique.

5.2 Évolution des sécheresses au barrage de Keddara (1951–2013)

L'analyse des conditions de sécheresse au barrage de Keddara sur la période historique constitue une étape clé pour évaluer la fiabilité des indices climatiques utilisés. En s'appuyant sur les données de réanalyse ERA5, couvrant la période 1951–2013, cette section vise à identifier les épisodes majeurs de sécheresse à différentes échelles temporelles (SPI-1, SPI-3, SPI-6 et SPI-12), reflétant respectivement les déficits météorologiques, agricoles et hydrologiques.

L'objectif est double : d'une part, décrire finement l'évolution des sécheresses dans le bassin versant de Keddara à partir de calculs empiriques du SPI ; d'autre part, confronter ces résultats aux observations historiques documentées sur le terrain et dans la littérature institutionnelle nationale. Ce croisement permet de valider l'utilisation du SPI comme outil pertinent de suivi et de caractérisation des sécheresses en Algérie, en particulier dans un contexte de forte variabilité interannuelle et de changement climatique.

Une attention particulière est portée à la décennie 1990–2000, période reconnue pour sa sécheresse généralisée, ainsi qu'aux cassures climatiques identifiées autour des années 1976 et 2000.

5.2.1 Interprétation des résultats calculés

SPI 1 : L'analyse de la sécheresse météorologique au niveau du barrage de Keddara, à partir de l'indice SPI-1 sur la période 1951–2013, révèle un total de 43 épisodes secs. Ces épisodes sont généralement courts, durant entre 1 et 5 mois, et surviennent majoritairement entre mai et septembre, période déjà marquée par une faible pluviométrie. Certains événements se distinguent par leur intensité élevée, comme ceux de 1995 et 2010, ce dernier étant le plus sévère de toute la série avec un SPI de $-2,18$. Les décennies 1950, 1980 et 2000 ont été particulièrement touchées, avec une fréquence accrue dans les années 1995 à 2008, suggérant une période de vulnérabilité accrue. Bien que les épisodes soient souvent ponctuels, leur répétition d'une année à l'autre peut avoir un impact notable sur la recharge du barrage, notamment en saison sèche.

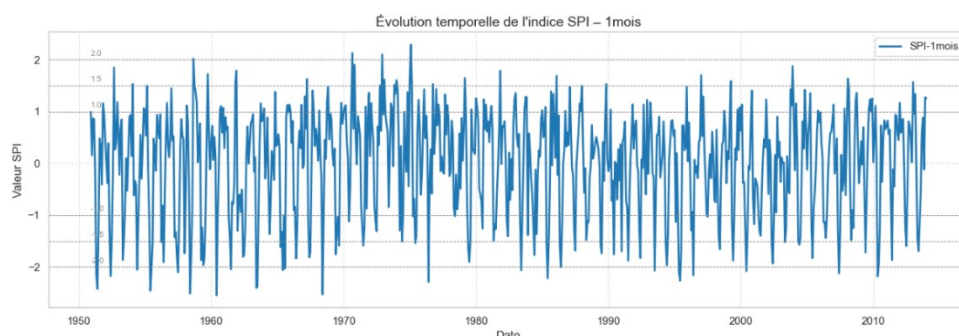


FIGURE 5.1 – Évolution Temporelle SPI-1 pour la période [1951–2013]

SPI-3 : L'analyse de la sécheresse agricole au barrage de Keddara à travers l'indice SPI-3, qui reflète l'impact sur l'humidité du sol de la région, met en évidence un total de 40 épisodes secs entre 1951 et 2012. Ces épisodes sont en grande majorité courts (1 à 3 mois), ce qui est cohérent avec la nature du SPI-3, centré sur les effets à moyen terme sur la végétation et les cultures. La grande majorité des sécheresses débute en septembre, période sensible marquant la fin de l'été, lorsque les sols sont déjà appauvris en eau. Plusieurs épisodes se distinguent par leur sévérité, notamment celui de 1954 avec une intensité exceptionnelle de $-2,89$, suivi par ceux de 1956, 1961 et 1978, tous supérieurs à $-2,0$. Cette récurrence en septembre souligne une vulnérabilité agricole saisonnière, où la faible recharge hydrique en fin d'été peut compromettre les semis et la croissance des cultures d'automne. Contrairement au SPI-1, les événements SPI-3 montrent une accumulation des déficits plus marquée, affectant directement la capacité des sols à retenir l'eau et à soutenir la végétation dans la région.

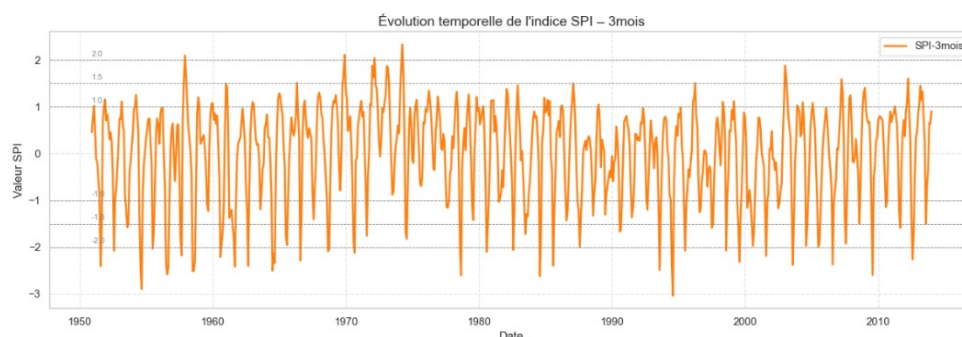


FIGURE 5.2 – Évolution Temporelle SPI-3 pour la période [1951–2013]

SPI-6 : L'analyse de la sécheresse hydrologique au barrage de Keddara, à partir de l'indice SPI-6, montre un total de 48 épisodes secs entre 1951 et 2012. Cet indice, basé sur l'accumulation des précipitations sur six mois, reflète mieux les impacts sur les débits des oueds qui alimentent le barrage. Les épisodes détectés sont souvent de durée moyenne (2 à 4 mois), et surviennent principalement entre

septembre et décembre, période cruciale pour la reprise du cycle hydrologique après l'été. Plusieurs événements se démarquent par leur intensité exceptionnelle, comme ceux de 1961 ($-2,95$), 1960 ($-2,80$), 2000 ($-2,41$) ou encore 1956 ($-2,66$), traduisant des déficits prolongés qui ont affecté la retenue du barrage. Les sécheresses SPI-6 sont donc plus représentatives des difficultés d'approvisionnement en eau de surface, car elles traduisent un manque cumulé à long terme, impactant aussi bien les usages agricoles qu'urbains. Ce type de sécheresse, bien que moins fréquent que la sécheresse météorologique, est généralement plus grave dans ses conséquences, car il touche directement les ressources en eau disponibles à l'échelle du bassin versant.

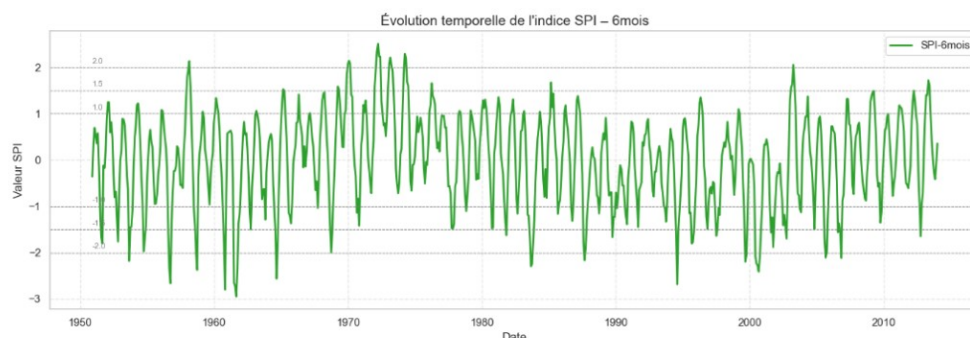


FIGURE 5.3 – Évolution Temporelle SPI-6 pour la période [1951–2013]

SPI-12 : permet d'évaluer les impacts sur le rechargement des réservoirs. Entre 1955 et 2007, 20 épisodes secs ont été identifiés, avec des durées variables allant de 1 à 19 mois. Les sécheresses SPI-12 sont moins fréquentes que celles à plus court terme, mais beaucoup plus graves car elles traduisent un manque persistant de précipitations affectant l'ensemble du cycle hydrologique. Certains épisodes sont particulièrement marquants, comme celui de 2000 à 2001, le plus long et le plus intense de toute la série (19 mois, intensité de $-3,17$), suivi par les périodes de 1961–1962 et 1983–1984, également très sévères (intensité $< -2,5$). Ces sécheresses prolongées ont entraîné une baisse importante des niveaux des retenues, compromettant l'approvisionnement en eau pour les usages domestiques, agricoles et industriels.

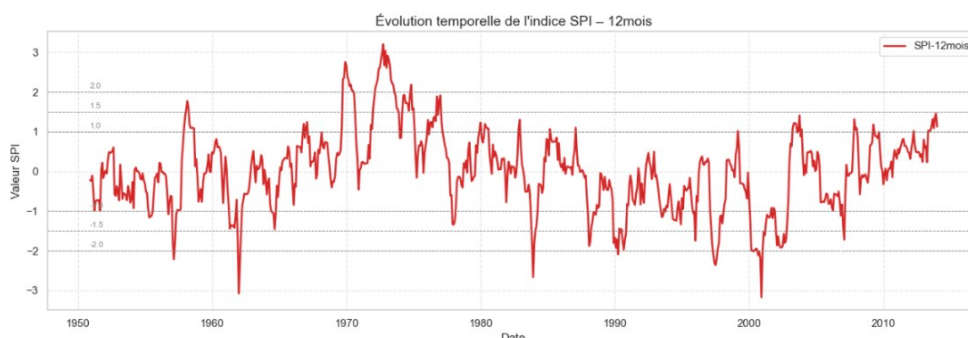


FIGURE 5.4 – Évolution Temporelle SPI-12 pour la période [1951–2013]

L'analyse détaillée des différents indices SPI sur la période 1951–2013 met en évidence la diversité des sécheresses affectant le bassin de Keddara, tant en termes de fréquence, d'intensité que de durée. Ces résultats, fondés sur une approche statistique multi-échelles, soulignent la capacité du SPI à discriminer les types de sécheresse selon leur temporalité et leurs impacts potentiels sur le système hydrique.

Cependant, pour juger pleinement de la robustesse de ces indicateurs et de leur pertinence dans le contexte algérien, il convient de confronter les résultats obtenus aux données historiques, aux constats de terrain ainsi qu'aux rapports nationaux. La sous-section suivante propose ainsi une validation

croisée entre les épisodes secs identifiés par le SPI et les grandes sécheresses documentées par les institutions publiques et la littérature scientifique, permettant de confirmer la cohérence des analyses et d'ancrer les résultats dans une réalité observée.

5.2.2 Validation des calculs du SPI à partir des constats historique

Les résultats issus de l'analyse multi-échelle des sécheresses au niveau du barrage de Keddara (SPI-1, SPI-3, SPI-6, SPI-12) confirment pleinement les constats établis par les études historiques et institutionnelles menées en Algérie. En effet, les périodes identifiées comme particulièrement sèches dans notre étude — notamment 1954–1956, 1960–1962, 1978, 1983–1984, 1995–2001 et 2010 — coïncident avec les grandes sécheresses hydrologiques et agricoles documentées à l'échelle nationale. Le Plan National de Lutte contre la Sécheresse rappelle d'ailleurs que l'Algérie a connu des épisodes majeurs en 1910, 1940, puis de manière persistante entre 1975 et 1980, ainsi qu'au début des années 2000, marqués par une forte baisse des apports en eau dans les barrages. Ces événements historiques trouvent un écho direct dans les indices SPI calculés au barrage de Keddara, qui traduisent une réduction significative des précipitations et une aggravation progressive du stress hydrique, à toutes les échelles temporelles.

Par ailleurs, les résultats confirment l'existence de la cassure pluviométrique autour de 1976, évoquée dans les travaux de l'ANRH (2017). Cette cassure désigne un changement brutal et durable du régime des précipitations à partir de cette date, marquant une transition vers un climat plus sec. Ainsi, notre analyse locale à partir des indices SPI valide, à l'échelle du bassin de Keddara, les grandes tendances de dégradation climatique observées à l'échelle nationale.

Ces constats sont également appuyés par les Première (2001) et Seconde (2010) Communications Nationales de l'Algérie sur les changements climatiques, qui dressent un état des lieux du climat passé et une projection à l'horizon 2020 et 2050. Il en ressort notamment :

- Un réchauffement d'environ 0,5 °C et une baisse de 10 % des précipitations entre les périodes 1931–1960 et 1961–1990 sur les régions méditerranéennes du pays.
- Une intensification et une plus grande fréquence des vagues de chaleur, pouvant survenir à tout moment de l'année.
- Des projections climatiques saisonnières annonçant une hausse continue des températures et une diminution des pluies à l'horizon 2020 et 2100.
- Une augmentation des événements climatiques extrêmes : sécheresses, pluies diluviennes, vagues de froid ou de chaleur, submersions marines, etc.
- Une perturbation des zones côtières et des écosystèmes marins.

Dans ce même cadre, le CRSTRA (Centre de Recherche Scientifique et Technique sur les Régions Arides), en partenariat avec l'Accord euro-méditerranéen sur les risques majeurs, a mené une étude approfondie sur les canicules en Algérie. En comparant les périodes 1951–1980 et 1981–2010, il a été constaté une hausse de l'intensité, de la fréquence et de la précocité des vagues de chaleur, particulièrement marquée vers le Nord-Est du pays. Ces résultats rejoignent les conclusions du GIEC, qui signalait dès 2007 un réchauffement au Maghreb de 1 à 2 °C depuis 1970, soit le double de la moyenne mondiale (0,74 °C).

Conclusion de la section 5.2 :

L'analyse historique des sécheresses au niveau du barrage de Keddara sur la période 1951–2013, à travers les indices SPI-1, SPI-3, SPI-6 et SPI-12, met en évidence la pertinence du SPI comme outil de diagnostic climatique multi-échelle. Les résultats montrent clairement que chaque indice permet de caractériser un type spécifique de sécheresse – météorologique, agricole ou hydrologique – en fonction de la durée d'accumulation choisie.

Les événements identifiés par le SPI concordent avec les sécheresses majeures recensées dans les

archives nationales et les rapports climatiques algériens, confirmant la fiabilité de l'indice dans le contexte local. De plus, la répartition temporelle et saisonnière des épisodes secs, ainsi que leur intensité croissante à partir des années 1970, confortent l'hypothèse d'une dégradation du régime pluviométrique dans la région, notamment après la cassure pluviométrique de 1976.

Ces résultats confirment non seulement la sensibilité du barrage de Keddara aux variations climatiques interannuelles, mais soulignent aussi la nécessité de prendre en compte des indicateurs intégrateurs comme le SPI dans les politiques de gestion de l'eau, d'alerte précoce et d'adaptation au changement climatique. Ils constituent ainsi une base de référence solide pour l'analyse des projections futures qui seront abordées dans la section suivante.

5.3 Évaluation comparative des modèles

Une évaluation des performances des deux modèles (XGBoost et CNN) pour les deux variables cibles (précipitations et SPI) est essentielle afin de déterminer quelle approche est la plus adaptée pour une projection future fiable.

5.3.1 Critères Numériques

Précipitation :

Le modèle CNN démontre une meilleure performance globale par rapport au modèle XGBoost pour la simulation des précipitations. Les meilleurs GCMs sous CNN (comme *IPSL-CM6A-LR_r13i1p1f1*, *MRI-ESM2-0_r7i1p1f1* et *CNRM-CM6-1_r12i1p1f2*) présentent des erreurs plus faibles (MSE autour de 1240–1250, MAE < 26 mm) et des SMAPE autour de 70 %, avec des coefficients de détermination $R^2 > 0,36$, ce qui reflète une meilleure capacité à expliquer la variabilité des précipitations. En comparaison, XGBoost atteint ses meilleures performances avec *MIROC-ES2H_r1i1p4f2*, mais ses métriques restent inférieures (MSE = 1644.83, MAE = 29.77 mm, $R^2 = 0,314$), et son SMAPE demeure élevé (96 % à 99 %) pour tous les GCMs testés, traduisant une moindre précision relative.

Par ailleurs, la capacité de généralisation est bonne pour les deux modèles, avec des écarts faibles entre les phases d'apprentissage et de test, ce qui indique une absence de surapprentissage. Toutefois, le CNN conserve une meilleure stabilité des performances entre les différents GCMs, tandis que XGBoost montre une plus grande variabilité (certains modèles affichant $R^2 < 0,27$ et MSE > 1800), ce qui suggère une sensibilité accrue au choix du GCM.

TABLEAU 5.1 – Performances du modèle CNN_sr sur la précipitation

Critère	Test (min – max)	Train (min – max)
MSE	1 241.19 – 1 584.58	1 464.92 – 1 698.09
MAE	25.90 – 29.82	28.07 – 32.75
R^2	0.200 – 0.374	0.343 – 0.434
RMSE	35.23 – 39.81	38.27 – 41.21
Adjusted R^2	0.197 – 0.373	0.342 – 0.432
SMAPE (%)	68.18 – 75.42	61.57 – 69.01

TABLEAU 5.2 – Performances du modèle Xgboost_sur la précipitation

Critère	Test (min – max)	Train (min – max)
MSE	1 644.83 – 1 832.50	1 241.60 – 1 539.00
MAE	29.77 – 31.37	22.33 – 27.10
R^2	0.236 – 0.314	0.358 – 0.483
RMSE	40.56 – 42.81	35.24 – 39.23
Adjusted R^2	0.231 – 0.310	0.357 – 0.482
SMAPE (%)	96.13 – 99.32	93.71 – 95.93

SPI-1 :

Le modèle CNN montre des performances globalement plus équilibrées et robustes que le modèle XGBoost pour la prévision de l'indice SPI-1. Les meilleurs membres GCMs sous CNN, tels que *CNRM-CM6-1_r18i1p1f2*, *IPSL-CM6A-LR_r13i1p1f1* ou *CNRM-CM6-1_r10i1p1f2*, atteignent des MSE test < 0.45 , des MAE ≈ 0.50 , et des R^2 test ≥ 0.52 , ce qui témoigne d'une capacité raisonnable à reproduire les variations de l'indice. Bien que le SMAPE reste élevé (souvent $> 95\%$), les performances restent stables entre apprentissage et test, ce qui indique une bonne généralisation et peu de surapprentissage.

À l'inverse, le modèle XGBoost présente une forte sensibilité au surapprentissage. Certains membres comme *GISS-E2-1-G_r20i1p1f2* ou *IPSL-CM6A-LR_r12i1p1f1* montrent des MSE d'entraînement très faibles (< 0.1), mais une forte dégradation en test (MSE > 0.60 , SMAPE $> 140\%$), révélant un manque de robustesse. Seuls quelques membres (par exemple *CNRM-CM6-1_r18i1p1f2* ou *IPSL-CM6A-LR_r16i1p1f1*) obtiennent des R^2 test > 0.56 , mais la majorité plafonne en dessous de 0.57, ce qui limite la fiabilité du modèle. De plus, le SMAPE élevé ($> 130\%$ pour tous les cas) illustre la difficulté du modèle à capter les faibles valeurs du SPI-1, essentielles pour identifier les épisodes de sécheresse.

TABLEAU 5.3 – Performances du modèle CNN sur l'indice SPI-1

Critère	Test (min – max)	Train (min – max)
MSE	0.3889 – 0.5128	0.4401 – 0.5366
MAE	0.4975 – 0.5936	0.5298 – 0.5871
R^2	0.4471 – 0.5805	0.4918 – 0.5832
RMSE	0.6236 – 0.7161	0.6632 – 0.7325
Adjusted R^2	0.4448 – 0.5788	0.4908 – 0.5824
SMAPE (%)	93.90 – 103.69	88.42 – 95.97

TABLEAU 5.4 – Performances du modèle XGBoost sur l'indice SPI-1

Critère	Test (min – max)	Train (min – max)
MSE	0.5234 – 0.7495	0.0729 – 0.4323
MAE	0.5861 – 0.6937	0.2033 – 0.5276
R^2	0.3901 – 0.5741	0.5487 – 0.9239
RMSE	0.7234 – 0.8657	0.2700 – 0.6575
Adjusted R^2	0.3861 – 0.5713	0.5480 – 0.9238
SMAPE (%)	137.33 – 141.30	132.89 – 138.94

SPI-3 :

Le modèle CNN s'impose comme plus performant que XGBoost pour la simulation du SPI-3, notamment en raison de sa capacité à capturer plus efficacement les dynamiques modérées à longues de sécheresse. Les GCMs les plus performants sous CNN (tels que *CNRM-CM6-1_r10i1p1f2*, *IPSL-CM6A-LR_r14i1p1f1*, *CAS-ESM2-0_r1i1p1f1*) atteignent des R^2 test supérieurs à 0.63, avec des MSE test compris entre 0.36 et 0.46, des MAE ≈ 0.49 , et un SMAPE inférieur à 81 %. Ces résultats traduisent une meilleure précision des prédictions, une généralisation plus stable, et une capacité plus forte à représenter les fluctuations climatiques sur trois mois.

À l'inverse, bien que XGBoost atteigne des R^2 test similaires (0.52 à 0.63), ses erreurs absolues (MAE ≈ 0.51) et RMSE (≈ 0.63 – 0.67) sont systématiquement plus élevées que celles du CNN. Le modèle souffre également d'un surapprentissage important, avec des écarts marqués entre les performances d'apprentissage (train $R^2 > 0.92$ pour certains GCMs) et celles de test (~ 0.58 à 0.62). Cette instabilité nuit à la robustesse des prévisions. De plus, les SMAPE très élevés (127 % à 133 %) reflètent une difficulté importante à prédire correctement les valeurs faibles du SPI-3, souvent critiques pour l'anticipation des sécheresses modérées.

TABLEAU 5.5 – Performances du modèle CNN sur l'indice SPI-3

Critère	Test (min – max)	Train (min – max)
MSE	0.3661 – 0.4584	0.3202 – 0.4049
MAE	0.4721 – 0.5345	0.4458 – 0.5018
R^2	0.5392 – 0.6319	0.6010 – 0.6844
RMSE	0.6050 – 0.6770	0.5659 – 0.6362
Adjusted R^2	0.5373 – 0.6304	0.6002 – 0.6837
SMAPE (%)	78.44 – 87.23	77.02 – 83.80

TABLEAU 5.6 – Performances du modèle XGBoost sur l'indice SPI-3

Critère	Test (min – max)	Train (min – max)
MSE	0.3735 – 0.4977	0.0498 – 0.4092
MAE	0.4893 – 0.5597	0.1633 – 0.5063
R^2	0.5199 – 0.6397	0.5894 – 0.9500
RMSE	0.6112 – 0.7055	0.2232 – 0.6397
Adjusted R^2	0.5167 – 0.6373	0.5887 – 0.9499
SMAPE (%)	127.13 – 132.98	133.04 – 138.41

SPI-6 :

Le modèle CNN présente une meilleure stabilité et une précision globalement plus élevée que le modèle XGBoost pour la modélisation du SPI-6. Les performances du CNN montrent des R^2 test compris entre 0,43 et 0,55, des MSE allant de 0,48 à 0,61, et des MAE autour de 0,54, avec des RMSE de 0,7 à 0,75, traduisant une capacité correcte à capturer les tendances des sécheresses prolongées. Malgré des valeurs de SMAPE élevées (87 % à 95 %), le CNN parvient à représenter convenablement la dynamique saisonnière du SPI-6, notamment pour les phases modérées à sévères. Des cas de surapprentissage modéré sont observés (train $R^2 > 0,6$ contre test $R^2 \approx 0,49$ – $0,52$), mais restent gérables.

Le modèle XGBoost, de son côté, montre des performances comparables mais légèrement inférieures, avec un R^2 test entre 0,42 et 0,51, et des MSE similaires (0,43 à 0,58). Toutefois, le MAE

est généralement plus élevé ($> 0,54$), et le SMAPE dépasse souvent les 138 %, ce qui indique une grande difficulté à représenter les faibles valeurs du SPI, essentielles pour détecter les débuts d'épisodes secs. De plus, le surapprentissage est plus marqué avec XGBoost, certains GCMs atteignant train $R^2 > 0,90$, mais chutant à test $R^2 < 0,52$, avec des écarts dépassant 0,35 entre les phases. Cela souligne une sensibilité accrue aux variations dans les données d'entrée et un besoin de régularisation plus rigoureuse.

TABLEAU 5.7 – Performances du modèle CNN sur l'indice SPI-6

Critère	Test (min – max)	Train (min – max)
MSE	0.4838 – 0.6159	0.3645 – 0.4291
MAE	0.5475 – 0.6140	0.4694 – 0.5177
R^2	0.4336 – 0.5511	0.5521 – 0.6196
RMSE	0.6956 – 0.7848	0.6037 – 0.6729
Adjusted R^2	0.4313 – 0.5534	0.5513 – 0.6188
SMAPE (%)	87.40 – 95.27	79.47 – 85.19

TABLEAU 5.8 – Performances du modèle XGBoost sur l'indice SPI-6

Critère	Test (min – max)	Train (min – max)
MSE	0.4326 – 0.5862	0.0717 – 0.4664
MAE	0.5310 – 0.6243	0.1929 – 0.5244
R^2	0.4208 – 0.5726	0.5323 – 0.9281
RMSE	0.6577 – 0.7656	0.2678 – 0.6829
Adjusted R^2	0.4170 – 0.5697	0.5315 – 0.9279
SMAPE (%)	137.8 – 142.4	135.8 – 140.3

SPI-12 :

Le modèle CNN démontre une très bonne aptitude à modéliser le SPI-12, grâce à sa capacité à capturer la persistance hydrologique sur de longues périodes. Les meilleures simulations atteignent des R^2 test $> 0,75$ (jusqu'à 0,95 pour IPSL-CM6A-LR_r19i1p1f1 ou CNRM-CM6-1_r12i1p1f2), avec des RMSE < 1 et des MSE modérés, traduisant une excellente capacité explicative. Le modèle montre aussi une bonne généralisation, avec peu d'écarts entre les phases d'entraînement et de test, sauf pour quelques membres isolés. Cependant, les SMAPE restent élevés ($> 125\%$), illustrant une difficulté persistante à prédire les valeurs faibles ou proches de zéro, caractéristiques des périodes normales ou des débuts/fin de sécheresse.

Le modèle XGBoost, quant à lui, offre une alternative compétitive, avec des R^2 test souvent $> 0,5$, atteignant 0,717 pour MIROC-ES2H_r11p4f2 et dépassant 0,65 pour plusieurs membres (IPSL-CM6A-LR, GISS-E2-1-G). Les erreurs sont bien contenues (MSE entre 0,25 et 0,50, RMSE $< 0,75$, MAE $\approx 0,43$ –0,48), indiquant une bonne précision sur le long terme. Cependant, le XGBoost est plus sensible au surapprentissage, comme en témoignent les écarts entre les scores d'apprentissage et de test (par ex. CNRM-CM6-1_r18i1p1f2). Le SMAPE est également plus élevé (145–151 %) que celui du CNN, montrant une performance relative moins stable sur les extrêmes.

TABLEAU 5.9 – Performances du modèle CNN sur l'indice SPI-12

Indicateur	Test (min – max)	Train (min – max)
MSE	0.4295 – 0.5357	0.3884 – 0.4645
MAE	0.5179 – 0.5803	0.4937 – 0.5392
R^2	0.4563 – 0.5641	0.5412 – 0.6164
RMSE	0.6554 – 0.7319	0.6232 – 0.6815
Adjusted R^2	0.4541 – 0.5624	0.5404 – 0.6157
SMAPE (%)	86.72 – 97.95	81.99 – 88.74

TABLEAU 5.10 – Performances du modèle XGBoost sur l'indice SPI-12

Indicateur	Test (min – max)	Train (min – max)
MSE	0.3842 – 0.5523	0.0178 – 0.4693
MAE	0.4948 – 0.6144	0.0896 – 0.5241
R^2	0.3938 – 0.5782	0.5360 – 0.9824
RMSE	0.6198 – 0.7431	0.1333 – 0.6851
Adjusted R^2	0.3897 – 0.5754	0.5352 – 0.9824
SMAPE (%)	134.93 – 139.86	137.22 – 141.43

5.3.2 Critères Graphiques

a) QQPLOT :

Précipitations

D'après les représentations visuelles QQ Plot , les modèles CNN et XGBoost présentent une dispersion importante des points, notamment pour les valeurs de précipitations supérieures à 100 mm, ce qui révèle une tendance systématique à sous-estimer les fortes précipitations. Aucun des deux modèles ne parvient à aligner correctement les prédictions sur la diagonale idéale ($y = x$), traduisant une précision limitée dans la reproduction des extrêmes. Le modèle CNN montre néanmoins une distribution légèrement plus dense et structurée des points par rapport à XGBoost, bien que les prédictions tendent à se rapprocher d'une droite qui s'écarte de la diagonale idéale , ce qui suggère une faible capacité à capturer la variabilité réelle des fortes valeurs.

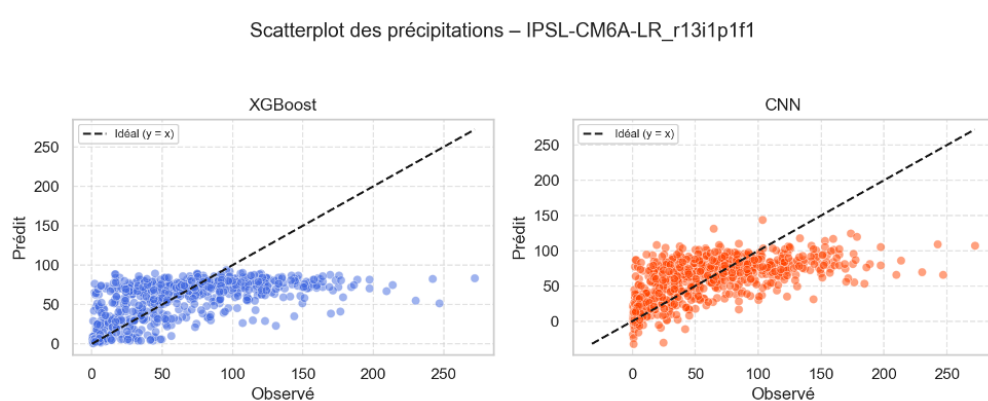


FIGURE 5.5 – QQPLOT de quelques exemples de GCMS de la variable précipitation

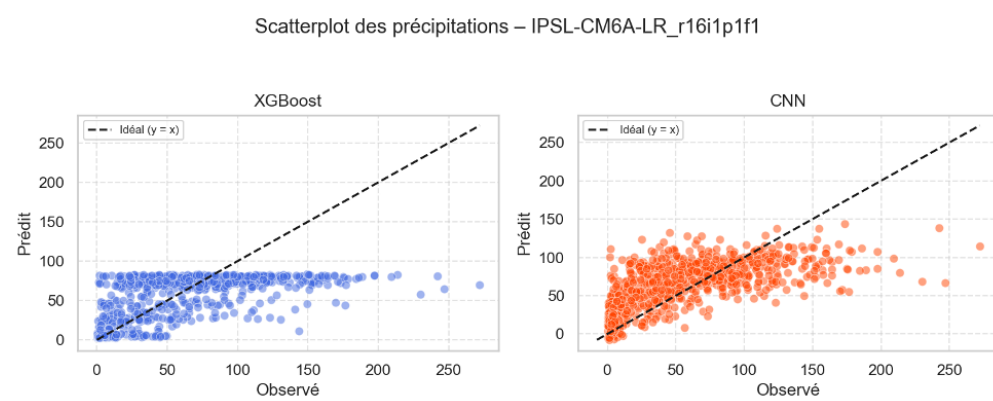


FIGURE 5.6 – QQPLOT de quelques exemples de GCMS de la variable précipitation

SPI

Les deux modèles ont tendance à s'écarter de la diagonale idéale ($y = x$), surtout pour les faibles et fortes valeurs de SPI, ce qui reflète une sous-estimation des extrêmes, en particulier dans les cas de sécheresse intense ($SPI < -2$) ou excès d'humidité ($SPI > 2$). Le CNN montre une meilleure capacité à suivre la variabilité réelle à mesure qu'on passe de SPI-1 à SPI-12, avec une dispersion plus réduite autour de la diagonale pour les longues échelles (notamment SPI-12). Le modèle XGBoost, quant à lui, présente des nuages de points plus dispersés, une tendance plus marquée à la compression des valeurs autour de zéro, et une incapacité plus nette à reproduire les extrêmes, quel que soit l'indice. Ainsi, le CNN semble mieux saisir la structure saisonnière et la persistance hydrologique pour les indices à plus longue échelle (SPI-6, SPI-12), tandis que XGBoost peine à généraliser les dynamiques climatiques, surtout en conditions extrêmes.

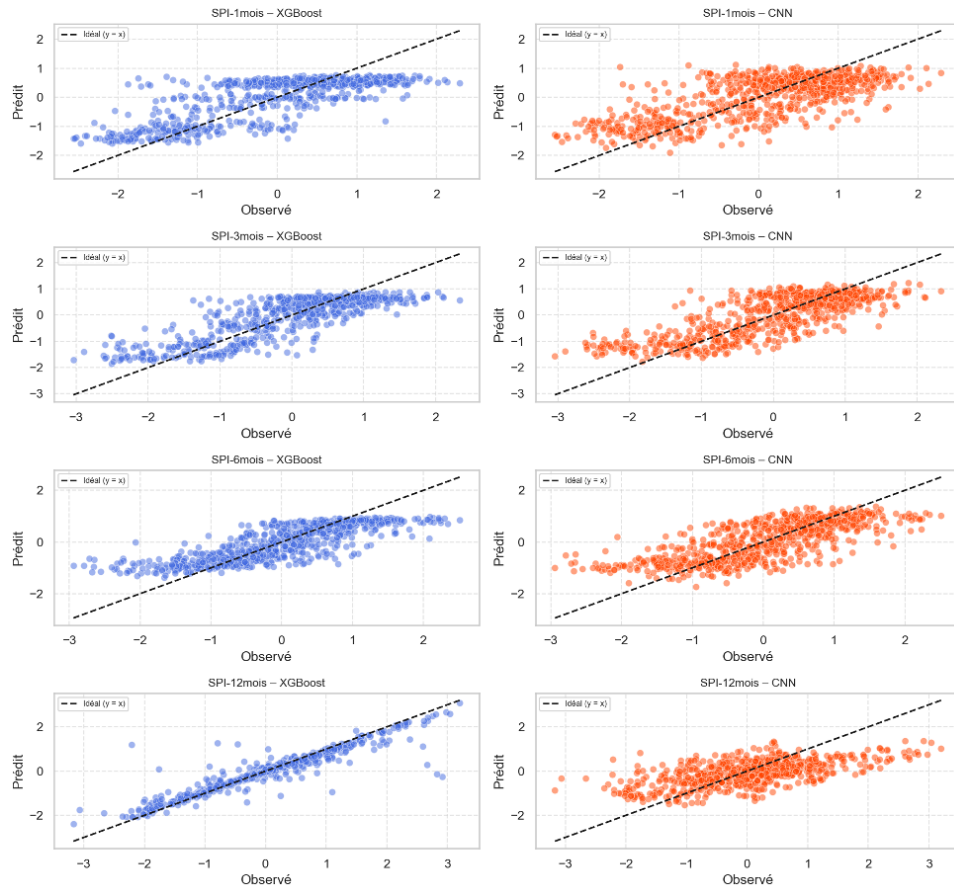


FIGURE 5.7 – QQPLOT de quelques exemples de GCMS de l'indice SPI

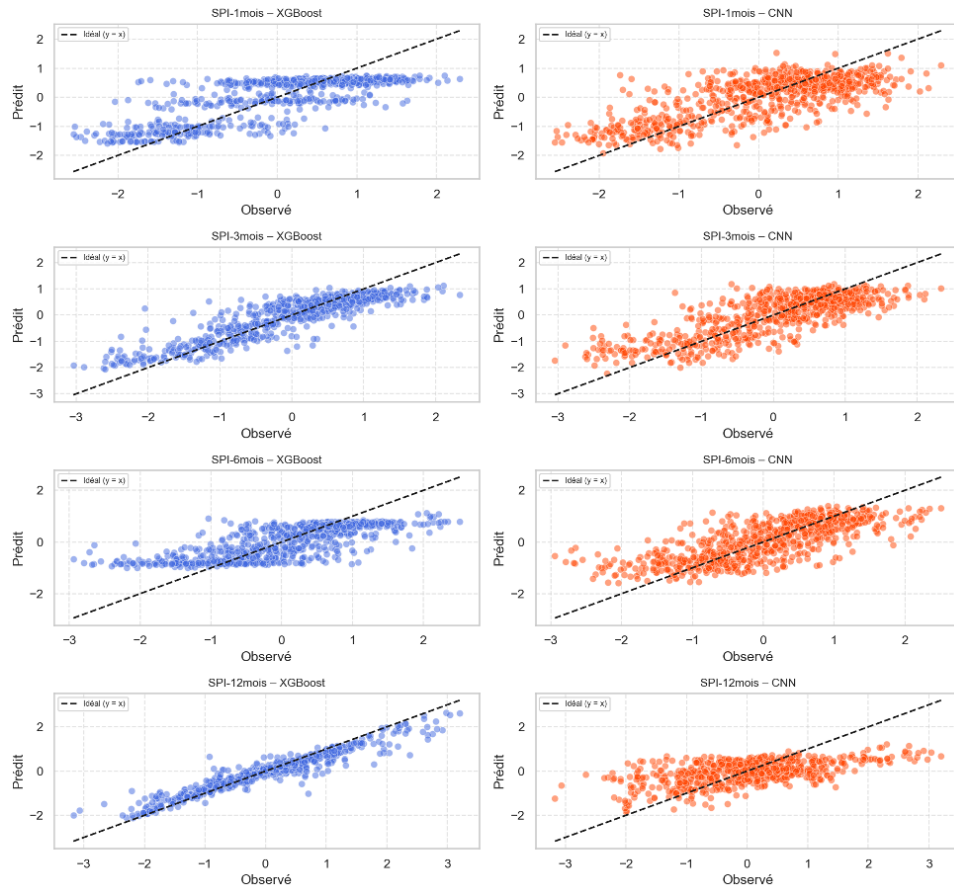


FIGURE 5.8 – QQPLOT de quelques exemples de GCMS de l'indice SPI

b) Évolution Temporelle :

Précipitations

Pour les précipitations, les deux modèles montrent des difficultés à reproduire fidèlement les pics élevés, mais le modèle CNN suit mieux les tendances saisonnières par rapport à XGBoost, qui affiche une forte sous-estimation systématique. On observe que CNN parvient à capturer certaines hausses et chutes brutales, même si l'amplitude reste réduite. En revanche, XGBoost donne une série plus lissée, très éloignée des observations, avec peu de réactivité aux épisodes intenses.

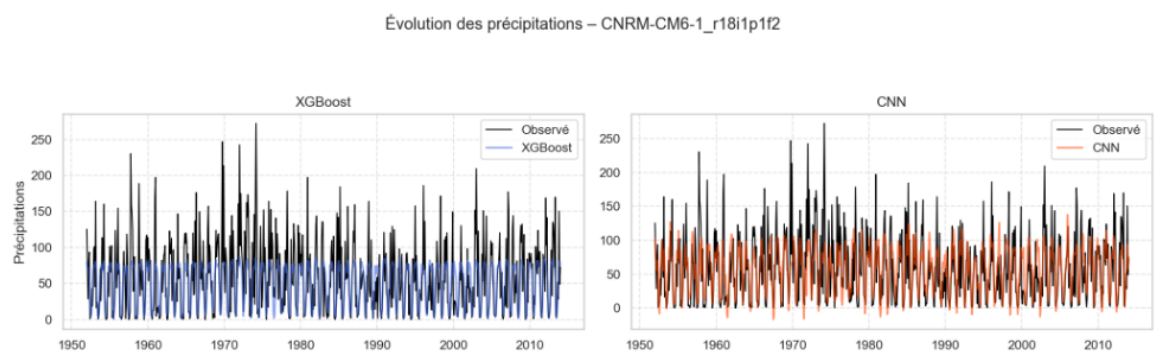


FIGURE 5.9 – Évolution temporelle de quelques exemples de GCMs pour la précipitation

Indices SPI (SPI-1 à SPI-12)

Pour les indices SPI, le CNN montre globalement une meilleure capacité à reproduire les variations de sécheresse, en particulier pour SPI-3, SPI-6 et SPI-12. L'alignement des courbes est plus proche de l'observation, et les cycles humides/sec sont bien représentés. À l'inverse, XGBoost tend à écraser les extrêmes, notamment pour SPI-6 et SPI-12, où les pics sont moins marqués et les phases sèches ou humides mal captées. On remarque que pour SPI-12, le CNN restitue beaucoup mieux la persistance des longues périodes sèches ou humides, ce qui est essentiel dans le suivi des sécheresses hydrologiques. XGBoost, lui, reste moins sensible aux tendances de fond sur le long terme.

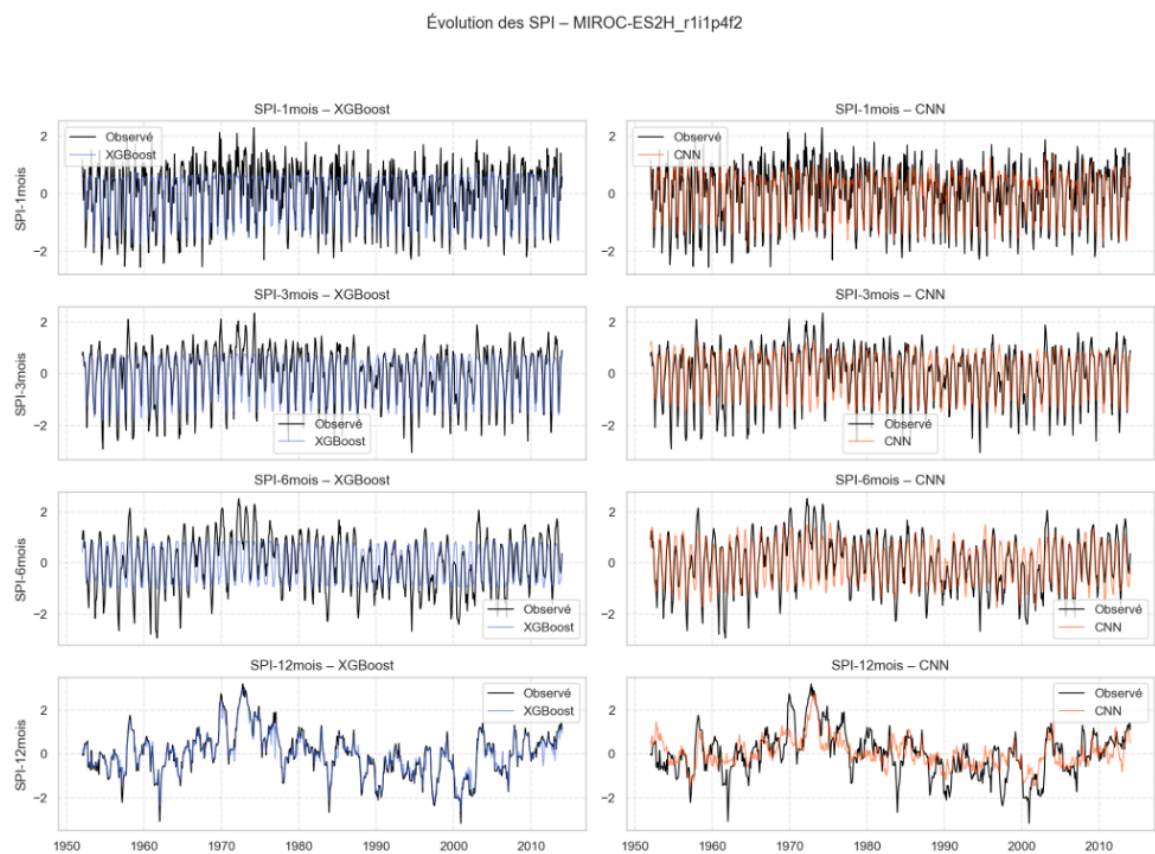


FIGURE 5.10 – Évolution temporelle de quelques exemples de GCMs pour l'indice SPI

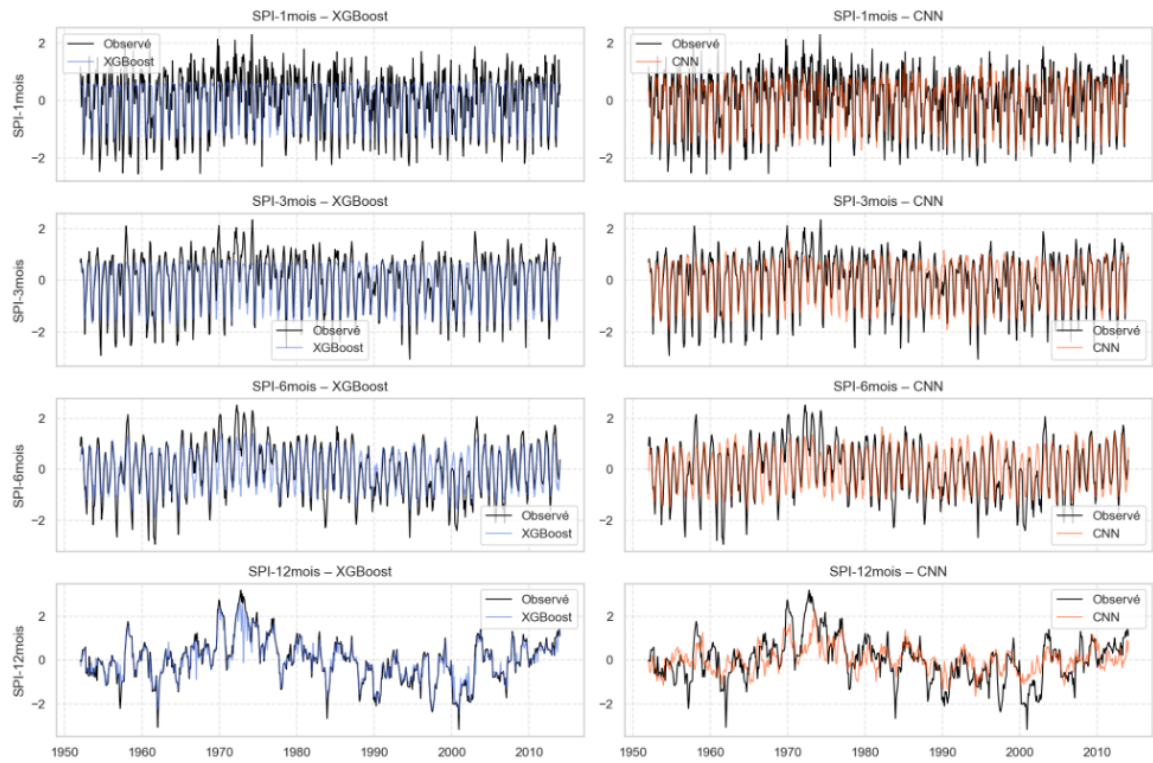


FIGURE 5.11 – Évolution temporelle de quelques exemples de GCMs pour l'indice SPI

5.3.3 Résultats de l'évaluation

Choix de la méthode :

L'analyse des résultats montre que le modèle CNN est plus performant que XGBoost pour la simulation des précipitations et des indices SPI. D'un point de vue numérique, le CNN présente des valeurs plus faibles d'erreur (MSE, MAE) et des scores de R^2 plus élevés, ce qui signifie qu'il explique mieux la variabilité des données. Il est aussi plus stable entre les différents modèles climatiques testés (GCMs), et moins sensible au surapprentissage, ce qui garantit une meilleure généralisation sur de nouvelles données. En résumé, le CNN est plus fiable et plus robuste pour les applications climatiques.

Choix de la variable cible :

Les précipitations présentent des performances limitées, avec des R^2 faibles ($< 0,4$), un SMAPE élevé ($> 70\%$) et une mauvaise représentation des valeurs extrêmes. Graphiquement, les modèles montrent une forte dispersion des points et une sous-estimation systématique des fortes précipitations, avec un éloignement notable par rapport à la diagonale idéale, ce qui limite leur capacité à représenter la variabilité naturelle. En comparaison, l'indice SPI, notamment le SPI-12, offre de meilleurs résultats en downscaling avec le CNN ($R^2 > 0,75$), une meilleure restitution des extrêmes (valeurs proches de -2 ou $+2$), et une structure graphique plus cohérente, bien que centrée autour de la moyenne. Malgré un SMAPE encore élevé, le SPI respecte mieux les variations saisonnières et la dynamique climatique, ce qui en fait une variable plus adaptée pour affiner les projections locales et analyser les épisodes de sécheresse.

5.4 Conclusion

Dans ce chapitre, une analyse des événements historiques de sécheresse a été réalisée à l'aide du calcul de l'indice SPI, et les résultats obtenus ont été validés par des constats historiques documentés. L'approche la plus pertinente pour réaliser des projections s'est avérée être l'application du downscaling directement sur les valeurs de SPI, en utilisant un algorithme de type CNN. Toutefois, afin d'améliorer la représentation des extrêmes après le downscaling, une correction par quantile mapping s'avère nécessaire. Cette étape d'ajustement statistique permettra d'affiner les résultats des projections futures, qui feront l'objet du chapitre suivant.

Chapitre 6

PROJECTIONS CLIMATIQUES FUTURES

6.1 Introduction

Dans un contexte de changement climatique mondial, l'anticipation des sécheresses futures représente un enjeu stratégique pour la planification et la gestion durable des ressources en eau, en particulier dans les régions semi-arides comme celle du barrage de Keddara. Après avoir confirmé, à travers l'analyse des données historiques ERA5 (1951–2013), la pertinence de l'indice SPI (Standardized Precipitation Index) pour caractériser les différents types de sécheresse — météorologique (SPI-1), agricole (SPI-3), hydrologique (SPI-6) et annuelle (SPI-12) — cette section s'attache à explorer l'évolution future de ces phénomènes à l'horizon 2100.

Les projections climatiques utilisées reposent sur un ensemble de modèles climatiques globaux (GCMs) issus du programme CMIP6. Afin de garantir une représentativité locale, les données brutes ont été régionalisées à l'échelle du bassin de Keddara à l'aide de la méthode de downscaling statistique, basées sur l'algorithme du CNN, puis ajuster les sorties via la méthode du quantile mapping.

Trois scénarios socio-économiques contrastés ont été retenus afin de tenir compte de l'incertitude liée aux trajectoires futures d'émissions :

- **SSP1–2.6** : scénario optimiste, fondé sur une transition mondiale vers le développement durable ;
- **SSP2–4.5** : scénario intermédiaire, marqué par des efforts modérés de réduction des émissions ;
- **SSP5–8.5** : scénario pessimiste, caractérisé par une forte dépendance aux énergies fossiles.

L'analyse s'intéresse à la fréquence, la durée et l'intensité des épisodes de sécheresse projetés entre 2025 et 2100. Elle vise à comparer les dynamiques induites par chacun des scénarios sur les différents types de sécheresse, pour orienter les stratégies d'adaptation et renforcer la résilience face aux stress hydriques à venir.

6.2 Évaluation du Quantile Mapping

Parmi les modèles présentant de bonnes performances, le MIROC se distingue en première position pour l'ensemble des paramètres sélectionnés, suivi des modèles CAS. Ces deux modèles, développés respectivement par le Japon et la Chine, montrent un fort potentiel.

Cependant, en raison de difficultés techniques liées à l'accès aux fichiers sur les sites de données, ces deux références ont dû être supprimées avant de pouvoir confirmer quel GCM était réellement le plus performant. C'est dans ce contexte que le modèle `IPSL-CM6A-LR_r14i1p1f1` a été retenu pour la suite de l'analyse.

Pour les projections futures, un ajustement statistique est appliqué à l'aide de la méthode du quantile mapping sur les valeurs de SPI générées par le modèle CNN, c'est-à-dire après le downscaling statistique.

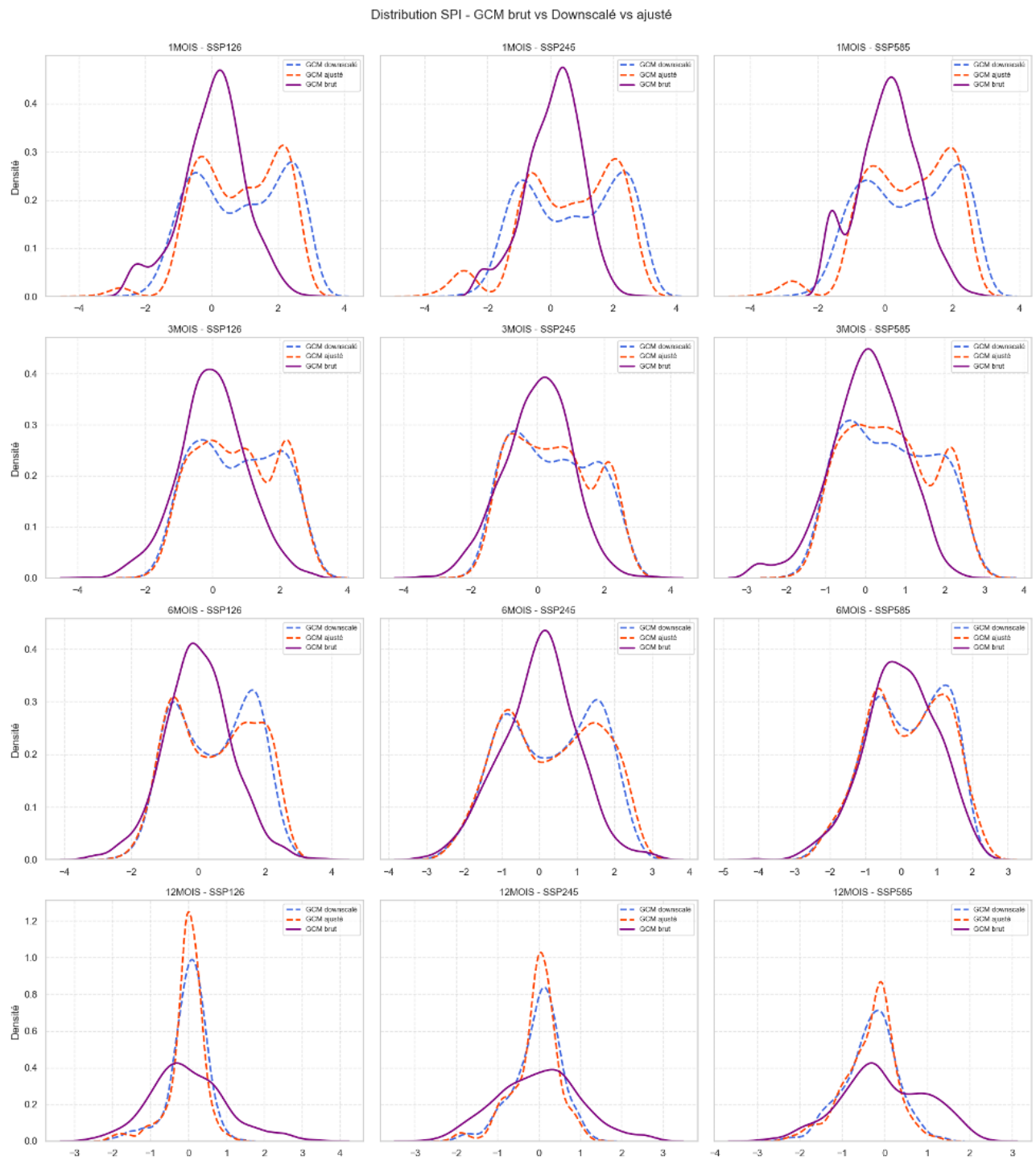


FIGURE 6.1 – Distribution comparative des données brutes (GCM), downscalées (CNN), et ajustées (QM)

Selon la figure ci-dessus, une amélioration notable de la distribution est observée : les sorties des GCMs, une fois downscalées puis corrigées, tendent à mieux se rapprocher du SPI simulé à partir des données brutes non downscalées. Cette tendance est visible pour les quatre échelles temporelles du SPI, quel que soit le scénario climatique considéré.

Ainsi, l'ajustement statistique appliqué après le downscaling s'avère essentiel pour améliorer la fiabilité des projections et garantir une représentation plus réaliste des conditions de sécheresse futures.

6.3 Description des scénarios climatiques (SSP)

Les Shared Socioeconomic Pathways (SSP) constituent une série de trajectoires prospectives utilisées dans les études climatiques pour explorer l'évolution possible des émissions de gaz à effet de serre en fonction de divers choix socio-économiques. Chacun de ces scénarios est associé à un niveau de forçage radiatif, exprimé en watts par mètre carré (W/m^2), projeté à l'horizon 2100, traduisant le potentiel de réchauffement global induit par les activités humaines.

- **SSP1-2.6** – Scénario durable (optimiste) :

Ce scénario, aligné sur l'ancien RCP2.6, repose sur une transition mondiale vers un développement durable. Il prévoit un forçage radiatif limité à $2,6 \text{ W/m}^2$ d'ici à 2100, en cohérence avec l'objectif de maintenir le réchauffement climatique en dessous de 2°C . Il suppose la mise en œuvre de politiques climatiques ambitieuses, une coopération internationale renforcée, une forte réduction des émissions de gaz à effet de serre, ainsi qu'un recours accru aux énergies renouvelables.

- **SSP2-4.5** – Scénario intermédiaire (tendance médiane) :

Ce scénario correspond à une trajectoire modérée, équivalente au RCP4.5. Il envisage une évolution où les tendances actuelles (croissance économique, démographie, urbanisation) se poursuivent, avec des efforts partiels de réduction des émissions. Le forçage radiatif atteint $4,5 \text{ W/m}^2$ en 2100, reflétant un compromis entre inertie des systèmes socio-économiques et politiques d'atténuation limitées, mais présentes.

- **SSP5-8.5** – Scénario intensif (pessimiste) :

Ce scénario représente une trajectoire de forte intensification des activités humaines, similaire à une version révisée du RCP8.5. Il repose sur une croissance économique rapide, alimentée majoritairement par les énergies fossiles, et sur une absence quasi totale de mesures de réduction des émissions. Le forçage radiatif atteint $8,5 \text{ W/m}^2$ à la fin du siècle, traduisant une accélération importante du réchauffement climatique. Ce scénario constitue la borne supérieure des projections et met en évidence les risques extrêmes en cas d'inaction climatique.

6.4 Projections futures de la sécheresse au barrage de Keddara

L'analyse suivante détaille les projections de sécheresse au niveau du barrage de Keddara, simulées à partir de l'indice SPI pour les quatre échelles temporelles (SPI-1, SPI-3, SPI-6 et SPI-12), et ce, selon chacun des trois scénarios climatiques retenus.

Chaque scénario (SSP1-2.6, SSP2-4.5 et SSP5-8.5) est analysé séparément afin de mettre en évidence les différences de trajectoire possibles en fonction des choix socio-économiques et des niveaux d'émissions de gaz à effet de serre. L'objectif est d'identifier, pour chaque situation, les périodes critiques, les intensités extrêmes, et les tendances évolutives entre 2025 et 2100, susceptibles d'affecter la disponibilité en eau à court, moyen et long termes.

Cette analyse différenciée constitue un outil essentiel pour anticiper les risques et adapter les politiques de gestion de l'eau selon les niveaux de vulnérabilité projetés dans chaque scénario.

6.4.1 Scénario SSP1–2.6

SPI1 : Sous le scénario optimiste SSP1–2.6, plusieurs épisodes de sécheresse météorologique courte, mais intenses sont attendus entre 2028 et 2097. Bien que la durée de ces sécheresses reste généralement limitée à un ou deux mois, certaines atteignent des niveaux très sévères, comme en 2028, 2031 ou 2083, avec des indices SPI proches ou inférieurs à -3 . Ces épisodes montrent que, même avec une stabilisation du climat global, des événements extrêmes ponctuels continueront de survenir, en particulier au printemps, période sensible pour les ressources en eau et l'agriculture.

SPI3 : les événements de sécheresse agricole ont été détectés au barrage de Keddara entre 2026 et 2097 à travers l'indice SPI-3. Ces épisodes, d'une durée de 1 à 3 mois, apparaissent majoritairement entre avril et juillet, période critique pour la croissance des cultures. Certains événements atteignent des intensités marquées, comme en 2044, 2062, 2066 ou 2080, avec des SPI inférieur à $-1,7$, traduisant un déficit d'humidité du sol sur plusieurs mois. Ces résultats montrent que même dans un contexte de réduction des émissions, les systèmes agricoles resteront vulnérables à des sécheresses saisonnières récurrentes.

SPI6 : Les séquences de sécheresse hydrologique ont été identifiées au barrage de Keddara entre 2026 et 2100 à travers l'indice SPI-6. Ces épisodes, de durée généralement comprise entre 2 et 5 mois, reflètent des déficits de précipitations prolongés, susceptibles d'affecter les écoulements des oueds et la recharge du barrage. Des événements intenses apparaissent notamment en 2028, 2033, 2055 et 2085 avec des valeurs de SPI inférieures à $-1,6$, indiquant un manque significatif d'eau à l'échelle du bassin.

SPI-12 : les résultats issus de l'indice SPI-12 révèlent plusieurs épisodes de sécheresse prolongée affectant le rechargement du barrage de Keddara. Parmi ceux-ci, l'épisode de 2028–2029 se distingue par sa durée (14 mois) et son intensité importante (SPI de $-2,01$), traduisant un déficit hydrique annuel marqué. D'autres événements notables apparaissent en 2068 (3 mois) et entre 2094 et 2095 (8 mois), avec des valeurs inférieures à $-1,5$, indiquant une pression persistante sur les ressources en eau stockées.

Récapitulatif – SSP1–2.6 :

Sous le scénario optimiste SSP1–2.6, qui suppose une réduction significative des émissions de gaz à effet de serre, les analyses montrent que le bassin de Keddara restera exposé à différents types de sécheresse jusqu'à la fin du siècle. Les épisodes de sécheresse météorologique (SPI-1), bien que généralement courts (1 à 2 mois), présentent parfois une intensité extrême ($\text{SPI} \leq -3$), notamment en 2028, 2031 et 2083, impactant directement les ressources immédiates en eau. La sécheresse agricole (SPI-3), souvent observée entre avril et juillet, atteint des seuils critiques en 2044, 2062, 2066 et 2080, traduisant un déficit d'humidité du sol en période clé pour les cultures. La sécheresse hydrologique (SPI-6), plus prolongée (jusqu'à 5 mois), est marquée par des événements notables en 2028, 2033, 2055 et 2085, compromettant les écoulements superficiels et la recharge des retenues. Enfin, la sécheresse à long terme (SPI-12) révèle des déficits hydriques annuels importants, comme celui de 2028–2029 (14 mois, $\text{SPI} = -2,01$), pouvant durablement affecter les réserves en eau. Ces résultats confirment qu'un climat plus stable ne garantit pas l'absence d'événements extrêmes, soulignant la nécessité d'une gestion intégrée, multisectorielle et anticipative des sécheresses à toutes les échelles temporelles.

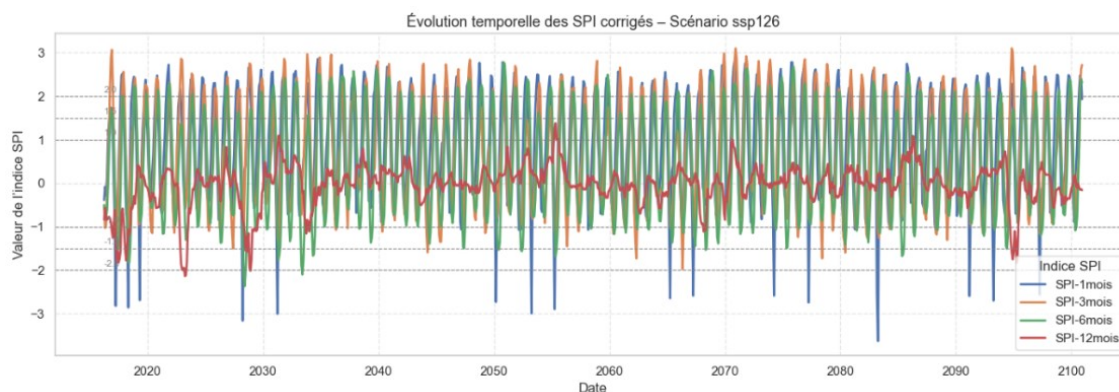


FIGURE 6.2 – Évolution temporelle pour les différents SPI selon le scénario SSP1-2.6

6.4.2 Scénario SSP2–4.5

SPI1 : La plupart épisodes de sécheresse se concentrent entre mars et juin . Plusieurs événements affichent des intensités extrêmes, avec des valeurs de SPI-1 inférieures à -3 , comme en 2028, 2041, 2052, 2055, 2060 ou encore 2098, traduisant des déficits pluviométriques brutaux mais marqués. Ce scénario révèle une fréquence élevée et une gravité croissante de ces épisodes.

SPI3 : Les séquences sèches agricoles, d’une durée de 1 à 3 mois, surviennent principalement au printemps et au début de l’été, période essentielle pour la croissance des cultures. Plusieurs événements montrent des intensités modérées à fortes, avec des valeurs de SPI-3 inférieures à $-1,5$, notamment en 2031, 2052, 2081 ou 2083, traduisant une baisse prolongée de l’humidité du sol.

SPI 6 : l’indice met en évidence de nombreux épisodes de sécheresse hydrologique au barrage de Keddara entre 2026 et 2100, avec des durées variant de 1 à 5 mois. Ces événements traduisent des déficits prolongés en précipitations susceptibles d’affecter les écoulements des oueds et le rechargement de la retenue. Plusieurs sécheresses intenses ont été enregistrées, notamment en 2035, 2036, 2073 et 2075, avec des valeurs de SPI-6 inférieures à $-2,3$, signalant une forte pression sur les ressources hydriques du bassin. La récurrence des épisodes modérés (valeurs comprises entre $-1,1$ et $-1,7$) tout au long du siècle souligne une variabilité hydrologique persistante, affectant directement la gestion de l’eau pour l’agriculture, l’alimentation en eau potable et l’industrie.

SPI12 : Des épisodes de sécheresse annuelle indiquant des périodes prolongées de déficit hydrique impactant l’accumulation d’eau dans le bassin versant. L’épisode de 2057–2058 se distingue par sa durée exceptionnelle de 11 mois et son intensité extrême ($\text{SPI} = -2,15$), traduisant un stress hydrique majeur sur le système de stockage. D’autres épisodes notables apparaissent entre 2028 et 2029 (5 mois, $\text{SPI} = -1,99$) et entre 2036 et 2037 (10 mois, $\text{SPI} = -1,86$), soulignant une récurrence inquiétante de périodes longues avec une recharge insuffisante.

Récapitulatif – SSP2–4.5 :

Sous le scénario SSP2–4.5, qui représente une trajectoire intermédiaire avec des efforts modérés d’atténuation, le barrage de Keddara pourrait faire face à une diversité d’épisodes de sécheresse entre 2026 et 2100. Des sécheresses météorologiques courtes mais intenses surviennent fréquemment au printemps, avec des valeurs de SPI-1 inférieures à -3 lors d’événements extrêmes (2028, 2041, 2052, 2060, 2098), traduisant une baisse significative des précipitations. En parallèle, les sécheresses agricoles (SPI-3), affectant la capacité de rétention d’eau des sols,

apparaissent chaque décennie, notamment au printemps et en début d'été, avec des impacts notables sur les cultures en 2031, 2052 et 2083. Sur le plan hydrologique (SPI-6), plusieurs épisodes prolongés mettent sous pression les ressources en eau du bassin, notamment en 2035, 2036, 2073 et 2075, avec des déficits sévères dépassant $-2,3$. Enfin, de longues sécheresses affectent le rechargement annuel des réservoirs, avec des déficits critiques en 2028–2029, 2036–2037 et surtout 2057–2058 (SPI-12 = $-2,15$). Ces résultats montrent que même dans un scénario modéré, les sécheresses resteront fréquentes, intenses et multiformes.

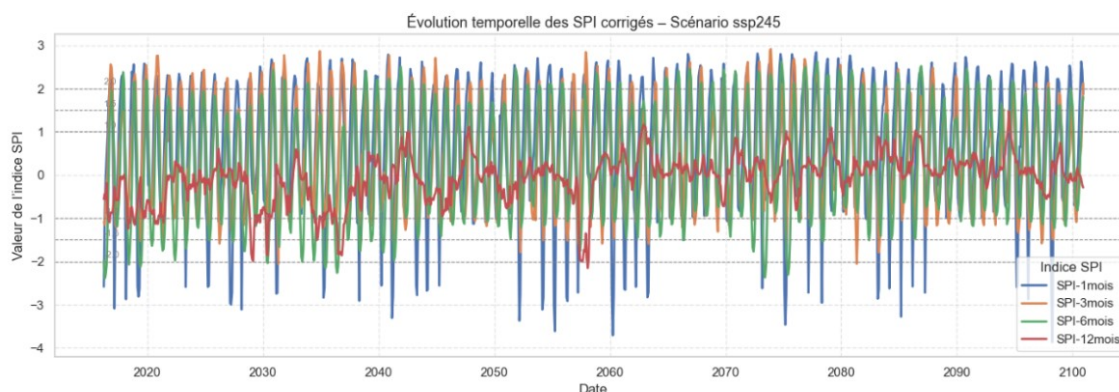


FIGURE 6.3 – Évolution Temporelle pour les différents SPI selon le scénario ssp245

6.4.3 Scénario SSP5–8.5

Le scénario SSP5–8.5 représente une trajectoire socio-économique à fortes émissions, caractérisée par une forte dépendance aux énergies fossiles, une croissance rapide et peu de politiques de réduction des émissions. Il correspond à la situation la plus critique du point de vue du réchauffement climatique. Les projections issues de l'indice SPI montrent une aggravation généralisée des épisodes de sécheresse au barrage de Keddara, à toutes les échelles temporelles.

SPI1 : Les événements pluviométriques déficitaires à court terme deviennent plus fréquents, intenses et marqués au barrage de Keddara entre 2026 et 2092. La majorité des événements se concentrent toujours entre mars et juin. De nombreux épisodes présentent des intensités extrêmes, avec des valeurs de SPI-1 inférieures à -3 comme en 2066 ($-3,71$), 2076 ($-3,18$), 2079 ($-3,46$) ou encore 2090 ($-3,54$). La répétition de ces épisodes extrêmes tout au long du siècle indique une pression croissante sur les écosystèmes et les systèmes de gestion de l'eau.

SPI-3 : Le stress hydrique pour la végétation, d'une durée de 1 à 3 mois, se produit principalement entre avril et juillet, période critique pour la croissance des cultures. Bien que la majorité des événements présentent des intensités modérées (SPI < -1), plusieurs épisodes dépassent le seuil de $-1,4$, comme en 2076 ($-1,83$), 2078 ($-1,45$) ou encore 2091 ($-1,31$), indiquant un stress hydrique marqué pour les sols cultivés.

SPI-6 : Les affaiblissements des apports en eau de surface, généralement concentrés entre avril et septembre, durent entre 2 et 5 mois et traduisent une diminution prolongée des précipitations, affectant directement les écoulements superficiels et les apports aux réservoirs. Plusieurs événements présentent une intensité élevée, avec des valeurs inférieures à -2 , notamment en 2025, 2035, 2048 ou encore 2053, traduisant un stress hydrique sévère. En plus de ces épisodes extrêmes, on observe une fréquence soutenue de sécheresses modérées à fortes (SPI entre $-1,0$ et $-1,7$), réparties sur l'ensemble du siècle. Cette répétition d'événements hydrologiques défavorables souligne une pression chronique sur les ressources en eau.

SPI-12 : Les carences hydrologiques prolongées, s'étalant souvent sur plusieurs mois consécutifs, traduisent des perturbations majeures dans le cycle de recharge des réservoirs. Certains événements, comme ceux de 2029–2031 (22 mois, $\text{SPI} = -2,37$) ou encore de 2049–2051 (18 mois, $\text{SPI} = -2,11$), présentent à la fois une durée exceptionnelle et une intensité critique, signalant une pression extrême sur les réserves stratégiques en eau. D'autres épisodes notables, de 9 à 12 mois, apparaissent régulièrement, notamment en 2032, 2045 ou 2053, témoignant d'un stress hydrique répété et persistant. Même les événements de plus courte durée (1 à 3 mois) présentent souvent une intensité significative, avec des valeurs de SPI inférieures à $-1,1$.

Récapitulatif – SSP5–8.5 :

Selon le scénario SSP5–8.5, représentant une trajectoire à fortes émissions, le barrage de Keddara est confronté à une intensification généralisée de tous les types de sécheresse entre 2025 et 2092. Les sécheresses météorologiques (SPI-1), souvent brèves mais extrêmement intenses, surviennent fréquemment entre mars et juin, avec des épisodes extrêmes dépassant $-3,5$, traduisant des déficits pluviométriques soudains. Les sécheresses agricoles (SPI-3), concentrées entre avril et juillet, deviennent plus nombreuses et marquées, affectant directement la croissance des cultures. Parallèlement, les sécheresses hydrologiques (SPI-6) s'installent de manière récurrente, avec des durées moyennes de 3 à 5 mois, perturbant les écoulements et les recharges des réservoirs. Enfin, les sécheresses annuelles (SPI-12) révèlent des déficits prolongés et parfois extrêmes, comme entre 2029 et 2031 ou 2049 et 2051, compromettant la durabilité des ressources en eau. Cette convergence d'impacts, à différentes échelles temporelles, souligne la vulnérabilité croissante du système hydrique local face au réchauffement climatique.

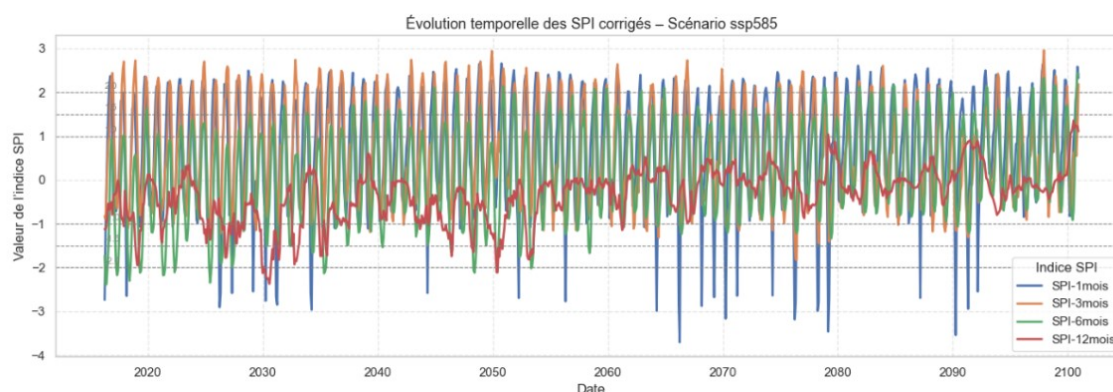


FIGURE 6.4 – Évolution Temporelle pour les différents SPI selon le scénario ssp585

6.5 Identification des dangers révélés par les résultats obtenus

Les projections de sécheresse réalisées pour le bassin versant de Keddara confirment les grandes tendances climatiques annoncées pour la région du Maghreb. Même dans le scénario optimiste SSP1-2.6, les sécheresses demeurent fréquentes, parfois d'intensité extrême. Cette observation corrobore les rapports du GIEC, qui anticipent une augmentation de la fréquence, de la durée et de la sévérité des sécheresses dans les régions méditerranéennes au cours du XXI^e siècle, et en particulier en Afrique du Nord [1].

6.5.1 Une vulnérabilité hydrique croissante

À Keddara, les différentes échelles de sécheresse (SPI-1, SPI-3, SPI-6, SPI-12) montrent un impact progressif et cumulé sur la disponibilité de l'eau :

- Baisse des précipitations : enregistrée à toutes les échelles temporelles, réduisant l'alimentation naturelle du bassin.
- Diminution des écoulements de surface : révélée par le SPI-6, elle affecte directement la capacité de collecte dans le barrage.
- Déficit de recharge annuel des réservoirs : mis en évidence par le SPI-12, ce qui compromet la résilience à long terme du système hydrique.

Selon les projections climatiques pour l'Algérie, les précipitations annuelles pourraient chuter jusqu'à 40% d'ici à la fin du siècle, entraînant un déséquilibre structurel entre l'offre et la demande en eau [1]. Le barrage de Keddara constitue ainsi un exemple local emblématique de ce déséquilibre attendu à l'échelle régionale.

6.5.2 Impacts sectoriels – Le cas critique de l'agriculture

La sécheresse agricole apparaît comme l'une des plus préoccupantes, tant par sa fréquence que par sa période d'occurrence. Les résultats montrent que les épisodes SPI-3 surviennent régulièrement entre avril et juillet, période décisive pour les cultures céréalières (blé, maïs, sorgho, millet). Ce déficit d'humidité du sol en pleine saison de croissance met en péril les rendements agricoles.

Les effets sont déjà visibles :

- En 2015, la faiblesse de la production locale a entraîné une hausse des importations de céréales de 14 % pour compenser le déficit et reconstituer les stocks.
- En 2016, la superficie emblavée a chuté de 13 %, et la productivité n'a pas dépassé 11,8 quintaux/ha, démontrant la forte sensibilité du secteur aux aléas climatiques.

Ces constats renforcent les alertes du GIEC, qui prévoit, pour le Maghreb, des pertes agricoles significatives si des mesures d'adaptation ne sont pas mises en œuvre rapidement.

TABLE 6.1 – Balance alimentaire céréalière (blé) en Algérie [1].

Élément	Unité	2014	2015	2016	2017
Production	1 000 T	2 436	2 657	2 440	2 437
Importations – Quantité	1 000 T	7 461	8 531	8 255	8 105
Variation de stock	1 000 T	117	900	400	0
Exportations – Quantité	1 000 T	3	7	6	10
Disponibilité intérieure	1 000 T	9 778	10 281	10 289	10 532
Superficie récoltée	ha	1 651 311	1 814 722	2 062 179	2 118 469
Rendement	T/ha	1,48	1,46	1,18	1,15

Les résultats obtenus dans le cadre de cette étude montrent une sécheresse persistante qui affecte la saison de croissance, renforçant ainsi les prévisions du GIEC et soulignant l'urgence de renforcer les stratégies d'adaptation agricole dans la région.

6.5.3 Analyse sectorielle des impacts de la sécheresse à Keddara et dans les régions semi-arides d'Algérie

Selon ce qui est indiqué dans le Plan national de sécheresse d'Algérie [62] :

La sécheresse associée à d'autres aléas climatiques tels que la canicule et les vagues de chaleur (ou de froid), joue toujours un rôle perturbateur et ses impacts sur les secteurs économiques sont souvent immédiats.

Sur le secteur des ressources en eau :

- La sécheresse provoque une baisse des apports en eau dans les lacs et les barrages, ce qui réduit les réserves disponibles pour l'irrigation agricole et l'approvisionnement des villes. En parallèle, le manque de pluie limite le renouvellement des nappes souterraines, ce qui pousse à une utilisation excessive de ces ressources.
- Avec l'augmentation de la demande en eau, notamment pour l'agriculture et les besoins domestiques, les pressions sur les nappes, qu'elles soient renouvelables ou fossiles, deviennent de plus en plus fortes. En plus, la qualité de l'eau se dégrade, car il y a moins d'eau pour diluer les polluants, ce qui augmente les risques de contamination.
- Enfin, ces tensions entraînent des perturbations dans la distribution de l'eau, pouvant aller jusqu'à des situations de pénurie. Face à cela, les autorités sont contraintes de prioriser certains usages de l'eau (comme l'alimentation des populations ou l'irrigation) et, dans les zones les plus touchées, de recourir à des solutions temporaires, comme la distribution d'eau par camions-citernes.

Sur l'agriculture, les forêts et la biodiversité :

- La baisse des précipitations, combinée à la rareté de l'eau et à l'élévation des températures, réduit l'humidité des sols, freine la croissance des plantes, diminue les rendements agricoles et menace la santé du bétail.
- Les écosystèmes naturels — forêts, maquis et steppes — deviennent plus vulnérables, ce qui favorise les feux de forêt, l'érosion éolienne et la dégradation des sols. Ces déséquilibres affectent la biodiversité, en provoquant la migration d'animaux sauvages, la disparition de certaines espèces, et des réactions écologiques comme le regroupement massif des criquets en cas de végétation appauvrie, entraînant des invasions acridiennes similaires à celles de 2004–2005.
- La répétition des sécheresses pourrait même étendre les zones de reproduction des criquets dans des régions jusque-là non affectées. Enfin, la faune sauvage et domestique devient plus vulnérable aux maladies, et la réduction des ressources alimentaires peut entraîner des cas de malnutrition animale et humaine.

Sur la santé, la population et l'économie :

- La sécheresse, souvent aggravée par les canicules, a des effets graves sur la santé publique. Elle peut entraîner de la déshydratation, la malnutrition, des épidémies, et une augmenta-

tion des maladies respiratoires et cardiovasculaires (comme l'asthme ou l'hypertension). Elle favorise aussi l'émergence de nouvelles maladies liées aux migrations d'insectes et de pathogènes, ainsi que des risques accrus de morsures d'animaux dangereux (serpents, scorpions...).

- La sécheresse a également des conséquences psychosociales : elle augmente les troubles mentaux (dépression, anxiété), les tensions sociales, et peut provoquer des conflits autour de l'eau et de la nourriture. Dans les cas extrêmes, elle pousse à la migration forcée, créant ainsi des réfugiés climatiques.
- Enfin, la sécheresse affecte aussi l'industrie et les secteurs économiques, en perturbant l'approvisionnement en eau et en énergie, ce qui peut entraîner une baisse de la production et des pertes économiques importantes.

6.6 Plan National Sécheresse (Algérie, 2019)

6.6.1 Définition du Plan National de la Sécheresse (PNS)

Le Plan National de la Sécheresse en Algérie constitue un cadre stratégique mis en place par l'État, à travers le Ministère de l'Agriculture, du Développement Rural et de la Pêche, notamment la Direction Générale des Forêts, avec l'appui des Nations Unies. Son objectif central est de doter les institutions nationales d'un dispositif spécifique, cohérent et réactif, capable de répondre efficacement à l'un des aléas climatiques les plus critiques du XXI^e siècle : la sécheresse.

Ce plan vise à renforcer la résilience des milieux naturels, des systèmes agricoles, des ressources en eau et des infrastructures, en mettant en œuvre des mesures d'anticipation, de prévention, d'atténuation et d'intervention, adaptées aux nouvelles réalités imposées par le changement climatique.

6.6.2 Objectifs principaux du PNS

1. Préparer l'ensemble des secteurs à la gestion proactive du risque sécheresse, en amont des crises.
2. Coordonner les interventions des institutions nationales et locales à travers un cadre structuré.
3. Réduire les impacts environnementaux, économiques et sociaux des épisodes de sécheresse.
4. Intégrer la sécheresse dans les politiques de développement durable et les plans d'adaptation au climat.

6.6.3 Recommandation institutionnelle

Le document recommande la création d'un groupe thématique spécifique "sécheresse" au sein du Comité National Climat (CNC), qui viendrait compléter les groupes déjà existants. Ce groupe jouerait un rôle clé dans la convergence entre le PNS (Plan National Sécheresse) et le PNC (Plan National Climat), deux cadres complémentaires à harmoniser pour une meilleure efficacité.

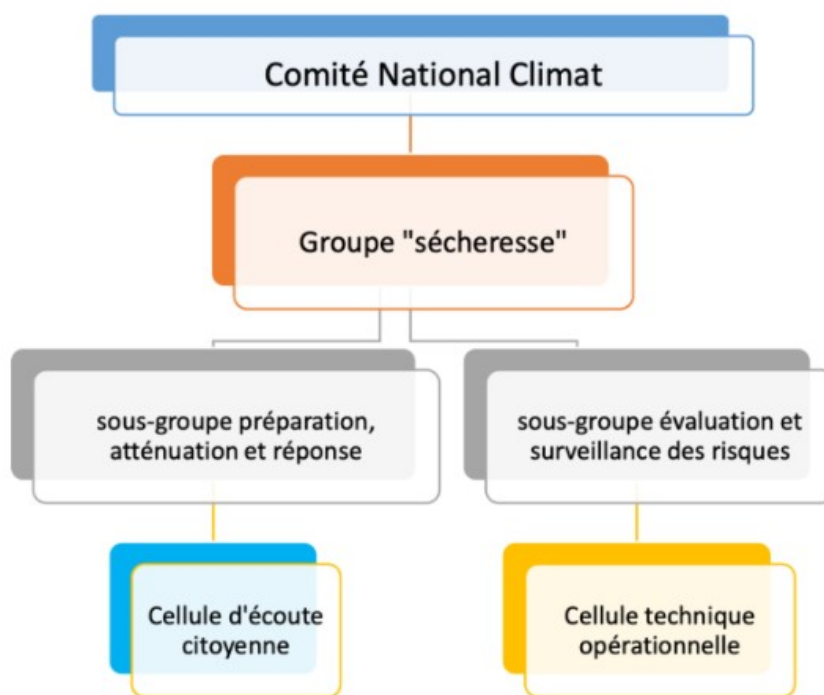


FIGURE 6.5 – Structure organisationnelle du groupe (sécheresse)

6.6.4 Missions du groupe thématique « sécheresse »

- Supervision et coordination du développement du Plan « sécheresse », en mobilisant les ressources intersectorielles (ministères, collectivités locales).
- Mise en œuvre du plan à tous les niveaux de
- Activation rapide des mécanismes du plan en période de sécheresse.
- Application des mesures d'atténuation, directement ou par délégation locale.
- Formulation de recommandations stratégiques, sur la base des retours d'expérience et du suivi de l'efficacité des actions entreprises.

6.7 Révision du Plan National de Sécheresse

Le rapport du GIEC (GIEC, 2023) a donné la voie pour faciliter les actions de résilience à la sécheresse à tous les niveaux de décision et de mise en œuvre. Les principales composantes pour une action effective sont :



Source: GIEC, 2023.

FIGURE 6.6 – Principales composantes pour une action résilient

- **Engagement politique :**

Pour faire face à la sécheresse, il est essentiel que les autorités politiques montrent une réelle volonté d'agir. Cela se traduit par l'adoption de politiques climatiques ambitieuses, l'intégration du risque climatique dans les plans de développement nationaux et locaux, ainsi que la mobilisation de financements durables pour mettre en œuvre ces actions.

- **Gouvernance inclusive :**

Une gouvernance efficace implique la participation de tous les acteurs concernés. Il est important d'associer les communautés locales, les groupes marginalisés, les femmes, les jeunes et les peuples autochtones aux processus de décision. Cette approche permet une meilleure appropriation des mesures de résilience et garantit qu'elles soient adaptées aux réalités du terrain.

- **Coopération internationale :**

La sécheresse ne connaît pas de frontières. Pour y faire face efficacement, les pays doivent coopérer entre eux. Cela passe par le partage des données climatiques, l'échange de bonnes pratiques, l'harmonisation des stratégies d'adaptation et la mobilisation de financements internationaux, notamment pour les pays les plus vulnérables.

- **Gestion effective à l'échelle de l'écosystème :**

Il est crucial d'adopter une gestion intégrée des ressources naturelles — telles que l'eau, les sols et les forêts — en prenant en compte l'ensemble des interactions au sein des écosystèmes. Cette approche aide à renforcer la résilience des territoires face aux effets du changement climatique, en particulier la sécheresse.

- **Renforcement des capacités :**

Le développement de compétences locales est une condition clé pour faire face à la sécheresse. Cela comprend la formation des acteurs sur le terrain, l'amélioration des outils de suivi et d'alerte précoce, ainsi que le soutien à la recherche et à l'expertise technique pour anticiper et réagir plus efficacement aux événements extrêmes.

Bien que le Plan National de Sécheresse (PNS) ait été publié en 2019, il n'a pas vraiment été mis en œuvre. Plusieurs problèmes freinent son application :

- Le groupe spécial sur la sécheresse prévu dans le Comité National Climat n'a jamais été créé.
- Il y a un manque de coordination entre les ministères et les collectivités locales.
- Les responsables politiques n'ont pas suffisamment agi pour transformer ce plan en actions concrètes sur le terrain.

Selon les critères du GIEC (2023), plusieurs faiblesses apparaissent clairement :

- Pas assez d'engagement politique : Le plan n'a pas de force légale : aucune loi ou décret ne l'impose comme priorité. Les actions des secteurs comme l'eau, l'agriculture ou l'environnement ne sont pas bien reliées entre elles, et le financement dépend presque uniquement du budget de l'État.
- Peu de participation des populations : Les citoyens, en particulier les femmes rurales, les jeunes, les agriculteurs ou les associations, ne sont pas vraiment associés aux décisions. Cela rend les mesures de lutte contre la sécheresse moins efficaces et moins bien acceptées.
- Ressources naturelles mal gérées ensemble : Chaque ressource est gérée séparément : l'eau par une institution, la forêt par une autre, etc. Il n'y a pas encore une vision globale pour gérer les zones touchées par la sécheresse, comme les bassins versants.
- Compétences locales à renforcer : Certains outils existent (bulletins d'alerte...), mais dans beaucoup de régions, le personnel manque de formation et les outils ne sont pas encore utilisés partout.

6.8 Perspectives pour l'Algérie à l'horizon 2030

L'année 2030 constitue une échéance stratégique, à la fois dans le cadre de l'**Accord de Paris sur le climat**, qui prévoit des bilans d'étape des engagements nationaux, et dans celui de l'**Agenda 2030 des Nations Unies**, qui fixe 17 Objectifs de Développement Durable (ODD). Pour l'Algérie, confrontée à une intensification des sécheresses, cette échéance est une opportunité de transition vers un modèle plus résilient et plus durable.

Si des initiatives de long terme, comme la création d'un groupe thématique interinstitutionnel sur la sécheresse, sont envisagées, elles demeurent lentes à se concrétiser. Or, le changement climatique, lui, n'attend pas. Il devient urgent de compléter ces démarches structurelles par des mesures concrètes à court terme, facilement déployables par les institutions existantes, sans attendre de réformes majeures. Cela implique de mobiliser trois leviers clés : **le secteur agricole, le système financier et le tissu social**.

Une attention particulière doit être portée au secteur agricole, qui est **le plus vulnérable face à la sécheresse**. Cette fragilité menace directement **la souveraineté alimentaire** du pays, tout en fragilisant l'économie nationale. Ce risque est d'autant plus préoccupant que l'Algérie affiche une volonté affirmée de réduire sa dépendance aux hydrocarbures et de faire de **l'agriculture un pilier central de son développement économique à l'horizon 2030**. Il est donc impératif que toute stratégie de gestion de la sécheresse accorde à ce secteur une priorité absolue, en mobilisant à la fois les outils financiers, techniques et politiques nécessaires à sa résilience.

6.8.1 Promouvoir l'agriculture de conservation des sols et L'agroforesterie

Le secteur agricole, principal utilisateur d'eau, est aussi le plus exposé aux effets de la sécheresse. Pour renforcer sa résilience, il est essentiel de diffuser les pratiques d'agriculture de conservation des sols, qui reposent sur trois principes fondamentaux :

- **Réduction du travail mécanique du sol** (labour minimal ou semis direct) afin de préserver la structure, favoriser l'infiltration de l'eau et la vie biologique du sol ;
- **Couverture végétale permanente**, même hors période de culture, pour limiter l'évaporation, maintenir l'humidité et prévenir l'érosion ;
- **Diversification des cultures** (rotations, associations), afin d'améliorer la structure du sol, limiter les maladies, et enrichir la biodiversité souterraine.
- **L'agroforesterie** est une pratique agricole qui consiste à associer des arbres et des cultures ou de l'élevage sur une même parcelle. Cette association peut se faire en bordure de champs (haies) ou à l'intérieur des parcelles (alignements d'arbres, arbres isolés). L'agroforesterie vise à combiner les bénéfices des arbres et ceux de l'agriculture, en créant un système plus diversifié et résilient.

Ces pratiques, soutenues par la FAO, permettent de mieux retenir l'eau dans les sols, de préserver leur fertilité et de sécuriser les rendements en conditions hydriques difficiles. Une politique d'appui financier (subventions, formations, accompagnement technique) portée par le ministère de l'Agriculture serait une incitation forte à leur adoption.

6.8.2 Mobiliser la finance pour l'adaptation agricole

L'adaptation à la sécheresse passe aussi par une mobilisation du secteur financier, encore peu engagé dans le soutien à l'agriculture résiliente. Il est nécessaire de développer une finance verte inclusive, impliquant banques, assurances et fonds d'investissement, autour de plusieurs axes :

-
- Création de **crédits agricoles spécifiques** pour l'équipement en techniques économes en eau (goutte-à-goutte, cuves de récupération, filets brise-vent...);
- Mise en place de **produits d'assurance sécheresse** adaptés au contexte algérien;
- Soutien aux **projets d'agriculture durable** à travers des mécanismes d'appui (fonds climat, subventions nationales, partenariats publics-privés).

Ce financement doit adopter une approche "chaîne de valeur", couvrant l'ensemble des acteurs agricoles (producteurs, transformateurs, transporteurs, distributeurs), pour construire un écosystème résilient face aux aléas climatiques. Mais cela nécessite une meilleure **identification des besoins locaux**, **un accès sécurisé à la terre et à l'eau**, et **un accompagnement technique ciblé**.

6.8.3 Renforcer le relais social pour la résilience hydrique

Enfin, la lutte contre la sécheresse doit être portée par un relai social fort et mobilisé. Cela passe par trois axes majeurs :

- **Cadre légal et réglementaire renforcé** : lois sur l'usage raisonné de l'eau, pénalités pour gaspillage, mais aussi mécanismes d'accès équitable aux ressources;
- **Éducation et sensibilisation** : intégration d'un module sur l'eau, le climat et les gestes écologiques dans les programmes scolaires, appui des médias, des ONG et des leaders locaux pour diffuser une culture de la sobriété hydrique;
- **Participation citoyenne** : inclusion active des familles, agriculteurs, collectivités locales et jeunes dans la gestion partagée de l'eau, afin de créer un climat de confiance et de coresponsabilité.

La résilience climatique ne pourra émerger que si elle repose sur **une conscience collective**, fondée sur la solidarité, la pédagogie, la justice sociale et l'ancrage territorial.

En résumé, l'horizon 2030 n'est pas une fin en soi, mais une étape décisive pour structurer une réponse rapide, intégrée et efficace aux sécheresses. En conjuguant **des pratiques agricoles adaptées**, un financement vert inclusif et une mobilisation sociale active, l'Algérie peut transformer la contrainte climatique en opportunité pour réinventer son modèle de développement.

6.9 Conclusion

L'analyse des sécheresses futures dans la région de Keddara, fondée sur les données climatiques projetées, confirme que même dans les scénarios les plus optimistes, l'Algérie restera exposée à des épisodes extrêmes de déficit hydrique. Ces résultats renforcent l'urgence d'agir dès aujourd'hui pour anticiper, adapter et atténuer les impacts à venir. Pourtant, malgré la

publication en 2019 d'un Plan National de la Sécheresse, sa mise en œuvre reste très limitée, freinée par un manque de coordination, de volonté politique et d'intégration intersectorielle.

Dans ce contexte, une stratégie à court terme, ciblée et inclusive à l'horizon 2030, devient indispensable. Elle doit reposer sur des actions simples, applicables rapidement, et portées par tous les acteurs : institutions, collectivités, société civile, et surtout le secteur agricole, particulièrement vulnérable.

CONCLUSION GÉNÉRALE

Dans cette étude prospective sur les sécheresses futures en Algérie, nous nous sommes d’abord concentrées sur l’amélioration de la fiabilité des projections climatiques à l’aide du downscaling statistique. L’objectif principal était de comparer les approches de désagrégation appliquées soit directement à l’indice SPI (Standardized Precipitation Index), soit aux précipitations brutes, afin d’identifier la méthode la plus pertinente pour la modélisation des sécheresses futures.

Il ressort que le downscaling direct du SPI, calculé à différentes échelles temporelles (SPI-1, SPI-3, SPI-6 et SPI-12), offre de meilleures performances que le downscaling des précipitations. Cette supériorité s’explique notamment par le fait que les deux séries de SPI – celle dérivée des données brutes (échelle globale) et celle issue des données downscalées (échelle régionale) – présentent une distribution statistique homogène grâce à leur normalisation intégrée dans le calcul du SPI. Cette cohérence statistique facilite une comparaison directe et améliore la robustesse des projections.

Deux approches ont été évaluées sur la période historique [1950—2014] :

- Le modèle XGBoost, une méthode d’apprentissage automatique,
- Le modèle CNN (Convolutional Neural Network), relevant du deep learning.

Les résultats montrent que le CNN surpasse XGBoost en termes de performance globale, notamment pour la reconstruction du signal climatique. Cependant, les deux modèles présentent une limitation commune : une incapacité à reproduire fidèlement les valeurs extrêmes de sécheresse, qui sont pourtant cruciales pour l’évaluation des risques. C’est pourquoi un ajustement post-downscaling, basé sur la technique du quantile mapping, a été intégré pour corriger les biais et rétablir les extrêmes climatiques, notamment dans les projections SPI futures couvrant la période [2026—2100].

L’analyse des projections pour le barrage de Keddara, basé sur le SPI à différentes échelles, révèle une tendance nette à l’aggravation des sécheresses futures, quels que soient les scénarios climatiques étudiés :

- Le scénario SSP1-2.6, le plus optimiste en termes de réduction des émissions, montre que malgré les efforts climatiques, des épisodes extrêmes de sécheresse printanière persistent. Cela compromet des périodes cruciales pour l’agriculture et la recharge hydrique.
- Le scénario SSP2-4.5 projette une fréquence soutenue de sécheresses agricoles et hydrologiques, perturbant les cycles de culture et la gestion des ressources en eau.
- Enfin, le scénario SSP5-8.5, basé sur une absence d’action climatique, anticipe des sécheresses extrêmes, récurrentes et multifformes. Celles-ci combinent les impacts météorologiques, agricoles, hydrologiques et sociaux, avec des valeurs de SPI-12 souvent inférieures à -2,5, signalant un risque majeur pour la sécurité hydrique, la production alimentaire et la résilience des écosystèmes.

Face à cette réalité, le Plan National de Lutte contre la Sécheresse de 2019 a été révisité à la lumière des critères du GIEC 2023, afin de proposer une vision stratégique à l'horizon 2030. Cette vision s'inscrit dans une dynamique internationale, faisant de 2030 une année charnière pour l'adaptation climatique. Les recommandations proposées à court terme visent principalement à protéger l'agriculture, un secteur particulièrement vulnérable, à travers :

- L'adoption de pratiques agricoles durables (agroécologie, rotation des cultures, gestion intégrée de l'eau),
- L'engagement en faveur du financement vert (fonds climatiques, investissements résilients),
- Et l'inclusion des acteurs sociaux, notamment les communautés rurales, dans la gouvernance de l'eau et du climat.

En conclusion, cette étude met en évidence la nécessité d'intégrer des outils de modélisation robustes, tels que le downscaling statistique ajusté, pour mieux anticiper les futurs climatiques régionaux et orienter les politiques publiques vers une résilience proactive face à la sécheresse.

Bibliographie

- [1] Fethi Lebdi and Abdourahman Maki. La sécheresse au maghreb : diagnostic, impacts et perspectives pour le renforcement de la résilience du secteur agricole, 2023. FAO international consultant, fonctionnaire technique terres et eaux à la FAO.
- [2] Mohamed Safar-Zitoun. Plan national sécheresse algérie : Lignes directrices en vue de son opérationnalisation. Technical report, Ministère de l’Agriculture, du Développement Rural et de la Pêche, Direction Générale des Forêts, Alger, Algérie, 2019. Consultant CNULCD, publié le 30 juin 2019.
- [3] Tércio Ambrizzi, Michelle Simões Reboita, Rosmeri Porfírio da Rocha, and Marta Llopart. The state of the art and fundamental aspects of regional climate modeling in south america. *Annals of the New York Academy of Sciences*, 1436(1) :98–120, 2018.
- [4] Bouchra El Jgham, Otman Abdoun, and Haimoudi El Khatir. Review of weed detection methods based on machine learning models. In *Proceedings of the International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD 2022)*, volume 713 of *Lecture Notes in Networks and Systems*, pages 576–586. Springer, 2023. First Online : 03 June 2023.
- [5] Batta Mahesh. Machine learning algorithms - a review. *International Journal of Science and Research (IJSR)*, 9(1), 2019. Independent Researcher.
- [6] Ahmedbahaaaldin Ibrahim Ahmed Osman, Ali Najah Ahmed, Ming Fai Chow, Yuk Feng Huang, and Ahmed El-Shafie. Extreme gradient boosting (xgboost) model to predict the groundwater levels in selangor, malaysia. *Civil Engineering*, xxxx.
- [7] Cifar-10 image classification in tensorflow. <https://www.geeksforgeeks.org/cifar-10-image-classification-in-tensorflow/>, 2024. Last updated : 20 September 2024.
- [8] A. Sabri-Khelifati, F. Attou, K. Ghennam, M. Hachemi, and Y. Moussaoui. Effect of climatic factors on the evolution of water oxygenation of the keddara dam (boumerdes, algeria). *Applied Ecology and Environmental Research*, 23(1) :1055–1070, 2025. Received 28th May 2024 ; accepted 4th Nov 2024.
- [9] Nations Unis. En quoi consistent les changements climatiques ? 2020. Publication décembre 2020. Disponible à : <https://www.un.org/fr/climatechange/what-is-climate-change#:~:text=Les%20changements%20climatiques%20d%C3%A9signent%20les,%C3%A0%20des%20eruptions%20volcaniques%20massives>.
- [10] Organisation internationale de normalisation. Changements climatiques : définition, fonctionnement et importance du phénomène. 2020. Publication décembre 2020. Disponible à : <https://www.iso.org/fr/changement-climatique>.
- [11] Changement climatique, dérèglement ou réchauffement climatique ? 2024. Publication mai 2024. Disponible à : <https://youmatter.world/fr/definition/changement-climatique-definition/#:~:text=Selon%20la%20d%C3%A9finition%20qu%27en,ses%20propri%C3%A9t%C3%A9s%20et%20qui%20persiste>.
- [12] Donald A. Wilhite. *Drought Assessment, Management, and Planning : Theory and Case Studies*. Springer, 1993.

- [13] Amin Zargar, Rehan Sadiq, Bahman Naser, and Faisal I. Khan. A review of drought indices. *Environmental Reviews*, 2011.
- [14] Donald A. Wilhite and Roger S. Pulwarty, editors. *Drought and Water Crises : Integrating Science, Management, and Policy*. CRC Press, Boca Raton, 2nd edition edition, 2017.
- [15] Donald A. Wilhite and Michael H. Glantz. Understanding the drought phenomenon : The role of definitions. *Water International*, 10(3) :111–120, 1985.
- [16] Thomas B. McKee, Nolan J. Doesken, and John Kleist. The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, Anaheim, CA, 1993. American Meteorological Society.
- [17] Richard R. Heim Jr. A review of twentieth-century drought indices used in the united states. *Bulletin of the American Meteorological Society*, 83(8) :1149–1166, 2002.
- [18] Yves Tramblay, Aristeidis Koutroulis, Luis Samaniego, Sergio M. Vicente-Serrano, Florence Volaire, Aaron Boone, Michel Le Page, Maria Carmen Llasat, Clement Albergel, Selmin Burak, Maxime Cailleret, Ksenija Cindrić Kalin, Hendrik Davi, Jean-Luc Dupuy, Peter Greve, Manolis Grillakis, Lahoucine Hanich, Lionel Jarlan, Nicolas Martin-StPaul, Jordi Martínez-Vilalta, Florent Mouillot, David Pulido-Velazquez, Pere Quintana-Seguí, Delphine Renard, Marco Turco, Murat Türkeş, Ricardo Trigo, Jean-Philippe Vidal, Alberto Vilagrosa, Mehrez Zribi, and Jan Polcher. Challenges for drought assessment in the mediterranean region under future climate scenarios. *Environmental Research Letters*, 15(10) :103001, 2020.
- [19] Anne F. Van Loon. Hydrological drought explained. *WIREs Water*, 2(4) :359–392, 2015.
- [20] A.-E. K. Vrochidou, I. K. Tsanis, M. G. Grillakis, and A. G. Koutroulis. The impact of climate change on hydrometeorological droughts at a basin scale. *Journal of Hydrology*, 476 :290–301, 2013.
- [21] Tarek Benhassen, Mohamed B. Abdallah, Sadok Khalfallah, and Mohamed Zairi. Climate change vulnerability, water resources and social implications in north africa. *Euro-Mediterranean Journal for Environmental Integration*, 5, 2020.
- [22] Yves Tramblay, Aristeidis Koutroulis, Luis Samaniego, Sergio M. Vicente-Serrano, Florence Volaire, Aaron Boone, Michel Le Page, Maria Carmen Llasat, Clement Albergel, Selmin Burak, Maxime Cailleret, Ksenija Cindrić Kalin, Hendrik Davi, Jean-Luc Dupuy, Peter Greve, Manolis Grillakis, Lahoucine Hanich, Lionel Jarlan, Nicolas Martin-StPaul, Jordi Martínez-Vilalta, Florent Mouillot, David Pulido-Velazquez, Pere Quintana-Seguí, Delphine Renard, Marco Turco, and Murat Türkeş. Challenges for drought assessment in the mediterranean region under future climate scenarios. *Earth-Science Reviews*, 210 :103348, 2020.
- [23] Zengchao Hao, Vijay P. Singh, and Youlong Xia. Seasonal drought prediction : Advances, challenges, and future prospects. *Reviews of Geophysics*, 56(1) :108–141, 2018.
- [24] Zengchao Hao, Vijay P. Singh, and Youlong Xia. Seasonal drought prediction : Advances, challenges, and future prospects. *Reviews of Geophysics*, 58(1) :108–141, 2020.
- [25] M. Bazza, M. Kay, and C. Knutson. *Drought characteristics and management in North Africa and the Near East*. Number 45 in FAO Water Reports. FAO, Land and Water Division, Rome, Italy, 2018. ref. 31.
- [26] Mohammed Achite, Nir Y. Krakauer, Andrzej Wałęga, and Tommaso Caloiero. Spatial and temporal analysis of dry and wet spells in the wadi cheliff basin, algeria. *Atmosphere*, 12(6) :798, 2021. Received : 14 May 2021 / Revised : 16 June 2021 / Accepted : 17 June 2021 / Published : 21 June 2021.
- [27] M. Svoboda and B. A. Fuchs. *Handbook of Drought Indicators and Indices*, volume 2 of *Integrated Drought Management Tools and Guidelines Series*. World Meteorological Organization (WMO) and Global Water Partnership (GWP), Geneva, 2016. WMO-No. 1173.

-
- [28] Douglas Maraun. Bias correction, quantile mapping, and downscaling : Revisiting the inflation issue. *Journal of Climate*, 26(6) :2137–2143, 2013.
 - [29] European Centre for Medium-Range Weather Forecasts (ECMWF). Reanalysis q&as. <https://www.ecmwf.int/en/about/media-centre/focus/reanalysis-qas>, 2020. Accessed : 2025-07-02.
 - [30] Intergovernmental Panel on Climate Change (IPCC). Guidelines on the use of gcm data. https://www.ipcc-data.org/guidelines/pages/gcm_guide.html, 2020. Accessed : 2025-07-02.
 - [31] Cynthia Rosenzweig and Ana Iglesias. Potential impacts of climate change on world food supply : Data sets from a major crop modeling study. <https://sedac.ciesin.columbia.edu/data/set/crop-climate-impact-rosenzweig-iglesias>, 1998. Data archived by SEDAC ; superseded by updated datasets. Accessed : 2025-07-02.
 - [32] W. Neil Adger, Iain Michael Brown, and Swenja Surminski. Advances in risk assessment for climate change adaptation policy. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 376(2121) :20180106, 2018.
 - [33] Marie Ekström, Michael R. Grose, and Penny H. Whetton. An appraisal of downscaling methods used in climate change research. *Wiley Interdisciplinary Reviews : Climate Change*, 6(3) :301–319, 2015.
 - [34] Noemi Brunner, Urs Beyerle, Andreas Weigel, Christoph Frei, and Reto Knutti. Evaluation of the empirical–statistical downscaling method episodes. *Climate Dynamics*, 52(1–2) :991–1026, 2019.
 - [35] Francisco J. Tapiador, Andrés Navarro, Raúl Moreno, José Luis Sánchez, and Eduardo García-Ortega. Regional climate models : 30 years of dynamical downscaling. *Atmospheric Research*, 235 :104785, 2020.
 - [36] J. Schmidli, C. M. Goodess, C. Frei, M. R. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith. Statistical and dynamical downscaling of precipitation : An evaluation and comparison of scenarios for the european alps. *Journal of Geophysical Research : Atmospheres*, 112(D4) :D04105, 2007.
 - [37] Claudia Teutschbein and Jan Seibert. Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dynamics*, 37(9–10) :2087–2105, 2011.
 - [38] Lauren E. Hay, Robert L. Wilby, and George H. Leavesley. A comparison of delta change and downscaled gcm scenarios for three mountainous basins in the united states. *Journal of the American Water Resources Association*, 36(2) :387–397, 2000.
 - [39] S. Jang and M. L. Kavvas. Downscaling global climate simulations to regional scales : Statistical downscaling versus dynamical downscaling. *Journal of Hydrologic Engineering*, 20(1) :04014035, 2015.
 - [40] G. Dayon, J. Boé, and E. Martin. Transferability in the future climate of a statistical downscaling method for precipitation in france. *Journal of Geophysical Research : Atmospheres*, 120(3) :1023–1043, 2015.
 - [41] John R. Lanzante, Keith W. Dixon, Mary Jo Nath, Carolyn E. Whitlock, and Dennis Adams-Smith. Some pitfalls in statistical downscaling of future climate. *Bulletin of the American Meteorological Society*, 99(4) :791–803, 2018.
 - [42] Douglas Maraun, Martin Widmann, José M. Gutiérrez, Sven Kotlarski, Richard E. Chandler, Elke Hertig, Joanna Wibig, Radan Huth, and Renate A.I. Wilcke. Value : A framework to validate downscaling approaches for climate change studies. *Earth’s Future*, 3(1) :1–14, 2015.
 - [43] Douglas Maraun. Bias correcting climate change simulations – a critical review. *Current Climate Change Reports*, 2(4) :211–220, 2016.
-

- [44] David Labeurthre, Anatole Reffet, and Anthony Schrapffer. An overview of statistical downscaling methods : Techniques, applications, and advances. SSRN preprint, December 2024.
- [45] Batta Mahesh. Machine learning algorithms – a review. *International Journal of Science and Research (IJSR)*, 9(1), 2019. Technical Report, PDF available.
- [46] Pei Wang. On defining artificial intelligence. *Journal of Artificial General Intelligence*, 10(2), January 2019. Published online 19 August 2019.
- [47] M. I. Jordan and T. M. Mitchell. Machine learning : Trends, perspectives, and prospects. *Science*, 349(6245) :255–260, 2015.
- [48] Rayan Alshamrani and Xiaogang Ma. Deep learning. In *Encyclopedia of Big Data*, pages 1–5. Springer, 2019. First Online : 06 September 2019.
- [49] Peter J. Huber. Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35(1) :73–101, March 1964.
- [50] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna : A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2623–2631. ACM, 2019.
- [51] Anastasia Borovykh, Sander Bohte, and Cornelis W. Oosterlee. Conditional time series forecasting with convolutional neural networks. *arXiv preprint arXiv :1703.04691v5*, September 2018. This version : September 18, 2018, [stat.ML].
- [52] Shaojie Bai, J. Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, pages 899–908, 2018.
- [53] Zheng Zhao, Weihai Chen, Xingming Wu, Peter C. Y. Chen, and Jingmeng Liu. LSTM network : a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 2020.
- [54] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. Empirical evaluation of rectified activations in convolution network. *arXiv preprint arXiv :1505.00853v2*, November 2015. [cs.LG], version 2.
- [55] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). *arXiv preprint arXiv :1511.07289*, November 2015.
- [56] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout : A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15 :1929–1958, 2014.
- [57] Boris N. Oreshkin, Dmitri Carpov, Nicolas Chapados, and Yoshua Bengio. N-beats : Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv :1905.10437*, 2019. [cs.LG], version 4.
- [58] Diederik P. Kingma and Jimmy Ba. Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*, 2014.
- [59] Davide Chicco, Matthijs J. Warrens, and Giuseppe Jurman. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *PeerJ Computer Science*, 7 :e623, 2021.
- [60] Michael H. Kutner, Christopher J. Nachtsheim, John Neter, and William Li. *Applied Linear Statistical Models*. McGraw-Hill/Irwin, Boston, MA, 5th edition, 2004.
- [61] Copernicus Climate Change Service (C3S). Era5 monthly averaged data on pressure levels from 1940 to present, 2019. Last accessed : 2025-07-03.
- [62] Mohamed Safar-Zitoun. Plan national sécheresse algérie : Lignes directrices en vue de son opérationnalisation, 2019. Consultant CNULCD, Alger, 30 juin 2019.