

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Hassiba Benbouali de Chlef

Faculté de génie civil et d'architecture

Département de l'hydraulique



THÈSE

Présentée pour l'obtention du diplôme de

DOCTORAT

Filière : Hydraulique

Spécialité : Ouvrages Hydraulique

Par

SALAH DIFI

Thème :

Amélioration de la prévision des débits des cours d'eau par des modèles hybrides dans le bassin du Cheliff

Soutenue le 22/06/2024, devant le jury composé de :

Karim EZZIANE	Professeur	Hassiba Benbouali, Chlef	Président
Yamina ELMEDDAHI	Professeur	Hassiba Benbouali, Chlef	Rapporteur
Salim HEDDAM	Professeur	20 Aout, Skikda	Co- Rapporteur
Mohamed RIABI	MCA	Hassiba Benbouali, Chlef	Examineur
Hamouda BOUTAGHANE	Professeur	Badji Mokhtar, Annaba	Examineur
Lakhdar DJEMILI	Professeur	Badji Mokhtar, Annaba	Invité

Dédicace

Je souhaite dédier ce travail à des personnes chères et précieuses dans ma vie :

À ma mère bien-aimée, Beya, ainsi qu'à mes frères et sœurs, dont le soutien constant m'a accompagné tout au long de cette thèse.

À mon cher père, Difi Bouraoui, dont la présence et les encouragements ont été une source d'inspiration inépuisable.

À mes amis et à toute ma famille Difi, dont le soutien indéfectible a été une source de réconfort et d'encouragement.

À toutes les personnes qui ont partagé mes espoirs et ont attendu avec impatience l'aboutissement de ce travail, je dédie également ces réalisations.

Difi Salah

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude envers le Dieu tout-puissant pour m'avoir accordé le courage, la patience et la santé nécessaires pour mener à bien ce travail. Sans Son aide et Sa bienveillance, rien de tout cela n'aurait été possible.

Je souhaite adresser mes remerciements les plus sincères à ma directrice de thèse, le Pr. ELMEDDAHI-Yamina, pour sa disponibilité, ses remarques constructives, ses critiques pertinentes qui ont contribué à l'amélioration de la qualité de ce travail, ainsi que pour sa pédagogie remarquable. De même, je tiens à exprimer ma profonde reconnaissance à mon Co-encadrant, le Pr. HEDDAM Salim, pour son encadrement attentif, ses conseils avisés et ses grandes qualités humaines. Je le considère comme l'un de mes mentors et je lui voue une grande estime. Merci infiniment, Professeur.

Je suis également honoré par la présence du Professeur EZZIANE Karim, qui a accepté de présider le jury de cette thèse. Mes remerciements vont également aux membres du jury qui ont accepté d'évaluer ce travail malgré leurs engagements professionnels. Leurs critiques constructives ne peuvent qu'enrichir la qualité de cette recherche. Je tiens à mentionner en particulier le Pr. BOUTAGHANE Hamouda de l'Université d'Annaba et le Dr. RIABI Mohamed de UHB Chlef.

Je saisis cette occasion pour exprimer ma reconnaissance envers les enseignants qui ont contribué à ma formation tout au long de mon parcours universitaire, en particulier le Pr. DJEMILI Lakhdar.

Je souhaite également remercier chaleureusement tous ceux qui ont contribué, de près ou de loin, à la réussite de ce travail, notamment le Dr. ZEROUALI Bilel et le Dr. BOUAMRANE Ali pour leur aide et leur collaboration précieuses, ainsi que mes co-auteurs, en particulier le Pr. OZGUR Kisi, le Dr. HEBAL Aziz, le Pr. VIJAY P. Singh et le Pr. SUNGWON Kim, pour leur collaboration fructueuse.

Enfin, je tiens à adresser une pensée affectueuse à ma famille, en particulier à mes parents et à toute la famille DIFI, pour leur soutien indéfectible, leurs encouragements constants et leur confiance inébranlable

Difi Salah

ملخص

أصبحت الحاجة المتزايدة للتنبؤات الدقيقة وفي الوقت المناسب لتدفق الأودية أمرًا بالغ الأهمية للإدارة الفعالة للموارد المائية والتنبؤ بمخاطر الفيضانات وتقييم تأثيرها على البيئة. تستكشف هذه الأطروحة تطبيق نماذج الذكاء الاصطناعي المتقدمة (AI) في التنبؤ بتدفق الأودية لتحسين دقة وموثوقية طرق التنبؤ الحالية.

في المرحلة الأولى من الدراسة، تم اقتراح نهج مبتكر للتنبؤ بتدفقات الأنهار اليومية والشهرية، بناءً على آلة التعلم القصوى (ELM) وخوارزمية الخفافيش الماورائية (Bat-ELM). تتم مقارنة أداء Bat-ELM بأداء ELM، وانحدار ناقل الدعم (SVR)، وانحدار العملية الغوسية (GPR)، والشبكة العصبية الإدراكية متعددة الطبقات (MLPNN)، والانحدار العام للشبكة العصبية (GRNN). تم تطبيق النماذج المقترحة باستخدام بيانات من ثلاث محطات هيدرومترية في حوض الشلف بالجزائر. توضح النتائج أن Bat-ELM يتفوق على النماذج المستقلة، خاصة في التنبؤ بالتدفقات العالية والمتطرفة في الخطوة اليومية والشهرية، كما أظهر نموذج Bat-ELM أعلى قيم معامل ارتباط بيرسون (R) وكفاءة ناش ساتكليف (NSE)، وأدنى القيم لجذر متوسط مربع الخطأ (RMSE) ومتوسط الخطأ المطلق (MAE).

في المرحلة الثانية من الدراسة، تم استكشاف إمكانية تحسين تنبؤات التدفق اليومي باستخدام نماذج التعلم الآلي مثل MLPNN و GPR و GRNN والذاكرة الطويلة قصيرة المدى (LSTM). يتم استخدام خوارزمية تحلل إشارة تحويل الموجات المنفصلة (MODWT) لتحليل التدفق اليومي (Q) والسلسلة الزمنية لهطول الأمطار (P) إلى عدة تسلسلات فرعية. تعمل هذه التسلسلات الناتجة كمتغيرات مدخلات لنماذج التعلم الآلي الأربعة. يعتمد تطوير النموذج الهجين على بيانات تم قياسها في محطتي قياس، بئر أولاد طاهر وسد بونتيا، الموزعتين في مستجمع مياه الشلف في شمال الجزائر. كشفت النتائج أن النماذج الهجينة المبنية على MODWT-GPR تحقق دقة تنبؤية عالية مع وصول معامل ارتباط بيرسون (R) إلى قيمة 0.99.

الكلمات مفتاحية

تدفق النهر، التنبؤ، النماذج الهجينة، تحويل الموجات، Bat-ELM، MODWT.

Résumé

Le besoin croissant de prévisions précises et opportunes du débit des cours d'eaux est devenu crucial pour une gestion efficace des ressources en eau, la prévision des risques d'inondations et l'évaluation de leur impact sur l'environnement. Cette thèse explore l'application de modèles avancés d'intelligence artificielle (IA) dans la prédiction du débit des cours d'eaux afin d'améliorer la précision et la fiabilité des méthodes de prédiction existantes. Dans la première phase de l'étude, une approche novatrice est proposée pour prédire les débits quotidiens et mensuels des cours d'eaux, basée sur la Machine d'Apprentissage Extrême (ELM) et l'algorithme métaheuristique de la chauve-souris (Bat-ELM). Les performances du Bat-ELM sont comparées à celles de l'ELM, de la Régression à Vecteur de Support (SVR), de la Régression par Processus Gaussien (GPR), du Réseau Neuronal à Perceptron Multicouche (MLPN), et du Réseau Neuronal de Régression Généralisée (GRNN). Les modèles proposés sont appliqués à l'aide de données provenant de trois stations hydrométriques (Arib Abda, El Ababsa et Oued Lili) dans le bassin du Cheliff, en Algérie. Les résultats démontrent que le modèle hybride Bat-ELM surpasse les modèles standards, en particulier dans la prévision des débits élevés et extrêmes au pas de temps journalier et mensuel. Le Bat-ELM a montré les valeurs les plus élevées du coefficient de corrélation de Pearson (R) et de l'efficacité de Nash-Sutcliffe (NSE), et les valeurs les plus faibles de l'erreur quadratique moyenne (RMSE) et de l'erreur absolue moyenne (MAE). Dans la deuxième phase de l'étude, la possibilité d'améliorer la prédiction de débit journalier a été explorée en utilisant des modèles d'apprentissage automatique tels que MLPNN, GPR, GRN et la Mémoire à Long Terme (LSTM). Un algorithme de décomposition du signal par Transformation d'Ondelettes Discrètes à Chevauchement Maximum (MODWT) est utilisé pour décomposer les séries temporelles de débits quotidiens (Q) et de précipitations (P) en plusieurs sous-séquences. Ces séquences obtenues servent de variables d'entrée pour les quatre modèles d'apprentissage automatique. Le développement du modèle hybride s'appuie sur des données mesurées à deux stations hydrométriques, Bir Ouled Taher et Ponteba Barrage, répartie dans le bassin de Cheliff. Les résultats révèlent que les modèles hybrides basés sur MODWT-GPR atteignant une précision prédictive élevée avec le coefficient de corrélation de Pearson (R) atteignant une valeur de 0,99.

Mots clés

Débit des cours d'eau, prédiction, modèles hybrides, transformée en ondelettes, BAT-ELM, MODWT

Abstract

The growing need for accurate and timely forecasts of river flow has become crucial for effective water resource management, flood risks prediction, and the assessment of their impact on the environment. This thesis explores the application of advanced artificial intelligence (AI) models in river flow prediction to enhance the accuracy and reliability of existing prediction methods. In the first phase of the study, an innovative approach is proposed to predict daily and monthly river flows, based on Extreme Learning Machine (ELM) and the bat algorithm metaheuristic (Bat-ELM). The performance of Bat-ELM is compared with that of ELM, Support Vector Regression (SVR), Gaussian Process Regression (GPR), Multilayer Perceptron Neural Network (MLPNN), and Generalized Regression Neural Network (GRNN). The proposed models are applied using data from three hydrometric stations (Arib Abda, El Ababsa, and Oued Lili) in the Cheliff basin, Algeria. The results demonstrate that the hybrid Bat-ELM model outperforms standard models, particularly in predicting high and extreme flows on daily and monthly scales, The Bat-ELM showed the highest values of the Pearson correlation coefficient (R) and Nash-Sutcliffe Efficiency (NSE), and the lowest values of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). In the second phase of the study, the possibility of improving daily flow predictions was explored using machine learning models such as MLPNN, GPR, GRNN, and Long Short-Term Memory (LSTM). A signal decomposition algorithm using Maximum Overlap Discrete Wavelet Transform (MODWT) is used to decompose daily flow (Q) and precipitation (P) time series into multiple sub-sequences. These obtained sequences serve as input variables for the four machine learning models. Model development is based on data measured at two-gauge stations, Bir Ouled Taher and Ponteba Barrage, distributed in the Cheliff watershed in northern Algeria. The results reveal that the hybrid models based on MODWT-GPR achieve high predictive accuracy with the Pearson correlation coefficient (R) reaching a value of 0.99.

Keywords

River flow, prediction, hybrid models, wavelet transform, BAT-ELM, MODWT.

Table des matières

Dédicaces	I
Remerciement.....	II
ملخص	III
Résumé.....	IV
Abstract	V
Sommaire	VI
Liste des figures	X
Liste des tableaux	XV
Liste des abréviations	XV III
Introduction générale	2

CHAPITRE I : Synthèse bibliographique

Introduction.....	7
I.1 Cycle hydrologique	8
I.2 Modélisation hydrologique.....	9
I.2.1 Définition du terme modèle	9
I.2.2 Qu'est-ce que la modélisation hydrologique ?	9
I.3 Objet et enjeux de la modélisation pluie-débit.....	11
I.4 Prévision hydrologique.....	12
I.4.1 Prévision des débits	12
I.4.2 Calibration d'un modèle de prévision.....	13
I.4.3 Validation d'un modèle de prévision	13
I.5 Classification des modèles hydrologiques	14
I.5.1 Modèles stochastiques	14
I.5.2 Modèles déterministes	14
I.5.3 Modèles à base physique	15
I.5.4 Modèles empiriques.....	15
I.5.5 Modèles analytiques	15
I.5.6 Modèles conceptuels.....	16
I.5.7 Modèles basés sur l'intelligence artificielle	16
I.6 Critères de performance du Modèle.....	17

I.6.1	Critère de Nash–Sutcliffe (NSE).....	18
I.6.2	Coefficient de corrélation (R)	18
I.6.3	La racine de l’erreur quadratique moyenne (RMSE)	18
I.6.4	L’erreur moyenne absolue (MAE)	19
	Conclusion.....	19

CHAPITRE II Etat de l'art

Introduction.....	21
II.1 Modèles à base d'intelligence artificielle pour la prévision du débit des cours d'eau	21
II.1.1 Le réseau neuronal à perceptron multicouche (MLPNN).....	22
II.1.2 Le réseau neuronal de régression généralisée (GRNN).....	27
II.1.3 Le machine à vecteur de support (SVM)	30
II.1.4 Régression par processus gaussien (GPR)	33
II.1.5 Machine d'apprentissage extrême (ELM)	36
II.1.6 Réseau récurrent à mémoire court et long terme (LSTM).....	38
Conclusion.....	41

CHAPITRE III Modèles d'intelligence artificielle

Introduction.....	43
III.1 Les Réseaux de Neurones Artificiels	43
III.1.1 Historique	43
III.1.2 Du neurone biologique au neurone artificiel.....	44
III.1.3 Fonction d’activation.....	46
III.1.4 Choix et structure d’un réseau de neurones artificiel	47
III.1.5 Apprentissage des Réseaux de Neurones	50
III.1.6 Algorithme de Rétro-Propagation.....	51
III.1.7 Conception d'un Réseau de Neurones Artificiels	51
III.2 Les perceptrons Multicouches MLPNN	52
III.3 Réseaux de neurones à fonction de régression générale (GRNN).....	54
III.4 Régression par processus gaussien (GPR).....	56
III.5 Machine d'apprentissage extrême (ELM)	57
III.6 Un réseau à mémoire à long terme (LSTM)	58
III.7 Optimisation des modèles ELM basés sur des algorithmes métaheuristiques avancés	61
III.7.1 Algorithme d’optimisateur des chauves-souris (Bat).....	61

III.7.2	Les caractéristiques d'écholocation des micro-chauves-souris.....	62
III.7.3	Mouvement des chauves-souris virtuelles.....	62
III.7.4	Pseudo-code de l'algorithme Bat.....	63
III.7.5	Efficacité de l'algorithme Bat	64
III.7.6	Les Machines A Vecteurs De Support (SVM).....	68
III.8	Transformation en ondelettes.....	70
III.8.1	Types de transformation en ondelettes	72
III.8.2	Transformée en ondelettes discrète à maximum de recouvrement	73
Conclusion.....		78

CHAPITRE IV Présentation de la zone d'étude

Introduction.....		80
IV.1	Présentation du bassin de Cheliff.....	80
IV.1.1	Critères de choix de zone d'étude	81
IV.1.2	Etude géomorphologique.....	88
IV.1.3	Etude du réseau hydrographique.....	99
Conclusion.....		102

CHAPITRE V Traitement des données utilisées et sélection des variables d'entrée des modèles

Introduction.....		104
V.1	Données en usage	104
V.1.1	Collecte des données	105
V.2	Analyse des écoulements.....	105
V.2.1	Variation des débits moyens journaliers.....	105
V.2.2	Variation mensuelle de l'écoulement	110
V.2.3	Variation interannuelle de l'écoulement.....	114
V.3	Analyse des données pluviométriques	117
V.3.1	Problème des lacunes.....	117
V.4	Séries chronologiques et leur utilisation	118
V.5	Différentes phases de la modélisation.....	119
V.6	Sélection des variables d'entrée pour les modèles.....	127
V.6.1	Fonction d'autocorrélation (ACF) et les d'autocorrélation partiels (PACF) et fonction de corrélation croisée (XCF)	127
Conclusion.....		133

CHAPITRE VI Prédiction des débits journaliers et mensuels par des techniques d'apprentissage automatique

Introduction.....	135
VI.1 La mise en œuvre des modèles d'apprentissage automatique proposée dans l'étude	135
VI.1.1 Acquisition des données pour les modèles.....	135
VI.1.2 Sélection des variables d'entrée par les méthodes ACF, PAC et XCF	137
VI.1.3 Critères de validation et de performances des modèles d'apprentissage automatique.....	138
VI.1.4 Développement des modèles d'apprentissage automatique	138
VI.2 Résultats de la modélisation et analyse comparative	140
VI.2.1 Résultats de la prédiction dans le pas journalier et évaluation des modèles	141
VI.2.2 Résultats de la prédiction à l'échelle mensuelle et évaluation des modèles.....	151
VI.2.3 Evaluations de performance des modèles de prédiction des débits dans le pas journalier et mensuelle par des critères des performances graphiques	159
VI.3 Discussion des résultats	165
Conclusion.....	169

CHAPITRE VII Amélioration de la prédiction des débits dans le pas journalier par des modèles hybrides ondelettes-intelligence artificielle

Introduction.....	173
VII.1 La mise en œuvre des modèles d'apprentissage automatique proposée dans l'étude	172
VII.1.1 Acquisition des données pour les modèles et sélection des variables d'entrée par les méthodes ACF, PAC et XCF.....	172
VII.1.2 Développement des modèles d'apprentissage automatique	175
VII.2 Résultats de la modélisation et analyse comparative	178
VII.2.1 Résultats de prédiction des débits journalières dans les deux stations d'étude.....	178
VII.3 Discussion des résultats	211
Conclusion.....	215
Conclusion générale	218
Références bibliographies	222

Liste des figures

Figure I.1 Processus de base du cycle hydrologique (Biancamaria, 2020).....	9
Figure I.2 Modules possibles d'un modèle hydrologique (Hingray et al., 2009).	11
Figure III.1 Représentation schématique d'un neurone biologique (Chekroun, 2009)	45
Figure III.2 Représentation schématique d'un neurone artificiel (Laqrichi, 2015).....	46
Figure III.3 Réseau de neurones non boucles (Salhi, 2015)	49
Figure III.4 Réseau de neurone bouclé (Chiang et al., 2004)	49
Figure III.5 Organigramme de conception d'un réseau de Neurones (Rafik, 2020)	52
Figure III.6 Structure du perceptron Multicouches (MLPNN).....	53
Figure III.7 Architecture d'un réseau de neurones GRNN	54
Figure III.8 Le modèle de régression par processus gaussien (GPR)	56
Figure III.9 Structure du modèle d'une machine d'apprentissage extrême (ELM)	58
Figure III.10 la structure de la modélisation LSTM (Bakhshi Ostadkalayeh et al., 2023).....	60
Figure III.11 Exemple d'écholocalisation (Charbonneau & Tremblay, 2010)	62
Figure III.12 Pseudo code de l'algorithme des chauves-souris (Yang, 2010)	64
Figure III.13 Organigramme du modèles hybride Bat-ELM proposé	67
Figure III.14 Un exemple de classification par SVM (Bouaziz, 2017)	69
Figure III.15 Structure d'un modèle de SVR.	70
Figure III.16 Organigramme de la structure de modélisation proposée pour la prévision des débits journaliers.	75
Figure IV.1 Situation géographique de Bassin de Cheliff au nord de l'Algérie.	81
Figure IV.2 Carte des sous bassins du Bas Cheliff et de la Mina.....	82
Figure IV.3 Situation géographique de sous bassin d'Oued Tiguiguest.	83
Figure IV.4 Carte des sous bassins du Haut et du Moyen Cheliff.....	84

Figure IV.5 Situation géographique de sous bassin d'Oued Ebda	85
Figure IV.6 Situation géographique de sous bassin d'Oued Rouina Zeddine	86
Figure IV.7 Situation géographique de sous bassin d'Oued Ras Ouahrane	87
Figure IV.8 Situation géographique de sous bassin d'Oued Cheliff Harreza.....	88
Figure IV.9 Courbe hypsométrique du bassin versant de Oued Rac Ouahrane	92
Figure IV.10 Courbe hypsométrique du bassin versant d'Oued Rouina Zeddine	93
Figure IV.11 Courbe hypsométrique du bassin versant d'Oued Ebda.....	94
Figure IV.12 Courbe hypsométrique du bassin versant d'Oued Cheliff Harreza.....	95
Figure IV.13 Courbe hypsométrique du bassin versant d'Oued Tiguiquest.....	96
Figure V.1 Variations des débits journaliers pour les stations hydrométriques : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda.....	108
Figure V.2 Variations des débits journaliers pour les stations hydrométriques : (d) Bir Ouled Tahar, et (e) Ponteba Barrage	109
Figure V.3 Variation des débits mensuels : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda	111
Figure V.4 Représentation graphique des coefficients d'hydraulicité mensuels : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda.....	114
Figure V.5 Représentation graphique des coefficients d'hydraulicité annuels : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda.....	117
Figure V.6 Localisation des stations pluviométriques et hydrométriques dans le bassin de Cheliff	118
Figure V.7 Séparation de la série des débits journalières des stations (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda	120
Figure V.8 Séparation de la série des débits journalières des stations (d) Bir Ouled Tahar et (e) Ponteba Barrage	121
Figure V.9 Séparation de la série des débits mensuels des stations (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda	122
Figure V.10 Autocorrélation (ACF), fonction d'autocorrélation partielle (PACF) et fonction de corrélation croisée (XCF) des séries de débit journalier pour les stations (a) Oued Lili, (b) El Ababsa et (c) Arib Abda	129

Figure V.11 Autocorrélation (ACF), fonction d'autocorrélation partielle (PACF) et fonction de corrélation croisée (XCF) des séries de débit mensuelles pour les stations (a) Oued Lili, (b) El Ababsa et (c) Arib abda.....	130
Figure V.12 Autocorrélation (ACF), fonction d'autocorrélation partielle (PACF) et fonction de corrélation croisée (XCF) des séries de débit mensuelles pour les stations (d) Bir Ouled Tahar et (e) Ponteba Barrage.....	132
Figure VI.1 Localisation des stations pluviométriques et hydrométriques dans le bassin de Cheliff.....	137
Figure VI.2 Organigramme des modèles hybrides BAT-ELM proposés	140
Figure VI.3 Histogramme de performance des algorithmes dans les six scénarios proposés pour la station d'Arib Abda au pas de temps journaliers (phase de validation)	141
Figure VI.4 Histogramme de performance des algorithmes dans les six scénarios proposés pour la station d'El Ababsa au pas du journaliers (phase de validation)	144
Figure VI.5 Histogramme de performance des algorithmes dans les six scénarios proposés pour la station d'Oued Lili au pas journaliers (phase de validation).....	147
Figure VI.6 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Arib Abda pendant la phase de test (pas de temps journalier)	149
Figure VI.7 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'El Ababsa pendant la phase de test (pas de temps journalier)	150
Figure VI.8 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Oued Lili pendant la phase de test (pas de temps journalier)	150
Figure VI.9 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Arib Abda pendant la phase de test (pas de temps mensuelle)	158
Figure VI.10 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'El Ababsa pendant la phase de test (pas de temps mensuelle).....	158
Figure VI.11 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Oued Lili pendant la phase de test (pas de temps mensuelle).....	159
Figure VI.12 Diagramme en boîte (Box plot) des débits prévus par les meilleurs modèles par rapport aux débits observées pendant la phase de validation dans les trois stations d'étude :(a) Pas journalier, (b) Pas mensuelle	160

Figure VI.13 Diagramme en violon (Violin plot) des débits prévus par les meilleurs modèles par rapport aux débits observés pendant la phase de validation dans les trois stations d'étude :(a) Pas journalier, (b) Pas mensuelle 161

Figure VI.14 Diagramme de Taylor des débits prévues par les meilleurs modèles par rapport aux débits observés pendant la phase de validation dans les trois stations d'étude :(a) Pas journalier, (b) Pas mensuelle 162

Figure VI.15 Diagramme radar (Radar chart) de meilleurs modèles pendant la phase de validation dans les trois stations d'étude :(a) Pas journalier, (b) Pas mensuelle 163

Figure VI.16 Variation temporelle des débits observés et calculé dans la phase de validation par le modèle (BAT-ELM) pour les trois stations d'étude : (a) Pas journalier, (b) Pas mensuelle..... 164

Figure VII.1 Localisation des stations pluviométriques et hydrométriques dans le bassin de Cheliff..... 173

Figure VII.2 Exemple de décomposition en ondelettes MODWT (sym4) de la série de débit d'ordre 7 pour le pas journalier dans la station de Bir Ould Taher 177

Figure VII.3 Organigramme des modèles hybrides ML-MODWT proposés 178

Figure VII.4 Graphique à barres des statistiques des critères de performance pour les meilleurs modèles ML autonomes pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation) .180

Figure VII.5 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GPR par rapport au modèle GPR optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation) 183

Figure VII.6 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-LSTM par rapport au modèle LSTM optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation) 185

Figure VII.7 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GRNN par rapport au modèle GRNN optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation) 187

Figure VII.8 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-MLPNN par rapport au modèle MLPNN optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation) 189

Figure VII.9 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles autonomes dans la station de Bir Ouled Tahar pendant la phase de validation 191

Figure VII.10 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles hybrides dans la station de Bir Ouled Tahar pendant la phase de validation	192
Figure VII.11 Résultats récapitulatifs des meilleurs modèles dans la station de Bir Ouled Tahar pendant la phase de validation:(a) Diagramme en boîte, (b) Diagramme en violon, (c) Diagramme de Taylor et (c) Diagramme radar	195
Figure VII.12 Variation temporelle des débits mesuré et calculé dans la phase de validation par le modèle MODWT-GPR3(sym4) pour la station de Bir Ouled Tahar	196
Figure VII.13 Graphique à barres des statistiques des critères de performance pour les meilleurs modèles ML autonomes pour la station de Ponteba Barrage (a : Apprentissage, b : Validation).197	
Figure VII.14 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GPR par rapport au modèle GPR optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation)	199
Figure VII.15 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-LSTM par rapport au modèle LSTM optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation).....	201
Figure VII.16 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GRNN par rapport au modèle GRNN optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation).....	203
Figure VII.17 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-MLPNN par rapport au modèle MLPNN optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation).....	205
Figure VII.18 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles autonomes dans la station de Ponteba Barrage pendant la phase de validation	207
Figure VII.19 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles hybrides dans la station de Ponteba Barrage pendant la phase de validation	208
Figure VII.20 Résultats récapitulatifs des meilleurs modèles dans la station de Ponteba Barrage pendant la phase de validation:(a) Diagramme en boîte, (b) Diagramme en violon, (c) Diagramme de Taylor et (c) Diagramme radar	210
Figure VII.21 Variation temporelle des débits mesuré et calculé dans la phase de validation par le modèle MODWT-GPR3(sym4) pour la station de Ponteba Barrage	211

Liste des tableaux

Tableau IV.1 Indices de forme des sous-bassins versants du secteur d'étude	90
Tableau IV.2 Répartition hypsométrique du bassin versant d'Oued Rac Ouahrane	91
Tableau IV.3 Répartition hypsométrique du bassin versant d'Oued Rouina Zeddine	92
Tableau IV.4 Répartition hypsométrique du bassin versant d'Oued Ebda.....	93
Tableau IV.5 Répartition hypsométrique du bassin versant d'Oued Cheliff Harrezal	94
Tableau IV.6 Répartition hypsométrique du bassin versant d'Oued Tiguiguest	95
Tableau IV.7 Altitudes caractéristiques des bassins versants étudiés.....	97
Tableau IV.8 Classification du relief selon l'indice de pente global et la dénivelée spécifique (Selon ORSTOM).....	98
Tableau IV.9 Indices de forme des sous bassins versants étudiés	99
Tableau IV.10 Densité de drainage des sous bassins versants étudiés	100
Tableau IV.11 Coefficient de torrentialité, Temps de concentration et Vitesse de concentration des sous bassins versants étudiés	102
Tableau V.1 Caractéristiques journaliers des stations hydrométriques utilisées	106
Tableau V.2 Caractéristiques mensuelles des stations hydrométriques utilisées.....	112
Tableau V.3 Variation mensuelle de débits et de coefficients d'hydraulicité	112
Tableau V.4 Variation interannuelle de débits et de coefficients d'hydraulicité	115
Tableau V.5 Caractéristiques statistiques des parties : Apprentissage, test et totale des données pour les stations Oued Lili, El Ababsa, et Arib Abda à l'échelle journalière	124
Tableau V.6 Caractéristiques statistiques des parties : Apprentissage, test et totale des données pour les stations Bir Ouled Tahar et Ponteba Barrage à l'échelle journalière.....	125
Tableau V.7 Caractéristiques statistiques des parties : Apprentissage, test et totale des données pour les stations d'étude à l'échelle mensuelle.....	126
Tableau V.8 Combinaisons des variables d'entrées de différents modèles pour la station d'Oued Lili, El Ababsa et Arib abda dans le pas journalier et mensuelle	131

Tableau V.9 Combinaisons des variables d'entrées de différents modèles pour la station de Bir Ouled Tahar et Ponteba Barrage.....	133
Tableau VI.1 Combinaisons des variables d'entrées de différents modèles pour la station d'Oued Lili, El Ababsa et Arib abda dans le pas journalier et mensuelle	137
Tableau VI.2 Résultats de prévision des différents modèles dans le pas journalier pour la station d'Arib Abda.....	143
Tableau VI.3 Résultats de prévision des différents modèles dans le pas journalier pour la station d'El Ababsa	146
Tableau VI.4 Résultats de prévision des différents modèles dans le pas journalier pour la station d'Oued Lili	148
Tableau VI.5 Résultats de prévision des différents modèles dans le pas mensuelle pour la station d'Arib Abda	152
Tableau VI.6 Résultats de prévision des différents modèles dans le pas mensuelle pour la station El Ababsa.....	154
Tableau VI.7 Résultats de prévision des différents modèles dans le pas mensuelle pour la station d'Oued Lili	156
Tableau VII.1 Combinaisons des variables d'entrées de différents modèles pour la station de Ponteba et Bir Ouled Tahar dans le pas journalier	175
Tableau VII.2 Résultats de prévision des différents modèles ML autonomes dans le pas journalier pour la station de Bir Ouled Tahar	181
Tableau VII.3 Résultats de prévision des modèles hybrides GPR-MODWT dans le pas journalier pour la station de Bir Ouled Tahar	184
Tableau VII.4 Résultats de prévision des modèles hybrides LSTM-MODWT dans le pas journalier pour la station de Bir Ouled Tahar	186
Tableau VII.5 Résultats de prévision des modèles hybrides GRNN-MODWT dans le pas journalier pour la station de Bir Ouled Tahar	188
Tableau VII.6 Résultats de prévision des modèles hybrides MODWT-MLPNN dans le pas journalier pour la station de Bir Ouled Tahar	190
Tableau VII.7 Résultats de prévision des différents modèles dans le pas journalier pour la station de Ponteba Barrage.....	198

Tableau VII.8 Résultats de prévision des modèles hybrides GPR-MODWT dans le pas journalier pour la station de Ponteba Barrage200

Tableau VII.9 Résultats de prévision des modèles hybrides LSTM-MODWT dans le pas journalier pour la station de Ponteba Barrage202

Tableau VII.10 Résultats de prévision des modèles hybrides GRNN-MODWT dans le pas journalier pour la station de Ponteba Barrage204

Tableau VII.11 Résultats de prévision des modèles hybrides MODWT-MLPNN dans le pas journalier pour la station de Ponteba Barrage206

Liste des abréviations

- AAC** : Artificial Ant Colony : Colonie de fourmis artificielles
- ACF** : Autocorrelation Function : Fonction d'autocorrélation
- PCA** : Principal Component Analysis : Analyse en composantes principales
- AI** : Artificial Intelligence : Intelligence artificielle
- ANFIS** : Adaptive Neuro-Fuzzy Inference System : Système adaptatif d'inférence neurofloue
- ANN** : Artificial Neural Networks : Réseaux de neurones artificiels
- ANRH** : Agence Nationale des Ressources Hydraulique
- ARIMA** : Autoregressive Integrated Moving Average models : Modèles autorégressifs intégrés et de moyenne mobile
- ARMA** : Auto Regressive Moving Average models : Modèles autorégressifs et de moyenne mobile
- BA** : Bat Algorithm : algorithme de chauve-souris
- BP** : Back Propagation : Rétro-propagation
- BR** : Bayesian Regularization : Régularisation bayésienne
- CWT** : Continuous Wavelet Transform : Transformée continue en ondelettes
- dbN** : Ondelettes Daubechies
- DWT** : Discrete Wavelet Transform : Ondelettes de type transformation discrète
- ELM** : Extreme Learning Machine : Machine d'apprentissage extrême
- EMD** : Empirical Mode Decomposition : Décomposition en mode empirique
- FFA** : Fruit Fly Algorithms : Algorithme de la mouche des fruits
- FFBP** : Feed Forward Back Propagation : algorithme de rétro-propagation du gradient
- GA** : Genetic Algorithm : Algorithme génétique
- GEP** : Gene Expression Programming : programmation d'expression génique
- GP** : Genetic Programming : Programmation génétique
- GR4J** : Modèle conceptuel Génie rural à 4 paramètres journalier
- GRNN** : Generalized Regression Neural Networks : Réseaux de neurones de régression généralisée
- GSA** : Gravitational Search Algorithm : Algorithme de recherche gravitationnelle
- GWO** : Grey Wolf Optimizer : Algorithme d'optimisation du loup gris
- LSSVM** : Least-Squares Support Vector Machines : Machines à vecteurs de support en moindres carrés
- M5Tree** : Model Tree : Modèles d'arbres

MAE : Mean Absolute Error : Erreur moyenne absolue

MARS : Multivariate Adaptive Regression Splines : spline de régression adaptative multivariée

MLP : Multi-Layer Perceptron : Perceptron multicouche

MLR : multiple linear regression : Régression linéaire multiple

NSE : Nash Sutcliffe criterion : Critère de Nash Sutcliffe

ONM : Office Nationale de la Météorologie

PACF : Partial Autocorrelation Function : Fonction d'autocorrélation partielle

PSO : Particle swarm optimization : Optimisation par essais des particules

Q : Débits

QPSO : Quantum Particle Swarm Optimization : Optimisation quantique par essais de particules

R : Coefficient de détermination

R² : Coefficient de corrélation

RBF : Radial Basis Function : Fonction de base radiale

RF : Random Forest : Forêt aléatoire

RMR : Régression Multiple Récursive

RMSE : Root Mean Square Error : Erreur moyenne quadratique

SFLA : Shuffled Frog-Leaping Algorithm : Algorithme du saut de grenouille traînante

SOM : Self organisation maps : Cartes auto-organisées

SHA : Spotted Hyena Algorithm : Hyène tachetée algorithme

SVM : Support Vector Machine : Machines à vecteurs de support

SVR : Support Vector Regression : Régression à vecteurs de support

WANN : Wavelet-Artificial Neural Networks models : Modèles ondelettes-réseaux de neurones artificiels

WT : wavelet Transform : transformation en ondelettes

WSVR : Wavelet Support Vector Regression models : Modèles Ondelettes- régression à vecteurs de support

INTRODUCTION GENERALE

Introduction Générale

L'eau, en tant que ressource fondamentale, sous-tend la création de toutes nos productions. Sans substitut possible, sa quantité, bien que renouvelable, est limitée. Bien que la Terre soit recouverte à hauteur de trois quarts d'eau, la part d'eau douce demeure relativement insignifiante (Westall and Brack 2018). Sur un total mondial d'environ 1.4 million de km³, 97.5 % est de l'eau salée dans les océans et les mers, tandis que seulement 2.5 % est de l'eau douce dans les rivières et les lacs (Magand 2014). De plus, une partie des ressources en eau douce se trouve aux pôles et sous terre, ce qui limite considérablement la quantité d'eau utilisable. Avec la croissance démographique, le développement économique et le changement climatique, la pression sur les ressources en eau douce et la compétition pour l'accès à ces ressources ne cessent d'augmenter (Shen et al. 2008). Cette situation laisse présager une crise mondiale de l'eau dans un avenir proche, positionnant ainsi la gestion des ressources en eau comme une priorité majeure sur l'agenda politique du 21^{ème} siècle pour les gouvernements et les institutions (Zerouali 2019).

En Algérie, la demande croissante en eau et le développement durable des ressources en eau constituent un défi de taille, comme l'indique un rapport de 2004 de la Commission nationale de l'eau (Messahel and Benhafid 2007). La Commission de l'Aménagement du Territoire et de l'Environnement du Conseil National Économique et Social estime qu'entre 15 et 20 milliards de mètres cubes d'eau par an seront nécessaires pour atteindre une sécurité alimentaire satisfaisante, dont 70% seront alloués à l'agriculture. Cette estimation, comparée à la capacité totale actuelle des 75 barrages algériens, estimée à 6.50 milliards de mètres cubes, pose un défi considérable (Masmoudi 2009).

Ainsi, une planification précise s'avère essentielle pour protéger, développer et utiliser les ressources en eau. L'étape cruciale de cette planification consiste à évaluer le potentiel actuel et futur de la ressource en eau (Zhang et al. 2024). De plus, une prévision précise et fiable du débit des cours d'eau, tant journaliers que mensuels, est nécessaire pour déterminer le potentiel de la source en eau en vue d'une gestion rationnelle, de la distribution des réservoirs, et de l'exploitation des centrales hydroélectriques, etc. Actuellement, un large éventail de méthodes de prévision a été développé, offrant des niveaux de précision variables (Dong et al. 2024 ; Gomaa et al. 2023). Ces méthodes vont des approches simples basées sur des équations linéaires à des modèles plus

complexes exploitant des techniques avancées telles que l'intelligence artificielle ([Dai et al. 2024](#) ; [Iamampai et al. 2024](#) ; [Elbeltagi et al. 2022](#) ; [Adamowski et al. 2012](#)).

L'estimation précise du débit d'un bassin versant est difficile en raison de la distribution chaotique, de la périodicité et des relations complexes et non linéaires entre les éléments du bassin versant ([Zerouali et al. 2022](#)). Les précipitations (durée et intensité), l'interception et le ruissellement, les caractéristiques du sol (humidité antérieure, type, épaisseur, texture et compacité), les caractéristiques du bassin versant (topographie et forme), l'utilisation des sols et la couverture végétale, ainsi que l'évapotranspiration, sont les principaux facteurs influençant le débit ([Bhasme and Bhatia 2024](#)). Trouver des algorithmes appropriés et robustes capables de prédire avec précision le débit des cours d'eau représente un défi majeur pour l'hydrologie des bassins versants ([Zuo et al. 2024](#)). En plus du manque de données sur le débit des cours d'eau dans les régions arides, les données météorologiques et autres données pertinentes sur les bassins versants montagneux, ainsi que sur pratiquement tous les bassins versants des pays en développement (tels que l'Algérie), sont également rares ([Aoulmi et al. 2021](#)). La plupart des bassins versants du nord de l'Algérie ont récemment été confrontés à de nombreuses inondations graves, ce qui souligne l'utilité de la modélisation du débit pour aider les décideurs à élaborer des plans de gestion à long terme des conditions de précipitations extrêmes intermittentes ([Abda et al. 2022](#)).

Actuellement, l'intelligence artificielle (IA) joue un rôle de plus en plus important dans le domaine de la prédiction hydrologique et de la gestion des ressources en eau ([Han and Morrison 2022](#)). L'intégration de l'IA dans les modèles hydrologiques permet d'améliorer la précision des prévisions en exploitant la capacité des algorithmes d'apprentissage automatique à analyser de vastes ensembles de données complexes et à détecter des schémas non linéaires difficiles à modéliser avec des approches traditionnelles ([Vázquez and Feyen 2010](#)).

Par exemple, les réseaux de neurones artificiels sont utilisés pour modéliser les relations entre les précipitations, l'évapotranspiration, l'occupation des sols et le débit des cours d'eau. Ces modèles peuvent prendre en compte une multitude de facteurs et apprendre à partir des données historiques pour fournir des prévisions précises du débit des cours d'eau, tant à court terme qu'à long terme ([Ghumman et al. 2011](#) ; [Tan et al. 2018](#) ; [Abda et al. 2022](#)). De même, les algorithmes d'apprentissage automatique tels que les machines à vecteurs de support (SVM) et les arbres de décision sont utilisés pour analyser les données hydrologiques et identifier les tendances, les cycles

et les anomalies. Ces techniques peuvent aider à prédire les variations saisonnières et à anticiper les périodes de sécheresse ou d'inondation, permettant ainsi une gestion plus efficace des ressources en eau (Liang et al. 2018). L'intégration de l'IA dans les modèles hydrologiques offre également la possibilité de prendre en compte les changements climatiques et environnementaux. Les modèles peuvent être formés avec des données climatiques historiques et des scénarios de changement climatique projetés pour évaluer l'impact sur les ressources en eau et développer des stratégies d'adaptation appropriées.

Le IA offre de nouvelles opportunités pour améliorer la précision des prévisions hydrologiques, optimiser la gestion des ressources en eau et répondre aux défis posés par les changements climatiques. Son intégration dans les modèles hydrologiques ouvre la voie à une gestion plus durable et efficace des ressources en eau pour les générations actuelles et futures.

La présente thèse aborde divers sujets de recherche liée au développement de modèles à base d'intelligence artificielle pour la prévision des débits journaliers et mensuels. Les objectifs et tâches de recherche abordés dans cette thèse sont les suivants :

1. Examiner la possibilité d'utiliser des modèles d'intelligence artificielle pour améliorer les performances des prévisions en utilisant le débit et les précipitations comme données d'entrée dans différents scénarios.
2. Comparer les performances des modèles standards sans hybridation, tels que les perceptrons multicouches (MLPNN), les réseaux de neurones à fonction de régression générale (GRNN), la régression par processus gaussien (GPR), les machines d'apprentissage extrême (ELM) et les machines à vecteurs de support (SVR), avec le modèle hybride à base des machines d'apprentissage extrême optimisé par l'algorithme Bat (ELM_Bat) dans le processus des prévisions des débits journaliers et mensuels.
3. Tentatives d'amélioration des performances des différents modèles à base d'intelligence artificielles par l'utilisation des techniques de décomposition par ondelettes (MODWT).
4. Évaluer l'impact de la transformation des ondelettes d'entrée sur la performance de différents types de modèles d'intelligence artificielle, y compris les réseaux à mémoire à long court terme (LSTM), les modèles GRNN, les GPR et les MLPNN dans le processus des prévisions des débits journaliers.

5. Identifier la fonction ondelette mère la plus appropriée pour transformer les variables d'entrée afin de les utiliser pour la prévision des débits journaliers.
6. Comprendre l'effet du choix des variables d'entrée sur la qualité des prévisions.
7. Et enfin, formuler des recommandations spécifiques pour la prévision des débits et présenter les modèles les plus fonctionnels pour ces prévisions.

Cette thèse est composée de sept chapitres. L'organisation générale et le contenu de chaque chapitre peuvent être décrits comme suit :

1. Le premier chapitre consiste en une synthèse bibliographique sur la prévision des crues, la modélisation pluie-débit et la classification des différents modèles hydrologiques, ainsi que leurs critères de performance.
2. Le deuxième chapitre introduit l'état de l'art des modèles d'intelligence artificielle, ainsi que les travaux de recherche relatifs à leur utilisation dans la prévision des débits fluviaux.
3. Le troisième chapitre présente une revue exhaustive et l'architecture des modèles d'apprentissage automatique, tels que les MLPNN, les GRNN, les GPR, les LSTM, les ELM, les SVR, le modèle hybride ELM_BAT, ainsi que la transformation en ondelettes MODWT.
4. Le quatrième chapitre décrit les caractéristiques morphologiques et hydrographiques du bassin versant sélectionné.
5. Le cinquième chapitre présente l'analyse, le traitement et les caractéristiques statistiques de l'ensemble des données utilisées dans ce travail. En outre, il utilise les méthodes d'autocorrélation simple, partielle et croisée pour sélectionner les variables d'entrée les plus pertinentes.
6. Le sixième chapitre est consacré à la présentation des résultats de la prévision des débits journaliers et mensuels en utilisant des modèles standards tels que MLPNN, GRNN, SVR, GPR, ELM, ainsi que le modèle hybride ELM_BAT, avec des discussions et une conclusion à la fin du chapitre.

7. Le septième et dernier chapitre expose les résultats obtenus par les modèles hybrides de transformation en ondelettes tels que MODWT-LSTM, MODWT-MLPNN, MODWT-GRNN et MODWT-GPR, et leur capacité à prédire les débits journaliers par rapport aux modèles autonomes, avec des discussions et une conclusion à la fin du chapitre.

Enfin, une conclusion générale qui récapitule les principales démarches et les résultats obtenus dans cette étude.

CHAPITRE I ***Synthèse bibliographique***

Introduction

La gestion des ressources en eau est le processus de planification, de développement, de distribution et de gestion de l'utilisation optimale des ressources en eau. Il s'agit d'un aspect important de la gestion du cycle de l'eau. Dans le domaine de l'hydrologie et de l'ingénierie des ressources en eau, des prévisions fiables et réalistes du débit des cours d'eau jouent un rôle crucial, car elles peuvent avoir une incidence directe sur le fonctionnement et la performance des barrages, la recharge/exploitation des eaux souterraines, la capacité de transport des sédiments des cours d'eau, la gestion des bassins versants, etc. Toutefois, il n'est pas facile de prédire avec précision le débit des cours d'eau en raison de la grande incertitude associée aux conditions climatiques et de la complexité de la collecte et du traitement des données spatiales et non spatiales. C'est pourquoi les hydrologues ont développé et adopté plusieurs types de techniques basées sur les données, allant de la modélisation stochastique traditionnelle des séries temporelles aux modèles hybrides modernes d'intelligence artificielle (IA) pour la prévision future du débit des cours d'eau.

Ce chapitre traite de la modélisation hydrologique et de ses applications. Il est utile de rappeler brièvement le cycle hydrologique, qui est à la base de toute modélisation hydrologique.

I.1 Cycle hydrologique

Le cycle de l'eau est lié à toute l'eau sur terre, c'est-à-dire au changement constant de son état et de sa localisation sous l'influence de son apport énergétique provenant du soleil et de l'effet de la gravité.([Marcinek, 2011](#))

Le cycle hydrologique se compose essentiellement de deux phases, la phase atmosphérique, qui décrit le mouvement de l'eau sous forme de gaz (vaporisations d'eau) et de liquide/solide (précipitations et neige) dans l'atmosphère.([Easton & Bock, 2015](#)) La deuxième phase est la phase terrestre, qui décrit le mouvement de l'eau sur la Terre, à travers la Terre et à l'intérieur de celle-ci. La phase terrestre est souvent décomposée en une phase d'eau de surface (ruissellement, débit des cours d'eau) et une phase d'eau souterraine (infiltration, percolation, recharge des aquifères). La **figure I.1** présente une interprétation graphique de l'interaction entre les différents composants.([Biancamaria, 2020](#)) ([Andersson et al., 2005](#)) ([Gupta, 2016](#))

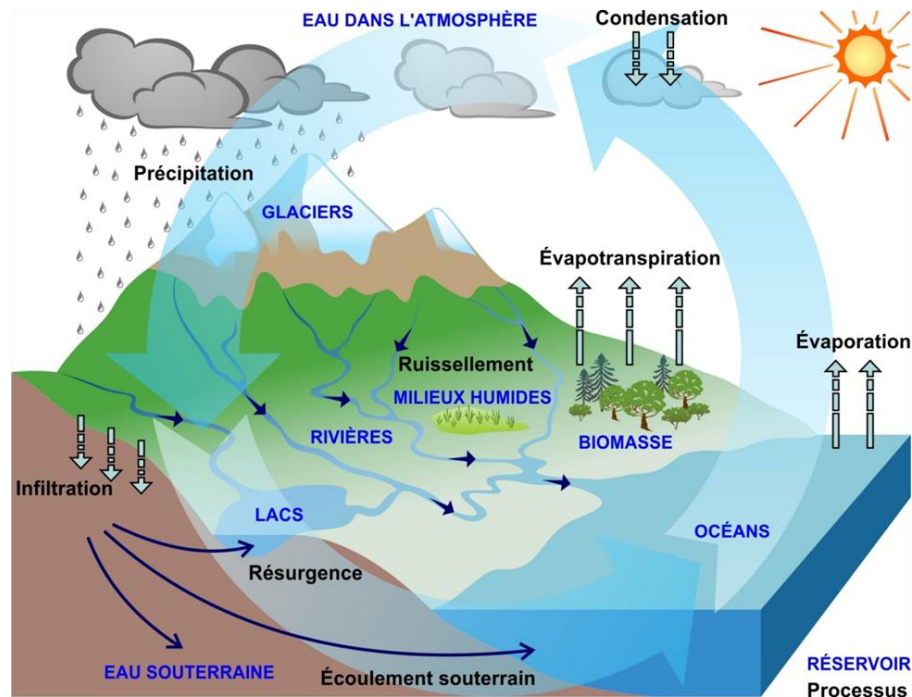


Figure I.1 Processus de base du cycle hydrologique (Biancamaria, 2020)

I.2 Modélisation hydrologique

I.2.1 Définition du terme modèle

Un modèle est une représentation simulée et simplifiée d'un système physique déterminé ainsi que des différents mécanismes qui permettent la compréhension de son comportement. La représentation du système peut être physique et fondée sur un modèle réduit du système. Elle est plus souvent mathématique. Le terme de modèle est fréquemment utilisé à la place de "système de modélisation". (Benoît Hingray et al., 2009).

I.2.2 Qu'est-ce que la modélisation hydrologique ?

La modélisation hydrologique c'est l'élaboration par des méthodes physiques ou mathématiques de modèles de simulation du cycle de l'eau et de son incidence sur un bassin hydrologique. Elle doit pouvoir décrire les différentes étapes de la transformation pluie-débit et en particulier les processus liés à la formation des crues. La modélisation hydrologique est effectuée à l'aide d'un outil appelé : modèle hydrologique (Gnouma, 2006).

Les modèles hydrologiques sont des outils importants pour améliorer la compréhension de la dynamique du cycle de l'eau dans le bassin versant, afin de mieux la gérer ou évaluer le comportement hydrologique.

Les modèles hydrologiques sont des outils mathématiques et informatiques utilisés pour représenter et simuler le comportement du cycle de l'eau dans un bassin versant, une région géographique ou un système hydrologique spécifique. Ces modèles sont conçus pour comprendre et prédire le mouvement de l'eau à travers différentes phases du cycle hydrologique, notamment l'évaporation, la précipitation, l'infiltration, le ruissellement, la transpiration des plantes, la fonte des neiges et le stockage dans les lacs ou les nappes phréatiques.

Ces modèles hydrologiques transforment des séries temporelles correspondant au climat du bassin versant (évapotranspiration, précipitations, etc.) en une série de débits (modèle pluie-débit) ou simulent l'ensemble du bilan hydrique (évapotranspiration, état hydrique du sol, ruissellement de surface, niveau de saturation, écoulement latéral des eaux souterraines, etc.) Cette transformation se fait généralement en deux phases, qui sont décrites de manière plus ou moins détaillée dans les différents modèles :

1. la phase de " Production " : la pluie brute qui tombe sur le bassin versant est divisée en une quantité d'eau stockée par le bassin versant, une quantité qui s'évapore et une quantité qui participe au débit,

2. la phase de " Transfert " : la quantité d'eau qui contribue au débit est transmise ou canalisée depuis son point de production vers l'exutoire du bassin ([Labbas, 2015](#)).

Les modèles hydrologiques sont souvent composés de différents modules interconnectés, chacun étant responsable de la représentation et de la simulation d'un aspect spécifique du cycle de l'eau. Ces modules sont conçus pour travailler de manière intégrée afin de capturer l'ensemble des processus hydrologiques. Ils sont reliés les uns aux autres : les entrées de l'un sont les sorties d'un autre ([figure I.2](#)). Le premier module transforme les données météorologiques pour produire les données de forçage d'entrée aux résolutions spatio-temporelles exigées par le modèle. Un module neige permet de traiter les phénomènes liés à la présence de précipitations solides. Un troisième module simule la distribution du débit d'eau entre le sol, la végétation et l'atmosphère. Ce module est assimilé à la "fonction de production" (entrée pluie brute, sortie pluie nette). L'eau infiltrée dans

le sol peut être gérée dans un autre module ou par le module précédent. Un module de routage transfère les volumes ruisselés et infiltrés vers l'exutoire. Ce module peut également prendre en compte la propagation dans la rivière et est alors appelé "fonction de transfert" (B Hingray et al., 2009).

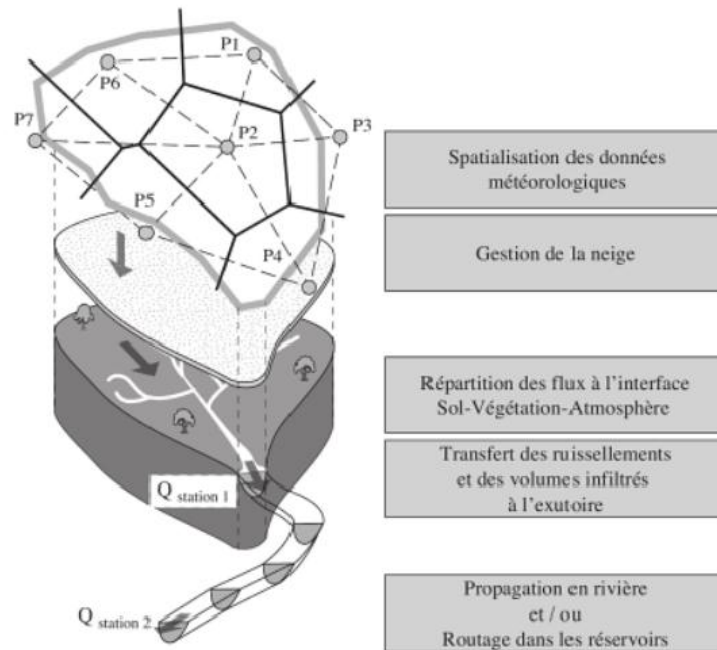


Figure I.2 Modules possibles d'un modèle hydrologique (Hingray et al., 2009).

I.3 Objet et enjeux de la modélisation pluie-débit

La modélisation pluie-débit peut réussir à répondre à de nombreuses problématiques centrées sur l'eau, telles que la gestion des ressources et des risques. (Michel, 1989) et (Refsgaard & Abbott, 1990) énumèrent les principales problématiques, comme suit :

- La simulation des débits, pour comblement des lacunes des ensembles de données, et reconstituer les débits historiques (les données pluviométriques sont souvent disponibles sur des périodes beaucoup plus longues que les débits) ou pour permettre un traitement statistique (Sefton & Howarth, 1998);
- Prédétermination des débits de crue ou d'étiage : le but est de pouvoir connaître la fréquence des débits de crue (au-dessus d'un seuil de risque) ou d'étiage (en dessous d'un débit réservé), et leur durée. Ces informations peuvent permettre de concevoir le

dimensionnement des structures et des réservoirs ou des aménagements dans le lit (mineur à majeur) du cours d'eau (Perrin, 2002);

- Influence des développements sur l'hydrologie : il est souhaitable que le bassin soit en mesure de prévoir les changements de réponse aux changements d'origine humaine dans les caractéristiques du bassin ou aux changements environnementaux.

Ces difficultés soulèvent deux aspects importants, celui de la valorisation des risques et celui de la gestion des ressources. La validité des réponses qui peuvent leur être apportées est conditionnée par la pertinence du modèle dans sa représentation du bassin par rapport aux objectifs visés. Les auteurs qui ont étudié ces problèmes à l'aide de modèles pluie-débit sont nombreux. Citons les travaux de (Cameron et al., 1999) et (Uhlenbrook et al., 1999) sur la prédétermination des crues, ceux de (Xu & Vandewiele, 1995) et (Yang et al., 1995) respectivement sur le dimensionnement et la gestion des réservoirs, ou encore ceux de (Da Ros & Borga, 1997) et (Yang & Michel, 2000) sur la prévision des débits.

I.4 Prévision hydrologique

La modélisation en tant qu'outil de prévision, fournit une anticipation des changements futurs du débit dans un cours d'eau. La prévision des crues consiste à estimer en temps réel le niveau futur du débit, le temps d'apparition et la durée d'une crue, en particulier la valeur de la pointe de crue pour un site spécifique, qui résulte des pluies et/ou des fontes de neige. Elle comporte trois aspects : les valeurs prévues, les objectifs de la prévision et les délais de prévision (Klemes, 1982).

I.4.1 Prévision des débits

La séparation entre la prévision et la prédiction est essentielle. La prévision concerne l'estimation des conditions futures à une date donnée, ou pendant un intervalle de temps spécifique, tandis que la prédiction représente l'estimation des conditions futures, sans référence à un moment spécifique (Lettenmaier & Wood, 1993). Les prédictions ne concernent pas toujours l'avenir, c'est-à-dire qu'il n'est limité par aucun temps, comme le passé ou le présent. Pour une gestion efficace des ressources en eau de surface et pour permettre une réponse adéquate et rapide dans des situations de crise, telles que les inondations, la prévision du débit des cours d'eau est nécessaire à l'échelle horaire, hebdomadaire, journalière et mensuelle. Ces prévisions peuvent être effectuées selon différentes approches.

I.4.2 Calibration d'un modèle de prévision

Pour arriver à une prévision de débit à partir d'observations passées et de prévisions météorologiques, on suppose qu'il existe un modèle f reliant ces observations et prévisions au débit Q au temps τ :

$$Q_\tau = f(Q_t, Q_{t-1}, \dots, Q_{t-r}, \hat{x}_{\tau-1}, \hat{x}_{\tau-2}, \dots, x_t, x_{t-1}, \dots, x_{t-r}, \varepsilon_\tau | \theta) \quad (\text{I.1})$$

ε_τ Correspond à un terme d'erreur tandis que $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ est un ensemble de m coefficients d'ajustement invariants dans le temps. Il est donc possible d'obtenir une prévision \hat{Q}_τ de Q_τ en estimant les coefficients d'ajustement et le terme d'erreur :

$$\hat{Q}_\tau = f(Q_t, Q_{t-1}, \dots, Q_{t-r}, \hat{x}_{\tau-1}, \hat{x}_{\tau-2}, \dots, x_t, x_{t-1}, \dots, x_{t-r}, \hat{\varepsilon}_\tau | \hat{\theta}) \quad (\text{I.2})$$

Dans laquelle $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m)$ et $\hat{\varepsilon}_\tau$ sont respectivement des estimateurs de θ et ε_τ . L'erreur (*i. e.* $\hat{\varepsilon}_\tau = 0$) est supposée être nulle dans la plupart des cas, mais il arrive que son estimation soit intrinsèquement liée à la prévision du débit. Un modèle de prévision doit être ajusté en utilisant les observations passées. Le modèle doit extraire les informations relatives au comportement du système modélisé de cette période d'étalonnage en éliminant le bruit. Si le modèle de prévision est trop flexible, le bruit peut être modélisé comme s'il s'agissait d'un comportement prévisible : sur la période de calibration, le modèle pourrait alors donner un très bon ajustement, puisqu'il reproduirait même les variations aléatoires du débit. En revanche, les prévisions faites par un tel modèle seraient alors, en comparaison, particulièrement mauvaises. Par conséquent, il faut essayer de minimiser le nombre de paramètres d'un modèle de prévision, tout en lui permettant d'être suffisamment flexible pour modéliser adéquatement le système (Shi et al., 2008).

Une approche importante pour vérifier si la modélisation extrait adéquatement les informations disponibles est d'observer l'auto-corrélation des différences entre les débits observés et modélisés. Une forte auto-corrélation indique que ce qui était considéré comme un bruit aléatoire contient des informations qui peuvent être incorporées dans le modèle (Bennis & Bruneau, 1993).

I.4.3 Validation d'un modèle de prévision

Dans le cadre de la construction d'un modèle de prévision, une phase importante consiste à évaluer la précision des prévisions. Une partie des données disponibles doit être conservée pour cette phase de validation. Ces données ne seront pas utilisées dans la phase de calibration. Pendant la

validation, différentes statistiques des prévisions peuvent être calculées, telles que leur variance et leur biais. La période de validation doit être suffisamment longue pour que ces statistiques soient suffisamment précises pour être utiles. Il faut donc déterminer les distributions d'échantillonnage des statistiques utilisées comme critères de qualité d'un modèle de prévision (Mroczkowski et al., 1997).

La robustesse fait référence à la capacité d'un modèle à fonctionner correctement dans différentes conditions et hypothèses en optimisant le nombre de paramètres du modèle en suivant l'évolution de ces paramètres de qualité, toujours calculés pendant la période de validation, en fonction de la complexité du modèle. Il est alors possible de déterminer à quel moment l'augmentation du nombre de paramètres conduit à une détérioration des critères de validation au lieu d'une amélioration des prévisions (Block et al., 2009).

I.5 Classification des modèles hydrologiques

La littérature fournit de nombreux modèles hydrologiques de bassins versants, en fonction de leurs spécificités et de leur domaine d'application. Une fois le modèle choisi, il faut évaluer sa capacité à représenter la réalité. Cela se fait généralement en comparant les résultats du modèle avec les observations. Il s'agit d'une étape très délicate (Berthier, 2005).

Par ailleurs, il existe presque autant de classifications que d'auteurs ayant traité le sujet, il serait donc difficile de classer ici les types de modèles hydrologiques. Rappelons à titre d'exemple les références suivantes : (Singh, 1995), (Singh & Woolhiser, 2002), (Clarke, 1973), (Beven, 1989), (Wurbs, 1998), (Ambroise, 1998), (Andréassian, 2002), (Oudin, 2004), etc.

I.5.1 Modèles stochastiques

Une modélisation stochastique est un modèle probabiliste permettant d'étudier un phénomène aléatoire au cours du temps. Les modèles stochastiques utilisent des variables aléatoires pour représenter l'incertitude du processus et génèrent des résultats différents à partir d'un ensemble de données d'entrée et de valeurs de paramètres lorsqu'ils fonctionnent dans des conditions similaires (Pechlivanidis et al., 2011).

I.5.2 Modèles déterministes

La modélisation est dite déterministe si aucune de ses variables n'est considérée comme aléatoire. Typiquement, ces modèles combinent chaque ensemble de variables de forçage (variables d'entrée

indépendantes du modèle, qui peuvent essentiellement être des mesures de précipitations), de variables d'état (variables utilisées pour caractériser l'état du système modélisé, telles que le niveau de remplissage des différents "réservoirs" d'eau dans le bassin versant, le taux de saturation du sol, la profondeur du sol, etc.) et de paramètres, ainsi qu'une valeur de réalisation unique pour les variables de sortie (essentiellement les débits simulés à l'exutoire du bassin versant) (MAFTEI, 2002).

I.5.3 Modèles à base physique

La fonction de ces modèles consiste à créer une relation purement mathématique basée sur les observations des précipitations et de leur débit dans un bassin hydrographique. Le bassin est traité comme une "boîte noire", sans faire d'hypothèses sur les processus internes qui contrôlent la transformation des précipitations en ruissellement. Les modèles de régression, les méthodes basées sur des fonctions de transfert telles que l'hydrogramme unitaire, les réseaux neuronaux, la programmation génétique, la logique floue, etc. entrent dans cette catégorie. Bien qu'assez simples à mettre en œuvre, ces modèles peuvent être utilisés pour la prévision des inondations. Cependant, ces modèles ne prédisent que le débit à l'exutoire du bassin versant. Ils nécessitent un grand nombre de données pour être calibrés et leurs paramètres sont difficiles à relier aux caractéristiques physiques du bassin. Par conséquent, leur pouvoir explicatif est très limité et ils ne nous aident pas à comprendre le comportement d'un bassin versant (McIntyre et al., 2007).

I.5.4 Modèles empiriques

Les modèles empiriques découlent des rapports observés entre les entrées et les sorties du système hydraulique considéré. A l'aide d'un ensemble d'équations développées et ajustées sur la base des données obtenues sur le système, ils expriment la relation entre les variables d'entrée et de sortie du système (relation pluie-débit). Cependant, un modèle empirique ne cherche pas à décrire les causes d'un phénomène hydrologique donné, ni à expliquer le fonctionnement du système, ce dernier est considéré comme une boîte noire (Perrin, 2002).

I.5.5 Modèles analytiques

Ce type de modèle implique des modèles dans lesquels les relations entre les variables de sortie et d'entrée ont été établies en analysant des séries de données mesurées. La pratique courante

Exemples incluent les modèles linéaires : les paramètres de ces modèles sont liés aux coefficients de corrélation entre les variables. Cependant, l'analyse des données peut conduire au choix de relations non linéaires entre les variables (Nascimento, 1995).

1.5.6 Modèles conceptuels

Généralement, les modèles conceptuels (boîte grise) fonctionnent sans aucun type de lois hydrologiques. C'est probablement la raison pour laquelle ils sont appelés conceptuels. Idéalement, le modèle conceptuel cherche à intégrer la complexité des processus en essayant de reproduire le concept physique du comportement du système d'une manière plus simple. Ambroise (1998) a considéré le modèle de bassin versant comme un assemblage de réservoirs d'humidité interconnectés (également connu sous le nom de modèle de réservoir), destiné à représenter plusieurs niveaux de stockage le long d'une dimension verticale. Bien que ce type de modèle ait un sens physique, il ne fournit pas une description exacte de tous les processus impliqués. Pour les applications techniques (dimensionnement des ouvrages, prévision des crues et des étiages, gestion des réservoirs) sur les bassins jaugés et non jaugés, le GR4J est un exemple de modèle conceptuel (Mouelhi, 2003).

1.5.7 Modèles basés sur l'intelligence artificielle

Les modèles hydrologiques basés sur des systèmes d'intelligence artificielle (IA) utilisent des algorithmes d'apprentissage automatique pour apprendre à partir de données historiques et identifier des modèles et des relations qui peuvent être utilisés pour prédire des événements hydrologiques futurs. Ils sont ainsi plus flexibles et adaptables aux conditions changeantes, et peuvent être utilisés avec des ensembles de données plus restreints. L'une des applications de la modélisation hydrologique basée sur l'IA les plus robustes du point de vue de l'avenir est la prévision des phénomènes météorologiques extrêmes, tels que les inondations et les sécheresses. Dans les modèles d'intelligence artificielle (IA) de prévision des débits, l'apprentissage se fait à partir de données historiques et l'identification de modèles (Dawson & Wilby, 2001). Ces données peuvent inclure des informations sur les précipitations, la température, la fonte des neiges et d'autres facteurs susceptibles d'influer sur le débit de cours d'eau. Le modèle d'IA utilise ensuite ces informations pour faire des prévisions sur le débit futur de cours d'eau. Il existe un certain nombre

de modèles d'IA différents qui peuvent être utilisés pour la prédiction du débit, chacun ayant ses propres forces et faiblesses (Vidal, 2005). Les modèles les plus courants sont les suivants

- **Les réseaux neuronaux artificiels (RNA)** : Les RNA sont un type de modèle d'apprentissage automatique capable d'apprendre à partir de données et d'identifier des modèles. Ils sont souvent utilisés pour la prédiction des débits car ils peuvent être entraînés sur de vastes ensembles de données historiques (Coulibaly et al., 1999).
- **Machines à vecteurs de support (SVM)** : Les SVM sont un autre type de modèle d'apprentissage automatique qui peut être utilisé pour la prédiction des débits. Ils sont souvent utilisés pour des tâches nécessitant une grande précision, comme la prévision des inondations (Shabri & Suhartono, 2012).
- **Forêts aléatoires** : Les forêts aléatoires sont un type de modèle d'apprentissage ensembliste qui combine les prédictions de plusieurs arbres de décision. Elles sont souvent utilisées pour des tâches qui nécessitent un équilibre entre précision et interprétabilité, comme la gestion des ressources en eau (Yu et al., 2017).

La précision des modèles d'IA pour la prédiction du débit peut varier en fonction du modèle spécifique et de la qualité des données d'entraînement (Hu et al., 2018). Toutefois, les modèles d'IA se sont révélés plus précis que les méthodes traditionnelles de prédiction des débits, telles que les modèles statistiques. Les modèles d'IA pour la prédiction des débits sont une nouvelle technologie prometteuse qui pourrait révolutionner la façon dont nous gérons les ressources en eau (Kim et al., 2021). Au fur et à mesure que la technologie de l'IA se développe, ces modèles deviendront encore plus puissants et polyvalents, ce qui en fera un outil essentiel pour la gestion des ressources en eau au 21^e siècle.

I.6 Critères de performance du Modèle

Dans le but de vérifier le fonctionnement du modèle et de préciser la pertinence et la performance du modèle utilisé, certaines méthodes peuvent être utilisées et permettent d'évaluer les qualités du modèle. Les critères de performance représentent un instrument utile pour estimer les résultats obtenus et analyser l'efficacité et l'efficience des modèles utilisés. La précision des modèles hydrologiques requiert au moins deux mesures : l'une dépend de la corrélation et l'autre de l'erreur absolue (Legates & McCabe Jr, 1999). Tous les modèles ont été évalués sur la base de quatre critères de performance.

Ces critères sont les suivants :

I.6.1 Critère de Nash–Sutcliffe (NSE)

Le coefficient d'efficacité du modèle de Nash-Sutcliffe (NSE) est utilisé pour évaluer la capacité de prédiction des modèles hydrologiques (Gerard & BERTHET, 2010). Il est défini comme suit :

$$NSE = 1 - \frac{\sum_{i=1}^N [Q_{iO} - Q_{iP}]^2}{\sum_{i=1}^N [Q_{iO} - \bar{Q}_{iO}]^2} \quad (I.3)$$

I.6.2 Coefficient de corrélation (R)

Le coefficient de corrélation de Pearson (R) représente le plus souvent l'outil le plus utilisé pour évaluer la performance des modèles hydrologiques (Wang et al., 2009). Ce coefficient est déterminé par le calcul de la régression linéaire entre les valeurs calculées (débit) et les valeurs observées ou mesurées (débit). Son expression est donnée par l'équation (I.4) : Le coefficient de corrélation est toujours défini entre -1 et 1.

$$R = \left[\frac{\frac{1}{N} \sum (Q_{iO} - \bar{Q}_{iO})(Q_{iP} - \bar{Q}_{iP})}{\sqrt{\frac{1}{N} \sum_{i=1}^n (Q_{iO} - \bar{Q}_{iO})^2} \sqrt{\frac{1}{N} \sum_{i=1}^n (Q_{iP} - \bar{Q}_{iP})^2}} \right] \quad (I.4)$$

I.6.3 La racine de l'erreur quadratique moyenne (RMSE)

L'erreur quadratique moyenne ou l'écart quadratique moyen est l'une des mesures les plus couramment utilisées pour évaluer la qualité des prédictions. Elle indique la distance entre les prédictions et les valeurs réelles mesurées à l'aide de la distance euclidienne (Chai & Draxler, 2014).

Pour calculer la RMSE, il faut calculer le résidu pour chaque point de données, calculer la moyenne des résidus et prendre la racine carrée de cette moyenne. La RMSE est couramment utilisée dans les applications d'apprentissage supervisé, car elle utilise et nécessite des mesures réelles à chaque point de données prédit (Atiquzzaman & Kandasamy, 2016).

L'erreur quadratique moyenne peut-être exprimée comme suit :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{iO} - Q_{iP})^2} \quad (I.5)$$

Dans le domaine de l'apprentissage automatique, il est extrêmement utile de disposer d'un chiffre unique pour évaluer les performances d'un modèle, que ce soit au cours de la formation, de la validation croisée ou du suivi après le déploiement. L'erreur quadratique moyenne est l'une des mesures les plus utilisées à cet effet. Il s'agit d'une règle de notation appropriée, intuitive et compatible avec certaines des hypothèses statistiques les plus courantes.

I.6.4 L'erreur moyenne absolue (MAE)

L'erreur absolue moyenne MAE correspond à la valeur moyenne de la différence absolue entre les valeurs calculées et les valeurs mesurées (Willmott & Matsuura, 2005). Une MAE peu élevée indique une bonne performance du modèle. Les valeurs observées et simulées sont en parfaite adéquation, ce qui produit une MAE=0. La MAE est calculée par l'équation (I.6) :

$$MAE = \frac{1}{N} \sum_{i=1}^N |Q_{iO} - Q_{iP}| \quad (I.6)$$

Où N le nombre de données, \bar{Q}_{iO} la moyenne du débit mesuré, \bar{Q}_{iP} la moyenne du débit mensuel prévu, Q_{iP} le débit mensuel prévu et Q_{iO} le débit mensuel mesuré.

Conclusion

Ce premier chapitre présente le contexte dans lequel s'inscrit notre recherche. La modélisation hydrologique pluie-débit. De nombreuses méthodes existent pour prédire les débits d'eau, celle qui est la plus répandue étant l'utilisation de modèles hydrologiques. Ces modèles constituent des outils informatiques qui utilisent des équations mathématiques pour simuler le débit de l'eau dans un bassin versant. Ils peuvent être utilisés pour prédire les débits des cours d'eau à court terme (quelques jours ou semaines) ou à long terme (plusieurs mois ou années). Le problème de la prévision hydrologique est très sensible et complexe, car les séries chronologiques de débit ne sont pas stationnaire et la relation entre les variables n'est pas linéaire.

CHAPITRE II Etat de l'art

Introduction

L'utilisation de l'intelligence artificielle (IA) s'est considérablement répandue depuis le milieu du XXe siècle, témoignant de son application diversifiée dans les domaines de l'ingénierie et des sciences. Les deux dernières décennies ont notamment été marquées par une expansion remarquable du développement et de l'application de diverses approches d'IA pour la prédiction des débits des cours d'eau. Globalement, l'IA a réalisé des avancées significatives dans la modélisation hydrologique non linéaire et la prise en compte de la complexité du bruit dans les ensembles de données.

Plusieurs types de modèles d'intelligence artificielle, qu'ils soient uniques ou hybrides, se basant sur des algorithmes d'optimisation ou métaheuristiques, ainsi que sur la transformation en ondelettes, sont utilisés pour la modélisation hydrologique et la prévision des débits des cours d'eau. Ce chapitre explore l'état de l'art dans l'application de l'IA à la prévision des débits des cours d'eau, en mettant l'accent sur l'utilisation de modèles d'IA simples et hybrides pour prédire les débits des cours d'eau à différentes échelles temporelles (journalière, mensuelle, etc.).

II.1 Modèles à base d'intelligence artificielle pour la prévision du débit des cours d'eau

Les réseaux neuronaux artificiels (RNA) sont largement utilisés dans le domaine de l'hydrologie pour la prévision du débit des cours d'eau. Les RNA constituent un type d'algorithme d'apprentissage automatique qui imite le fonctionnement des neurones biologiques, permettant ainsi de résoudre le phénomène de non-linéarité associé aux données de séries temporelles. Un modèle RNA peut apprendre à approximer des relations complexes entre des données d'entrée et de sortie. Dans le cas de la prévision du débit, les données d'entrée pourraient inclure des mesures historiques du débit, des données météorologiques et d'autres données environnementales, tandis que les données de sortie seraient les valeurs prédites du débit. Les modèles RNA ont démontré leur efficacité dans la prédiction du débit des cours d'eau à différentes échelles de temps, allant du court terme (heures ou jours) au long terme (mois ou années). Ces modèles sont également capables de saisir les relations non linéaires entre le débit et d'autres variables.

[Tareke et Awoke \(2023\)](#) ont utilisé un réseau neuronal artificiel pour étudier et effectuer des prévisions à long terme du débit et de la sécheresse hydrologique en Éthiopie, visant à prédire le

débit et la sécheresse hydrologique futurs de 2026 à 2099. Les futures données de précipitations, mises à l'échelle à partir de modèles climatiques régionaux (MCR), ont été utilisées comme données d'entrée pour prédire le futur débit d'eau. Les résultats obtenus ont montré que le RNA est un outil efficace pour prédire le débit des cours d'eau dans les régions où une bonne corrélation entre les précipitations et le débit des cours d'eau existe. [Kisi et al. \(2022\)](#) ont utilisé le modèle RNA en combinaison avec le modèle d'arbre radial (M5 RM5Tree), le modèle d'arbre M5 (M5Tree), le réseau de neurones à fonction de base radiale (RBFNN) et la régression spline adaptative multivariée (MARS) pour la prévision du débit journalier dans un climat froid, en utilisant des données provenant de deux stations suédoises. [Nourani et al. \(2019\)](#) ont utilisé le modèle hybride d'ondelettes M5Tree et l'ont comparé au modèle RNA pour prévoir le ruissellement des rivières sur une base journalière et mensuelle. [Tongal et Booij \(2018\)](#) ont comparé le modèle RNA, le machine à vecteur de support (SVM) et la forêt aléatoire (RF) pour la prévision du débit des cours d'eau en utilisant des données provenant de quatre rivières aux États-Unis. [Sivapragasam et al. \(2014\)](#) ont cherché à démontrer que les prévisions mensuelles peuvent être améliorées de manière significative si les variables d'entrée des modèles RNA sont choisies avec soin, même si le nombre de modèles d'entraînement est plus faible. [Corzo et Solomatine \(2007\)](#) ont utilisé trois techniques différentes de séparation des débits pour améliorer les performances de prévision d'un modèle RNA et ont constaté que l'RNA modulaire donnait de meilleurs résultats qu'un modèle RNA global. [Wu et al. \(2005\)](#) ont exploité deux modèles de prévision hydrologique RNA pour fournir des prévisions de débit à court terme en temps quasi réel. [Halff et al. \(1993\)](#) ont développé un RNA utilisant les hyétogrammes des précipitations observées comme données d'entrée et les hydrogrammes enregistrés par le U.S. Geological Survey (USGS) à Bellvue, Washington, comme données de sortie, afin de prédire le débit.

II.1.1 Le réseau neuronal à perceptron multicouche (MLPNN)

[Hassan et al. \(2022\)](#) ont introduit les modèles SVR et le réseau neuronal à perceptron multicouche (MLPNN) comme solution pour la prédiction efficace du débit des réservoirs. Dans cette étude, trois paramètres d'entrée influençant le débit du réservoir ont été pris en compte : le niveau de stockage, les précipitations et l'évaporation, qui doivent être inclus dans les deux modèles. Différents paramètres du modèle ont été testés, tels que les types de noyaux dans le SVR et le nombre de couches cachées et de neurones dans le MLPNN. Les deux modèles ont prouvé leur

capacité, mais le MLPNN avec deux couches cachées et 4 neurones dans chaque couche a surpassé le SVR après avoir été testé à l'aide de quatre tests de performance différents. [Waqas et al. \(2021\)](#) ont utilisé des techniques basées sur l'intelligence artificielle (IA) telles que le Boost d'arbre (TB), la Forêt d'arbres décisionnels (DTF), l'Arbre décisionnel simple (SDT) et le MLPNN pour prédire le débit du bassin fluvial de la Jhelum. [Rezaie-Balf and Kisi \(2018\)](#) ont examiné les performances de trois méthodes informatiques douces différentes, le MLPNN, le système d'apprentissage extrême élagué de manière optimale (OP-ELM) et la régression polynomiale évolutive (EPR), dans la prévision du débit journalier des cours d'eau. Les données de trois stations différentes, SoleymanTange, PerorichAbad et Ali Abad, situées sur la rivière Tajan en Iran, ont été utilisées. Le modèle MLPNN a été utilisé pour déterminer les combinaisons d'entrées optimales de chaque station en appliquant les critères d'évaluation. [Yonaba et al. \(2010\)](#) ont cherché à répondre à la question de la sélection de la fonction de transfert. Trois fonctions de transfert non linéaires limitées par -1 et 1 ont été sélectionnées pour être testées, sur la base d'un examen de la littérature : la sigmoïde d'Elliott, la sigmoïde bipolaire et la sigmoïde tangente. Elles ont été utilisées pour concevoir des réseaux neuronaux à perceptrons multicouches pour la prévision du débit des cours d'eau à plusieurs étapes pour cinq bassins versants différents et des délais d'exécution de 1 à 5 jours. Les résultats ont confirmé que la sigmoïde tangente était la fonction de transfert la plus pertinente pour la prévision des débits, par rapport à la sigmoïde bipolaire (logistique) et à la sigmoïde d'Elliott. [Mutlu et al. \(2008\)](#) ont développé deux algorithmes, notamment le MLPNN et le réseau neuronal à fonction de base radiale (RBFNN), pour la prédiction du débit des cours d'eau. Les auteurs ont indiqué que le modèle MLPNN était plus performant que le modèle RBFNN. [Shamseldin et al. \(2007\)](#) ont proposé une étude comparative de trois méthodes de prévision de type réseau de neurones en combinant les débits de rivière simulés par différents modèles pluie-débit. Ces structures de réseau sont : le réseau neuronal simple (SNN), RBFNN et MLPNN. Les données quotidiennes de huit bassins versants, situés dans différentes parties du monde et présentant des conditions hydrologiques et climatiques différentes, ont été utilisées pour comparer les performances de ces trois méthodes. Dans le cas de chaque bassin versant, chaque méthode de combinaison de réseaux neuronaux utilise de manière synchronisée les débits fluviaux simulés de quatre modèles pluie-débit fonctionnant en mode de conception sans mise à jour pour produire les débits fluviaux combinés. Deux de ces quatre modèles sont des boîtes noires, les deux autres étant des modèles conceptuels. En termes d'indice d'efficacité du modèle de Nash-Sutcliffe, la méthode

de combinaison MLPNN est généralement plus performante que les deux autres méthodes de combinaison testées.

II.1.1.1 Optimisation à l'aide de la combinaison du modèle MLPNN avec des algorithmes d'optimisation et métaheuristiques

Xu et al. (2023) ont focalisé leur article sur l'introduction d'un modèle innovant capable de prédire avec une grande précision le débit des cours d'eau. Par conséquent, un réseau neuronal perceptron multicouche, dont le nombre optimal de neurones est spécifié à l'aide de métriques d'évaluation, est considéré pour la simulation. Le réseau MLPNN utilisé a été optimisé à l'aide de différents optimiseurs, à savoir l'optimisation par essais de particules, l'algorithme génétique, l'optimisation par loup gris, l'optimisation chaotique par loup gris, l'optimisation avancée par loup gris et l'optimisation adaptative par loup gris chaotique. Selon les résultats statistiques, l'optimisation chaotique adaptative du loup gris est présentée comme l'optimiseur le plus performant. Gomaa et al. (2023) ont examiné l'application d'un modèle hybride de perceptron multicouche optimisé par essaimage de particules (MLPNN-PSO) dans l'analyse de la relation entre les précipitations et le ruissellement, ainsi que dans la prédiction du débit de ruissellement. Les modèles de réseau neuronal à fonction de base radiale gaussienne (GRNN) et de régression par processus gaussien (GPR) ont été utilisés comme procédure de comparaison pour évaluer la performance et la supériorité du modèle MLPNN-PSO. Les résultats montrent que le modèle MLPNN-PSO est plus performant que les modèles GRNN et GPR. Sahoo et al. (2022) se sont concentrés sur l'application de deux modèles hybrides robustes pour la prévision des inondations dans deux stations de jaugeage du bassin de la rivière Barak, dans le district de Cachar de l'Assam, en Inde, et sur l'évaluation de leur précision pour diverses combinaisons d'entrées hydrologiques. Les deux modèles hybrides intègrent le MLPNN et la machine à vecteurs de support (SVM) avec une méta-heuristique inspirée de la nature, l'optimiseur à loup gris (GWO). Les résultats concluants indiquent que l'utilisation du GWO en tant qu'optimiseur conduit à un développement concevable de la performance des modèles MLPNN et SVM pour la prévision des inondations. Khosravi et al. (2022) démontrent les premières applications d'un nouvel algorithme d'apprentissage profond optimisé d'un réseau neuronal convolutif (CNN) utilisant l'algorithme métaheuristique BAT (CNN-BAT). En utilisant les pouvoirs de prédiction de quatre algorithmes bien connus comme références - perceptron multicouche (MLPNN-BAT), système d'inférence neuro-floue adaptatif (ANFIS-BAT), régression vectorielle de soutien (SVR-BAT) et forêt aléatoire (RF-BAT), le modèle CNN-

BAT est testé pour la prédiction du débit journalier dans le bassin versant de Korkorsar, dans le nord de l'Iran. Le résultat obtenu montre que l'utilisation du BAT en tant qu'optimiseur conduit à un développement concevable de la performance des modèles CNN, MLPNN, ANFIS et SVM pour la prévision des débits des cours d'eaux. [Sammen et al. \(2021\)](#) ont mis au point un nouveau modèle de prévision combinant la technique d'optimisation des tournesols (SFA) en tant qu'algorithme évolutionnaire et l'algorithme MLPNN pour prévoir le débit des stations de Jam Seyed Omar et de Muda Di Jeniang en Malaisie. Les résultats de l'algorithme MLPNN-fleur de soleil (SFA) ont été comparés à ceux du MLPNN classique et de deux autres modèles MLPNN hybrides, à savoir l'optimisation par essaims de particules (MLPNN-PSO) et l'algorithme génétique (MLPNN-GA). En conclusion, cette recherche a montré que la combinaison d'un MLPNN avec des algorithmes d'optimisation améliorerait la précision du modèle MLPNN autonome, l'intégration de l'algorithme SFA étant la plus efficace. [Panahi et al. \(2021\)](#) ont utilisé le modèle Bayesian copula-based mean model (CBMA) comme version améliorée du modèle BMA pour prédire le débit des rivières Golok, Kelantan, Lanas et Nenggiri en Malaisie. Les auteurs ont mis en œuvre l'algorithme d'optimisation archimédienne (AOA) pour entraîner le modèle MLPNN. La capacité du modèle MLPNN-AOA a été comparée à celle de l'algorithme MLPNN-BAT, de l'algorithme MLPNN-PSO et de l'algorithme MLPNN-firefly (MLP-FFA). [Mohammadi et al. \(2020\)](#) ont montré que les modèles d'apprentissage automatique (ML) avaient d'excellentes performances pour simuler les séries temporelles de débit de quatre rivières au Canada et aux États-Unis. Ils ont mis en œuvre quatre types différents de modèles basés sur des données, à savoir le modèle bi-linéaire, MLPNN, le MLP MLPNN-PSO et le MLPNN-PSO couplé à l'optimiseur multivers (MLP-PSOMVO). Leurs résultats se situent dans une fourchette R2 comprise entre 0,90 et 0,99, et ils ont abouti à des modèles ML capables de comprendre les phonèmes SF, puis d'obtenir une simulation de ruissellement appropriée. [Pham et al. \(2020\)](#) ont démontré l'efficacité du remplacement de l'algorithme de descente du gradient (GDA) par un algorithme d'optimisation avancé, à savoir l'algorithme IWD (Intelligent Water Drop), pour améliorer les performances du modèle MLPNN pour la prédiction du débit des cours d'eau. Les auteurs rapportent que l'algorithme IWD a effectivement contribué à améliorer le modèle MLPNN.

II.1.1.2 Amélioration de la performance du modèle MLPNN par transformation en ondelettes des variables d'entrée

Gomaa et al. (2023) ont évalué, dans leur étude, l'application de la décomposition en modes empiriques - transformation de Hilbert-Huang (EMD-HHT) en conjonction avec les modèles GPR et MLPNN-PSO dans l'analyse de la relation entre les précipitations et le ruissellement, ainsi que dans la prédiction du taux de ruissellement. Les résultats obtenus prouvent que cette combinaison donne des résultats prometteurs dans la prédiction du ruissellement, le modèle MLP-PSO-EMD montrant une plus grande précision que le modèle GPR-EMD. Agarwal et al. (2022) ont employé les modèles d'ondelettes intégrées, de MLPNN, de réseau neuronal à retardement (TDNN) et de réseau neuronal à mémoire gamma (GMNN) pour prédire les fluctuations horaires du niveau de la rivière, y compris la variable de la variation du taux d'emmagasinement. Les chercheurs ont d'abord utilisé la transformée en ondelettes discrètes pour décomposer la série temporelle du débit d'eau en composantes à basse et haute fréquence. Ensuite, chaque composante a été prédite séparément à l'aide des modèles MLPNN, TDNN et GMNN. Les performances des modèles hybrides WMLPNN, WTDNN et WGMNN proposés ont été comparées à celles des modèles MLPNN, TDNN et GMNN. Les résultats de l'étude viennent confirmer la supériorité des performances des modèles hybrides dans la prévision des débits fluviaux. Shoaib et al. (2018) ont comparé les performances des différents modèles de réseaux neuronaux artificiels (RNA) de type feedforward basés sur les ondelettes. Les types de RNA feedforward utilisés dans l'étude comprennent le MLPNN, le réseau neuronal feedforward généralisé (GFFNN), le RBFNN, le réseau neuronal modulaire (MNN) et les modèles de réseaux neuronaux neuro-flous (NFNN). La transformation en ondelettes discrète (DWT) est utilisée dans la présente étude pour décomposer les données pluviométriques d'entrée à l'aide de la fonction d'ondelette db8. Les résultats de l'étude suggèrent que les modèles WNN sont plus performants dans la prévision des débits des cours d'eaux. Shoaib et al. (2016) ont exploré pour la première fois le potentiel des modèles de réseaux neuronaux récurrents décalés dans le temps (TLRNN) couplés à des ondelettes pour la prévision du ruissellement à partir de données pluviométriques. DWT est employée dans cette étude pour décomposer les données pluviométriques d'entrée à l'aide de six des fonctions d'ondelettes les plus couramment utilisées. L'étude a montré que les modèles TLRNN dynamiques couplés aux ondelettes peuvent être considérés comme une alternative aux modèles MLPNN statiques couplés aux ondelettes. Shoaib et al. (2014) se sont concentrés sur l'évaluation des effets de 23 fonctions

d'ondelettes mères sur la performance des modèles de ruissellement pluvial par réseau de neurones artificiels hybrides basés sur les ondelettes. Les modèles hybrides de MLPNN et de RBFNN sont développés dans cette étude en utilisant à la fois les types de transformation d'ondelettes continues et d'ondelettes discrètes. Le réseau neuronal à perceptron multicouche à transformation en ondelettes discrètes (DWT-MLPNN) et les modèles à fonction de base radiale à transformation en ondelettes discrètes (DWT-RBFNN) au niveau de décomposition neuf avec la fonction d'ondelette db8 présentent les meilleures performances. [Kişi \(2008b\)](#) a proposé l'application d'une technique de neuro-ondelettes pour la modélisation des débits mensuels. Le modèle de neuro-ondelettes est amélioré en combinant deux méthodes, la transformée en ondelettes discrète et le perceptron multicouche, pour la prévision du débit d'un mois à l'avance et les résultats sont comparés à ceux du MLPNN, de la régression multilinéaire (MLR) et des modèles autorégressifs (AR). Les données mensuelles de débit de deux stations, la station Gerdelli sur la rivière Canakdere et la station Isakoy sur la rivière Goksudere, dans la région orientale de la mer Noire en Turquie, sont utilisées dans l'étude. Les résultats de la comparaison ont révélé que le modèle suggéré pouvait augmenter la précision des prévisions et était plus performant que les modèles MLPNN, MLR et AR

II.1.2 Le réseau neuronal de régression généralisée (GRNN).

[Zounemat-Kermani et al. \(2021\)](#) ont évalué différents types de réseaux neuronaux conventionnels, à savoir le réseau neuronal de régression généralisée (GRNN), le RBNN et le MLPNN, dans la prévision du débit journalier de la Tamise, située au Royaume-Uni. Les modèles sont comparés d'un point de vue mathématique, statistique et diagnostique sur la base des résultats prévus pour dix séries temporelles différentes. Les résultats indiquent que tous les modèles de réseaux neuronaux prévoient de manière acceptable le débit journalier. [Akçakoca and Apaydin \(2020\)](#) ont examiné la modélisation des débits du ruisseau Bektas. Les données météorologiques quotidiennes (précipitations, température, temps d'ensoleillement, humidité relative) mesurées dans la région de Kangal et les débits différés d'un jour ont été utilisés à des fins de modélisation. Les prévisions de débit sont simulées à l'aide du modèle GRNN. Pour révéler la différence entre le modèle GRNN et d'autres ANN, les mêmes données ont également été utilisées dans le modèle de réseau neuronal à progression directe (FFNN). [Muhammad Adnan et al. \(2019\)](#) ont évalué la précision de cinq techniques informatiques douces pour la prédiction du débit mensuel du bassin de la rivière Gilgit par une méthode de validation croisée. Les cinq techniques évaluées étaient le FFNN, le RBFNN,

le GRNN, le système d'inférence adaptatif neuro-flou avec partition en grille (Anfis-GP) et le système d'inférence adaptatif neuro-flou avec regroupement soustractif (Anfis-SC). L'interaction entre la température et le débit a été prise en compte dans l'étude. Luo et al. (2019) ont proposé un cadre hybride de prévision du débit, qui intègre l'analyse factorielle, la décomposition des séries temporelles, la régression des données et la suppression des erreurs. Les modèles de régression vectorielle de support (SVR) et GRNN sont intégrés à la décomposition saisonnière et tendancielle pour établir des prévisions mensuelles du débit. Des méthodes de correction d'erreur par autorégression et par combinaison de plusieurs modèles sont utilisées pour garantir la précision. Les données mensuelles de débit de Shigu et Xiangjiaba, en Chine, de 1961 à 2009, sont utilisées pour évaluer la méthode proposée. Diop et al. (2018) ont testé l'utilité de deux approches, le modèle SVR et GRNN, pour prédire le débit quotidien dans le bassin supérieur du fleuve Sénégal à la station de Bafing Makana en Afrique de l'Ouest. Yaseen et al. (2016) ont exploré le potentiel d'une méthode relativement nouvelle basée sur les données, à savoir la méthode de machine d'apprentissage extrême (ELM), pour la prévision des débits mensuels du fleuve Tigre, en Irak. Une étude comparative est menée pour évaluer les performances de l'ELM par rapport à d'autres modèles basés sur des données : SVR et GRNN. Danandeh Mehr et al. (2015) ont appliqué l'algorithme de réseau neuronal à rétropropagation feed-forward (FFBP) avec deux autres techniques de RNA, à savoir les algorithmes de GRNN et RBFNN, dans le but de prédire le débit dans les bassins versants à faible pluviométrie au pas de temps mensuel de deux stations successives sur la rivière Çoruh, en Turquie. Mehr et al. (2014) ont employé différents modèles de réseaux neuronaux artificiels, à savoir feedforward back propagation (FFBP), GRNNRBFNN comme outils de comparaison afin d'examiner la capacité de la technique de programmation génétique linéaire (LGP) à prédire les débits mensuels à partir des données de terrain enregistrées dans deux stations de hydrométries sur la rivière Çoruh, en Turquie. Besaw et al. (2010) ont utilisé un réseau GRNN et une stratégie de réseau de contre-propagation (CPN) pour développer et tester une méthode de prévision du débit des rivières non jaugées à l'aide de deux réseaux neuronaux pilotés par des données climatiques et hydrologiques accessibles au public. Kişi (2008a) a démontré l'application de différentes techniques de RNA pour l'estimation des débits mensuels. Trois techniques ANN différentes, à savoir les réseaux FFNN, GRNN et RBFNN, sont utilisées pour la prévision des débits. Kişi and Kerem Cigizoglu (2007) ont étudié l'utilisation de MLPNN, RBFNN et de GRNN pour la prévision des séries de débits intermittents. Cigizoglu (2005b) a

comparé le modèle GRNN avec le modèle de rétropropagation à erreur directe (FFBP) pour la prévision du débit moyen mensuel. La performance des prévisions du GRNN s'est avérée supérieure à celle des méthodes FFBP, statistique et stochastique en termes de critères de performance sélectionnés. Cigizoglu (2005a) a étudié l'utilisation de MLPNN et GRNN pour la prévision des débits intermittents. Islam et al. (2001) ont proposé le modèle GRNN pour prévoir les séries temporelles chaotiques. L'approche utilise la théorie de l'espace des phases pour reconstruire la trajectoire d'évolution du mouvement, qui est utilisée comme entrée. Contrairement à la méthode de prédiction non linéaire (NLP), où le poids de l'état projeté est le même, le GRNN utilise des poids inégaux. L'état projeté le plus proche est plus pondéré que l'état projeté le plus éloigné, ce qui constitue une approximation raisonnable dans l'espace des phases. Les performances du GRNN sont d'abord vérifiées sur une série temporelle chaotique artificielle, puis sur une série temporelle hydrologique réelle. Les résultats indiquent que les performances du GRNN sont comparables à celles du NLP.

II.1.2.1 Optimisation à l'aide de la combinaison du modèle GRNN avec des algorithmes d'optimisation ou des métaheuristiques

Yin et al. (2016) ont cherché à tester la performance du réseau neuronal de régression générale avec rotation (RGRNN), en tant qu'amélioration du modèle GRNN, dans la prévision des débits moyens mensuels de la rivière Heihe, en Chine. La performance du modèle RGRNN a été comparée à celle du modèle GRNN, du modèle de rétropropagation à erreur directe (FFBP) et du modèle de comptabilisation et de routage de l'humidité du sol (SMAR). Les résultats indiquent que le modèle GRNN combiné pourrait être un outil utile de prévision du débit des rivières pour les régions arides et semi-arides sélectionnées en Chine. Ng et al. (2009) examinent différentes techniques d'estimation pour compléter les observations manquantes dans les séries hydrologiques quotidiennes extrêmes. Des modèles GRNN sont proposés pour l'estimation des observations manquantes, leur configuration d'entrée étant déterminée par une approche d'optimisation par algorithme génétique (GA). L'efficacité de la technique GRNN-GA a été obtenue par des analyses comparatives des performances de la technique proposée par rapport aux techniques existantes. Sur la base des résultats de ces analyses comparatives, en particulier dans le cas de la rivière anglaise (Canada), la technique GRNN-GA s'est avérée être une méthode très compétitive par rapport aux techniques de réseaux neuronaux artificiels existantes. En outre, sur la base des critères de l'erreur quadratique moyenne et de l'erreur absolue, une analyse comparative détaillée impliquant GRNN-

GA, les k-voisins les plus proches et l'imputation multiple pour compléter les enregistrements manquants de la rivière Saugeen (Canada), a également montré que GRNN-GA était supérieure lorsqu'elle était évaluée par rapport à d'autres techniques concurrentes.

II.1.2.2 Amélioration de la performance du modèle GRNN par transformation en ondelettes des variables d'entrée

Tayyab et al. (2016) ont étudié la certitude du débit mensuel en appliquant des réseaux neuronaux artificiels aux cours supérieurs du fleuve Yangtze, en particulier le GRNN. Pour surmonter le problème de non-linéarité du débit, les réseaux neuronaux artificiels ont été intégrés à la transformée en ondelettes discrète (DWT). Il s'avère que la technique de décomposition DWT permet d'améliorer les résultats des prévisions par rapport aux réseaux neuronaux artificiels appliqués séparément. De plus, tous les modèles ont été appliqués séparément aux valeurs maximales, ce qui a également montré que le modèle intégré avait une plus grande capacité à capturer les valeurs maximales. Kişi (2011) a examiné un modèle hybride de réseau neuronal de régression généralisée combiné à la technique des ondelettes (WGRNN) pour la prévision des débits mensuels. Le modèle WGRNN est obtenu en combinant deux méthodes, DWT et GRNN, pour prévoir les débits un mois à l'avance. Les données de débit mensuel de deux stations, la station de Gerdelli sur la rivière Canakdere et la station d'Isakoy sur la rivière Goksudere, dans la région orientale de la mer Noire en Turquie, ont été utilisées. Les résultats de la comparaison ont montré que le WGRNN surpassait les modèles GRNN et FFNN dans la prévision des débits mensuels.

II.1.3 Le machine à vecteur de support (SVM)

Saravanan et al. (2023) ont utilisé cinq modèles d'apprentissage automatique : la machine à vecteur de support (SVR), la forêt aléatoire (RF), le modèle M5 élagué (M5P), le MLPNN et la régression linéaire (LR) pour simuler le débit un jour à l'avance dans le sous-bassin de Pranhita (bassin de Godavari), en Inde. Les résultats soulignent l'importance des modèles avancés d'apprentissage automatique dans les applications de modélisation du débit des cours d'eau. Ikram et al. (2022) ont développé une nouvelle approche, la stratégie d'évolution et d'adaptation de la matrice de covariance (CMAES), pour améliorer la précision de sept modèles d'apprentissage automatique : la machine d'apprentissage extrême (ELM), le réseau élastique (EN), la régression des processus gaussiens (GPR), SVR, le SVR des moindres carrés (LSSVR), l'extrême gradient boosting (XGB) et le RBFNN, dans la prédiction du débit des cours d'eau. Les résultats montrent que la SVR a

fourni des résultats plus précis que les autres méthodes. [Parisouj et al. \(2020\)](#) ont utilisé trois algorithmes d'apprentissage automatique, SVR, ANN et ELM, pour prédire le débit de quatre rivières aux États-Unis. Les auteurs ont souligné que le SVR était meilleur que les modèles utilisés. [Tongal et Booij \(2018\)](#) ont simulé le débit de quatre rivières aux États-Unis en utilisant le modèle SVR, les réseaux ANN et RF en fonction des précipitations (P), de la température (T) et de l'évapotranspiration potentielle (PET). Les résultats de la simulation montrent que l'utilisation de la méthode de séparation des flux de base permet d'améliorer les performances de simulation des modèles d'apprentissage automatique. [Yin et al. \(2018\)](#) ont utilisé SVR, la régression cannelée multivariée adaptée (MARS) et l'arbre de modèle M5 (M5Tree) pour prévoir les données sur le débit des rivières dans une région montagneuse semi-aride et écologiquement importante du bassin versant de Pailugou, dans le nord-ouest de la Chine. Les résultats de ces modèles ont montré que la performance du M5Tree peut être considérée comme supérieure à celle des modèles SVR et MARS lorsqu'il est appliqué à un problème de prévision du débit des rivières à plusieurs horizons de prévision. [Kalra et al. \(2013\)](#) ont étudié les effets de la NAO (Oscillation du Nord Atlantique), de l'ENSO (El Niño-Southern Oscillation) et de la SST (Température de surface de la mer) sur le débit du bassin supérieur du fleuve Colorado. Ils ont utilisé le modèle SVM pour prédire le débit. L'identification de la meilleure combinaison d'entrées a été l'un des défis de leur étude. Il a été conclu que la combinaison de la NAO et de la SST fournissait de bonnes estimations. [Asefa et al. \(2006\)](#) ont utilisé SVM pour prédire le débit saisonnier et horaire à plusieurs échelles. Les résultats de ces modèles ont montré une performance intéressante dans la prédiction des débits des cours d'eau en temps réel et spécifiques. [Sivapragasam et Liong \(2005\)](#) ont dévisagé la plage de débit de trois régions (régions de débit faible, moyen et élevé) et les attributs sont décidés sur la base du processus hydrologique sous-jacent de la région de débit. Un modèle de prévision du débit est appliqué pour chaque région de débit, en utilisant uniquement les données historiques de débit comme données d'entrée. L'approche proposée est mise en œuvre dans le bassin versant de Tryggevælde (Danemark) pour les jours de 1 et 3, en SVM, qui donne des résultats prometteurs, en particulier pour les débits élevés dans un modèle de jour de 3

II.1.3.1 Optimisation à l'aide de la combinaison du modèle SVM avec des algorithmes d'optimisation ou des métaheuristiques

[Samantaray and Sahoo \(2023\)](#) ont souligné que l'objectif principal de leurs recherches est d'explorer le potentiel du nouveau modèle hybride PSR-SVM-FFA (intégration de la reconstruction

de l'espace de phase avec la machine à vecteur de support et l'algorithme firefly) et d'évaluer ses performances par rapport au réseau conventionnel de fonction de base radiale, au SVM et au SVM-FFA hybride pour prédire le débit en tenant compte des données provenant de quatre stations de jaugeage dans le bassin de la rivière Mahanadi, en Inde. Le PSR est appliqué pour extraire des informations et des caractéristiques des séries temporelles de débit et améliorer la précision du modèle hybride SVM-FFA. Les résultats montrent que PSR-SVM-FFA fournit des prédictions robustes. [He et al. \(2023\)](#) ont proposé dans leur étude une méthode intégrée de prédiction de la moyenne des modèles bayésiens qui combine des algorithmes d'intelligence artificielle, notamment le réseau neuronal à mémoire longue et courte (LSTM), le réseau neuronal à unités récurrentes (GRU), le réseau neuronal récurrent (RNN), le réseau neuronal à rétropropagation (BP), la régression linéaire multiple (MLR), la régression par forêt aléatoire (RFR), la régression AdaBoost (ABR) et le modèle SVR pour la prévision des débits fluviaux à moyen terme. En particulier, l'algorithme de recuit simulé (SA) est utilisé pour optimiser les hyperparamètres du modèle. L'application pratique du modèle proposé montre de bonnes performances de prédiction. [Adnan et al. \(2022\)](#) ont intégré le modèle SVM avec l'algorithme SA et l'algorithme d'optimisation de la mouche de mai (MOA) pour prédire le débit dans le bassin de la rivière Helium. [Malik et al. \(2020\)](#) ont optimisé SVR par six algorithmes méta-heuristiques, l'optimisation Ant Lion (SVR-ALO), l'optimiseur Multi-Verse (SVR-MVO), l'optimiseur SpottedHyena (SVR-SHO), l'optimisation Harris Hawks (SVR-HHO), l'optimisation ParticleSwarm (SVR-PSO) et l'optimisation bayésienne (SVR-BO) pour prédire le débit journalier dans le bassin versant de Naula, dans l'État d'Uttarakhand, en Inde. Les résultats montrent que le nouvel algorithme HHO a surpassé les autres algorithmes méta-heuristiques en termes de précision de prédiction. [Zaini et al. \(2018\)](#) ont développé un modèle basé sur SVM et son hybridation avec l'optimisation par l'algorithme PSO pour prévoir le débit journalier à court terme de la rivière dans le bassin versant supérieur de Bertam situé dans les Cameron Highland, en Malaisie. Le résultat obtenu montre que le modèle hybride SVM-PSO est plus performant que SVM. [Zhu et al. \(2016\)](#) ont effectué une prévision du débit du fleuve Yangtze en utilisant une machine à vecteur de support et plusieurs méthodes de décomposition des séries temporelles. [Kalteh \(2015\)](#) a utilisé des modèles de régression par algorithme génétique et vecteur de support (GA-SVR) pour prévoir le débit mensuel de deux rivières et a obtenu de bonnes performances.

II.1.3.2 Amélioration de la performance du modèle SVM par transformation en ondelettes des variables d'entrée

Tikhamarine et al. (2019) ont utilisé un nouvel algorithme méta-heuristique appelé optimiseur de loup gris (GWO) dans le but d'améliorer la précision de la SVR dans la prévision des débits mensuels. L'approche proposée est comparée à d'autres méthodes évolutionnaires, telles que l'optimisation par essaims de particules, l'évolution complexe mélangée et l'optimisation multi-verse, qui sont utilisées pour régler les paramètres du SVR. De plus, les méthodes proposées ont également été combinées avec la transformée en ondelettes. Selon les résultats obtenus, les modèles hybrides présentent de meilleures performances dans la prédiction des débits mensuels par rapport aux versions uniques.

II.1.4 Régression par processus gaussien (GPR)

Kumar et al. (2021) ont développé quatre techniques de régression des processus gaussiens (GPR) : le noyau polynomial, le noyau de fonction de base radiale, le noyau polynomial normalisé et le noyau PUK, qui ont été utilisées pour la prédiction du débit journalier des cours d'eau. En comparant les quatre modèles développés, les résultats suggèrent que le GPR-PUK est plus précis et plus stable que les autres modèles, ce qui en fait un outil utile pour aider les utilisateurs d'eau, les décideurs et les planificateurs du développement à gérer les ressources en eau et à parvenir à un développement durable. Niu et Feng (2021) ont examiné le potentiel de cinq méthodes d'intelligence artificielle pour prédire les séries de débits journaliers dans deux grands réservoirs hydroélectriques en Chine, à savoir le ANN, le système d'inférence floue adaptatif basé sur les neurones (ANFIS), ELM, GPR et SVM. Les résultats suggèrent que les cinq méthodes d'intelligence artificielle peuvent produire des résultats de prévision satisfaisants, tandis que les méthodes SVM, GPR et ELM peuvent produire de meilleures performances que ANN et ANFIS à la fois dans les phases d'apprentissage et de test. Thapa et al. (2020) ont mis en œuvre la régression par processus gaussien, SVR et le modèle non linéaire autorégressif exogène (NARX), qu'ils ont comparés à un modèle LSTM pour modéliser le débit induit par la fonte des neiges dans un bassin de l'Himalaya. Les résultats de cette étude démontrent l'applicabilité des modèles d'apprentissage automatique pour prédire les débits induits par la fonte des neiges dans les bassins versants montagneux où les données sont rares. Zhu et al. (2019) ont introduit GPR comme nouveau modèle basé sur l'IA pour la prévision du débit. Le GPR s'est avéré plus précis que le modèle basé sur le

GRNN. Sun et al. (2014) ont mis en œuvre le modèle GPR pour la prévision du débit des cours d'eau. La capacité prédictive de l'algorithme GPR est testée pour une prévision d'un mois à l'avance en utilisant la base de données d'expérience d'estimation des paramètres du modèle (MOPEX), qui comprend des séries temporelles hydrométéorologiques à long terme recueillies dans 438 bassins à travers les États-Unis de 1948 à 2003. Les comparaisons avec les modèles de régression linéaire et de réseau neuronal artificiel indiquent que l'algorithme GPR surpasse les deux méthodes de régression dans la plupart des cas.

II.1.4.1 Optimisation à l'aide de la combinaison du modèle GPR avec des algorithmes d'optimisation ou des métaheuristiques

L. Wang et al. (2023) ont proposé un modèle de prévision des probabilités intégrant des données multi-résolutions afin d'améliorer les prévisions de débit à court terme. Dans le modèle proposé, l'analyse du spectre singulier (SSA) est utilisée pour traiter les données de débit multi-résolutions afin d'éliminer le bruit caché. Ensuite, SVR est utilisée pour la modélisation, et les méthodes de recherche de grille (GS) et de validation croisée (CV) sont employées pour déterminer les paramètres SVR optimaux. Enfin, GPR est utilisée pour la fusion non linéaire et la prédiction probabiliste. Pour vérifier l'efficacité du modèle proposé, les données de débit du pont Pingchuan et de la station Gaoya avec deux résolutions différentes sont collectées, et plusieurs modèles et indices pertinents sont utilisés pour l'analyse comparative et l'évaluation globale. Les résultats expérimentaux montrent que le modèle proposé surpasse de manière significative les modèles comparés pertinents, ce qui indique que le modèle proposé réduit efficacement l'influence des signaux d'interférence sur la modélisation et utilise pleinement les informations fournies par les données à différentes résolutions pour améliorer la prédiction du débit. En outre, le modèle proposé fournit une estimation fiable de l'incertitude en plus d'une prédiction précise des points. S. Wang et al. (2023) ont étudié l'efficacité d'un modèle hybride de régression par processus gaussien et d'algorithme de recherche coopérative (CSA) pour la prévision de séries de données hydrologiques non stationnaires. L'approche CSA évite le défaut de convergence prématurée du GPR en déterminant efficacement les combinaisons de paramètres appropriées dans l'espace du problème. Plusieurs modèles traditionnels d'apprentissage automatique sont établis pour évaluer la validité de la méthode GPR-CSA proposée dans trois stations hydrologiques de Chine. Différents résultats expérimentaux montrent que le modèle GPR-CSA développé peut prédire avec précision le ruissellement non linéaire et surpasse les modèles traditionnels développés en termes de divers

indicateurs statistiques. Par conséquent, un modèle GPR entraîné par CSA peut fournir une efficacité d'entraînement satisfaisante et une performance de simulation robuste pour la prévision du ruissellement.

II.1.4.2 Amélioration de la performance du modèle GPR par transformation en ondelettes des variables d'entrée

[Ghasempour et al. \(2021\)](#) ont proposé dans leur étude des méthodes hybrides pour prévoir les débits mensuels de la rivière Housatonic sur la période 1941-2018. En intégrant les nouvelles méthodes de prétraitement telles que la décomposition en mode empirique d'ensemble (EEMD) et la décomposition en mode variable (VMD) dans le processus d'apprentissage de deux méthodes d'intelligence artificielle, le modèle de la machine d'apprentissage extrême à noyau (KELM) et le GPR. Dans l'ensemble, les résultats ont montré que les modèles hybrides augmentaient la précision de la modélisation de 20 à 40 % par rapport aux modèles autonomes. [Sun et al. \(2020\)](#) ont cherché dans leur étude à améliorer la précision de la prévision des séries de vitesse du vent et de débit des rivières en développant une architecture innovante de prévision de l'incertitude. Cette architecture combine la méthode de décomposition des données. Dans l'architecture conçue, la Décomposition Adaptative des Modes de Variation (AVMD) est d'abord appliquée pour extraire les informations implicites de la série temporelle originale. Ensuite, la forêt aléatoire est utilisée pour sélectionner les entrées appropriées pour chaque mode. Le GPR est piloté par divers réseaux neuronaux : ELM, Réseaux neuronaux à rétropropagation (BP), GRNN et RBFNN. L'efficacité et l'applicabilité de l'approche proposée sont vérifiées à l'aide de données instables sur la vitesse du vent et le débit des cours d'eau, et comparées à onze modèles apparentés. Les résultats indiquent que le modèle proposé améliore non seulement la précision des prévisions déterministes, mais fournit également davantage d'informations probabilistes pour la prise de décision. [Chamani and Roushangar \(2020\)](#) ont cherché à améliorer la performance du modèle GPR en examinant les débits aux stations transversales sur la rivière Arkansas aux États-Unis, en utilisant des modèles combinés de GPR, ELM et de décomposition en mode empirique de l'ensemble complet (CEEMD). À cette fin, les débits quotidiens et mensuels ont d'abord été prédits à l'aide des modèles GPR et ELM. Ensuite, les séries temporelles de débit ont été divisées en cages par la méthode CEEMD, et ces sous-classes ont été introduites dans la régression du processus gaussien et la modélisation ELM pour simuler le débit. Les résultats ont montré que l'approche CEEMD améliorerait considérablement les performances des modèles susmentionnés. Les résultats obtenus par le modèle hybride GPR pour

prédire le débit dans la première, deuxième et troisième station avec le prétraitement CEEMD ont été réduits de 34, 27 et 32 pour cent, respectivement, par rapport à ceux du modèle GPR sans prétraitement.

II.1.5 Machine d'apprentissage extrême (ELM)

[Sulugodu and Deka \(2019\)](#) ont ciblé dans leur recherche le rendement des données pluviométriques satellitaires CHIRPS par rapport aux données pluviométriques quadrillées IMD et le développement de divers modèles de prévision du débit journalier dans le bassin de Nethravathi, au Karnataka, en Inde. Différentes combinaisons d'entrées sont développées et des méthodes telles que ANFIS, GRNN, PSO-ANN et la machine d'apprentissage extrême (ELM) sont appliquées pour la prévision du débit à l'échelle temporelle. Les résultats indiquent que la précipitation CHIRPS a montré une meilleure performance en comparaison avec les données IMD. L'ELM a exprimé un effet amélioré par rapport à toutes les autres méthodes. [Lima et al. \(2016\)](#) ont étudié la capacité du modèle ELM à modéliser le débit entrant des réservoirs à l'échelle mensuelle en Chine. Le modèle ELM a fourni des résultats précis par rapport à plusieurs autres modèles d'IA bien établis. [Deo and Şahin \(2016\)](#) ont utilisé le modèle ELM pour prédire le débit des cours d'eau. La comparaison de la précision du modèle ELM avec celle du modèle ANN a montré que le modèle ELM donnait de meilleurs résultats que le modèle ANN. La première modélisation de débit dans un cours d'eau à l'aide d'ELM a été réalisée dans l'étude de [Siqueira et al. \(2014\)](#). Les auteurs ont développé un modèle ELM pour capturer la non-linéarité associée au débit saisonnier de l'hydroélectricité brésilienne. Leurs résultats ont montré le potentiel du modèle proposé et ont mis l'accent sur plusieurs études hydrologiques par la suite.

II.1.5.1 Optimisation à l'aide de la combinaison du modèle ELM avec des algorithmes d'optimisation ou des métaheuristiques

[Adnan et al. \(2021\)](#) ont développé et proposé une nouvelle méthode hybride pour la prévision du débit mensuel des cours d'eau dans le bassin versant de Mangla, dans le nord du Pakistan, en intégrant l'algorithme PSO et GWO au modèle ELM, sous le nom d'ELM-PSOGWO. La méthode proposée a été comparée à l'ELM autonome, à l'hybride ELM-PSO et à l'hybride binaire PSOGSA. Les résultats indiquent que le modèle ELM-PSOGWO peut être recommandé pour la prévision du ruissellement mensuel. [Yaseen et al. \(2020\)](#) ont examiné l'intégration d'un algorithme métaheuristique récemment exploré (SSA) avec le modèle ELM pour prévoir le débit mensuel des

rivières. Ils ont utilisé vingt ans de séries chronologiques de données sur le débit du fleuve Tigre à la station de Bagdad, en Irak. En résumé, l'étude confirme que le modèle SSA-ELM est un modèle de données intelligent qualifié pour la prédiction du débit mensuel du fleuve Tigre, en Irak. [Jiang et al. \(2020\)](#) ont développé un modèle hybride qui intègre la théorie du chaos et une machine d'apprentissage extrême avec des paramètres optimaux sélectionnés à l'aide d'une optimisation par essaim particulaire améliorée (ELM-IPSO) pour l'analyse et la prévision du ruissellement mensuel. La méthode développée avait une plus grande précision de prédiction comparée à une méthode d'autorégression, un réseau neuronal artificiel, une machine d'apprentissage extrême avec algorithme génétique et avec l'algorithme PSO, suggérant que l'ELM-IPSO est une méthode efficace pour la prédiction des débits mensuels. [Adnan et al. \(2019\)](#) ont examiné la capacité de prédiction et d'estimation d'une nouvelle méthode heuristique, le modèle de machine d'apprentissage extrême élagué de manière optimale (OP-ELM), pour les débits journaliers des stations de Fujiangqiao et de Shehang sur la rivière Fujiang. La précision de la prédiction de la méthode OP-ELM est comparée à d'autres modèles, comme le système d'inférence neuro-floue adaptatif - optimisation par essaims de particules (ANFIS-PSO), MARS et M5Tree. Les résultats globaux indiquent que la méthode OP-ELM peut être utilisée avec succès pour prédire et estimer le débit journalier. [Lima et al. \(2016\)](#) ont amélioré un modèle ELM développé dans une nouvelle version appelée machine d'apprentissage extrême séquentielle en ligne (OS-ELM) pour la prévision de plusieurs échelles de débit de rivière au Canada. Le même modèle OS-ELM a été développé pour la prévision des inondations dans le cadre d'une surveillance horaire du débit des cours d'eau. Les résultats du modèle ELM amélioré ont démontré une performance de prédiction notable ([Yadav et al., 2016](#))

II.1.5.2 Amélioration de la performance du modèle ELM par transformation en ondelettes des variables d'entrée.

[Siddiqi et al. \(2021\)](#) ont créé un modèle hybride en combinant l'ELM et l'ANN avec une ondelette pour prédire le débit mensuel moyen du barrage de Tarbela sur la rivière Indus. Les modèles hybrides se sont révélés plus performants que les modèles individuels pour la prédiction du débit. [Li and Cheng \(2014\)](#) ont proposé un modèle combinant les réseaux neuronaux à ondelettes et l'ELM (WNN-ELM) pour la prévision des débits à un mois d'échéance. La transformée en ondelettes à trous est utilisée pour décomposer la série temporelle originale des débits en plusieurs sous-séries. Ces sous-séries sont ensuite utilisées comme entrées pour les SLFN couplés à

l'algorithme ELM (SLFNs-ELM) ; la sortie est le débit observé à l'étape suivante. À des fins de comparaison, les SLFNs-ELM et la machine à vecteurs de support (SVM) sont également utilisés. Les résultats indiquent que le SLFNs-ELM est légèrement plus performant que le SVM pour l'estimation des débits de pointe, et que le modèle proposé WNN-ELM fournit des prévisions plus précises.

II.1.6 Réseau récurrent à mémoire court et long terme (LSTM)

[Vu et al. \(2023\)](#) ont utilisé plusieurs modèles d'apprentissage profond, notamment le MLPNN, le réseau neuronal convolutif (CNN), LSTM et le transformateur, pour prédire le débit des cours d'eau à différentes échelles de temps sur la base des variables météorologiques et des indices climatiques observés (c'est-à-dire le débit, le niveau de l'eau, les précipitations et la température) de 1979 à 2019. Les résultats indiquent que la LSTM est très performante pour la prévision des débits pendant les saisons sèches et humides, tandis que le transformateur n'est pas recommandé pour la prévision à long terme, en particulier pendant la saison sèche. [Tounsi et al. \(2023\)](#) ont visé à évaluer les performances spatio-temporelles des techniques basées sur l'apprentissage automatique pour simuler le débit des cours d'eau à l'échelle continentale en utilisant le modèle LSTM. L'ensemble de données utilisé est dérivé de l'expérience d'estimation des paramètres du modèle (MOPEX), qui englobe 438 bassins versants à travers les États-Unis. MOPEX possède le plus long enregistrement de données (55 ans) comparé à d'autres ensembles de données, ce qui le rend très approprié pour l'entraînement des LSTM. L'impact de l'incorporation de la fraction de verdure (GF) de la végétation dans le modèle LSTMGF a été évalué. Les résultats indiquent que la LSTMGF est plus performante que la LSTM dans les zones où la variation saisonnière de la couverture végétale est importante. De plus, les résultats montrent que l'utilisation de données étendues (MOPEX) a renforcé les performances de la LSTM. [Hosseinzadeh et al. \(2023\)](#) ont utilisé des modèles de prévision basés sur l'apprentissage automatique, notamment RFR, LSTM, la moyenne mobile intégrée autorégressive saisonnière (SARIMA) et Facebook Prophet (PROPHET), pour prévoir 24 mois à l'avance le débit naturel du site de Lees Ferry, situé dans la partie inférieure du bassin supérieur du fleuve Colorado (UCRB) aux États-Unis. [Rahimzad et al. \(2021\)](#) et [Hu et al. \(2020\)](#) ont mené des études comparant la précision des modèles LSTM pour les prévisions horaires et quotidiennes des débits à d'autres modèles basés sur des données tels que LR, MLPNN et SVR. Les résultats des deux études indiquent que le modèle LSTM a atteint la plus grande précision,

même jusqu'à 6 heures avant la prédiction du débit. [Kratzert et al. \(2018\)](#) ont étudié le potentiel du modèle LSTM pour la modélisation du ruissellement d'un certain nombre de bassins versants avec des observations météorologiques. Les résultats ont montré que LSTM était capable de prédire le ruissellement avec une précision comparable au modèle de référence bien établi. [Hsu et al. \(1997\)](#), [Carriere et al. \(1996\)](#) et [Rumelhart et al. \(1985\)](#) ont réalisé les premières études utilisant les RNN pour modéliser les débits des cours d'eau. Les premiers auteurs ont testé l'utilisation des RNN dans des conditions de laboratoire et ont démontré leur utilisation potentielle pour des applications basées sur des événements.

II.1.6.1 Optimisation à l'aide de la combinaison du modèle LSTM avec des algorithmes d'optimisation ou des métaheuristiques

[Vatanchi et al. \(2023\)](#) ont proposé des techniques avancées d'apprentissage profond pour prévoir le débit des cours d'eau. L'application de ces méthodes à la prévision du débit à long terme est une question qui a reçu moins d'attention. Quatre modèles, dont les systèmes d'inférence neuro-floue adaptatifs (ANFIS), ANN, la mémoire bidirectionnelle à court terme (BiLSTM) et l'hybride réseau neuronal convolutif (CNN) - unité récurrente gérée (GRU) - LSTM, sont appliqués pour prévoir le débit journalier à long terme sur le fleuve Colorado aux États-Unis. Les auteurs montrent que les modèles BiLSTM et CNN-GRU-LSTM ne sont pas plus performants que les autres ; une comparaison de ces modèles avec ANN et ANFIS montre clairement que leur performance n'est pas meilleure. [Kilinc \(2022a\)](#) a proposé une approche hybride intégrant les réseaux de mémoire à long terme (LSTM) et l'algorithme PSO pour la prévision du débit quotidien des cours d'eau. À cette fin, l'auteur a employé trois stations hydrologiques dans l'étude le long du bassin de l'Oronte, respectivement Karasu, Demirköprü et Samandağ. Des méthodes statistiques telles que la régression linéaire et le modèle plus classique de moyenne mobile intégrée autorégressive (ARIMA) ont été utilisées au cours du processus de comparaison pour évaluer les performances de la méthode proposée et démontrer sa capacité prédictive supérieure. Les résultats ont démontré l'efficacité et la supériorité du modèle hybride LSTM-PSO proposé pour la prévision des débits de cours d'eau quotidiens. [Kilinc and Haznedar \(2022\)](#) ont proposé une approche hybride, intégrant des réseaux de mémoire à long et court terme (LSTM) et un algorithme génétique (GA), pour prévoir le débit quotidien de l'Euphrate en Turquie. La régression linéaire a été utilisée dans le cadre du processus de comparaison pour évaluer les performances de la méthode proposée et démontrer sa capacité prédictive supérieure. La comparaison des résultats de la prévision des débits

journaliers a révélé l'efficacité et les performances supérieures du modèle LSTM-GA. [Lin et al. \(2021\)](#) ont développé un modèle hybride, le modèle DIFF-FFNN-LSTM, pour prédire le débit horaire. Le modèle comprend trois composantes, la différence de premier ordre (DIFF), le FFNN et LSTM. Appliqué au bassin d'Andun en Chine, le modèle DIFF-FFNN-LSTM proposé donne de très bons résultats dans la prévision du débit horaire. La comparaison avec cinq autres modèles d'apprentissage automatique (d'une complexité ou d'une structure de modèle similaire) et quatre modèles statistiques montre la supériorité du modèle DIFF-FFNN-LSTM proposé. [Feng et al. \(2021\)](#) ont proposé un nouveau modèle amélioré de LSTM appelé LN-LSTM-PSO en intégrant la normalisation des couches (LN), le réseau LSTM et l'optimisation par l'algorithme PSO afin d'améliorer la précision des prévisions. L'application du modèle amélioré proposé est illustrée à l'aide de données hydrologiques et météorologiques du bassin de la rivière Jiulong dans la province de Fujian, en Chine, afin de tester la précision du modèle. Les résultats indiquent que le modèle proposé est très précis. Le modèle amélioré de prévision du ruissellement surpasse un modèle de régression à vecteur de support, un réseau neuronal artificiel, un réseau neuronal récurrent et un réseau à mémoire à long terme.

II.1.6.2 Amélioration de la performance du modèle LSTM par transformation en ondelettes des variables d'entrée

[Chakraborty and Biswas \(2023\)](#) ont appliqué des modèles de prévision du débit basés sur les réseaux de neurones artificiels (ANN) et les réseaux de mémoire à long terme (LSTM) à trois stations hydrologiques, TeestaBazaar, Domohani et Mekhliganj, dans le bassin de la rivière Teesta, en Inde, pour différents délais. La transformée en ondelettes a été combinée aux modèles ANN et LSTM pour développer deux modèles hybrides : le réseau neuronal artificiel basé sur les ondelettes (WANN) et le modèle WLSTM. Les résultats obtenus ont prouvé que les modèles hybrides proposés sont largement efficaces pour la prédiction des flux. [Ni et al. \(2020\)](#) ont développé deux modèles hybrides, basés sur un réseau de mémoire à long terme (LSTM), pour la prévision des débits mensuels et des précipitations. L'un des modèles, WLSTM, applique un algorithme de transformée en ondelettes pour décomposer les séries, et l'autre, convolutif LSTM (CLSTM), associe un réseau neuronal convolutif pour extraire les caractéristiques temporelles. La précision de prédiction du WLSTM et du CLSTM a été comparée à celle du MLPNN et du LSTM. Les résultats indiquent que le LSTM est applicable à la prédiction de séries temporelles, mais que le WLSTM et le CLSTM sont des alternatives supérieures. [Liu et al. \(2020\)](#) ont utilisé un réseau

neuronal profond pour prédire le débit mensuel de la station hydrologique de Hankou sur le fleuve Yangtze. Cette méthode combine l'algorithme de décomposition en mode empirique (EMD) avec une architecture de mémoire à long terme de l'encodeur-décodeur (En-De-LSTM). La technique EMD peut lire et décomposer les données originales en plusieurs composantes de fréquence différentes, ce qui permet au modèle de faire des prédictions à long terme plus efficacement. Le réseau neuronal En-De-LSTM basé sur la LSTM peut rapprocher la prédiction de la valeur observée du débit de pointe grâce à la lecture, à l'entraînement, à la mémorisation des informations utiles et à l'oubli des données inutiles. Les résultats ont démontré la fiabilité de cette méthode pour prédire le débit mensuel des rivières.

Conclusion

Dans ce chapitre, on a exposé les techniques utilisées pour la prédiction des débits des cours d'eau à différents pas de temps, en se basant sur les modèles d'intelligence autonome et les modèles hybrides les plus couramment utilisés dans la littérature. Dans le présent travail, nous intéressons particulièrement au développement des différents modèles d'intelligence artificielle proposés pour la prédiction des débits dans le bassin du Cheliff, au nord de l'Algérie.

CHAPITRE III Modèles d'intelligence artificielle

Introduction

Les réseaux neuronaux artificiels (ANN), l'apprentissage automatique (ML) et les machines à vecteurs de support de régression (SVR) sont des outils puissants qui peuvent être utilisés pour les prévisions hydrologiques. Grâce à leurs propriétés de modélisation non linéaires et flexibles, ces outils possèdent un potentiel important pour modéliser et prévoir des séries temporelles non linéaires et non stationnaires par rapport à d'autres techniques. Dans l'ensemble, les algorithmes métaheuristiques et les transformations en ondelettes constituent un outil prometteur pour améliorer les performances des modèles d'intelligence artificielle pour la prévision hydrologique. Toutefois, il est important d'examiner attentivement le problème spécifique avant de les appliquer. Pour cette raison, des techniques d'intelligences artificielles autonomes et des modèles hybrides basés sur l'algorithme métaheuristique des chauves-souris (Bat) et transformée en ondelettes discrète à maximum de recouvrement (MODWT) ont été utilisées dans cette étude pour développer des modèles de prévision de débit à l'échelle mensuelle et journalière.

III.1 Les Réseaux de Neurones Artificiels

III.1.1 Historique

Le développement du concept des systèmes neuronaux artificiels (ANN) s'articule autour de quatre phases principales (Park & Lek, 2016):

- **Préhistoire (1943-1958) :** Warren McCulloch et Walter Pitts créent le premier neurone artificiel en 1943. Ce modèle simple de neurone permettait d'effectuer des opérations logiques simples. Au cours des années suivantes, d'autres chercheurs ont développé des modèles plus complexes de neurones et de réseaux neuronaux (Abraham, 2002).
- **Premiers développements (1959-1969) :** En 1959, Frank Rosenblatt a mis au point le Perceptron, un réseau neuronal à une seule couche qui pouvait être entraîné à classer des modèles simples. Le Perceptron a constitué une avancée majeure, mais il s'est avéré par la suite incapable d'apprendre des modèles plus complexes (Peters, 2005).
- **L'effondrement de la recherche sur les réseaux neuronaux (1970-1986) :** L'échec du Perceptron a entraîné un déclin de la recherche sur les réseaux neuronaux. Toutefois, certains chercheurs ont continué à travailler sur le problème et, à la fin des années 1970, de

nouvelles techniques d'entraînement des réseaux neuronaux ont été mises au point (Schmidhuber, 2015).

- **Renaissance des réseaux neuronaux (1987-aujourd'hui) :** À la fin des années 1980, les réseaux neuronaux ont connu un regain de popularité. Cela est dû en partie au développement de nouveaux algorithmes de formation, tels que la rétropropagation, et à la disponibilité d'ordinateurs plus puissants. Aujourd'hui, les réseaux neuronaux sont utilisés dans une grande variété d'applications, notamment la reconnaissance d'images, le traitement du langage naturel et la traduction automatique (Tan & Lim, 2018).

III.1.2 Du neurone biologique au neurone artificiel

III.1.2.1 Neurone biologique

Un neurone biologique est l'unité structurelle et fonctionnelle du système nerveux. C'est une cellule qui reçoit et transmet des signaux à différentes parties du corps. Les neurones sont généralement classés en trois types selon leur fonction (Sharp et al., 1993) :

- **Les neurones sensoriels :** Ces neurones détectent les stimuli de l'environnement, tels que la lumière, le son et le toucher.
- **Neurones moteurs :** Ces neurones contrôlent les muscles et d'autres organes.
- **Interneurones :** Ces neurones relient les neurones sensoriels et les neurones moteurs.

Chaque neurone se compose de trois parties principales :

- **Le corps cellulaire :** C'est la plus grande partie du neurone. Il contient le noyau et d'autres organites.
- **Les dendrites :** Ce sont les prolongements ramifiés du corps cellulaire. Elles reçoivent des signaux d'autres neurones.
- **L'axone :** il s'agit d'un prolongement long et fin du corps cellulaire. Il transmet les signaux à d'autres neurones. L'axone d'un neurone est recouvert d'une gaine de myéline. Cette dernière est constituée de cellules graisseuses qui isolent l'axone et contribuent à accélérer la transmission des signaux.

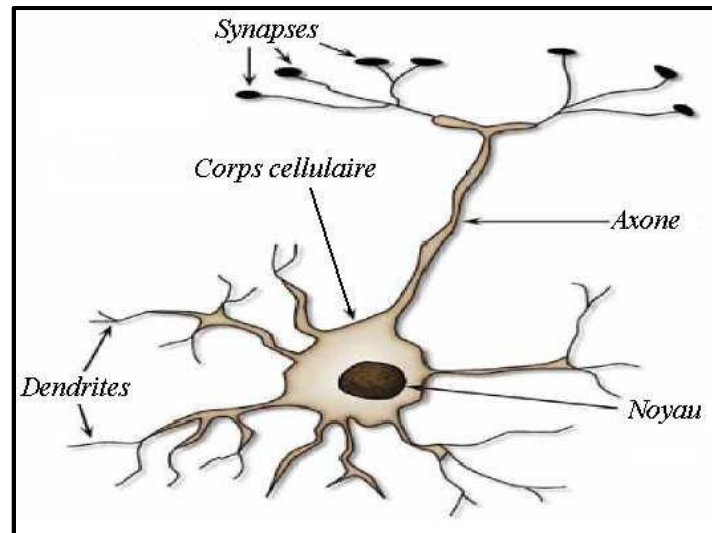


Figure III.1 Représentation schématique d'un neurone biologique (Chekroun, 2009)

Les neurones communiquent entre eux par l'intermédiaire de synapses. Une synapse est un petit espace entre l'axone d'un neurone et la dendrite d'un autre neurone. Lorsqu'un signal électrique atteint l'extrémité de l'axone, il provoque la libération de substances chimiques appelées neurotransmetteurs. Ces derniers traversent la synapse et se lient aux récepteurs de la dendrite de l'autre neurone. Cette liaison des neurotransmetteurs aux récepteurs entraîne la génération d'un signal électrique dans l'autre neurone (Solaiman & Richard, 2003) (Figure III.1).

III.1.2.2 Le Neurone Artificiel

Un neurone artificiel est un modèle mathématique d'un neurone biologique. C'est l'unité de base d'un réseau neuronal artificiel. Les neurones artificiels s'inspirent du fonctionnement des neurones biologiques, mais ils sont beaucoup plus simples. Un neurone artificiel se compose de trois parties principales (Dreyfus, 1998):

- **Les entrées** : Les entrées sont les signaux que le neurone reçoit d'autres neurones.
- **Les poids** : Les poids sont les coefficients qui sont multipliés par les entrées.
- **Biais** : le biais est une valeur constante qui est ajoutée à la somme pondérée des entrées.

La somme pondérée des entrées passe par une fonction d'activation pour produire la sortie du neurone. La fonction d'activation détermine la manière dont le neurone réagit à ses entrées. La sortie d'un neurone artificiel peut être utilisée comme entrée pour d'autres neurones ou comme sortie du réseau. Les neurones artificiels sont connectés les uns aux autres dans un réseau. Ces connexions entre les neurones sont appelées synapses. Ces derniers ont des poids qui sont ajustés

au cours du processus d'apprentissage. Le processus d'apprentissage est la manière dont le réseau neuronal artificiel apprend à effectuer une tâche. Le réseau est formé sur un ensemble de données contenant les sorties souhaitées pour un ensemble donné d'entrées. Le réseau ajuste ses poids et ses biais jusqu'à ce qu'il produise les sorties souhaitées pour les données d'apprentissage (Parizeau, 2004)(Figure III.2).

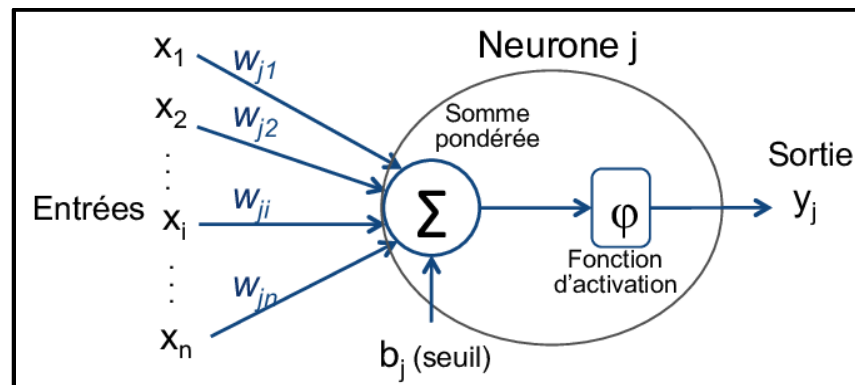


Figure III.2 Représentation schématique d'un neurone artificiel (Laqrichi, 2015)

Où $x_1, x_2, x_3 \dots x_n$ sont les entrées, $W_{1j}, W_{2j}, W_{3j} \dots W_{nj}$ sont les poids, φ est la fonction d'activation. Mathématiquement, comme l'illustre la figure III.2, chaque neurone reçoit des entrées sous forme vectorielle puis il calcule une somme pondérée de ses entrées pour que le résultat passe ensuite par la fonction d'activation afin de créer une sortie.

III.1.3 Fonction d'activation

Dans un réseau de neurones artificiels, la fonction d'activation est une fonction mathématique qui est appliquée à la sortie d'un neurone. La fonction d'activation permet au neurone de calculer sa sortie en fonction de ses entrées. Il existe de nombreuses fonctions d'activation différentes, chacune ayant ses propres avantages et inconvénients (Sharma et al., 2017). Certaines des fonctions d'activation les plus courantes sont les suivantes :

- **La fonction sigmoïde** : La fonction sigmoïde est une fonction non linéaire qui prend une valeur comprise entre 0 et 1. Elle est souvent utilisée dans les réseaux de neurones de classification.
- **La fonction tangente hyperbolique** : La fonction tangente hyperbolique est une fonction non linéaire qui prend une valeur comprise entre -1 et 1. Elle est souvent utilisée dans les réseaux de neurones de régression.

- **La fonction ReLU** : La fonction ReLU est une fonction non linéaire qui est égale à x si x est positif et à 0 si x est négatif. Elle est souvent utilisée dans les réseaux de neurones profonds.

Le choix de la fonction d'activation dépend de la tâche que le réseau de neurones est censé effectuer. Par exemple, si le réseau de neurones doit effectuer une tâche de classification, il faut utiliser une fonction d'activation qui prend une valeur comprise entre 0 et 1. Si le réseau de neurones doit effectuer une tâche de régression, il faut utiliser une fonction d'activation qui prend une valeur comprise entre -1 et 1 (Pratiwi et al., 2020).

III.1.4 Choix et structure d'un réseau de neurones artificiel

III.1.4.1 Choix d'un réseau de neurones artificiel

Le choix d'un réseau de neurones artificiels dépend de plusieurs facteurs, notamment de la méthode d'apprentissage utilisée et de l'objectif à atteindre (May et al., 2011). Les principaux types de réseaux de neurones artificiels sont les suivants :

- **Réseaux de neurones feedforward** : Les neurones sont organisés en couches, et les informations circulent de la couche d'entrée à la couche de sortie, sans boucle de rétroaction. Ce type d'architecture est couramment utilisé dans des tâches telles que la classification d'images ou la prédiction de valeurs numériques (Svozil et al., 1997).
- **Réseaux de neurones récurrents (RNN)** : Les RNN sont utilisés pour des tâches telles que la traduction automatique ou la génération de texte (Sherstinsky, 2020).
- **Réseaux de neurones convolutifs (CNN)** : Ils sont utilisés principalement pour le traitement d'images et de vidéos. Les CNN sont basés sur des filtres qui permettent de détecter des motifs spécifiques dans les données (Kattenborn et al., 2021).
- **Réseaux de neurones auto-organisés** : Ils sont utilisés pour la visualisation de données multidimensionnelles et l'apprentissage non supervisé. Ces réseaux sont capables d'apprendre et de représenter les caractéristiques intrinsèques des données sans étiquettes (Carpenter & Grossberg, 1991).

Le choix du réseau de neurones dépendra donc de l'objectif de l'application, des données à traiter et des performances attendues. Il est important de bien choisir les données, leur nature et leur pertinence pour un bon apprentissage.

III.1.4.2 Structure d'un réseau de neurones artificiel

La structure d'un réseau de neurones artificiel peut varier selon plusieurs paramètres, dont la topologie des connexions entre les neurones, la fonction d'agrégation, la fonction de seuillage et l'algorithme d'apprentissage (Pedrycz, 1993). On distingue généralement deux principales structures de réseau :

✚ Les réseaux de neurones feedforward (ou non bouclés) :

Les réseaux de neurones non bouclés sont un type de réseau de neurones artificiels dans lequel les neurones ne sont connectés que dans une seule direction, de l'entrée à la sortie. Ils sont l'un des types de réseaux de neurones les plus simples et les plus courants (Yuan et al., 2020) (Figure III.3). La structure d'un réseau de neurones feedforward est composée de plusieurs couches de neurones, où chaque couche est connectée à la couche suivante. Les neurones de la première couche sont connectés aux neurones de la deuxième couche, et ainsi de suite. La couche d'entrée est connectée aux neurones de la couche cachée, et la couche cachée est connectée aux neurones de la couche de sortie. Les neurones d'une couche sont connectés par des synapses, qui ont des poids associés. Les poids sont ajustés pendant le processus d'apprentissage pour que le réseau de neurones puisse effectuer la tâche souhaitée (Svozil et al., 1997).

Le processus d'apprentissage d'un réseau de neurones feedforward est appelé entraînement supervisé. Pendant l'entraînement, le réseau de neurones reçoit un ensemble de données d'entraînement, qui contient des données d'entrées et de sorties. Le réseau de neurones utilise les données d'entrée pour générer des sorties. Les sorties sont ensuite comparées aux sorties souhaitées, et les poids sont ajustés pour réduire l'erreur. Ce processus est répété jusqu'à ce que le réseau de neurones puisse générer des sorties suffisamment proches des sorties souhaitées. Une fois qu'un réseau de neurones feedforward a été formé, il peut être utilisé pour générer des sorties pour de nouvelles entrées (Bebis & Georgiopoulos, 1994). Les sorties générées par le réseau de neurones sont basées sur les données d'entrée et de sortie utilisés pour l'entraînement.

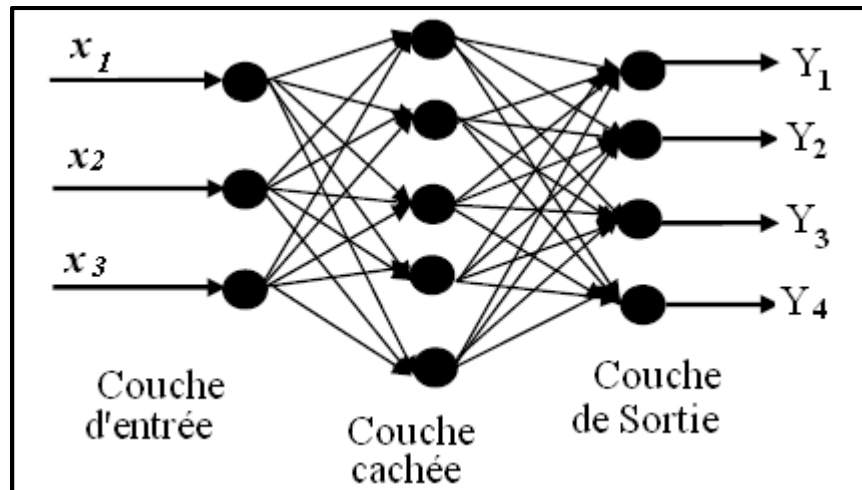


Figure III.3 Réseau de neurones non bouclés (Salhi, 2015)

✚ Les réseaux de neurones récurrents (ou bouclés, ou dynamiques) :

Un réseau de neurones récurrent (RNN) est un type de réseau de neurones artificiels dans lequel les connexions entre les neurones forment des boucles, permettant ainsi aux informations de circuler dans les deux sens. Les RNN sont constitués d'unités (neurones) interconnectées interagissant non-linéairement et pour lequel il existe au moins un cycle dans la structure (Salehinejad et al., 2017). La structure d'un RNN est composée de plusieurs couches de neurones, chacune ayant un nombre variable de neurones et étant connectée aux couches adjacentes (Medsker & Jain, 2001). Les RNN sont capables de prendre en compte l'historique des données grâce à leurs boucles de rétroaction, ce qui leur permet de modéliser des dépendances temporelles. Les RNN peuvent être entraînés à l'aide d'algorithmes tels que la rétropropagation du gradient, qui ajuste les poids des connexions pour minimiser l'erreur de prédiction (Figure III.4).

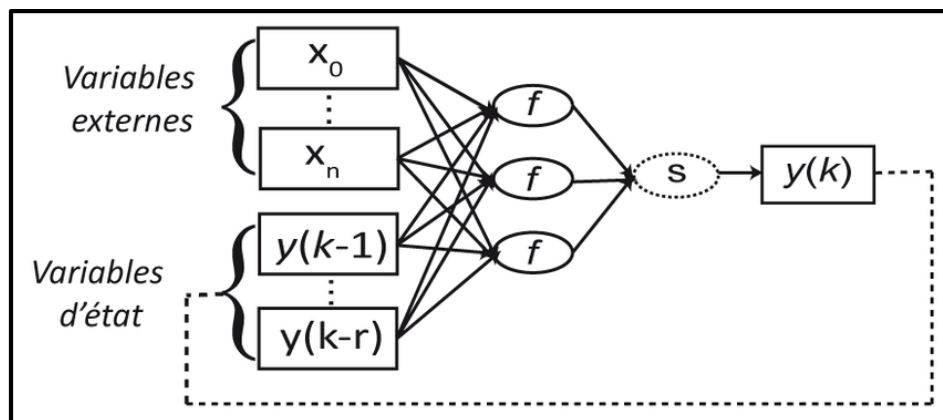


Figure III.4 Réseau de neurone bouclé (Chiang et al., 2004)

III.1.5 Apprentissage des Réseaux de Neurones

L'objectif de l'apprentissage des réseaux de neurones est de minimiser l'écart entre les sorties réelles du réseau et les sorties désirées, en modifiant les paramètres du réseau par étapes successives (Sinitsin et al., 2020). L'apprentissage vise à rendre les sorties du réseau aussi proches que possible des sorties souhaitées pour les exemples utilisés lors de l'apprentissage. Les techniques d'apprentissage des réseaux de neurones sont des algorithmes d'optimisation qui cherchent à ajuster les paramètres du réseau pour réduire l'erreur entre les prédictions du réseau et les sorties désirées. L'apprentissage des réseaux de neurones se déroule par étapes itératives, appelées itérations. À chaque itération, les exemples d'entraînement sont présentés au réseau, les prédictions sont comparées aux sorties désirées, et les poids des connexions sont ajustés en fonction de l'erreur. Ce processus est répété jusqu'à ce que le réseau atteigne un niveau de performance satisfaisant. Au fur et à mesure de l'apprentissage, la sortie du réseau de neurones s'adapte de manière plus précise aux données, ce qui lui permet de généraliser et de faire des prédictions précises sur de nouvelles données. L'apprentissage permet au réseau de capturer les motifs et les relations dans les données d'entraînement, ce qui lui permet de faire des inférences sur des données similaires (Hinton, 1992). Il existe de nombreux types de règles d'apprentissage qui peuvent être regroupées en trois catégories (Singh & Chauhan, 2009) : l'apprentissage supervisé, non supervisé, et renforcé.

III.1.5.1 Apprentissage Supervisé

L'apprentissage par réseau neuronal supervisé est un type d'apprentissage automatique dans lequel le réseau est formé sur un ensemble de données comprenant la sortie souhaitée pour chaque entrée. Le réseau apprend à faire correspondre les données d'entrée à la sortie souhaitée en ajustant ses poids et ses biais (Cunningham et al., 2008; Osisanwo et al., 2017).

III.1.5.2 Apprentissage non Supervisé

L'apprentissage non supervisé des réseaux neuronaux est un type d'apprentissage automatique dans lequel un réseau neuronal est formé sur des données non étiquetées. Le réseau neuronal n'a pas connaissance de la sortie souhaitée pour les données d'entrée, il doit donc découvrir par ses propres moyens des modèles et des relations dans les données (Osisanwo et al., 2017).

III.1.5.3 Apprentissage Renforcé

L'apprentissage renforcé des réseaux neuronaux est une méthode d'apprentissage qui utilise des récompenses et des pénalités pour former un modèle de réseau neuronal (Lin, 1992). Cette technique rappelle l'apprentissage supervisé, à la différence toutefois qu'au lieu de fournir au réseau les résultats souhaités, il reçoit un rang (ou score) correspondant à une mesure de la performance du réseau après quelques itérations.

III.1.6 Algorithme de Rétro-Propagation

L'algorithme de rétropropagation est un algorithme largement utilisé pour l'apprentissage des réseaux neuronaux de type feedforward. Il consiste à calculer le gradient de la fonction de perte par rapport à chaque poids via la règle de la chaîne, en calculant le gradient couche par couche. L'algorithme est utilisé pour former efficacement un réseau neuronal par le biais d'une méthode appelée règle de la chaîne. En termes simples, après chaque passage en avant d'un réseau, la rétropropagation effectue un passage en arrière tout en ajustant les paramètres du modèle (poids et biais) (Rojas & Rojas, 1996). Le principe de l'algorithme de rétropropagation repose sur des opérations algébriques linéaires courantes, telles que l'addition de vecteurs, la multiplication d'un vecteur par une matrice, etc. L'algorithme de rétropropagation considère tous les neurones du réseau de la même manière et calcule leurs dérivés à chaque passage en arrière (Leonard & Kramer, 1990). Il est utilisé pour calculer rapidement les dérivés et constitue un outil mathématique important pour améliorer la précision des prédictions dans les domaines de l'exploration de données et de l'apprentissage automatique. Les poids sont mis à jour rétroactivement, de la sortie à l'entrée, afin de minimiser autant que possible la différence entre les sorties souhaitées et les sorties effectives du système.

III.1.7 Conception d'un Réseau de Neurones Artificiels

La conception d'un réseau de neurones artificiels implique plusieurs étapes (Palmer et al., 2006) (Figure III.5), telles que :

- Définir l'objectif du réseau de neurones et les données d'entrée et de sortie attendues.
- Choisir l'architecture du réseau, c'est-à-dire le nombre de couches et le nombre de neurones dans chaque couche.
- Choisir la fonction d'activation pour chaque neurone.

- Initialiser les poids des connexions entre les neurones.
- Diviser les données en ensembles d'entraînement, de validation et de test.
- Entraîner le réseau de neurones sur l'ensemble d'entraînement en utilisant un algorithme d'apprentissage, tel que la rétropropagation du gradient.
- Valider le réseau de neurones sur l'ensemble de validation pour ajuster les hyperparamètres, tels que le taux d'apprentissage et le nombre d'itérations.
- Tester le réseau de neurones sur l'ensemble de test pour évaluer sa performance.

Il est important de noter que la conception d'un réseau de neurones peut varier en fonction de la nature de la tâche à accomplir et de la complexité des données. Les étapes ci-dessus sont donc générales et peuvent être adaptées en fonction des besoins spécifiques (Frasconi et al., 1998).

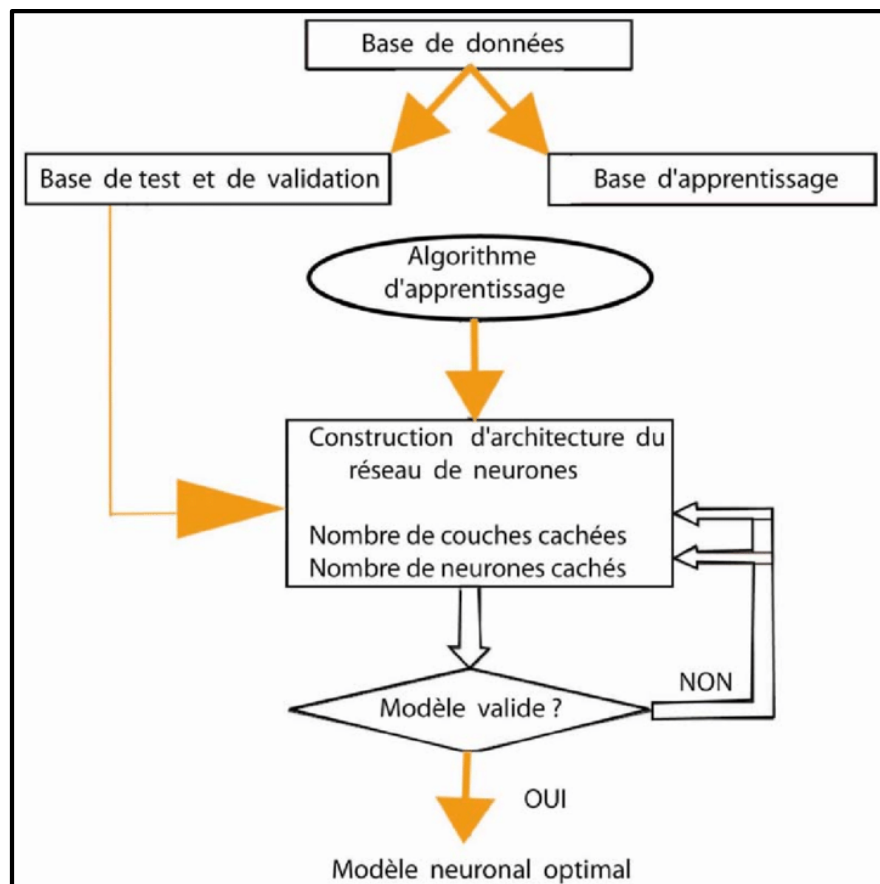


Figure III.5 Organigramme de conception d'un réseau de Neurones (Rafik, 2020)

III.2 Les perceptrons Multicouches MLPNN

L'architecture du modèle MLPNN, basée sur le réseau neuronal biologique humain, est construite avec trois types de couches (c'est-à-dire entrée, cachée et sortie) et également avec des neurones

comme unités de traitement dans chaque couche, cette technique peut capturer la relation non linéaire entre les sorties et les entrées. Les données d'entrée sont transmises au réseau de manière progressive (Figure III.6). La relation entre les réseaux neuronaux est assurée par des poids de connexion adaptés entre les couches correspondantes. Les neurones indépendants sont associés aux neurones des couches successives. En outre, les neurones indépendants additionnent les signaux d'entrée, puis délivrent les signaux de sortie à l'aide d'une fonction d'activation puissante. Les fonctions d'activation les plus utilisées sont les modèles non linéaires de la sigmoïde et de la tangente hyperbolique. L'ajustement des poids de connexion est également effectué lors de la phase d'apprentissage sur la base de l'algorithme de rétropropagation (BP) en propageant l'erreur vers les couches arrière afin de réduire la variance des valeurs observées et calculées (Kim et al., 2012; Shoaib et al., 2019; Zounemat-Kermani et al., 2020).

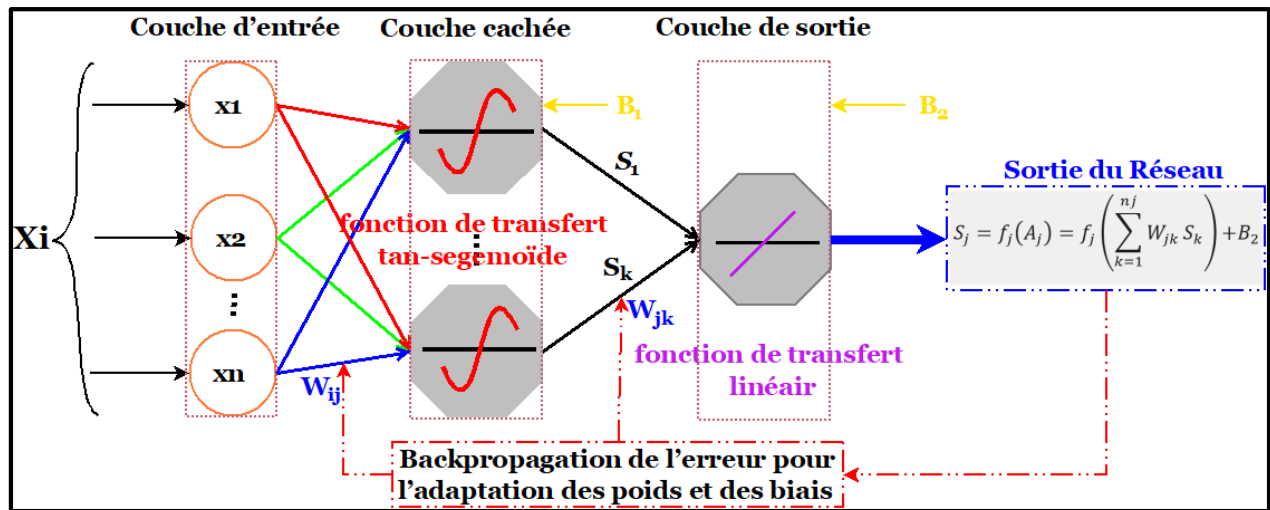


Figure III.6 Structure du perceptron Multicouches (MLPNN)

Un neurone d'un perceptron multicouche se caractérise par trois concepts : son état, ses connexions avec d'autres neurones, et sa fonction d'activation.

$$S_j = f_j(A_j) = f_j \left(\sum_{k=1}^{n_j} W_{jk} S_k \right) + B_2 \quad (\text{III. 1})$$

Avec :

S : L'ensemble des états possibles des neurones ;

S_j : L'état du neurone j , où $S_j \in S$;

f_j : Fonction d'activation associée au neurone j ;

A_j : Activité du neurone j ;

W_{jk} : Poids de la connexion entre les neurones j et k ;

B_2 : Le biais

La construction d'un réseau de neurones multicouches consiste à choisir le nombre de neurones et leurs fonctions de transfert, les connexions déterminées par les neurones, et à fixer les poids de ces derniers. L'attrait de cette approche réside dans le fait qu'en ne changeant que les poids des connexions, on peut utiliser une vaste gamme de fonctions différentes, et qu'il existe une méthode générique pour régler les poids des connexions, appelée règle d'apprentissage.

III.3 Réseaux de neurones à fonction de régression générale (GRNN)

Le modèle de réseau de neurone à fonction de régression générale (GRNN), développé par Specht (2002), est un réseau neuronal feedforward supervisé. Il s'agit d'un réseau neuronal probabiliste et d'une forme modifiée du réseau neuronal à fonction de base radiale (RBFN). Comme la montre la [figure III.7](#), le modèle GRNN se compose de quatre couches, à savoir la couche d'entrée, la couche cachée et la couche de sommation qui contient deux types de neurones différents : le premier type est le numérateur Nu et le second type est le dénominateur De , ainsi que la couche de sortie. Les neurones des couches d'entrée, cachée et de sommation sont entièrement connectés, tandis que le nœud de la couche de sortie n'est connecté qu'à certains nœuds de la couche de sommation (Cigizoglu & Alp, 2006; Diop et al., 2018; Ma et al., 2022).

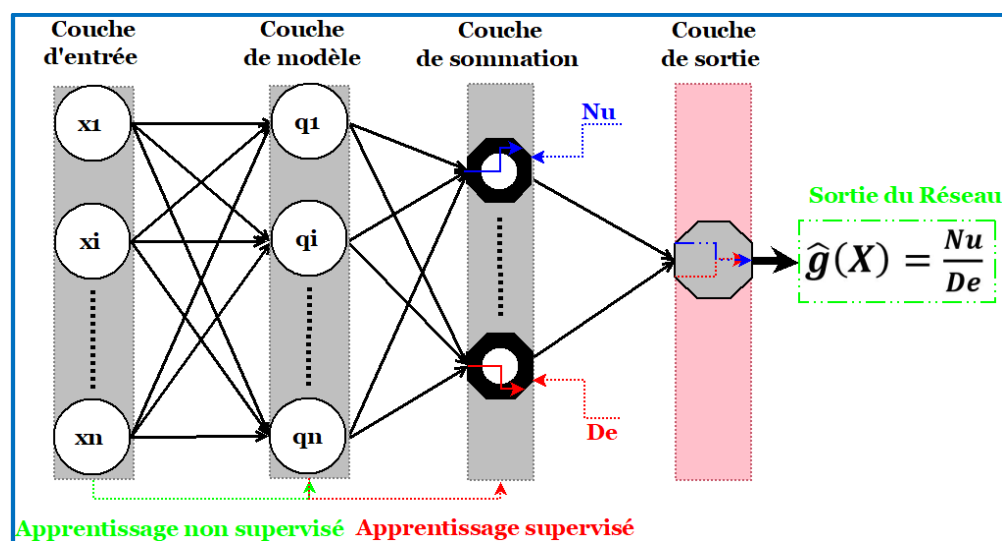


Figure III.7 Architecture d'un réseau de neurones GRNN

En supposant une variable aléatoire X et la fonction de densité de probabilité (PDF) conjointe, $g(X, y)$, la régression de Y sur x peut s'écrire sous la forme de l'équation (III. 2) :

$$f[y|X] = \frac{\int_{-\infty}^{+\infty} yg(X, y)dy}{\int_{-\infty}^{+\infty} g(X, y) dy} \quad (\text{III. 2})$$

Où $f[y|X]$ représente l'espérance conditionnelle de la sortie y correspondant à la variable d'entrée X . La PDF conjointe, $g(X, y)$, n'est généralement pas identifiée et peut être fournie par l'estimateur à noyau gaussien :

$$\hat{g}[X|y] = \frac{1}{n(2\pi)^{(q+1)/2}(\sigma)^{(q+1)}} \sum_{i=1}^n \exp\left[-\frac{D_i^2}{2\sigma^2}\right] \exp\left[-\frac{(y - y_i)^2}{2\sigma^2}\right] \quad (\text{III. 3})$$

$$D_i^2 = (X - X^i)^T (X - X^i) \quad (\text{III. 4})$$

Où n est le nombre de données observées, q est la taille de la variable aléatoire X et σ est le paramètre de lissage. Le modèle GRNN conventionnel peut être formulé à l'aide des équations (III. 2) et (III. 3) et la version générale du GRNN peut être réalisée comme suit :

$$\hat{g}(X) = \frac{\sum_{i=1}^n y^j \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \quad (\text{III. 5})$$

Avec

$$Nu(\text{Numérateur}) = \sum_{i=1}^n y^j \exp\left(-\frac{D_i^2}{2\sigma^2}\right) \quad (\text{III. 6})$$

Et

$$De(\text{Dénominateur}) = \sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right) \quad (\text{III. 7})$$

Dans laquelle $\hat{g}(X) = \frac{Nu}{De}$ est la fonction d'estimation de la probabilité pour $y(x)$.

III.4 Régression par processus gaussien (GPR)

Le modèle de régression par processus gaussien (GPR) est une technique de modélisation probabiliste robuste. Il s'agit d'un modèle de régression non paramétrique qui utilise le concept de lissage spatial (Rasmussen & Williams, 2006). L'avantage du cadre des processus gaussiens (GP) est que, sans connaissance préalable des données et des dépendances fonctionnelles, les modèles de bruit peuvent être formés à l'aide de transactions matricielles. Dans l'approche du maximum de vraisemblance ou l'approche bayésienne, les hyperparamètres qui contrôlent la forme du GP peuvent être prédits à partir des données, ce qui conduit finalement à une formation de "détermination automatique de la pertinence" (Zhu et al., 2019). Dans l'ensemble de données présentés, $S = (a_i, z_i)_{i=1}^n$, avec a comme données d'entrée pour le modèle GP et z comme données de sortie. Un GP définit la fonction C comme une distribution gaussienne conjointe, où $C(a_1), C(a_2), \dots, C(a_n)$, peut constituer un ensemble de variables aléatoires (Figure III.8).

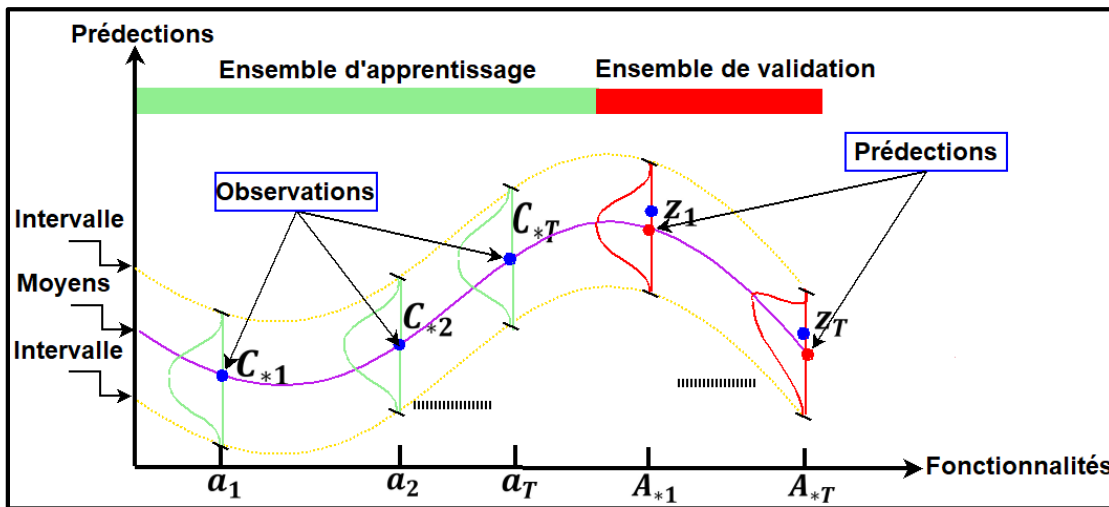


Figure III.8 Le modèle de régression par processus gaussien (GPR)

De plus, un GP est entièrement défini par sa fonction moyenne $m(a)$ et sa fonction de $k = (a, a')$.

Par conséquent, le GP s'écrit comme suit :

$$C(a) \sim GP(m(a), k = (a, a')) \tag{III. 8}$$

Pour tenir compte du bruit dans la valeur mesurée z , nous pouvons construire le modèle standard du problème GPR (Karbasi, 2018; Rasmussen & Williams, 2006):

$$z_i = C(a_i) + \epsilon_i, i = 1 \dots, n \quad (\text{III. 9})$$

Le bruit ϵ suit une distribution gaussienne indépendante et identiquement distribuée, de moyenne zéro et de variance σ^2 :

$$\epsilon \sim (0, \sigma^2) \quad (\text{III. 10})$$

III.5 Machine d'apprentissage extrême (ELM)

Une machine d'apprentissage extrême (ELM) est un algorithme de formation pour un réseau neuronal feedforward à couche unique (SLFN) introduit par (Huang, Chen, et al., 2006; Huang, Zhu, et al., 2006). L'ELM a été proposé dans le but d'améliorer la vitesse de l'algorithme de formation pour le réseau neuronal artificiel (RNA) autonome et de le rendre très rapide. Cette amélioration a été obtenue en divisant les paramètres du modèle en deux parties (Figure III.9) : (i) les poids et les biais reliant les neurones d'entrée aux neurones cachés sont initialisés de manière aléatoire et (ii) les poids et les biais reliant les couches cachées aux couches de sortie sont déterminés analytiquement à l'aide d'un algorithme prédéfini minimisant la fonction de perte carrée. Étant donné un ensemble de données d'apprentissage U comportant un ensemble d'entrées x_i et de sorties y_i avec une taille d'échantillon N comme suit (Li et al., 2021; Ren et al., 2021; Zehai et al., 2021):

$$U = [(x_i, y_i)] \in R^N \quad (\text{III. 11})$$

$$x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{iL}) \quad (\text{III. 12})$$

$$y_i = (y_{i1}, y_{i2}, y_{i3}, \dots, y_{iM}) \quad (\text{III. 13})$$

Où x et y sont les variables d'entrée et de sortie, N est le nombre d'échantillons d'apprentissage, L est la dimension de l'entrée et M est la dimension de la sortie. L'expression mathématique de l'algorithme ELM peut être exprimée comme suit (Sun & Yang, 2022; Wang et al., 2021):

$$y_i = f(x_i) = \sum_k^Z W_k \left[g \cdot \sum_{j=1}^L W_{jk} \times x_{ij} + \beta_k \right] \quad (\text{III. 14})$$

Où z désigne le nombre de neurones cachés, g correspond à la fonction d'activation sigmoïde ; W_{jk} est la matrice de poids entre les neurones d'entrée et les neurones cachés, W_k est le poids de sortie

entre les neurones cachés et les neurones de sortie, β_k le biais du neurone caché, et x_{ij} est la variable d'entrée i liée au neurone caché j . La couche cachée H peut être exprimée comme suit (Sun & Yang, 2022; Wang et al., 2021; Zehai et al., 2021):

$$H = \left\{ \begin{array}{ccc} \left[g \cdot \sum_{j=1}^L W_{j1} \times x_{1j} + \beta_1 \right] & \cdots & \left[g \cdot \sum_{j=1}^L W_{jz} \times x_{1j} + \beta_z \right] \\ & \ddots & \\ \left[g \cdot \sum_{j=1}^L W_{j1} \times x_{nj} + \beta_1 \right] & \cdots & \left[g \cdot \sum_{j=1}^L W_{jz} \times x_{nj} + \beta_z \right] \end{array} \right\} \quad (III. 15)$$

Les poids de sortie du modèle ELM peuvent être calculés comme suit (Song et al., 2021) :

$$\beta = H^+ Y \quad (III. 16)$$

Où H^+ désigne l'inverse généralisé de Moore-Penrose de la matrice H (Sun & Yang, 2022; Wang et al., 2021; Zehai et al., 2021).

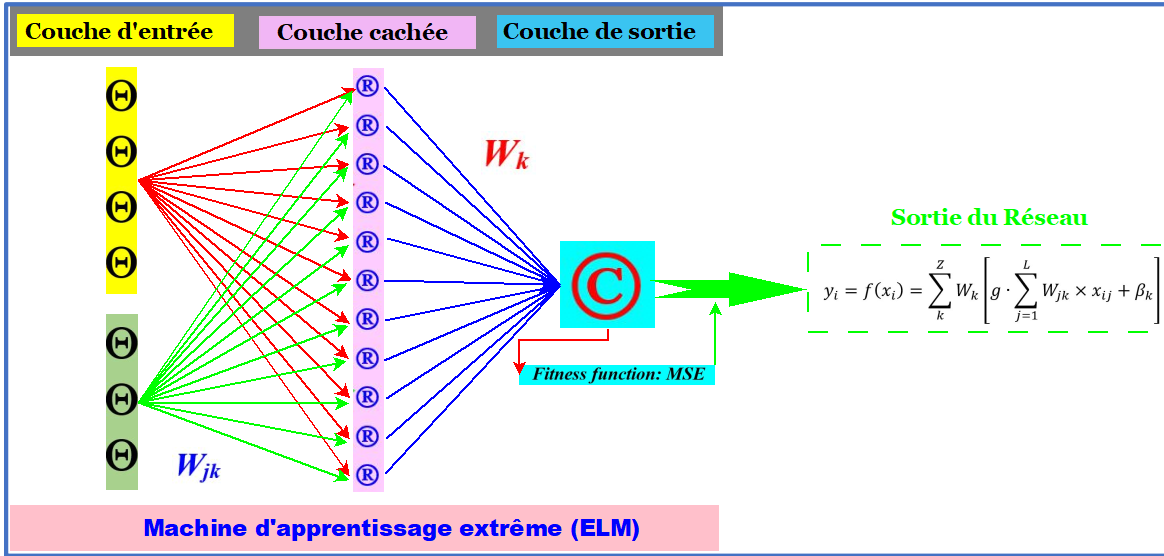


Figure III.9 Structure du modèle d'une machine d'apprentissage extrême (ELM)

III.6 Un réseau à mémoire à long terme (LSTM)

Les modèles universels à base d'ANN ne peuvent pas utiliser les informations acquises lors du pas de temps précédent pour modéliser le problème du pas de temps actuel. Ce phénomène constitue

l'une des principales lacunes des modèles universels basés sur les réseaux neuronaux. Un réseau neuronal récurrent (RNN) tente d'apporter une solution à ce problème en utilisant les informations transmises d'une étape à l'étape suivante du système. Le RNN associe donc les informations antérieures au problème actuel. En outre, l'utilisation d'informations antérieures peut faciliter la compréhension du problème actuel ([Abdel-Nasser & Mahmoud, 2019](#)).

Un RNN a été un choix précieux pour la modélisation des dynamiques temporelles et séquentielles. Cependant, étant donné qu'il a des problèmes de disparition et d'explosion, il est difficile pour un RNN universel d'entraîner la dynamique à long terme ([Gers & Schmidhuber, 2001](#)). ([Hochreiter & Schmidhuber, 1997](#)) ont proposé la mémoire à long terme (LSTM), une version étendue du réseau neuronal récurrent standard (RNN), pour améliorer la mémoire en augmentant les interactions supplémentaires par module (ou cellule). Étant donné que la LSTM combine une dynamique à temps discret dans le système interne du RNN, la LSTM est plus performante pour trouver les dépendances à long terme ([Elman, 1990](#)). En outre, la LSTM est un système récurrent puissant conçu pour surmonter les problèmes d'explosion du gradient lors de l'apprentissage de la fiabilité à long terme. Globalement, ce problème peut être évité en utilisant un carrousel d'erreurs constantes (CEC) ([Fischer & Krauss, 2018](#); [Sherstinsky, 2020](#); [Van Houdt et al., 2020](#)).

La théorie principale de LSTM peut être définie comme une cellule de mémoire qui peut contenir l'information organisée dans le temps par des unités de porte spéciales. Les niveaux cachés de la LSTM comprennent des architectures plus complexes et plus difficiles, et caractérisent la réponse de l'information comme une série de modules simples répétés par rapport à ceux du RNN standard. En outre, l'architecture de la LSTM se compose d'un bloc de mémoire avec des portes d'entrée, d'oubli et de sortie ([Cheng et al., 2017](#); [Han et al., 2019](#); [Zakhrouf et al., 2023](#)). La porte d'entrée contrôle le signe d'activation dans la cellule de mémoire. Cette division de la LSTM forme la situation dans laquelle l'information contenue doit être améliorée. La porte d'oubli vise à réorganiser les cellules de mémoire en ignorant les données précédentes lorsque les parties spéciales de l'état de la cellule doivent être modifiées par des informations plus récentes à l'aide d'une fonction d'activation. En outre, la porte de sortie forme et choisit les données qui seront diffusées pour transférer les activations au dernier nœud du niveau de sortie ([Hua et al., 2019](#); [Shahid et al., 2020](#)). La conception du modèle LSTM est indiquée à la [figure III.10](#).

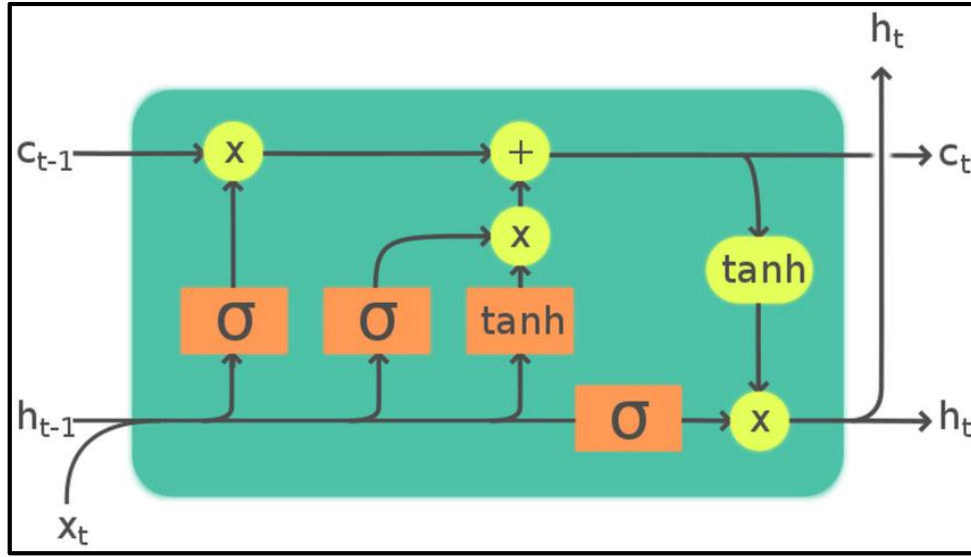


Figure III.10 la structure de la modélisation LSTM (Bakhshi Ostadkalayeh et al., 2023)

Soit $X = \{X_1, X_2, \dots, X_N\}$ représente une entrée séquentielle et $y = \{y_1, y_2, \dots, y_N\}$ la chaîne de sortie, la porte des oubliés est une condition principale qui affiche si les données actuelles doivent être oubliées ou remémorées. En fonction d'une période de temps t , il est possible d'effectuer le calcul suivant (Fang et al., 2021) :

$$f_t = \sigma(W_{fX}X_t + W_{fh}h_{t-1} + b_f) \quad (\text{III. 17})$$

Où W_{fX} et W_{fh} sont respectivement la matrice de poids de l'oubli et la matrice de poids de l'oubli-caché, b_f est le biais de la porte de l'oubli, et σ est la fonction sigmoïde. La porte d'entrée i_t conclut la mise à jour des données et \tilde{C}_t mémorise les données actualisées :

$$i_t = \sigma(W_{iX}X_t + W_{ih}h_{t-1} + b_i) \quad (\text{III. 18})$$

$$\tilde{C}_t = \tanh(W_{CX}X_t + W_{Ch}h_{t-1} + b_C) \quad (\text{III. 19})$$

Où W_{iX} et W_{Ch} représentent la matrice de poids, les vecteurs de biais de la porte d'entrée et l'état de la cellule de mise à jour sont représentés par b_i et b_C , en série. Après cela, le nouvel état de la cellule de mémoire c_t est mis à jour comme suit :

$$c_t = f_t * c_{t-1} + i_t * \tilde{C}_t \quad (\text{III. 20})$$

Où c_{t-1} est l'état antérieur de la cellule de mémoire et où $*$ indique le produit élément par élément. À la fin, la porte de sortie gère les activations de sortie. La couche cachée envoyée à un stade ultérieur est spécifiée comme suit :

$$h_t = o_t * \tanh(c_t) \quad (\text{III. 21})$$

$$o_t = \sigma(W_{oX}X_t + W_{oh}h_{t-1} + b_o) \quad (\text{III. 22})$$

Où la matrice de poids de sortie apparaît via W_{oX} , W_{oh} est la matrice de poids cachée en sortie et b_o est le biais de la porte de sortie.

III.7 Optimisation des modèles ELM basés sur des algorithmes métaheuristiques avancés

La méthode moderne des algorithmes d'optimisation s'inspire la plupart du temps de la nature, et l'algorithme qui en découle peut-être très varié. Néanmoins, tous les algorithmes tendent à être utilisés avec certaines des caractéristiques essentielles pour le calcul de la principale formule modernisée. Un très grand nombre d'algorithmes heuristiques et métaheuristiques (Lones, 2011) ont été dérivés sur la base des caractéristiques des systèmes biologiques et physiques de la nature. Par exemple, l'optimisation des essaims de particules (Kennedy & Eberhart, 1995) a été développée sur la base du comportement en essaim des oiseaux et des poissons, tandis que le recuit simulé (Kirkpatrick et al., 1983) est basé sur le processus de trempe du métal.

III.7.1 Algorithme d'optimisateur des chauves-souris (Bat)

En 2010, Xin-She Yang a développé un algorithme métaheuristique inspiré de la nature des chauves-souris, appelé Bat Algorithm (Yang, 2010). Les micro-chauves-souris jouent un rôle crucial dans l'élaboration des principales caractéristiques de cet algorithme. Dans l'élaboration des principales caractéristiques de cet algorithme (Griffin, 2002). Les deux principaux paramètres qui sont utilisés dans cet algorithme sont les taux d'impulsion et l'émission, et les valeurs de ces deux paramètres peuvent être modifiées. Peuvent être ajustées. L'algorithme Bat utilise également la méthode de réglage de la fréquence, pour élargir la variété des solutions présentes dans la population, même si c'est en en même temps. Il utilise un zoom automatique qui tente d'ajuster l'exploration et l'exploitation tout au long du processus en imitant la méthode de l'algorithme Bat. L'exploration et l'exploitation tout au long du processus en imitant la variation du flux des

battements de cœur et du volume sonore des chauves-souris pendant la chasse de leurs proies. L'écholocation des micro-chauves-souris est une capacité unique qui leur permet de repérer leurs proies et de séparer les différentes espèces de chauves-souris. Différentes sortes d'insectes dans l'obscurité totale. Ces excellentes caractéristiques rendent l'algorithme plus efficace avec un excellent démarrage rapide.

III.7.2 Les caractéristiques d'écholocation des micro-chauves-souris

- Toutes les micro chauves-souris utilisent l'écholocation pour trouver le sujet de leur choix (nourriture ou proie) et les obstacles. Ce comportement leur permet également de percevoir la distance et d'acquérir la capacité de distinguer la nourriture, les proies ou les obstacles.
- Les chauves-souris volent aléatoirement avec une vitesse à n'importe quelle position avec une fréquence fixe, en variant la longueur d'onde et l'intensité sonore pour rechercher des proies (Griffin, 2002).
- L'intensité sonore des micro-chauves-souris peut varier et passer d'une valeur positive importante à une valeur constante minimale (Griffin, 2002).
- La gamme de fréquences correspond à une gamme de longueurs d'onde.

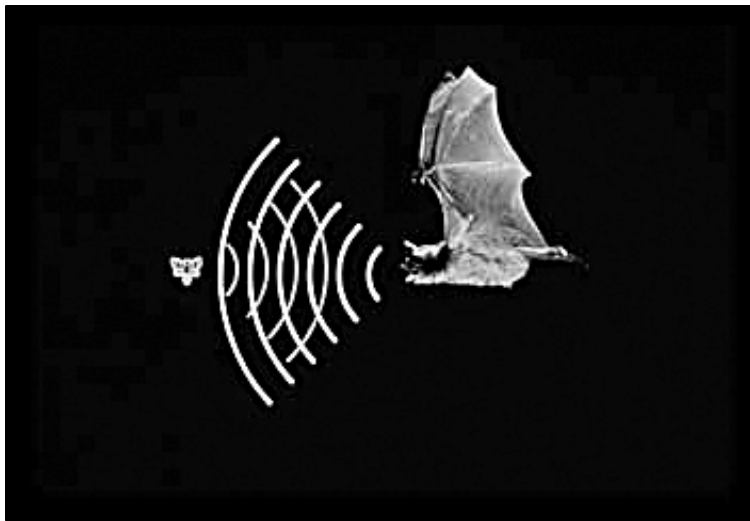


Figure III.11 Exemple d'écholocation (Charbonneau & Tremblay, 2010)

III.7.3 Mouvement des chauves-souris virtuelles

Les chauves-souris volent aléatoirement avec une vitesse v_i à la position x_i avec une fréquence fixe f_{min} , en faisant varier la longueur d'onde λ et l'intensité sonore A_0 pour chercher des proies. Elles peuvent ajuster automatiquement la longueur d'onde (ou la fréquence) de leurs impulsions

émises. Longueur d'onde (ou fréquence) de leurs impulsions émises et ajuster le taux d'émission des impulsions $\gamma \in [0, 1]$, en fonction de la proximité de leur cible.

Dans les simulations, les chauves-souris virtuelles ont leurs positions x_i et leurs vitesses v_i dans un espace de recherche d - dimensionnel au pas de temps t . dimensionnel au pas de temps t ; qui sont données par (Yang, 2010):

$$f_i = f_{min} + (f_{max} - f_{min})\beta \quad (\text{III. 23})$$

$$v_i^t = v_i^{t-1} + (x_i^{t-1} - x_*)f_i \quad (\text{III. 24})$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (\text{III. 25})$$

Où $\beta \in [0, 1]$ est un vecteur aléatoire tiré d'une distribution uniforme. Ici x_* est le meilleur emplacement global actuel (solution) qui est situé après avoir comparé toutes les solutions parmi tous les n chauves-souris.

Après avoir sélectionné la meilleure solution actuelle parmi les autres solutions, l'équation (III. 26) est utilisée pour générer une nouvelle solution pour chaque chauve-souris (Yang, 2010).

$$x_{new} = x_{old} + \varepsilon A^t \quad (\text{III. 26})$$

L'intensité sonore moyenne de toutes les chauves-souris $A_t = \langle A_i^t \rangle$ est et $\varepsilon \in [-1, 1]$ désigne un nombre aléatoire.

III.7.4 Pseudo-code de l'algorithme Bat

Sur la base d'approximations et d'une idéalisation des caractéristiques de l'écholocalisation, les étapes de base de l'algorithme de la chauve-souris (BA) peuvent être résumées sous forme de pseudo-code comme suit (Figure III.12) (Yang, 2010) :

L'algorithme des chauves-souris	
1	Définir la fonction objective : $f(x), x = (x_1, \dots, x_d)^T$
2	Initialiser la population de chauves-souris $x_i (i = 1, 2, \dots, n)$ et v_i
3	Définir la fréquence des impulsions f_i à x_i
4	Tant que ($t < \text{Nombre maximal d'itérations}$) fait
5	Générer de nouvelles solutions par l'ajustement des fréquences, et mettre à jour les vitesses et les positions/solutions
6	Si ($rand > r_i$)
7	Sélectionner une solution parmi les meilleures solutions
8	Générer une solution locale autour de la meilleure solution sélectionnée
9	Fin si
10	Générer une nouvelle solution en volant aléatoirement
11	Si ($rand > A_i$) et $f(x_i)$
12	Accepter de nouvelles solutions
13	Fin si
14	Classer les chauves-souris et trouver la meilleure solution courante x_*
15	Fin tant que
16	Résultats et visualisations.

Figure III.12 Pseudo code de l'algorithme des chauves-souris (Yang, 2010)

III.7.5 Efficacité de l'algorithme Bat

La réalisation de l'algorithme des chauves-souris répond à différents objectifs. Dans de nombreuses applications, l'algorithme des chauves-souris permet d'atteindre un certain objectif en utilisant une quantité minimale de matériaux. Les problèmes d'optimisation non linéaire ne sont pas aussi difficiles à résoudre que les problèmes d'optimisation topologique. Les résultats suggèrent que l'algorithme Bat s'est avéré capable de résoudre les problèmes d'optimisation non linéaire et topologique. À partir des caractéristiques clés et des équations de mise à jour, l'efficacité de l'algorithme Bat peut-être résumée dans les points suivants (Yang & He, 2013):

III.7.5.1 Réglage de la fréquence

Les algorithmes Bat (Yang, 2010) utilisent l'écholocation et l'accord de fréquence pour résoudre différents types de problèmes. En réalité, il n'est pas possible d'utiliser directement l'écholocation, c'est pourquoi les variations de fréquence sont utilisées pour résoudre les

problèmes. Les variations de fréquence sont donc utilisées pour résoudre les problèmes. L'optimisation par essaims de particules et la recherche harmonique possèdent certaines fonctionnalités essentielles qui sont tirées des caractéristiques de l'accord de fréquence. Ces caractéristiques uniques placent l'algorithme Bat dans une position plus avancée que les autres techniques d'intelligence en essaim.

III.7.5.2 Zoom Automatique

L'algorithme Bat (Yang, 2010) peut effectuer un zoom automatique, ce qui est utile pour trouver des solutions prometteuses en zoomant sur la région. Au début des itérations, l'algorithme Bat a un taux de convergence rapide grâce au zoom automatique qui permet de passer d'une région à l'autre, il a un taux de convergence rapide grâce au zoom automatique qui aide à passer de l'exploration à l'exploitation locale intensive, ce qui donne à l'algorithme Bat un taux de convergence élevé. L'exploration à l'exploitation locale intensive, ce qui donne à l'algorithme Bat un avantage significatif sur les autres algorithmes métaheuristiques.

III.7.5.3 Contrôle des paramètres

L'algorithme Bat (Yang, 2010) a plus de paramètres à contrôler. Certains de ces paramètres peuvent être initialisés avec des nombres aléatoires, mais d'autres paramètres sont définis en fonction de l'expérience. Parfois, un plus grand nombre de paramètres rend l'utilisation de l'algorithme Bat difficile pour les débutants, mais il est utile pour augmenter l'efficacité de l'algorithme Bat. Les valeurs des paramètres varient à chaque itération. Lorsque l'algorithme s'approche d'une solution optimale, cette technique de contrôle des paramètres solution optimale, cette technique de contrôle des paramètres permet de passer automatiquement de l'exploration à l'exploitation. De l'exploration à l'exploitation automatiquement.

III.7.5.4 Le modèle hybride Bat-ELM

Les algorithmes d'optimisation métaheuristiques (MOA) inspirés des processus naturels des insectes et des animaux pour améliorer les performances des modèles d'apprentissage automatique (ML) ont suscité un intérêt mondial au cours des dernières années, et l'ELM est l'un des nombreux modèles qui ont été combinés avec les MOA. Parmi les méthodes ELM hybrides, nous pouvons citer : l'ELM basé sur l'algorithme de la luciole (FFA-ELM) (Akhavan-Amjadi, 2020), l'ELM basé

sur l'algorithme de la colonie d'abeilles artificielle (ABC-ELM) (Alshamiri et al., 2016), l'ELM basé sur l'optimisation par essaims de particules (PSO-ELM) (Alshamiri et al., 2016), l'ELM basé sur l'optimisation hybride améliorée de la mite et de la flamme (HMFA-ELM) (Zhang et al., 2021) et l'ELM basé sur l'optimisation par colonies de fourmis (AC-ELM) (Sheoran et al., 2020).

Pour améliorer les performances du modèle ELM, nous avons adopté dans la présente étude l'algorithme Bat, ce qui peut être réalisé en prenant en compte les paramètres du modèle, c'est-à-dire les poids et les biais. Comme la montre la **figure III.13**, chaque chauve-souris peut être considérée comme une solution des paramètres ELM et sélectionnée sur la base de l'évolution de la fonction d'aptitude calculée à partir des sorties mesurées et prédites. L'organigramme du modèle hybride Bat-ELM est présenté dans la **figure III.13**.

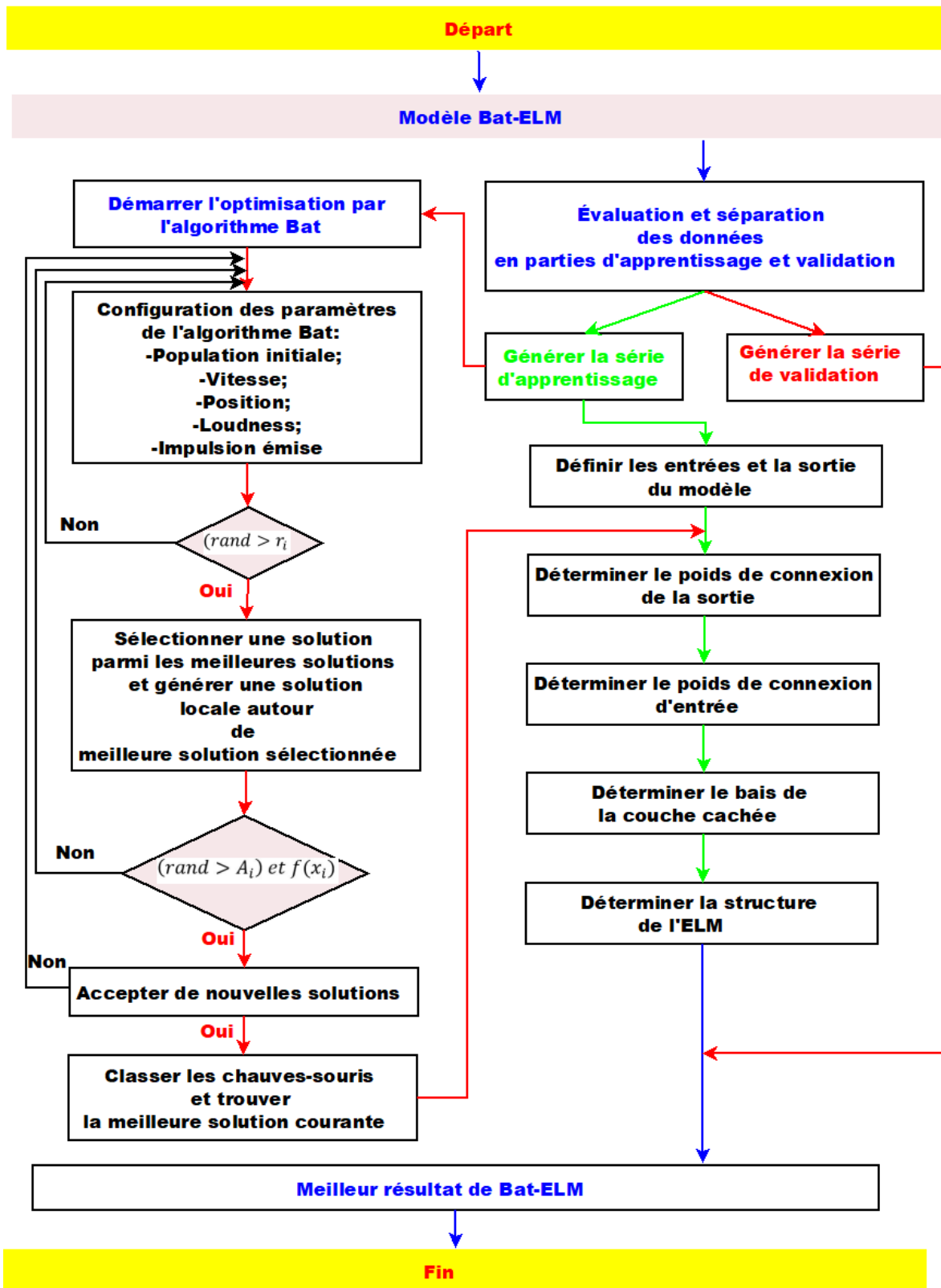


Figure III.13 Organigramme du modèles hybride Bat-ELM proposé

III.7.6 Les Machines A Vecteurs De Support (SVM)

III.7.6.1 Introduction

Les machines à vecteurs de support sont des machines d'apprentissage supervisé basées sur la théorie de l'apprentissage statistique qui peuvent être utilisées pour la reconnaissance des formes et la régression introduite par (Cortes & Vapnik, 1995). La théorie de l'apprentissage statistique permet d'identifier assez précisément les facteurs à prendre en compte pour apprendre avec succès certains types d'algorithmes simples. Toutefois, les applications du monde réel nécessitent généralement des modèles et des algorithmes plus complexes (tels que les réseaux neuronaux), ce qui les rend beaucoup plus difficiles à analyser sur le plan théorique. En effet, un SVM peut être considéré comme un algorithme linéaire dans un espace à haute dimension. Le SVM fonctionne par l'identification de l'hyperplan qui sépare le mieux les deux classes de données. L'hyperplan est une ligne ou un plan dans un espace à N dimensions qui divise les données en deux régions, chaque région contenant tous les points de données d'une classe (Figure III.14). L'algorithme SVM recherche l'hyperplan qui présente la marge maximale, c'est-à-dire la distance entre l'hyperplan et les points de données les plus proches de chaque classe. Les points de données les plus proches de l'hyperplan sont appelés vecteurs de support. Ces points sont les plus importants pour le modèle SVM, car ils déterminent la position de l'hyperplan.

Le SVM est un puissant algorithme d'apprentissage automatique qui peut être utilisé pour résoudre une grande variété de problèmes. Il est particulièrement bien adapté aux problèmes où les données sont linéairement séparables, mais il peut également être utilisé pour des données non linéairement séparables à l'aide de fonctions noyau.

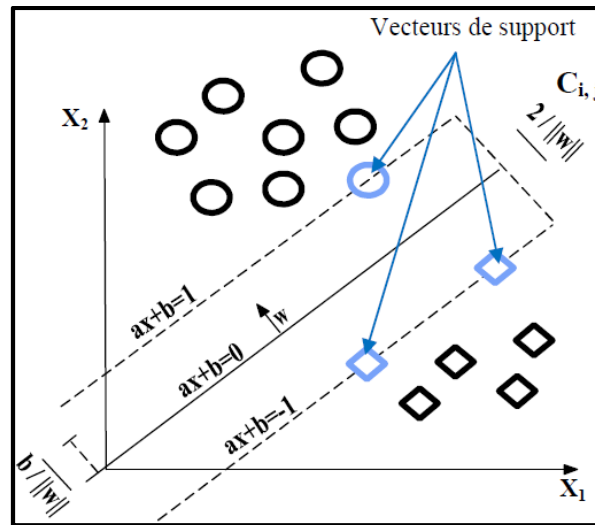


Figure III.14 Un exemple de classification par SVM (Bouaziz, 2017)

III.7.6.2 Les SVM pour la Régression (SVR)

(Smola, 1996) a ajouté des fractions de régression à la version originale du modèle SVM et a créé une nouvelle version du SVM pour résoudre les problèmes de régression et de prévision appelée régression à vecteur de support (SVR). L'idée principale du modèle SVR est de trouver la fonction de dépendance $f(x)$ qui utilise toutes les données (inputs X_i / output y_i) en ce qui concerne la précision la plus minimale possible (Smola, 1996). La figure III.15 donne la configuration structurelle du modèle SVR.

La fonction de régression du SVR est exprimée sous la forme :

$$f(x) = w * \phi(x) + b \quad (\text{III. 27})$$

Où ; f est la fonction de régression, ϕ est la fonction de transfert, w et b sont respectivement le poids et le biais. Le problème de régression peut être déclaré comme ;

$$\text{Minimiser} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (\text{III. 28})$$

$$\text{En fonction de} \quad \begin{cases} y_i - f(x) \leq \varepsilon + \xi_i \\ f(x) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, 3 \dots, N \end{cases} \quad (\text{III. 29})$$

Où : C est le paramètre de pénalité, ξ_i et ξ_i^* correspondent aux variables de relâchement et ε est la valeur limite.

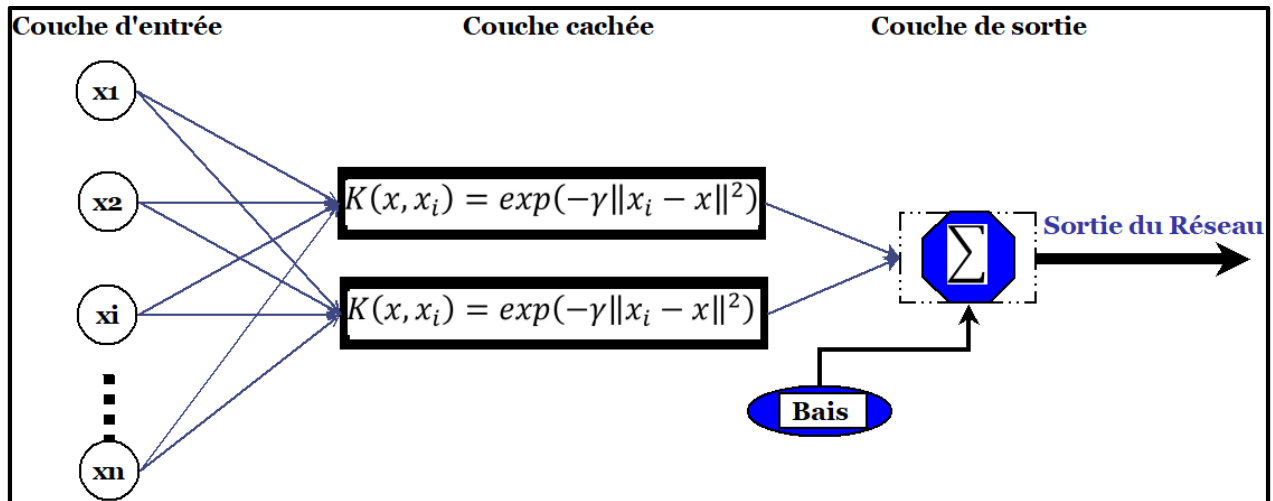


Figure III.15 Structure d'un modèle de SVR.

En utilisant les multiples lagrangiens, le problème d'optimisation se convertit largement en programmation quadratique et une solution de fonction de régression non linéaire peut être donnée comme suit :

$$f(x) = \sum_{i=1}^N (a_i - a_i^*) K(x, x_i) + b \quad (\text{III. 30})$$

Où $K(x, x_i)$ est la fonction de noyau et a_i, a_i^* sont les variables doubles

Il existe plusieurs fonctions du noyau, telles que la fonction de base linéaire, polynomiale, fonction de base radiale (RBF) et sigmoïde. Les fonctions disponibles du noyau peuvent être résumées comme suit :

- Linéaire $K(x, x_i) = x_i x$
- Polynomiale $K(x, x_i) = (l + x_i x)^d$
- Fonction de base radiale $K(x, x_i) = \exp(-\gamma \|x_i - x\|^2)$
- Sigmoïde $K(x, x_i) = \tanh(\alpha x_i x + \delta)^2$

Où : d, γ, α et δ sont les paramètres de la fonction de noyau.

III.8 Transformation en ondelettes

Grossmann et Morlet ont introduit la transformée en ondelettes (WT) en 1984 (Meyers et al., 1993). Elle à été inspiré des travaux de Jean Morlet, un géophysicien qui avait mis au point une fonction ressemblant à une ondelette, appelée "ondelette de Marr" (Kronland-Martinet, 1988). L'ondelette

de Grossmann-Morlet est une ondelette plus générale qui peut être utilisée pour analyser une plus grande variété de signaux. WT est une outille puissante pour analyser les signaux qui ont à la fois des composantes temporelles et fréquentielles. Il s'agit d'une généralisation de la transformée de Fourier (FT), qui décompose un signal en ses fréquences constitutives (Ho et al., 1995). Cependant, la transformée en ondelettes présente l'avantage de pouvoir localiser à la fois le temps et la fréquence des composantes d'un signal. La transformée en ondelettes a été utilisée dans un large éventail d'applications, notamment l'analyse des données, le traitement des signaux, l'imagerie et la compression des données.

La transformée en ondelettes permet de capturer les propriétés des séries temporelles qui offrent un avantage de maîtriser le problème de non-stationnarité rencontré par l'intelligence artificielle simple ou les modèles linéaires (Shoaib et al., 2014). Le WT décrit une fonction oscillatoire de longueur finie. À mesure que la transformation est calculée pour chaque composante spectrale, la largeur de la fonction d'ondelette varie, ce qui permet à la transformation en ondelettes d'analyser des données non stationnaires. La taille variable de la fenêtre dans la transformée en ondelettes permet d'obtenir des informations plus précises sur le temps et la fréquence d'un signal. Dans la transformation en ondelettes, la fonction d'ondelettes est utilisée pour décomposer les données de la série temporelle en différentes composantes à différents niveaux de résolution (Manimaran et al., 2005). La fonction d'ondelette $\psi(t)$, appelée ondelette mère, a une énergie finie et est mathématiquement définie comme suit :

$$\int_{-\infty}^{+\infty} \Psi(t) dt = 0 \quad (\text{III. 31})$$

Où $\Psi_{ab}(t)$ est la fonction d'ondelette et peut être obtenue par l'équation suivante :

$$\Psi_{ab}(t) = |a|^{-\frac{1}{2}} \Psi\left(\frac{t-b}{a}\right) \quad a, b \in R, A \neq 0 \quad (\text{III. 32})$$

Les paramètres a et b sont des nombres réels ; a est le paramètre d'échelle ou de fréquence ; b est le paramètre de translation. Par conséquent, WT est une fonction de deux paramètres a et b . Le paramètre a est interprété comme un facteur d'expansion ($a > 1$) ou de contraction ($a < 1$) de la fonction d'ondelette $\Psi(t)$ correspondant à différentes échelles (Zhang & Zhang, 2019). Ce paramètre b est interprété comme une translation ou un décalage temporel de la fonction $\Psi(t)$. Pour analyser les données temporelles, il est possible de procéder soit par translation (déplacement

de l'ondelette à différents endroits le long de l'axe temporel des données), soit par contraction ou étirement de l'ondelette (appelée mise à l'échelle ou dilatation) afin de transformer les données considérées.

III.8.1 Types de transformation en ondelettes

III.8.1.1 Transformée continue en ondelettes

CWT est l'outil d'analyse temps-fréquence le plus couramment utilisé, qui utilise une famille de fonctions d'ondelettes pour décomposer un signal dans le domaine temps-fréquence (Aguiar-Contraria & Soares, 2014; Grossmann et al., 1990). Elle hérite et développe l'idée de localisation de la STFT, mais contrairement à la STFT, CWT peut fournir une résolution temporelle élevée et une résolution basse fréquence dans les hautes fréquences, ainsi qu'une résolution haute fréquence et une résolution temporelle basse dans les basses fréquences en ajustant les paramètres d'échelle et de translation. Formellement, la CWT est définie comme suit :

$$CWT_{a,b}(t) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{a}} \Psi^* \left(\frac{t-b}{a} \right) dt \quad (\text{III. 33})$$

L'expression $CWT_{a,b}$ représente le coefficient d'ondelette. * correspond au conjugué complexe de la fonction, la fonction d'ondelette ou ondelette mère se référant à $\Psi(t)$. Toute la gamme de signaux est analysée par la fonction d'ondelette en utilisant deux paramètres, à savoir a et b . Les paramètres a et b sont respectivement appelés paramètres de dilatation (échelle) et de translation (position). Alors que $f(t)$ représente les données temporelles originales à transformer. L'ondelette $\Psi(t)$ et la fonction $f(t)$ de l'équation (III.33) sont intégrées sur l'ensemble de la plage de données temporelles, ce qui, d'un point de vue mathématique, s'appelle une convolution. En CWT, la fonction d'ondelette d'analyse $\Psi(t)$ compare les données temporelles à transformer à différentes échelles et à différents endroits (Sinha et al., 2005).

III.8.1.2 Transformée discrète en ondelettes

La transformée en ondelettes discrète (DWT) a été introduite pour la première fois à la fin des années 1980 par (Daubechies, 1988) et (Mallat, 1989). DWT est une implémentation utilisant un ensemble discret d'échelles et de translations d'ondelettes obéissant à certaines règles. Cette transformée décompose le signal en un ensemble d'ondelettes mutuellement orthogonales.

L'analyse DWT consiste à effectuer une comparaison locale d'un signal avec des motifs d'ondelettes, à la manière d'un microscope mathématique permettant de zoomer sur le signal à différentes échelles. Les bases orthonormales d'ondelettes permettent une analyse multirésolution basée sur des algorithmes de décomposition et de reconstruction très rapides pour un signal discret fini (Daubechies, 1992). Il s'agit de fonctions produites par le processus de dilatation et de translation d'une fonction d'ondelette mère $\Psi(t)$, qui est donnée comme suit :

$$\Psi_{m,n} \left(\frac{t-b}{a} \right) = a_0^{-\frac{m}{2}} \left(\frac{t-nb_0a_0^m}{a_0^m} \right) \quad (\text{III. 34})$$

Où m et n sont des entiers qui régissent respectivement l'échelle/la dilatation et la translation de l'ondelette ; a_0 est un pas d'échelle fin spécifié supérieur à 1 ; et b_0 est le paramètre d'emplacement et doit être supérieur à zéro.

III.8.2 Transformée en ondelettes discrète à maximum de recouvrement

La transformée en ondelettes discrète à recouvrement maximal (MODWT) présente des caractéristiques similaires à celles de la transformée en ondelettes discrète (DWT) en ce sens que des filtres passe-bas et passe-haut sont appliqués au signal d'entrée à chaque niveau (Lark & Webster, 2001). Toutefois, la transformée MODWT ne décime pas les coefficients et le nombre de coefficients d'ondelettes et de mise à l'échelle est identique au nombre d'observations d'échantillons à chaque niveau de la transformée (Percival & Walden, 2000). En d'autres termes, les coefficients MODWT sont le résultat d'une simple modification de l'algorithme pyramidal utilisé dans le calcul des coefficients DWT, qui consiste à ne pas réduire l'échantillonnage de la sortie à chaque échelle et à insérer des zéros parmi les coefficients dans les filtres d'échelonnement et d'ondelettes (Barzegar et al., 2021). C'est pourquoi la TCDM est également appelée TCD non décimée, TCD stationnaire, TCD invariante en translation et TCD invariante dans le temps.

Le MODWT décompose la série temporelle X_t en une composante d'approximation ($A_{j,t}$) à l'aide d'un filtre passe-bas $\tilde{g}_{j,1} = g_{j,1}/2^{j/2}$ et en une composante de détail ($D_{j,t}$) à l'aide d'un filtre passe-haut ($\tilde{h}_{j,1} = h_{j,1}/2^{j/2}$) où $\tilde{g}_{j,1}$ et $\tilde{h}_{j,1}$ sont j^{th} de MODWT (Seo et al., 2017). Selon Percival et Walden (2000), le MODWT est donné par les équations ci-dessous :

$$X = \sum_{j=1}^L D_j + A_{j0}, \quad (\text{III. 35})$$

$$D_{j,t} = \sum_{l=1}^{n-1} \tilde{h}_{j,l}^0 W_{j,t+1 \bmod n'} \quad (\text{III. 36})$$

$$A_{j,t} = \sum_{l=1}^{n-1} \tilde{g}_{j,l}^0 V_{j,t+1 \bmod n'} \quad (\text{III. 37})$$

Le MODWT perd en orthogonalité et en efficacité de calcul. Le MODWT présente certains avantages par rapport au DWT :

- Le MODWT peut traiter n'importe quelle taille d'échantillon.
- Les coefficients lisses et détaillés de l'analyse multirésolution MODWT sont associés à des filtres à phase nulle.
- Elle est invariante par rapport à la transformation, puisqu'un déplacement du signal ne modifie pas la structure des coefficients de la transformée en ondelettes.
- Elle produit un estimateur de variance d'ondelette asymptotiquement plus efficace que la DWT.

III.8.2.1 Les modèles hybrides ondelettes-intelligence artificielle.

La présente étude propose une comparaison de quatre modèles ML et une évaluation de leurs performances dans le cadre de la modélisation et prédiction des séries temporelle (prédiction des débits journalière des cours d'eaux). Les quatre modèles seront respectivement le GPR, le LSTM, le GRNN et le MLPNN. Les performances des modèles seront évaluées selon deux scénarios : (i) modèles simples sans décomposition, c'est-à-dire GPR, LSTM, GRNN, MLPNN, (ii) modèles hybrides avec algorithme MODWT, c'est-à-dire GPR_MODWT, LSTM_MODWT, GRNN_MODWT, MLPNN_MODWT. L'organigramme du cadre de modélisation développé est résumé dans la [figure III.16](#).

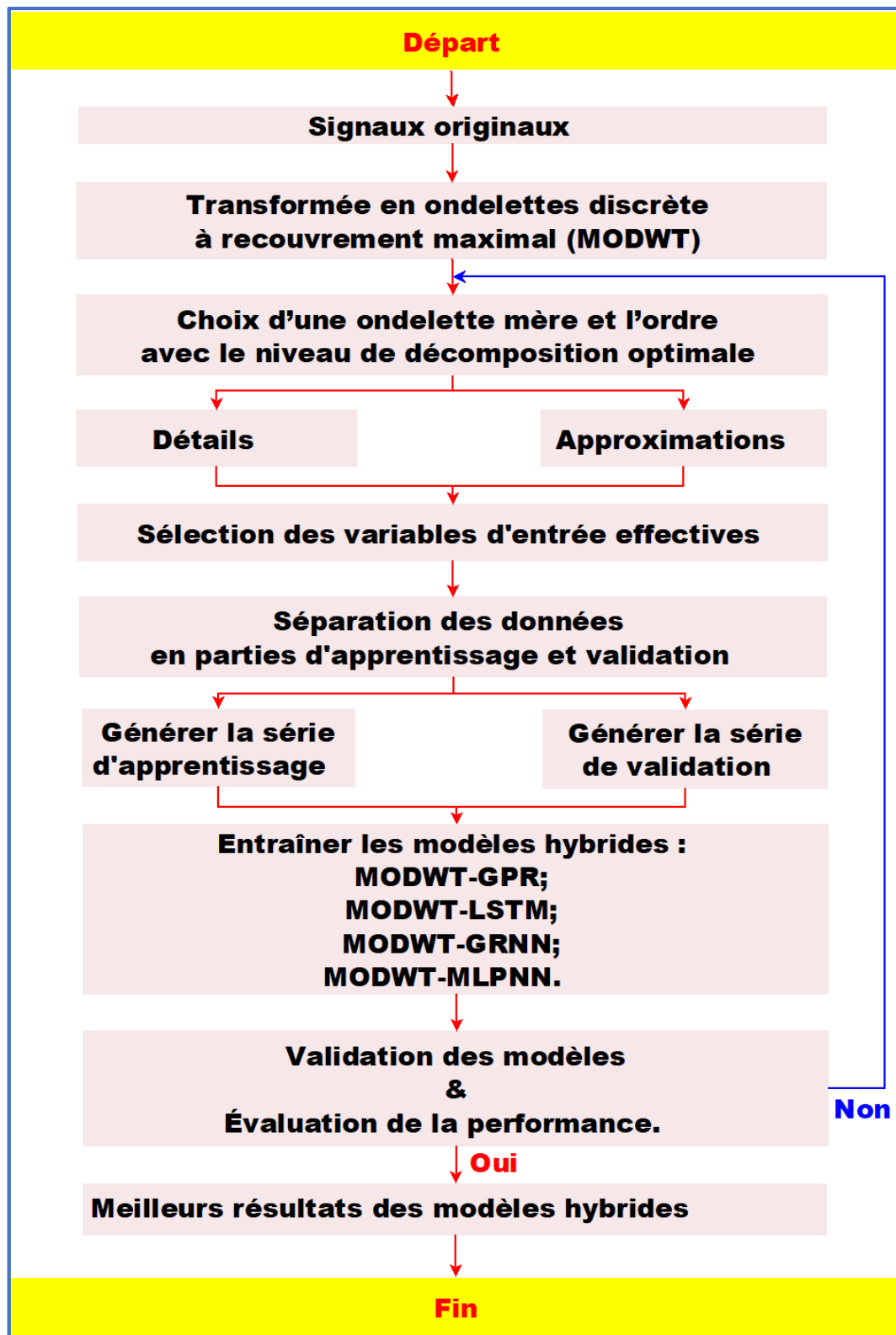


Figure III.16 Organigramme de la structure de modélisation proposée pour la prévision des débits journaliers.

Dans la modélisation du débit des cours d'eau à l'aide de MODWT, MODWT décompose les signaux d'entrée originaux (données quotidiennes sur le débit des cours d'eau et les précipitations) en sous-signaux (composantes de détail et d'approximation). Les sous-signaux sont ensuite utilisés comme entrées dans les modèles simples, GPR, LSTM, GRNN et MLPNN. Comme la montre la [figure III.16](#), les approches de modélisation de l'état des cours d'eau basées sur le MODWT consistent en un algorithme en trois étapes ([Seo et al., 2017](#)). L'algorithme est décrit comme suit :

- Étape 1. Décomposition des signaux d'entrée originaux en sous-signaux (composantes de détail et d'approximation) à l'aide de MODWT.
- Étape 2. Sélection des entrées efficaces parmi les sous-signaux.
- Étape 3. Apprentissage et test de modèles uniques, GPR, LSTM, GRNN et MLPNN, à l'aide des entrées effectives.

III.8.2.2 Choix d'une ondelette mère et l'ordre avec le niveau de décomposition optimale

Il existe de nombreuses familles d'ondelettes différentes, chacune avec ses propres avantages et inconvénients. Certaines familles d'ondelettes sont plus adaptées à la décomposition de signaux à haute fréquence, tandis que d'autres sont plus adaptées à la décomposition de signaux à basse fréquence. Certaines familles d'ondelettes sont également plus adaptées à la reconstruction de signaux avec un certain niveau de précision. D'après l'analyse de la littérature, la limitation de la DWT ou de la MODWT réside dans le bon choix de l'ondelette mère. Ainsi, dans cette recherche, les ondelettes mères les plus utilisées ont été employées et évaluées comme haar, Daubechies, Symlet, Coiflets et Fejer-Korovkin ([Barzegar et al., 2021](#); [Percival & Walden, 2000](#); [Seo et al., 2017](#)).

- **Ondelettes de Haar :**

Toute discussion sur les ondelettes commence par l'ondelette de Haar, la première et la plus simple. L'ondelette de Haar est discontinue et ressemble à une fonction en escalier ([Guf & Jiang, 1996](#)). Elle représente la même ondelette que l'ondelette db1 de Daubechies.

- **Ondelettes de Daubechies :**

Ingrid Daubechies, l'une des étoiles les plus brillantes du monde de la recherche sur les ondelettes, a inventé ce que l'on appelle les ondelettes orthonormées à support compact, rendant ainsi possible l'analyse par ondelettes discrètes ([Lina & Mayrand, 1995](#)). Les noms des ondelettes de la famille

Daubechies s'écrivent db_N , où N est l'ordre et db le "nom de famille" de l'ondelette. L'ondelette db_1 , comme mentionné ci-dessus, est la même que l'ondelette de Haar. Elles s'écrivent db_N , où $N = 1, 2, \dots, 45$

- **Ondelettes de Coiflets :**

Ces ondelettes sont conçues pour être plus localisées dans le domaine temporel que les ondelettes de Haar. Elles portent le nom de Jean-Pierre Antoine et Ingrid Daubechies, qui les ont introduites pour la première fois en 1991. Les Coiflets sont orthogonales, ce qui signifie qu'elles sont linéairement indépendantes. Elles sont également plus localisées dans le domaine temporel que les ondelettes de Haar, ce qui en fait un bon choix pour les applications où la localisation temporelle est importante, comme le traitement des signaux (Zaghden et al., 2006). Elles s'écrivent $coif_N$, où $N = 1, 2, \dots, 5$.

- **Ondelettes de Symlets :**

Les symlets sont des ondelettes presque symétriques proposées par Daubechies pour modifier la famille db . Les propriétés des deux familles d'ondelettes sont similaires (Chavan et al., 2011). Elles s'écrivent sym_N , où $N = 2, 3, \dots, 45$.

- **Les ondelettes de Fejer-Korovkin :**

Sont une famille d'ondelettes basées sur les noyaux de Fejer. Les noyaux de Fejer sont une famille de fonctions continues, à support compact et orthogonales (Hidayat et al., 2018). Les ondelettes de Fejer-Korovkin sont construites par dilatation et translation des noyaux de Fejer. Elles ont un support compact, sont orthogonales et ont de bonnes propriétés d'approximation. Elles s'écrivent fk_N , où $N = 4, 6, 8, 14, 18, 22$.

Le MODWT peut être utilisé pour effectuer une analyse multi-résolution (MRA) d'un signal (Chun et al., 2007). MRA est une technique qui décompose un signal en un ensemble de sous-bandes, chacune représentant le signal à une fréquence et une échelle de temps différentes (Doucoure et al., 2016). Le MODWT est bien adapté à MRA car il utilise un banc de filtres non décimés, ce qui signifie que la longueur du signal reste constante à chaque niveau de décomposition. Cela permet de reconstruire le signal original à partir de ses coefficients décomposés, ce qui est nécessaire pour MRA.

Pour décomposer les signaux d'entrée à l'aide de MODWT, le niveau de décomposition doit être déterminé au préalable (Chidepudi et al., 2023). Le niveau L a été déterminé sur la base de l'équation (III.38) :

$$L = \text{int}[\log(n)] \quad (\text{III.38})$$

Où ; L est le niveau de décomposition et N est la taille de la série chronologique. Int ; est la partie entière de $\text{Log}(N)$,

La longueur du signal peut influencer le niveau de décomposition approprié, Si le signal est très court, le niveau de décomposition doit être limité pour éviter une sur-décomposition.

Conclusion

Pour les prévisions hydrologiques, la sélection entre les techniques autonomes d'intelligence artificielle et les techniques hybrides dépend d'un certain nombre de facteurs, notamment la complexité du problème, les ressources informatiques, la disponibilité et la qualité des données, la sélection des algorithmes, le réglage des paramètres et la possibilité d'interprétation. Une compréhension approfondie de ces facteurs est nécessaire pour sélectionner l'approche la plus appropriée à un problème spécifique de prévision hydrologique. Ce chapitre a été consacré à la présentation des différents modèles d'intelligence artificielle qui sont appliqués dans cette recherche pour la modélisation des débits journalier et mensuelles dans le bassin de Cheliff en Algérie.

CHAPITRE IV Présentation de la zone d'étude

Introduction

L'Algérie est divisée en cinq bassins hydrographiques (Oranie-Chott Chergui, Chélif-Zahrez, Algérois-Hodna-Soummam, Constantinois Seybousse-Mellegue et le bassin de Sahara), Ces bassins regroupent 17 grands bassins versants selon le découpage de l'Agence Nationale des Ressources en Eau.

Le domaine d'étude concerne le bassin de Cheliff au nord de l'Algérie ainsi que ces dérivés sous bassins étudiés, une présentation de ces caractéristiques géographiques, climatologique hydrographique sera abordé dans ce chapitre.

IV.1 Présentation du bassin de Cheliff

Le bassin du Cheliff, situé au nord de l'Algérie, correspond à une cuvette intra montagneuse, est compris entre les longitudes 33°91' et 36°58' E et les latitudes 0°12' et 3°87' N. D'une superficie approximative de 43 750 km², il couvre 77% de la superficie totale du bassin hydrographique Cheliff-Zahrez. Celui-ci est limité par la mer Méditerranée au nord, les hautes plaines au sud, le bassin algérien à l'est et le bassin oranais à l'ouest. Cette région bénéficie d'un climat méditerranéen semi-aride avec des étés chauds et des hivers froids. Les précipitations présentent une grande variabilité avec une tendance à la baisse (600 mm/an) du nord au sud et (300 mm/an) de l'est à l'ouest. La température moyenne annuelle diminue généralement du nord (18,7°C) au sud (14,2°C). La zone d'étude est drainée par le fleuve Cheliff, qui s'écoule sur une longueur de 750 km avant de se jeter dans la mer Méditerranée près de Mostaganem. La région hydrographique du bassin de Cheliff compartimenté naturellement en trois sous bassins versants (**figure IV.1**) :

- Le cheliff en amont de bougazoul ;
- Le haute et le moyen de cheliff ;
- Le bas Chélif et de Mina.

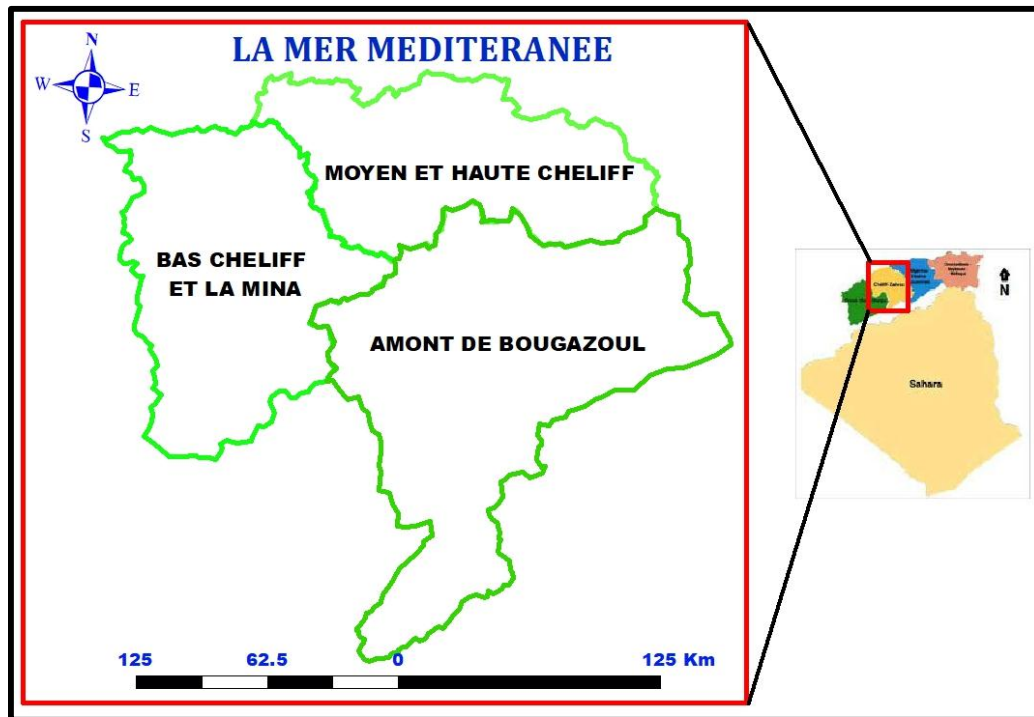


Figure IV.1 Situation géographique de Bassin de Cheliff au nord de l'Algérie.

Cette étude portera essentiellement à cinq sous bassins versant situées dans le grand bassin de Cheliff. Selon la classification de l'ANRH, ces bassins sont répartis entre le bassin Bas Chélif et de la Mina et le bassin de Haut moyen Cheliff.

IV.1.1 Critères de choix de zone d'étude

On a choisi les deux bassins de Haut et Moyen Cheliff et Bas Cheliff et de la Mina pour la disponibilité des données aux pas journaliers et mensuelles grâce aux stations pluviométrique et hydrométrique située dans la zone d'étude.

IV.1.1.1 Bas Chélif et de la Mina

Le bassin du Bas Chélif et de la Mina se situe dans la partie nord-ouest de l'Algérie. Il fait partie du grand bassin versant du Cheliff et constitue l'un des principaux affluents de l'Oued Cheliff. D'une superficie de 13150 km², il est limité par le Cheliff moyen à l'Est, le bassin de la Macta à l'Ouest, le massif du Dahra au Nord et le Chott Ech Chergui au Sud. Le bassin du Haut et Moyen Cheliff comprend 13 sous-bassins versants codifiés de 0124 à 0136, celui-ci contient un sous-bassin versant lié à la zone d'étude, le bassin de Tiguiguest ([figure IV.2](#)).

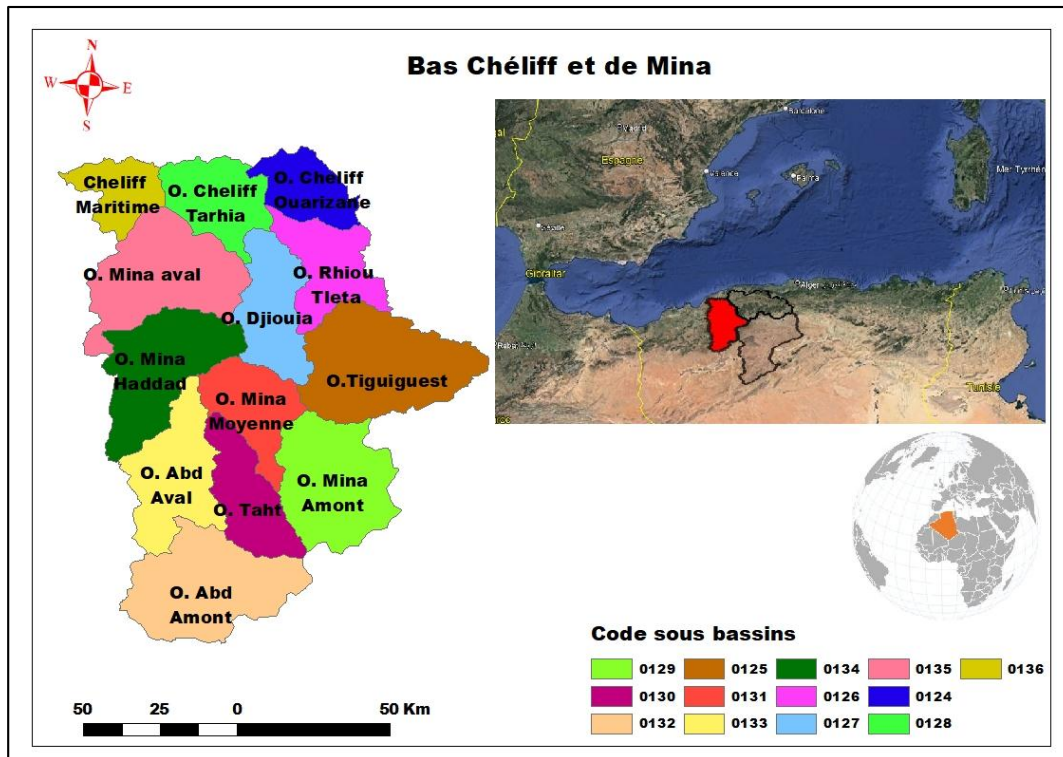


Figure IV.2 Carte des sous bassins du Bas Cheliff et de la Mina

✚ Bassin versant d'Oued Tiguiguest (0125)

Située dans la partie nord-ouest du pays, son territoire est très étendu. Il est compris entre les longitudes 1° et 2°0' E et entre les latitudes 35° et 35°50' N. Il est limité au nord par le sous bassin de Oued Rhiou Tleta, à l'ouest par les sous bassins d'Oued Djiouia et Oued Mina Moyenne, à l'est par le sous bassin d'Oued Sly et au sud par le sous bassin de Oued Mina Amont (figure IV.3). L'oued Tiguiguest s'étend sur une superficie de 1617.56 km². Des massifs montagneux, parallèles à la côte, séparés par des zones de basse altitude ou par des plaines. Les conditions climatiques de cette région sont tempérées et généralement chaudes. Les précipitations moyennes dans ce bassin sont plus élevées en hiver qu'en été. La température moyenne annuelle est de l'ordre de 16,6°C. Les précipitations annuelles moyennes sont de 487 mm.

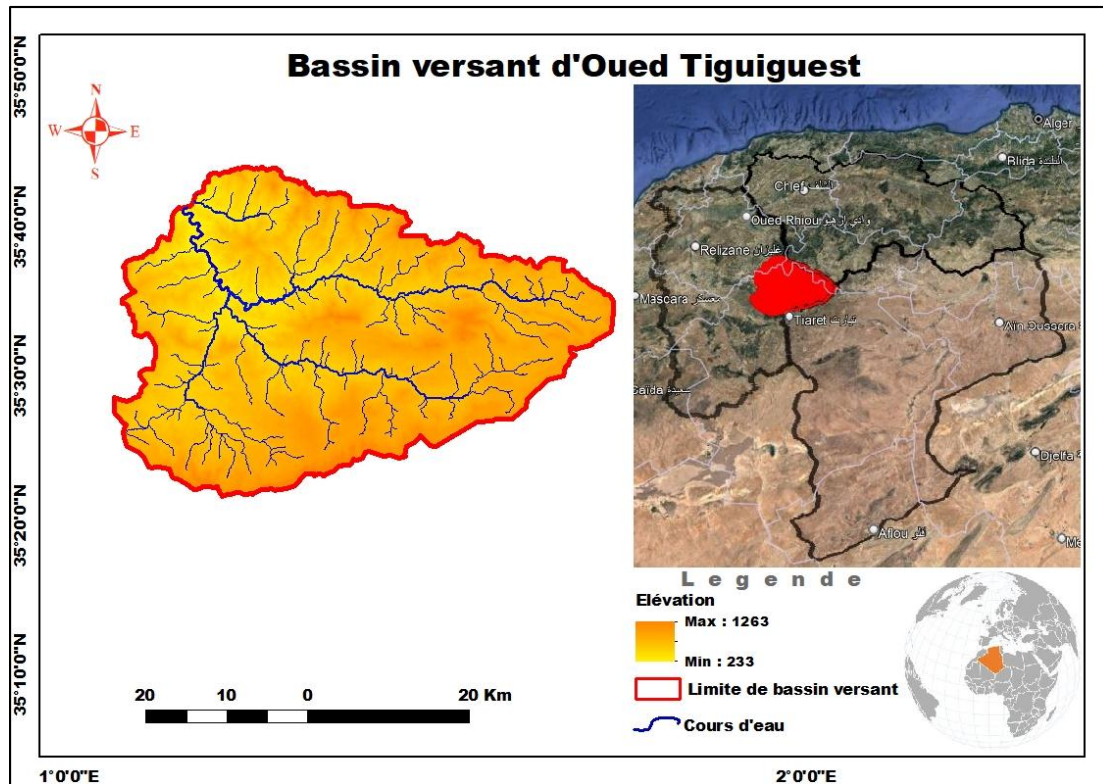


Figure IV.3 Situation géographique de sous bassin d'Oued Tiguiguest.

IV.1.1.2 Haut et Moyen Cheliff

Avec une superficie de 10 930 km², le bassin versant du Haut et Moyen Cheliff est situé dans la partie centre-ouest de l'Algérie du Nord. Il s'étend de l'étranglement qui le sépare du sous-bassin en amont de Boughzoul jusqu'au seuil de Boukadir, où le bassin inférieur du Cheliff et de la Mina. Le bassin du Haut et Moyen Cheliff est composé de 11 sous-bassins codifiés de 0113 à 0123, il comprend quatre sous-bassins versants concernant la zone d'étude, le bassin du Cheliff Harreza, Oued Ebda, Oued Rouina Zeddine et le sous-bassin de Oued Ras Ouahrane (figure IV.4).

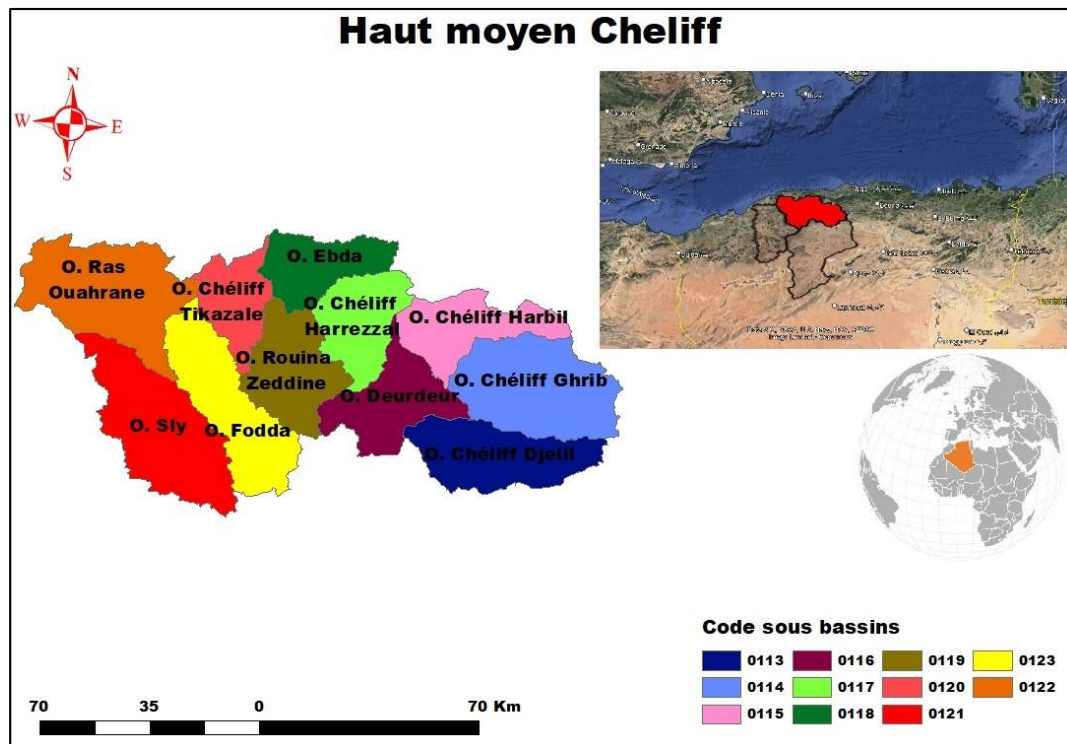


Figure IV.4 Carte des sous bassins du Haut et du Moyen Cheliff

✚ Bassin versant d'Oued Arib-Ebda (0118)

Dans le Sud de la ville de Cherchell et dans le Nord de la ville d'Ain-Defla, le bassin s'étend sur une superficie de 661.45 km². Il est compris entre les longitudes 1° et 2°3' E et entre les latitudes 36° et 36°30' N. Il est formé d'une série de montagnes qui constituent une unité géographiquement relativement homogène de l'Atlas tellien septentrional (figure IV.5). Les altitudes de ce bassin sont modérées, atteignant rarement 1 500 m. De par sa position géographique, il est soumis à un climat tempéré semi-aride.

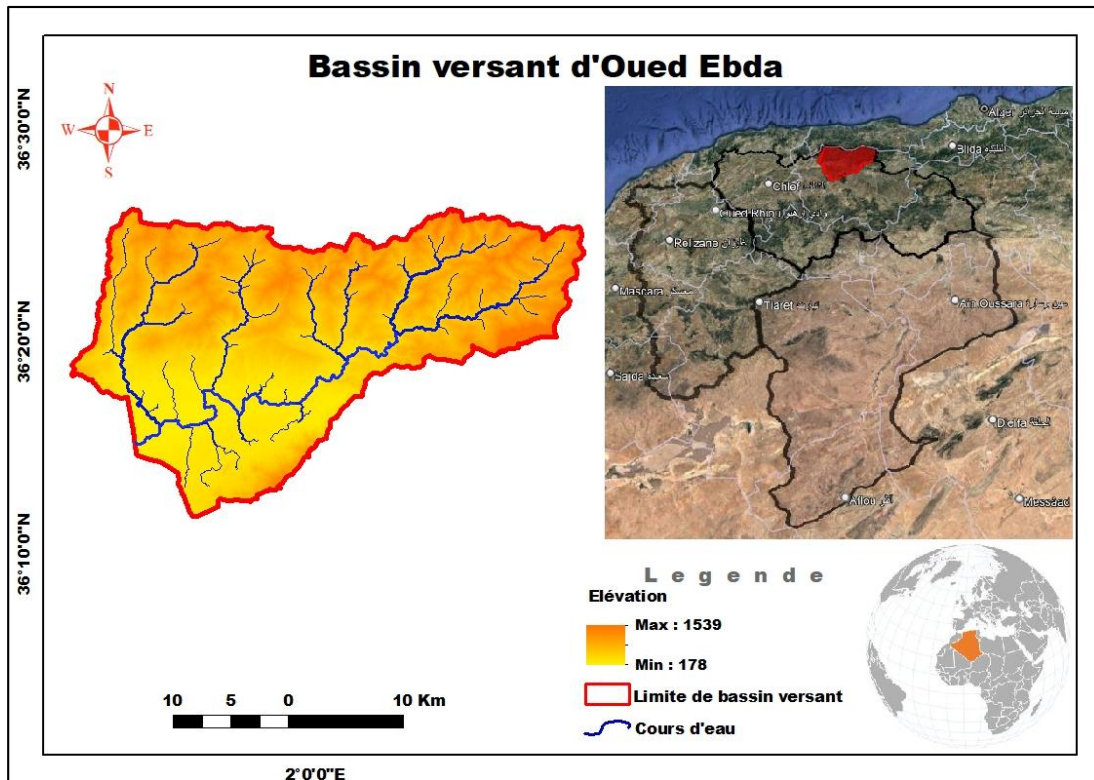


Figure IV.5 Situation géographique de sous bassin d'Oued Ebda

✚ Bassin versant d'Oued Rouina Zeddine (0119)

D'une superficie de 891.46 km², le bassin de Rouina est situé au sud du village de Rouina et au nord de la ville de Theniet el Had. Il est compris entre la longitude 1°40' et 2°10' E et entre la latitude 35°50' et 36°10' N. Il appartient essentiellement à l'ensemble des bassins situés dans la zone de dépression à relief relativement large entre la chaîne de l'Atlas Tellien septentrional et l'Atlas Tellien méridional (figure IV.6). Il est caractérisé par une succession de dépressions comblées par des formations quaternaires et des reliefs miocènes du Cheliff (Miocène inférieur et Miocène supérieur), avec une altitude maximale de 1700m.

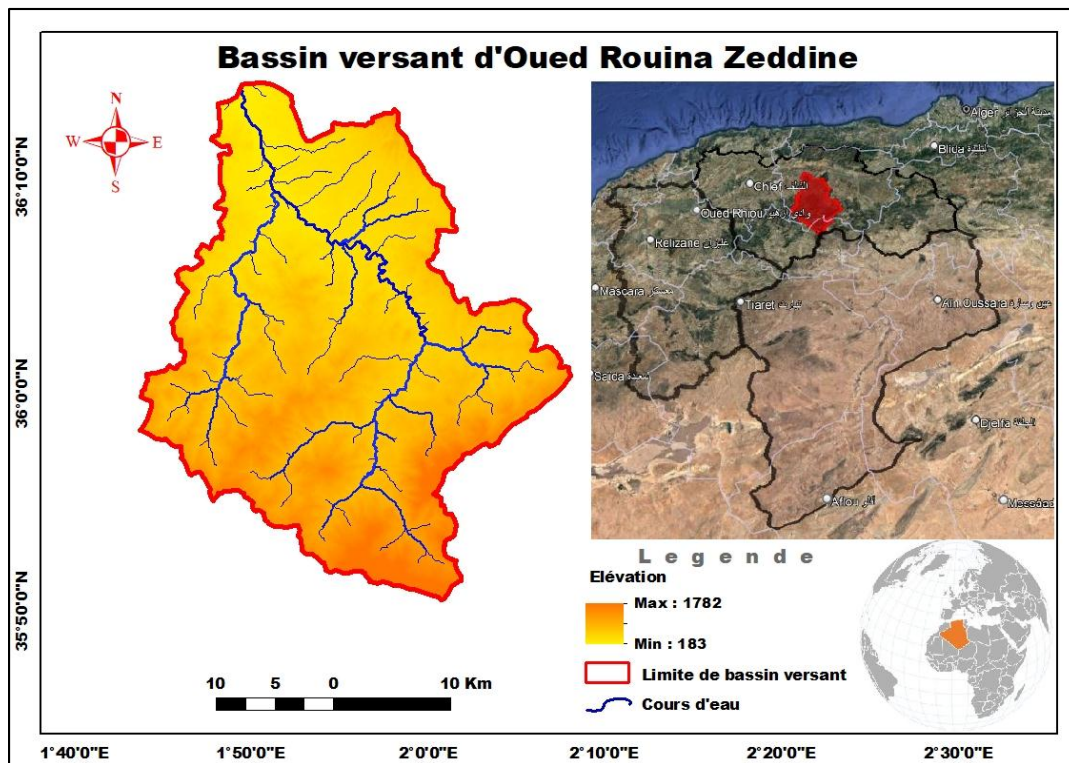


Figure IV.6 Situation géographique de sous bassin d'Oued Rouina Zeddine

✚ Bassin versant d'Oued Ras Ouahrane (0122)

Le bassin versant de l'Oued Ouahrane occupe une superficie de 1438.45 km² et fait partie de la partie septentrionale du grand bassin du Cheliff (figure 1). Il est compris entre les longitudes 1° et 1°3' E et entre les latitudes 36° et 36°24' N. L'Oued Ouahrane est un affluent mineur de l'Oued Cheliff. Ce bassin (figure IV.7) est surveillé par trois stations pluviométriques et une station hydrométrique. Le bassin de l'Oued Ouahrane est limité à l'est par le bassin de l'Oued Fodda, à l'ouest par le bassin de l'Oued Ras, au nord par le bassin de l'Oued Allala et au sud par le bassin de l'Oued Sly.

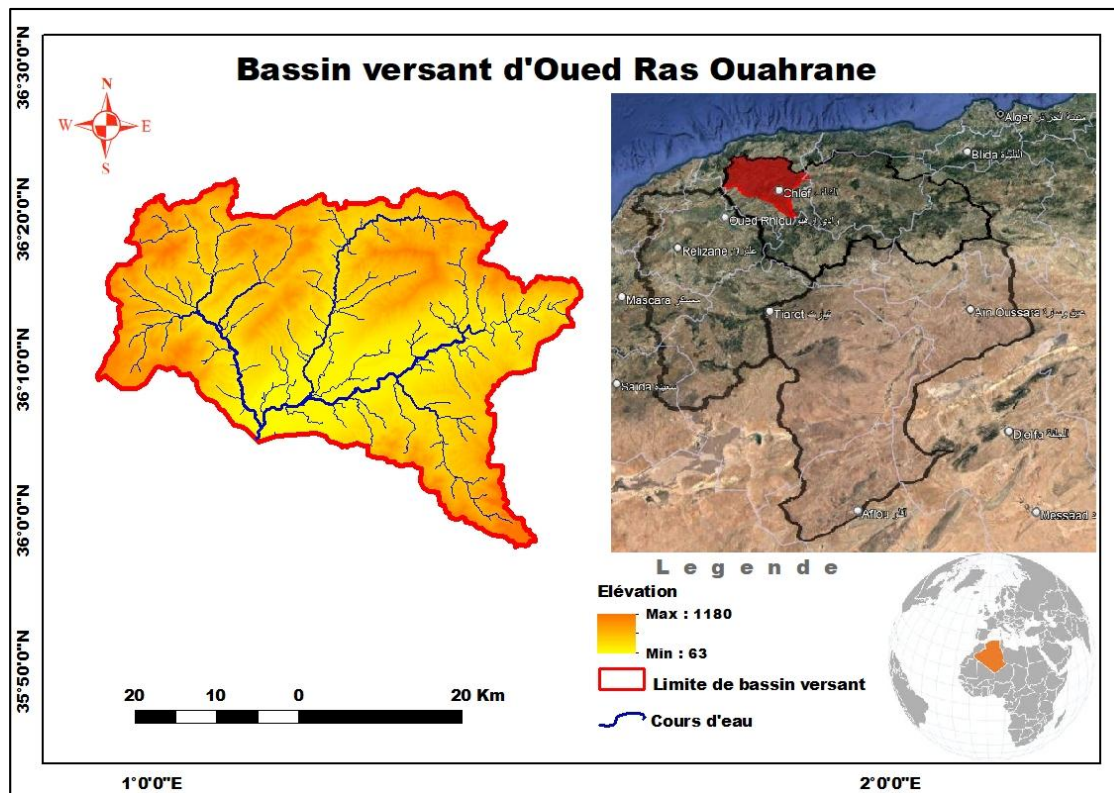


Figure IV.7 Situation géographique de sous bassin d'Oued Ras Ouahrane

✚ Bassin versant d'Oued Cheliff Harreza

Le bassin versant de d'Oued Cheliff Harreza fait partie du bassin de l'oued Cheliff. Situé au nord-ouest de l'Algérie, entre 02°00' et 02°24' de longitude Est, et entre 36°00' et 36°30' de latitude Nord. Le sous bassin d'Oued Cheliff Harreza s'étend sur une superficie de 756.70 km² et un périmètre de 142 km (figure IV.8).

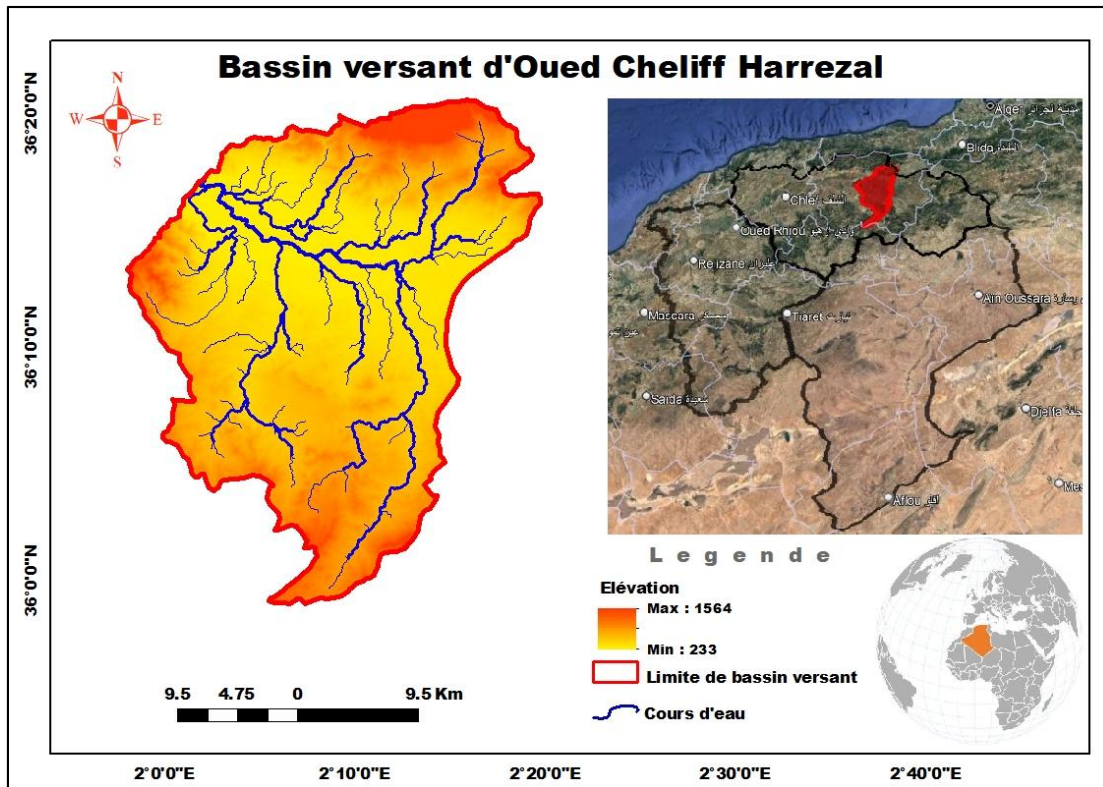


Figure IV.8 Situation géographique de sous bassin d'Oued Cheliff Harrezal

IV.1.2 Etude géomorphologique

IV.1.2.1 Paramètres géométriques

a. Superficie et périmètre

Les principaux paramètres des sous-bassins concernant la zone d'étude sont la superficie et le périmètre, lesquels sont obtenus par la technique de la digitalisation ([Tableau IV.1](#)).

b. Paramètres de forme

- *Indice de compacité et de forme*

Le paramètre K_c , appelé indice de compacité de Gravelius, peut exprimer une influence significative sur l'écoulement. La forme du bassin détermine la forme de l'hydrogramme de crue, où un bassin allongé ne réagira pas de la même manière qu'un bassin de forme compacte. On l'obtient par l'équation suivante :

$$K_c = 0.28 \frac{P}{\sqrt{S}} \quad (\text{IV } 1)$$

Où :

P : Périmètre du bassin versant (km) ;

S : Superficie du bassin versant (km²).

Selon la valeur de K_c , le bassin versant à une forme bien déterminée. Ainsi :

Si $K_c \leq 1$: implique que le bassin versant à une forme circulaire.

Si $K_c = 1.12$: implique que le bassin versant à une forme carrée.

Si $K_c > 1.12$: implique que le bassin versant à une forme allongée.

- **Rectangle équivalent**

Sur le plan géométrique, le rectangle équivalent : également appelé rectangle de Gravelius, correspond à une transformation du bassin versant de manière strictement géométrique. Celui-ci prend une forme rectangulaire tout en gardant la même surface, le même périmètre, le même indice de compacité et donc la même répartition hypsométrique.

Où : L et l sont respectivement la longueur et la largeur du rectangle équivalent ont données par les expressions suivantes :

$$L = \frac{K_c \sqrt{S}}{1.12} \left(1 + \sqrt{1 - \left(\frac{1.12}{K_c} \right)^2} \right) \quad (\text{IV } 2)$$

$$l = \frac{K_c \sqrt{S}}{1.12} \left(1 - \sqrt{1 - \left(\frac{1.12}{K_c} \right)^2} \right) \quad (\text{IV } 3)$$

Son périmètre est défini par l'équation suivante :

$$P = 2 * (L + l) \quad (\text{IV } 4)$$

La superficie du rectangle équivalent est définie par :

$$S = (L * l) \quad (\text{IV } 5)$$

Tableau IV.1 Indices de forme des sous-bassins versants du secteur d'étude

Basin	Code		Superficie (Km ²)	Périmètre (Km)	Indice de compacité Kc	Rectangle équivalent	
	sous- bassin	Sous- bassin				Longueur (Km)	Largeur (Km)
Bas Cheliff et La Mina	0125	Tiguiguest	1617.56	197.10	1.37	77.74	20.81
	0117	Cheliff	756.70	142.46	1.45	58.24	12.99
		Harrezal					
Moyen et Haut Cheliff	0118	Ebda	661.45	139.10	1.51	58.18	11.37
	0119	Rouina	891.46	164.97	1.55	69.69	12.79
		Zeddine					
	0122	Ras	1438.45	224.52	1.66	97.51	14.75
Ouahrane							

Globalement, la zone d'étude présente des formes allongées à assez allongées permettant un écoulement lent, le relief étant cependant un facteur décisif.

IV.1.2.2 Etude de relief

Le relief a une influence significative sur l'écoulement des cours d'eau, et plusieurs paramètres géographiques et morphologiques entrent en jeu. Le relief est déterminé par les indices dont caractéristiques suivantes :

a. Courbe hypsométrique

Elle représente la répartition de la surface du bassin versant en fonction de son altitude. L'abscisse indique la surface (ou le pourcentage de surface) du bassin versant qui se situe au-dessus (ou au-dessous) de l'altitude donnée en ordonnée, indiquant ainsi la surface du bassin versant ou le pourcentage de surface, à partir d'une certaine altitude.

En utilisant un modèle numérique de terrain conjointement avec un système d'information géographique (SIG) tel que ArcGIS, ainsi que le logiciel Global Mapper, les données relatives aux surfaces partielles, aux surfaces cumulées et aux altitudes correspondantes pour chacun des cinq sous-bassins ont été collectées et compilées. Les résultats de cette collecte sont présentés dans les tableaux [IV.2](#) à [IV.6](#).

Tableau IV.2 Répartition hypsométrique du bassin versant d'Oued Rac Ouahrane

N° Classe	Classe	Altitude	Surfaces partielles		Surfaces cumulées	
	d'altitude	Moyenne.				
	$H_i - H_{i-1}$	$H_i(m)$	$S_i(Km^2)$	$S_i(\%)$	$\sum S_i(Km^2)$	$\sum S_i(\%)$
1	63-100	81.5	131.74	9.16	131.74	9.16
2	100-200	150	295.38	20.53	427.12	29.69
3	200-300	250	326.05	22.67	753.17	52.36
4	300-400	350	312.85	21.75	1066.02	74.11
5	400-500	450	225.82	15.70	1291.84	89.80
6	500-600	550	116.44	8.09	1408.28	97.90
7	600-700	650	19.77	1.37	1428.05	99.27
8	700-800	750	5.59	0.39	1433.64	99.66
9	800-900	850	2.95	0.21	1436.59	99.87
10	900-1000	950	1.44	0.10	1438.03	99.97
11	1000-1100	1050	0.35	0.02	1438.38	99.99
12	1100-1180	1140	0.13	0.01	1438.51	100.00

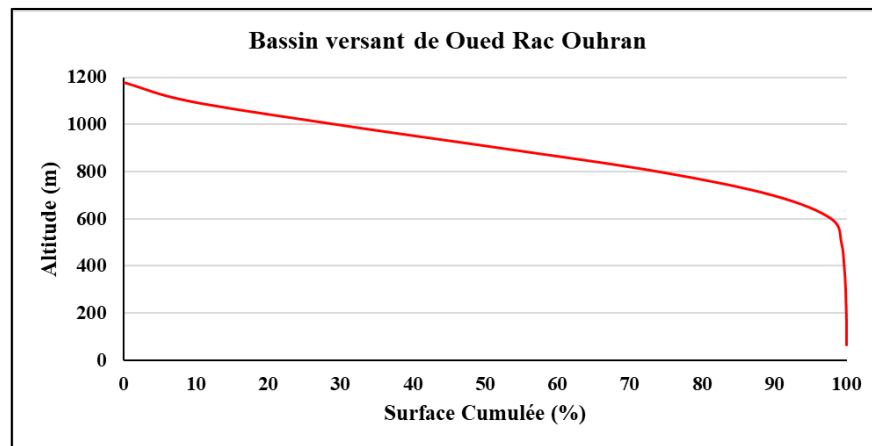


Figure IV.9 Courbe hypsométrique du bassin versant de Oued Rac Ouahrane

Tableau IV.3 Répartition hypsométrique du bassin versant d'Oued Rouina Zeddine

N° Classe	Classe d'altitude	Altitude Moyenne.	Surfaces partielles		Surfaces cumulées	
	$H_i - H_{i-1}$	$H_i(m)$	$S_i(Km^2)$	$S_i(\%)$	$\sum S_i(Km^2)$	$\sum S_i(\%)$
1	183-200	191.5	1.30	0.15	1.30	0.15
2	200-300	250	70.91	7.96	72.21	8.10
3	300-400	350	110.38	12.39	182.59	20.49
4	400-500	450	122.80	13.78	305.39	34.27
5	500-600	550	143.12	16.06	448.50	50.33
6	600-700	650	121.03	13.58	569.53	63.91
7	700-800	750	88.21	9.90	657.74	73.81
8	800-900	850	62.91	7.06	720.66	80.87
9	900-1000	950	61.80	6.94	782.46	87.81
10	1000-1100	1050	45.85	5.14	828.30	92.95
11	1100-1200	1150	26.06	2.92	854.36	95.88
12	1200-1300	1250	14.46	1.62	868.82	97.50
13	1300-1400	1350	9.57	1.07	878.39	98.58
14	1400-1500	1450	6.80	0.76	885.19	99.34
15	1500-1600	1550	2.84	0.32	888.03	99.66
16	1600-1700	1650	2.24	0.25	890.26	99.91
17	1700-1782	1741	0.82	0.09	891.08	100.00

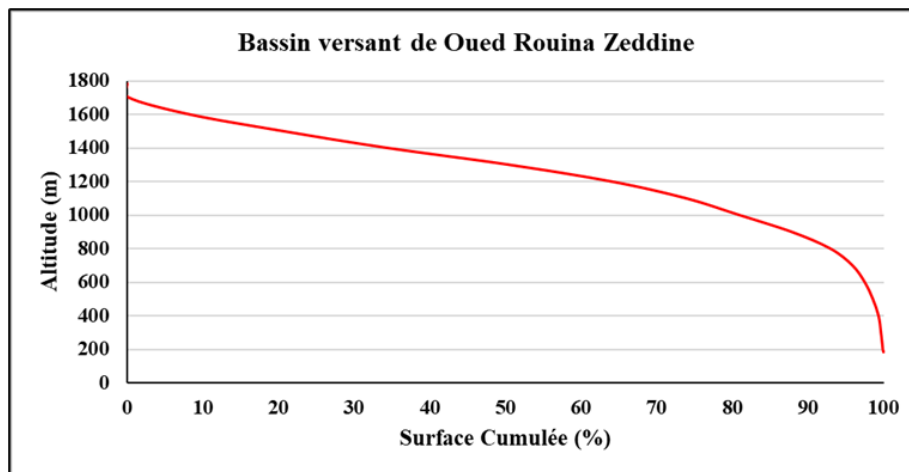


Figure IV.10 Courbe hypsométrique du bassin versant d'Oued Rouina Zeddine

Tableau IV.4 Répartition hypsométrique du bassin versant d'Oued Ebda

N° Classe	Classe	Altitude	Surfaces partielles		Surfaces cumulées	
	d'altitude	Moyenne.	$S_i(Km^2)$	$S_i(\%)$	$\sum S_i(Km^2)$	$\sum S_i(\%)$
	$H_i - H_{i-1}$	$H_i(m)$				
1	178-200	189	12.51	1.89	12.51	1.89
2	200-300	250	128.98	19.50	141.49	21.40
3	300-400	350	63.58	9.61	205.07	31.01
4	400-500	450	64.99	9.83	270.06	40.84
5	500-600	550	77.89	11.78	347.95	52.62
6	600-700	650	89.99	13.61	437.94	66.23
7	700-800	750	89.27	13.50	527.21	79.73
8	800-900	850	70.51	10.66	597.72	90.39
9	900-1000	950	38.26	5.79	635.98	96.18
10	1000-1100	1050	14.82	2.24	650.81	98.42
11	1100-1200	1150	4.47	0.68	655.28	99.09
12	1200-1300	1250	1.67	0.25	656.95	99.35
13	1300-1400	1350	2.06	0.31	659.01	99.66
14	1400-1500	1450	2.05	0.31	661.06	99.97
15	1500-1538	1519.5	0.21	0.03	661.27	100.00

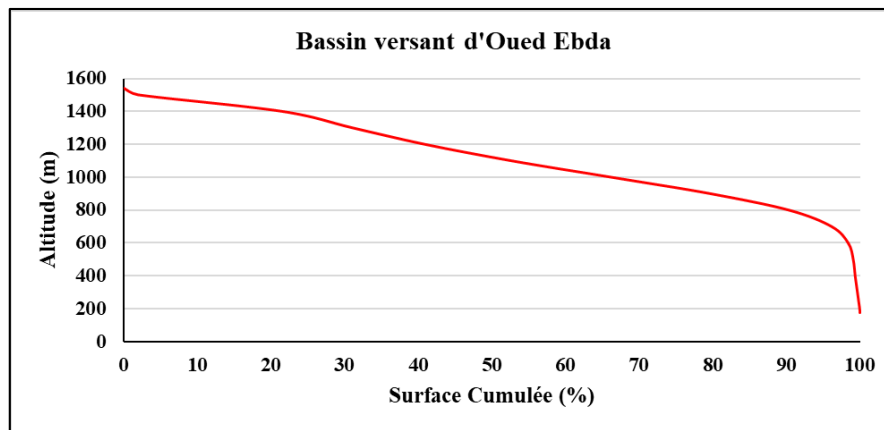


Figure IV.11 Courbe hypsométrique du bassin versant d'Oued Ebda

Tableau IV.5 Répartition hypsométrique du bassin versant d'Oued Cheliff Harreza

N° Classe	Classe	Altitude	Surfaces partielles		Surfaces cumulées	
	d'altitude	Moyenne.	$S_i(Km^2)$	$S_i(\%)$	$\sum S_i(Km^2)$	$\sum S_i(\%)$
	$H_i - H_{i-1}$	$H_i(m)$				
1	203-300	251.5	217.82	28.80	217.82	28.80
2	300-400	350	154.20	20.39	372.03	49.19
3	400-500	450	137.57	18.19	509.59	67.38
4	500-600	550	98.40	13.01	608.00	80.39
5	600-700	650	71.41	9.44	679.41	89.83
6	700-800	750	34.51	4.56	713.92	94.40
7	800-900	850	15.66	2.07	729.58	96.47
8	900-1000	950	6.91	0.91	736.50	97.38
9	1000-1100	1050	3.80	0.50	740.30	97.88
10	1100-1200	1150	4.08	0.54	744.38	98.42
11	1200-1300	1250	4.98	0.66	749.36	99.08
12	1300-1400	1350	4.79	0.63	754.14	99.71
13	1400-1500	1450	1.93	0.26	756.08	99.97
14	1500-1564	1532	0.23	0.03	756.30	100.00

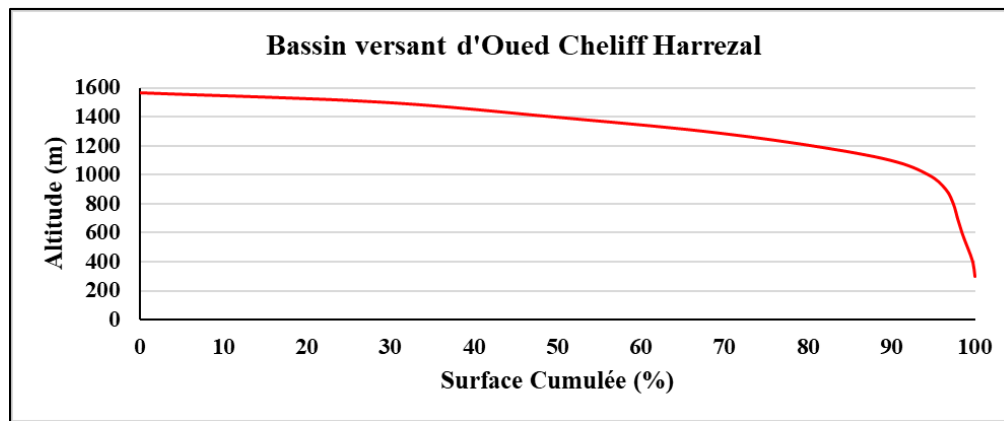


Figure IV.12 Courbe hypsométrique du bassin versant d'Oued Cheliff Harrezal

Tableau IV.6 Répartition hypsométrique du bassin versant d'Oued Tiguiguest

N° Classe	Classe	Altitude	Surfaces partielles		Surfaces cumulées	
	d'altitude	Moyenne.	$S_i(Km^2)$	$S_i(\%)$	$\sum S_i(Km^2)$	$\sum S_i(\%)$
	$H_i - H_{i-1}$	$H_i(m)$				
1	233-300	266.5	9.40	0.58	9.40	0.58
2	300-400	350	46.54	2.88	55.94	3.46
3	400-500	450	102.52	6.34	158.46	9.80
4	500-600	550	221.88	13.72	380.34	23.52
5	600-700	650	329.54	20.38	709.88	43.89
6	700-800	750	363.69	22.49	1073.58	66.38
7	800-900	850	309.36	19.13	1382.94	85.51
8	900-1000	950	175.15	10.83	1558.09	96.34
9	1000-1100	1050	44.32	2.74	1602.40	99.08
10	1100-1200	1150	14.17	0.88	1616.57	99.95
11	1200-1263	1231.5	0.76	0.05	1617.33	100.00

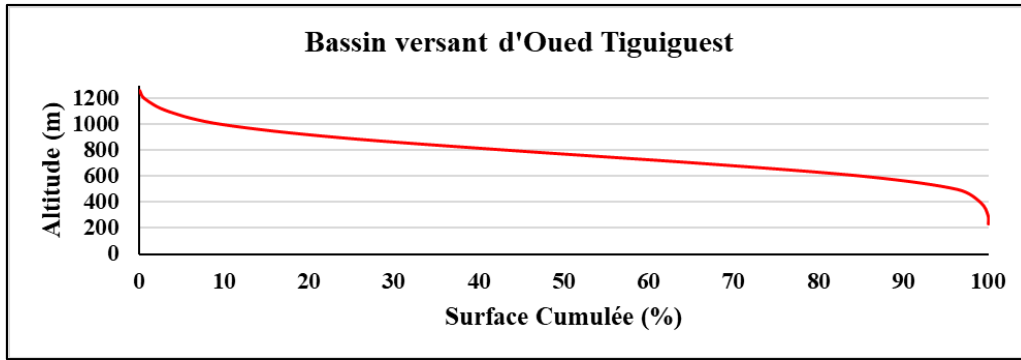


Figure IV.13 Courbe hypsométrique du bassin versant d'Oued Tiguiguest

b. Altitudes caractéristiques

Dans un bassin versant, la connaissance des variations d'altitude ainsi que de l'altitude moyenne est essentielle pour estimer un certain nombre de paramètres très importants qui sont strictement liés à la nature géologique du terrain et à la topographie du bassin versant. Dans le cadre des études hydrologiques des bassins versants, les altitudes les plus employées restent celles lues sur les courbes hypsométriques, à savoir :

- Altitude minimale (H_{min})
- Altitude maximale (H_{max})
- Altitude à 95% de surface ($H_{95\%}$)
- Altitude à 5% de surface ($H_5\%$)
- Altitude médiane ($H_{50\%}$)
- Altitude moyenne (H_{moy})

H_{moy} Est calculée à partir de la relation suivante :

$$H_{moy} = \frac{1}{S} \sum_{i=1}^n S_i \left(\frac{H_i + H_{i-1}}{2} \right) \quad (\text{IV } 6)$$

Où :

H_{moy} Est l'altitude moyenne du bassin versant exprimée en m, S est la superficie totale du bassin versant exprimée en km^2 , S_i est la superficie partielle comprise entre deux altitudes successives H_i et H_{i-1} exprimée en km^2 , est le nombre de courbes de niveau.

c. La dénivelée (D)

On définit la dénivelée entre $H_5\%$ et $H_{95}\%$

$$D = H_5\% - H_{95}\% \quad (\text{IV } 7)$$

Le tableau suivant récapitule les principales caractéristiques des élévations ([tableau IV.7](#)) :

Tableau IV.7 Altitudes caractéristiques des bassins versants étudiés

Sous-bassins versants	$H_{min}(m)$	$H_{max}(m)$	$H_{95\%}(m)$	$H_5\%(m)$	$H_{50\%}(m)$	$H_{moy}(m)$	$D(m)$
Tiguigues	233.00	1263.00	411.59	1122.01	776.52	749.82	710.42
t							
Cheliff	203.00	1564.00	747.94	1525.00	1519.60	898.82	777.06
Harrezal							
Ebda	178.00	1539.00	514.63	1477.35	1187.56	850.57	962.72
Rouina	183.00	1782.00	557.81	1651.80	1386.08	951.91	1094.0
Zeddine							0
Ras	63.00	1180.00	413.14	1117.45	981.77	601.79	704.31
Ouahrane							

d. Indice de pente global (I_g)

I_g Correspond aux proportions entre l'altitude entre 5% et 95% de la surface du bassin versant avec la longueur du rectangle équivalent Cela permet de déduire la nature du relief qu'il est défini par :

$$I_g = \frac{D}{L} \quad (\text{IV } 8)$$

Où : L est longueur du rectangle équivalent et D est la dénivelée.

e. Indice de pente de roche (I_p)

L'indice de pente I_p proposé par Roche (1963) permet de prendre en compte le facteur pente sur le régime hydrologique. I_p Est la moyenne de la racine carrée des pentes mesurées sur le rectangle

équivalent, et pondérée par les surfaces comprises entre 2 courbes de niveau il est calculé par l'expression :

$$I_p = \frac{\sum S_i \sqrt{n_i}}{S} \quad (\text{IV } 9)$$

Où : (n_i) est la pente d'une tranche de surface (S_i) comprise entre deux courbes de niveau.

f. Dénivelée spécifique (D_S)

Permet de définir les différents types de relief des bassins versants par une classification quel que soit leurs superficies, selon Office de la Recherche Scientifique et Technique Outre-Mer (ORSTOM) D_S est donnée par l'expression suivante :

$$D_S = I_p * \sqrt{S} \quad (\text{IV } 10)$$

La classification des reliefs selon D_S est dénotée dans le tableau suivant (tableau IV.8) :

Tableau IV.8 Classification du relief selon l'indice de pente global et la dénivelée spécifique (Selon ORSTOM)

<i>Classification du relief</i>	<i>Selon I_g (m/km)</i>	<i>Selon D_S (m)</i>
Relief très faible	$I_g < 0.002$	$D_S < 10$
Relief faible	$0.00 < I_g < 0.005$	$10 < D_S < 25$
Relief assez faible	$0.005 < I_g < 0.01$	$25 < D_S < 50$
Relief modéré	$0.01 < I_g < 0.02$	$50 < D_S < 100$
Relief assez fort	$0.02 < I_g < 0.05$	$100 < D_S < 250$
Relief fort	$0.05 < I_g < 0.5$	$250 < D_S < 500$
Relief très fort	$I_g > 0.5$	$D_S > 500$

Après l'application numérique on trouve les résultats suivants (**tableau IV.9**) :

Tableau IV.9 Indices de forme des sous bassins versants étudiés

Sous-bassins versants	I_g	I_p	D_s	Relief selon D_s
Tiguiguest	13,21	12.77	513,84	Relief très fort
Cheliff Harrezal	23,87	23.67	650,97	Relief très fort
Ebda	23,73	24.07	619,02	Relief très fort
Rouina Zeddine	21,38	16.89	640,85	Relief très fort
Ras Ouahrane	11,31	11.30	428,71	Relief fort

IV.1.3 Etude du réseau hydrographique

IV.1.3.1 Paramètre hydrographique

- **Densité de drainage (D_d)**

Elle est déterminée comme étant le rapport entre la longueur totale des cours d'eau de tout ordre (L_x) sur la surface du bassin versant (S).

La densité de drainage D_d est donnée par l'équation suivante :

$$D_d = \frac{\sum l_i}{S} \quad (\text{IV } 11)$$

Où :

D_d : Densité de drainage en Km/Km²

$\sum l_i$: Longueur cumulée de tous les thalwegs du bassin en Km,

S : Superficie du bassin en Km².

- **La densité de drainage permanent (D_{dp})**

Les valeurs de la densité permanente de drainage D_{dp} offrent une meilleure appréciation des variations de débit à l'échelle d'un bassin. Effectivement, les valeurs de D_{dp} varient fortement d'un bassin à l'autre et dépendent étroitement de deux facteurs principaux : la géologie et la pluviométrie.

$$D_{dp} = \frac{\sum l_{ip}}{S} \quad (\text{IV } 12)$$

Où :

D_{dp} : Densité de drainage permanente en Km/Km²

$\sum l_{ip}$: Longueur cumulée de tous les Oueds permanents du bassin en Km.

- **La densité de drainage temporaire (D_{dt})**

$$D_{dt} = \frac{\sum l_{it}}{S} \quad (\text{IV } 13)$$

Où :

D_{dt} : Densité de drainage temporaire en Km/Km²

$\sum l_{it}$: Longueur cumulée de tous les Oueds temporaires du bassin en Km.

Les résultats obtenus sont résumés dans le **tableau IV.10**.

Tableau IV.10 Densité de drainage des sous bassins versants étudiés

Sous-bassins versants	L_p	L_t	L	D_{dp}	D_{dt}	D_d	D_{dp}/D_d	D_{dt}/D_d
	km	km	km	km/km ²	km/km ²	km/km ²	%	%
Tiguiguest	545.87	527.04	1072.91	0.34	0.33	0.66	51	49
Cheliff								
Harreza	107.45	251.62	359.07	0.14	0.33	0.47	30	70
Ebda								
Ebda	74.6	392.84	467.44	0.11	0.59	0.71	16	84
Rouina								
Zeddine	177.73	575.18	752.91	0.2	0.65	0.84	24	76
Ras								
Ouahrane	171.76	988.26	1160.02	0.12	0.69	0.81	15	85

- **Coefficient de torrentialité (C_t)**

C'est le produit de la densité de drainage par la fréquence des talwegs élémentaires.

$$C_t = N_1 \frac{D_d}{S} \quad (\text{IV } 14)$$

Où :

C_t : coefficient de torrentialité

D_d : densité de drainage

N_1 : nombre de talwegs d'ordre « 1 »

S : surface du bassin versant

- **Temps de concentration (T_c)**

C'est le temps que met une goutte d'eau tombée au point hydrauliquement le plus éloigné pour arriver à l'exutoire.

On va utiliser la formule de GIANDOTTI pour déterminer ce paramètre

$$T_c = \frac{4\sqrt{S} + 1.5L_p}{0.8\sqrt{H_{moy} - H_{min}}} \quad (\text{IV } 15)$$

Où :

T_c : temps de concentration (heures)

S : superficie du bassin versant (km²)

L_p : longueur du talweg principal (km)

H_{moy} : altitude moyenne du bassin versant (m)

H_{min} : altitude moyenne du bassin versant (m)

- **Vitesse de concentration (V_c)**

C'est la vitesse moyenne de propagation de la crue, elle est calculée par la formule suivante :

$$V_c = \frac{L_p}{T_c} \quad (\text{IV } 16)$$

Où :

V_c : vitesse de concentration (km/h)

L_p : longueur du talweg principal (km)

T_c : temps de concentration (h)

Les résultats obtenus sont résumés dans le **tableau IV.11**.

Tableau IV.11 Coefficient de torrentialité, Temps de concentration et Vitesse de concentration des sous bassins versants étudiés

Sous-bassins versants	C_t	T_c (Heure)	V_c (Km/h)
Tiguiguest	21.30	51.66	10.57
Cheliff Harreza	18.06	12.85	08.36
Ebda	12.80	10.35	07.21
Rouina Zeddine	70.50	17.40	10.21
Ras Ouahrane	71.22	22.04	07.79

Conclusion

En conclusion, ce chapitre a permis de dresser un portrait détaillé des caractéristiques morphologiques des cinq sous-bassins versants étudiés dans le bassin du Cheliff. Nous avons examiné divers aspects tels que l'altitude, la pente, la topographie du bassin versant, ainsi que le réseau hydrographique. Ces éléments sont essentiels pour comprendre la dynamique hydrologique de la région. De plus, nous avons souligné que la zone d'étude est largement influencée par un climat semi-aride, ce qui a des répercussions importantes sur les ressources en eau et sur la gestion des bassins versants.

CHAPITRE V Traitement des données utilisées et
sélection des variables d'entrée des modèles

Introduction

Cette étude se concentre principalement sur l'application de modèles basés sur l'intelligence artificielle pour prédire le débit (Q) dans le bassin du Cheliff, à des pas de temps journaliers et mensuels. La qualité des données sélectionnées est cruciale pour garantir la performance optimale des modèles. Les conditions météorologiques jouent un rôle prépondérant dans la variation du débit des cours d'eau, avec les précipitations liquides en tant que principal facteur d'influence, les régimes hydrologiques distincts, les précipitations journalières maximales, et surtout, les épisodes d'inondations. En outre, les caractéristiques physiques telles que l'altitude, ainsi que les données climatiques telles que l'évapotranspiration, la température, l'humidité et la vitesse du vent, ont un impact sur le débit et le régime hydrologique du bassin.

Dans ce chapitre, nous allons définir, analyser et traiter l'ensemble des données collectées à partir de différentes sources, qui seront utilisées dans le développement des modèles proposés. Les données précipitations comprennent les totaux annuels, mensuelles et saisonnières.

V.1 Données en usage

Pour un bassin versant donné, les stations météorologiques et hydrométriques fournissent des mesures ponctuelles essentielles. Ces stations, installées dans des conditions spécifiques, forment un réseau d'observation crucial. En combinant les données provenant des pluviomètres et des stations hydrométriques, il est possible d'obtenir des informations précieuses sur la dynamique de l'eau à l'intérieur du bassin versant. En surveillant de manière continue les niveaux de précipitations et les débits d'eau, il devient alors possible d'améliorer la compréhension des flux hydriques dans le bassin versant, ce qui est indispensable pour prendre des décisions éclairées en matière d'utilisation et de gestion des ressources en eau.

Cette étude se divise en deux chapitres d'application dédiés à la prévision du débit des cours d'eau. Dans le premier chapitre, nous examinons l'efficacité des modèles d'apprentissage automatique (ML) autonomes ainsi qu'un modèle d'apprentissage extrême (ELM) optimisé grâce à un algorithme métaheuristique (Bat-ELM) pour prédire le débit des cours d'eau à l'échelle journalière et mensuelle. Dans le second chapitre, nous explorons des modèles ML autonomes ainsi que des modèles ML hybrides qui reposent sur des techniques de décomposition du signal appuyées par la

transformée en ondelettes à chevauchement maximal (MODWT) pour prédire le débit des cours d'eau à l'échelle journalière.

V.1.1 Collecte des données

Les données de cinq stations hydrométriques situées dans le bassin de Cheliff, qui est contrôlé par l'Agence Nationale des Ressources Hydriques (ANRH). La répartition de ces stations est entre le bassin de Bas Cheliff et de la Mina et le bassin de Haut et Moyen Cheliff.

La base de données utilisée dans l'application des modèles est constituée des données de précipitations (P), et de débits (Q) collectées auprès d'ANRH d'Alger.

V.2 Analyse des écoulements

Les données sur le débit des cours d'eau, collectées sur plusieurs années, révèlent des variations systématiques des principaux critères géographiques et physiques qui influent sur le débit des cours d'eau. Ainsi, le régime hydrologique d'un cours d'eau synthétise l'ensemble de ses caractéristiques hydrologiques et leur mode de variation. Comprendre l'évolution des données hydrométriques recueillies par une station hydrométrique est donc essentiel, car elles sont utilisées à la fois pour évaluer les apports liquides à différentes échelles de temps et pour analyser la variabilité potentielle du débit.

L'évolution des débits à différentes échelles, les bilans hydrologiques, les régimes fluviaux, les intensités extrêmes de crue et d'étiage et leur fréquence sont autant de paramètres pris en compte pour illustrer l'intérêt de l'hydrologie dans les principaux problèmes liés à la gestion de l'eau tels que le contrôle des régimes hydrologiques.

V.2.1 Variation des débits moyens journaliers

Les variations quotidiennes du débit dans un bassin versant peuvent être influencées par divers facteurs, notamment les précipitations, la végétation, le développement, la géologie et d'autres caractéristiques du bassin versant. La variabilité temporelle et spatiale de ces facteurs peut rendre difficile la caractérisation du régime d'écoulement d'un bassin versant.

Les séries des débits journaliers utilisées pour cette étude nous ont été fournies par l'Agence Nationale des Ressources Hydrauliques (ANRH). Le choix des stations utilisées a été fait selon la disponibilité des séries les plus complètes.

Une série de 3 année successive de débits journaliers pour les trois bassins versants Tiguiguest, Cheliff Harrezal et Oued Abda (du 01 Septembre 1995 au 31 Août 1998 pour la station hydrométrique d'Oued Lili, 01 Septembre 1989 au 31 Août 1992 pour la station d'El Ababsa et du 01 Septembre 1986 au 31 Août 1989 pour la station d'Arib Abda). Pour les deux bassins versants de Rouina Zeddine et Ras Ouahrane une série de 10 année successive de débits journaliers (du 01 Septembre 2000 au 31 Août 2010 pour la station hydrométrique de Bir Ouled Tahar et du 01 Septembre 1982 au 31 Août 1992 pour la station de Ponteba Barrage) Selon la disponibilité des données (Données successives complètes et pas des lacunes).

Les caractéristiques des stations hydrométriques choisies sont représentées dans le **tableau V.1**.

Tableau V.1 Caractéristiques journaliers des stations hydrométriques utilisées

Station	Code de Station	Oued	Latitude (N)	Longitude (E)	Min (m ³ /s)	Max (m ³ /s)	Moy (m ³ /s)	Ecart (m ³ /s)	Taille de Série (mois)
Oued Lili	012501	Tiguiguest	35° 52'	1° 26'	0.00	37.42	0.67	1.67	1096
El Ababsa	011715	Cheliff Harrezal	36° 16'	2° 08'	0.00	7.61	0.09	0.40	1096
Arib Abda	011801	Ebda	36° 32'	2° 03'	0.00	82.45	1.70	5.90	1096
Bir Ouled Tahar	011905	Rouina Zeddine	36° 6'	1° 57'	0.00	25.84	0.36	1.22	3652
Ponteba Barrage	012203	Ras Ouahrane	36° 25'	1° 52'	0.00	221.34	6.05	18.12	3653

En effet, il est nécessaire de procéder à des études de débits journaliers, notamment en raison de l'importance des variations des régimes fluviaux sur des périodes très courtes (24 heures) à l'échelle du bassin versant, ainsi que de leur impact sur le transport des matières solides. Le ruissellement

pluvial se développe surtout lorsque l'intensité de la pluie dépasse la capacité d'infiltration du sol, d'où le colmatage rapide par l'eau et l'apparition des premiers pics de crue (Bouzeria., 2018). Les variations des débits journaliers au niveau des cinq stations étudiées sont présentées dans (Figure V.1 et V.2). Le débit maximal à la station d'Oued Lili est de 37,42 m³/s, 7,61 m³/s à El Ababsa, 82,45 m³/s à Arib Abda, 25,84 m³/s à Bir Ouled Tahar et 222,34 m³/s au barrage de Ponteba.

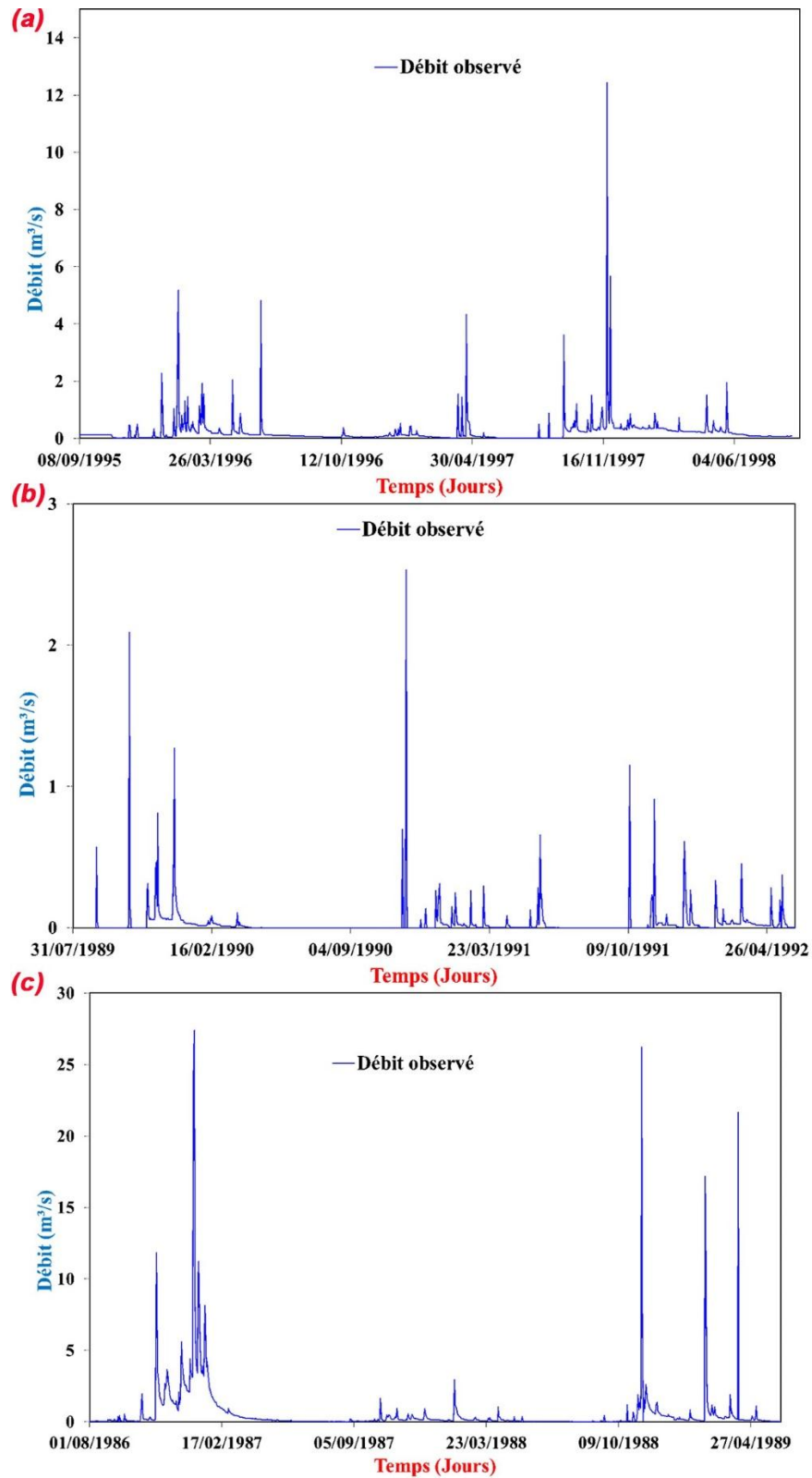


Figure V.1 Variations des débits journaliers pour les stations hydrométriques : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda

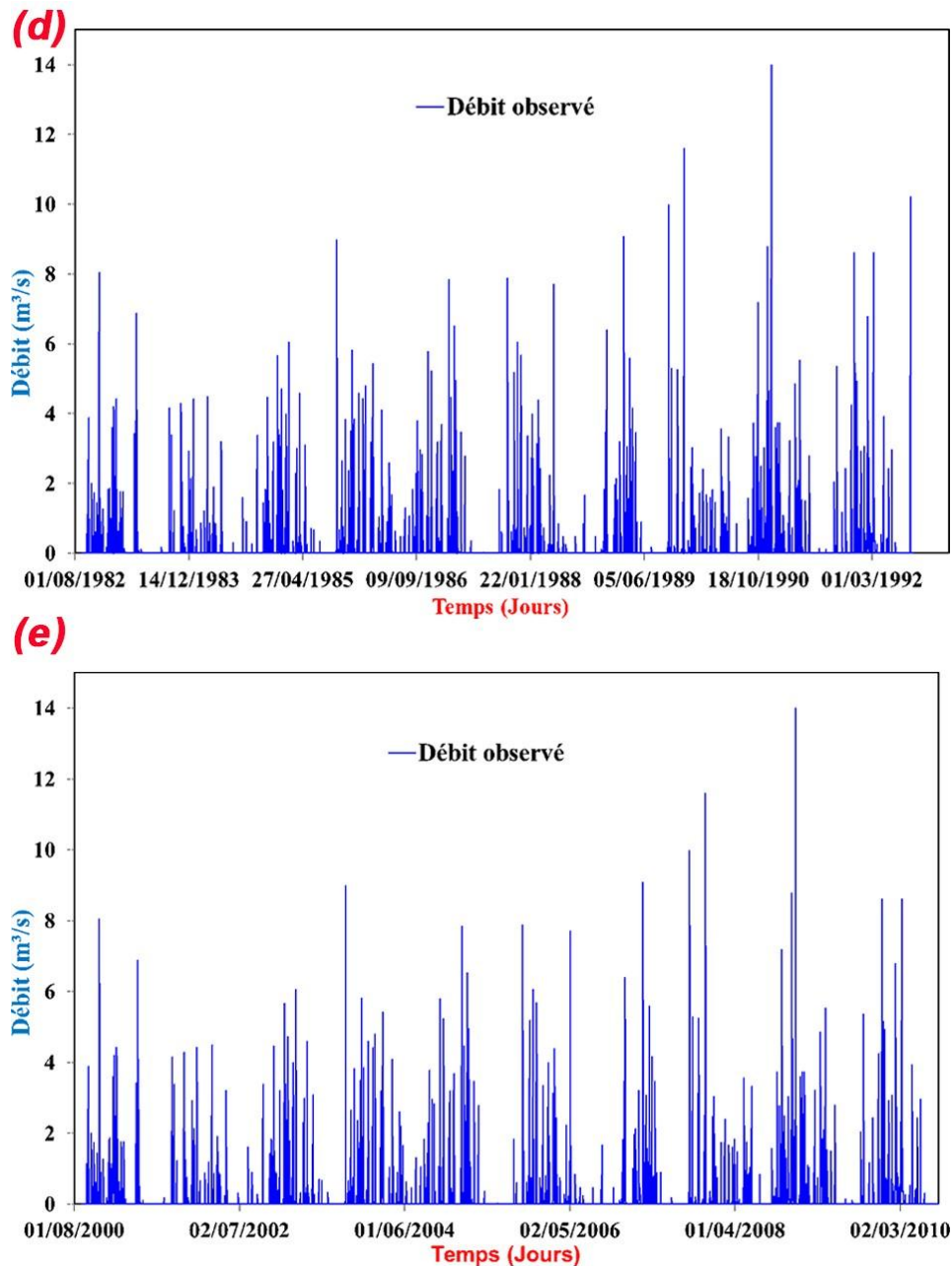


Figure V.2 Variations des débits journaliers pour les stations hydrométriques : (d) Bir Ouled Tahar, et (e) Ponteba Barrage

Les stations de Arib Abda, Oued Lili et El Ababsa furent utilisées pour la modélisation proposée dans cette recherche pour la prédiction des débits journaliers en utilisant des modèles d'apprentissage automatique autonomes et un modèle d'apprentissage extrême optimisé par un algorithme métaheuristique (Bat-ELM).

Ainsi que La station de Bir Ouled Tahar et Ponteba Barrage furent utilisées pour la 2^{-ème} modélisation proposée dans cette recherche pour la prédiction des débits journaliers en utilisant des modèles ML autonomes et des modèles ML hybrides basés sur des techniques de décomposition du signal soutenues par la transformée en ondelettes à chevauchement maximal (MODWT) pour prédire le débit des cours d'eau à l'échelle journalière.

V.2.2 Variation mensuelle de l'écoulement

Dans un bassin versant, les changements de précipitations et de température peuvent avoir une incidence significative sur les variations mensuelles du débit des cours d'eau. Les précipitations sont positivement liées au débit des cours d'eau, tandis que la température peut affecter le débit des cours d'eau en modifiant la fonte des neiges et l'évapotranspiration. La variabilité du climat à différentes échelles de temps peut également contrôler les variations mensuelles du débit.

Les équipements de la station hydrométrique permettent d'enregistrer en continu la hauteur du cours d'eau en fonction du temps. La transformation de la hauteur du plan d'eau en débit passe par la courbe de tarage qui traduit la relation entre la hauteur d'eau (cm) et le débit de ruissèlement (m^3/s).

Le sous bassin versant de l'oued de Tiguiguest est contrôlé par la station d'Oued Lili (code 012501) versant Bas Chélif et de Mina, l'Oued Harrezal est contrôlé par la station de d'El Ababsa (code 011715) ainsi que la station d'Arib Abda (code 011801) qui contrôle l'Oued Ebda dans le bassin versant Haut moyen Cheliff.

Une série de 27 années de débits mensuels de Septembre 1975 jusqu'à Aout 2002 pour la station d'Oued Lili, 34 années de débits mensuels de Septembre 1973 jusqu'à Aout 2007 pour la station d'El Ababsa et une série de 12 ans de débits mensuels de Septembre 1983 jusqu'à Aout 1995 pour la station d'Arib Abda.

La présentation de la variation temporelle des séries des débits mensuels pour chacune des stations est donnée par **la figure V.3**.

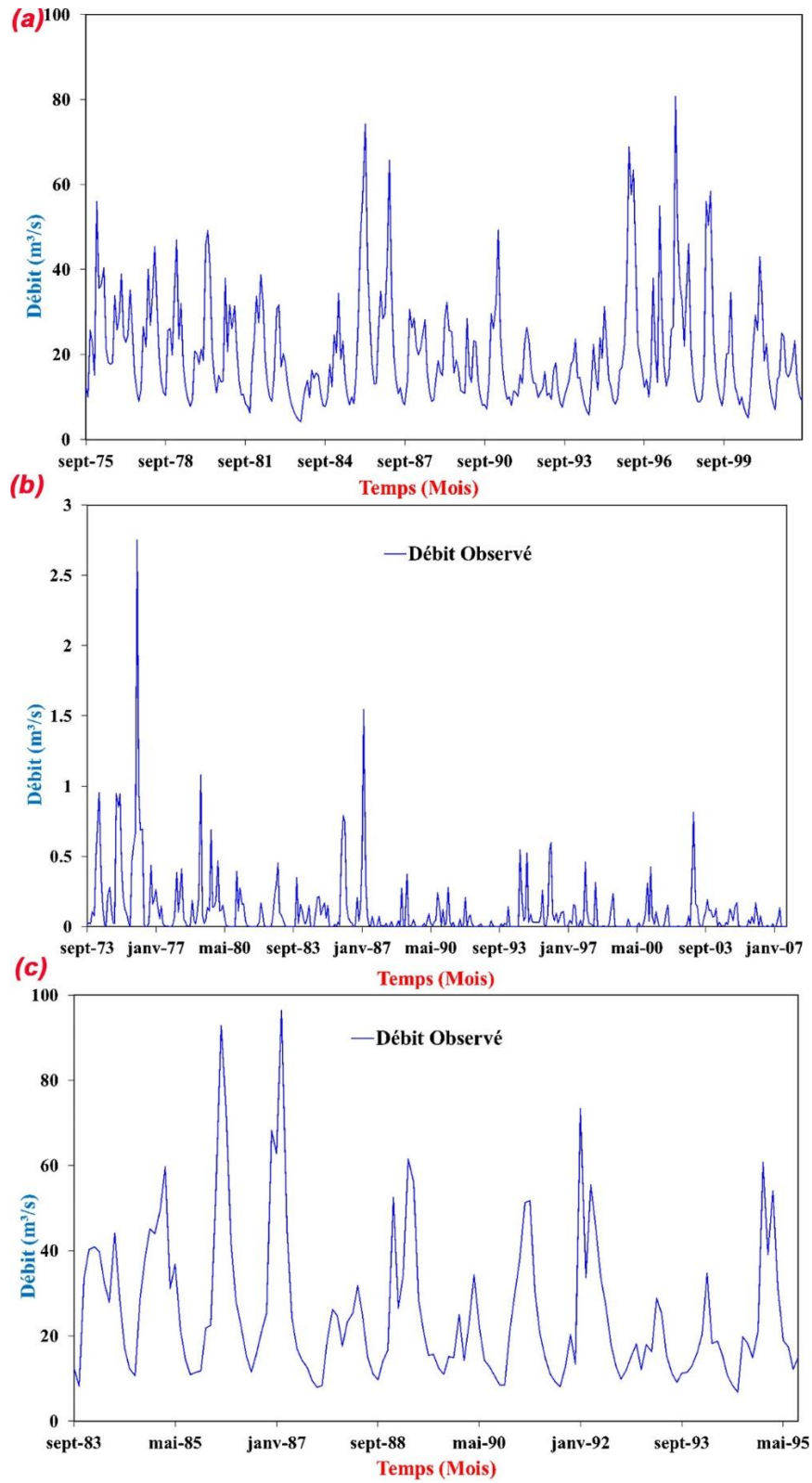


Figure V.3 Variation des débits mensuels : (a) Oued Lili, (b) El Ababsa, et (c) Arrib Abda

Ces trois séries temporelles furent utilisées pour la modélisation proposée dans cette recherche pour la prévision des débits mensuels en utilisant des modèles d'apprentissage automatique autonomes et un modèle d'apprentissage extrême optimisé par un algorithme métaheuristique (Bat-ELM). Les caractéristiques des stations hydrométriques choisies sont représentées dans le **tableau V.2**.

Tableau V.2 Caractéristiques mensuelles des stations hydrométriques utilisées

Station	Code de Station	Oued	Latitude (N)	Longitude (E)	Min (m ³ /s)	Max (m ³ /s)	Moy (m ³ /s)	Ecart (m ³ /s)	Type de Série (mois)
Oued Lili	012501	Tiguiguest	35° 52'	1° 26'	0.000	3.792	0.363	0.545	324
El Ababsa	011715	Cheliff Harrezal	36° 16'	2° 08'	0.000	2.750	0.110	0.234	408
Arib Abda	011801	Ebda	36° 32'	2° 03'	0.000	19.379	1.569	2.810	144

Le coefficient hydraulique ($C_{h,m}$) définit la variation du régime d'écoulement d'un cours d'eau à l'échelle mensuelle. Celui-ci correspond au rapport entre le débit mensuel moyen du mois considéré et le débit moyen interannuel obtenu par traitement à partir de la longue série que constitue l'observation des cours d'eau. Ainsi, la période de hautes eaux est caractérisée par un écoulement supérieur au seuil de référence 1, et contrairement à la période de basses eaux, le coefficient d'écoulement est inférieur à la valeur 1.

La distribution du coefficient d'hydraulicité mensuel au cours de l'année permet de définir la période de décembre à avril comme une période creuse, en raison de l'importance des précipitations qui tombent au cours de cette période (**figure V.4**). Pendant cette période, les débits moyens mensuels varient de 0,08 à 0,77 m³/s (pour la station hydrométrique d'Oued Lili), de 0,01 à 0,30 m³/s (pour la station d'El Ababsa) et de 0,04 à 3,97 m³/s (pour la station d'Arib Abda), le mois de février étant le plus humide et le mois de juillet le plus sec (**tableau V.3**).

Tableau V.3 Variation mensuelle de débits et de coefficients d'hydraulicité

	Station de Oued Lili	Station de El Ababsa	Station de Arib Abda

Mois	Q_{moy} (m ³ /s)	A (Hm ³)	$C_{h.m}$	Q_{moy} (m ³ /s)	A (Hm ³)	$C_{h.m}$	Q_{moy} (m ³ /s)	A (Hm ³)	$C_{h.m}$
Septembre	0.19	0.50	0.53	0.07	0.17	0.60	0.11	0.29	0.07
Octobre	0.26	0.71	0.73	0.09	0.23	0.80	1.25	3.34	0.79
Novembre	0.44	1.14	1.21	0.10	0.25	0.88	0.94	2.45	0.60
Décembre	0.37	0.99	1.02	0.10	0.27	0.91	1.91	5.12	1.22
Janvier	0.57	1.54	1.59	0.16	0.42	1.43	3.39	9.07	2.16
Février	0.77	1.99	2.12	0.30	0.77	2.69	3.97	10.29	2.53
Mars	0.63	1.69	1.74	0.23	0.63	2.13	3.88	10.40	2.48
Avril	0.40	1.03	1.10	0.14	0.37	1.30	1.83	4.75	1.17
Mai	0.31	0.82	0.85	0.08	0.22	0.73	1.04	2.79	0.66
Juin	0.19	0.50	0.53	0.03	0.07	0.24	0.36	0.93	0.23
Juillet	0.08	0.21	0.22	0.02	0.04	0.15	0.09	0.25	0.06
Août	0.14	0.37	0.38	0.01	0.04	0.13	0.04	0.11	0.03
Moyenne	0.36	0.96	1.00	0.11	0.29	1.00	1.57	4.15	1.00
Min.	0.08	0.21	0.22	0.01	0.04	0.13	0.04	0.11	0.03
Max.	0.77	1.99	2.12	0.30	0.77	2.69	3.97	10.40	2.53
Ecart-type	0.21	0.56	0.58	0.09	0.23	0.78	1.46	3.87	0.93

La période des hautes eaux commence dès Septembre et Octobre pour les trois bassins versants à la station d'Oued Lili, El Ababsa et d'Arrib Abda, Mais les coefficients mensuels de débits tombent rapidement au-dessous de l'unité pendant le mois de Novembre et Décembre avant d'augmenter à nouveau en Janvier et en Février, le maximum est atteint en Mars et Avril. Par ailleurs, la période de basses eaux est décalée dans le temps, elle ne commence qu'au mois de Juin du fait de l'importance relative des précipitations de printemps, la moyenne mensuelle minimale est enregistrée en Juillet.

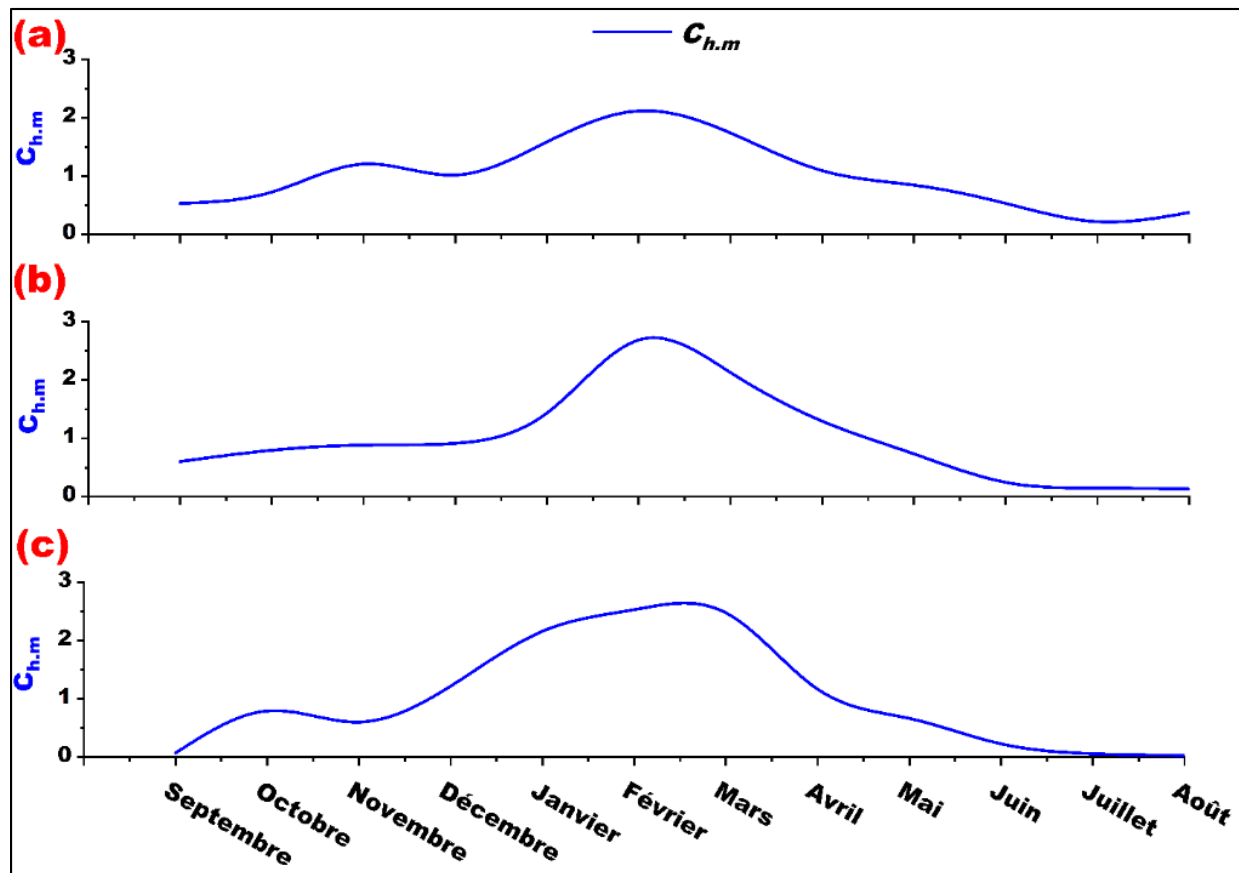


Figure V.4 Représentation graphique des coefficients d'hydraulicité mensuels : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda

V.2.3 Variation interannuelle de l'écoulement

La variation du régime d'écoulement se fait d'une année à l'autre, avec une amplitude de débit importante pour les trois stations hydrométriques ; le bassin versant de Tiguiguest, dont l'exutoire est la station d'Oued Lili, a des apports liquides annuels de $11,44 \text{ Hm}^3$, ce qui représente un débit moyen de $0,36 \text{ m}^3/\text{s}$ ainsi qu'une superficie contrôlée de $1\,617,56 \text{ km}^2$. Tandis que le bassin versant du Cheliff Harrezal, rejeté à la station d'El Ababsa, a un débit moyen de $0,11 \text{ m}^3/\text{s}$ et reçoit $3,47 \text{ Hm}^3$ annuellement, pour une superficie de drainage de $756,70 \text{ km}^2$, ainsi que le bassin versant de l'Oued Ebda, rejeté à la station d'Arib Abda, a un débit moyen de $1,57 \text{ m}^3/\text{s}$ et reçoit $49,48 \text{ Hm}^3$ annuellement, pour une superficie drainée de $661,45 \text{ km}^2$ ([tableau V.4](#)) ; la représentation graphique des coefficients hydrauliques annuels est présentée dans [la figure V.5](#).

Tableau V.4 Variation interannuelle de débits et de coefficients d'hydraulicité

Année	Station de Oued Lili			Station de El Ababsa			Station de Arib Abda		
	Q_{moy} (m ³ /s)	A (Hm ³)	$C_{h.a}$	Q_{moy} (m ³ /s)	A (Hm ³)	$C_{h.a}$	Q_{moy} (m ³ /s)	A (Hm ³)	$C_{h.a}$
1972-1973	/	/	/	0.24	7.61	2.19	/	/	/
1973-1974	/	/	/	0.36	11.20	3.22	/	/	/
1974-1975	0.35	10.97	0.96	0.57	18.00	5.18	/	/	/
1975-1976	0.41	13.06	1.14	0.12	3.89	1.12	/	/	/
1976-1977	0.20	6.17	0.54	0.11	3.47	1.00	/	/	/
1977-1978	0.53	16.63	1.45	0.17	5.34	1.54	/	/	/
1978-1979	0.39	12.15	1.06	0.17	5.43	1.56	/	/	/
1979-1980	0.23	7.25	0.63	0.10	3.04	0.87	/	/	/
1980-1981	0.36	11.29	0.99	0.03	0.84	0.24	/	/	/
1981-1982	0.26	8.05	0.70	0.11	3.39	0.98	/	/	/
1982-1983	0.14	4.43	0.39	0.08	2.40	0.69	1.23	38.91	0.79
1983-1984	0.09	2.83	0.25	0.10	3.00	0.86	3.28	103.39	2.09
1986-1987	0.80	25.08	2.19	0.21	6.56	1.89	1.72	54.22	1.10
1987-1988	0.38	12.04	1.05	0.24	7.60	2.19	3.30	104.10	2.10
1988-1989	0.24	7.65	0.67	0.01	0.45	0.13	0.41	13.05	0.26
1989-1990	0.24	7.61	0.67	0.06	2.04	0.59	1.46	46.14	0.93
1990-1991	0.17	5.35	0.47	0.02	0.72	0.21	0.26	8.09	0.16
1991-1992	0.11	3.33	0.29	0.07	2.35	0.68	1.70	53.58	1.08
1992-1993	0.20	6.24	0.55	0.04	1.15	0.33	2.95	92.99	1.88
1993-1994	0.06	1.77	0.15	0.01	0.22	0.06	0.44	13.87	0.28
1994-1995	0.09	2.85	0.25	0.02	0.53	0.15	0.32	10.17	0.21
1995-1996	1.21	38.18	3.34	0.14	4.31	1.24	1.75	55.22	1.12
1996-1997	0.69	21.91	1.92	0.16	5.00	1.44	/	/	/
1997-1998	0.29	9.10	0.80	0.05	1.68	0.48	/	/	/
1998-1999	1.04	32.79	2.87	0.08	2.65	0.76	/	/	/
1999-2000	0.33	10.48	0.92	0.03	1.09	0.31	/	/	/
2000-2001	0.08	2.54	0.22	0.01	0.22	0.06	/	/	/
2001-2002	0.62	19.59	1.71	0.09	2.98	0.86	/	/	/

2002-2003	0.30	9.41	0.82	0.02	0.76	0.22	/	/	/
2003-2004	/	/	/	0.14	4.35	1.25	/	/	/
2004-2005	/	/	/	0.07	2.28	0.66	/	/	/
2005-2006	/	/	/	0.05	1.67	0.48	/	/	/
2006-2007	/	/	/	0.04	1.32	0.38	/	/	/
2007-2008	/	/	/	0.02	0.61	0.18	/	/	/
Moyenne	0.36	11.44	1.00	0.11	3.47	1.00	1.57	49.48	1.00
Min.	0.06	1.77	0.15	0.01	0.22	0.06	0.26	8.09	0.16
Max.	1.21	38.18	3.34	0.57	18.00	5.18	3.30	104.10	2.10
Ecart-type	0.29	9.11	0.80	0.11	3.59	1.03	1.12	35.46	0.72

Dans le bassin versant du Tiguiguest, l'année la plus humide est révélée en 1995/1996. Au cours de cette année, la station de l'Oued Lili a enregistré un volume d'eau important (38,18 Hm³), avec un débit moyen annuel de 1,21 m³/s pour la station de l'Oued Lili. Cette valeur est principalement due aux pluies torrentielles. Alors que. L'année la plus humide en 18 ans d'observation.

En 1974/1975, année la plus pluvieuse pour la station d'El Ababsa, avec un débit entrant annuel de l'ordre de 18,00 Hm³) et un débit moyen annuel de 0,57 m³/s. Ces apports importants peuvent conduire à des crues exceptionnelles, caractérisées par des débits de pointe extrêmement élevés, qui provoquent généralement des inondations graves et dangereuses.

La station d'Arib Abda se caractérise par trois années hydrologiques très humides au cours de la période d'observation, notamment 1983/1984, 1987/1988 et 1992/1993, 1987/1984 étant la plus importante de ces années, avec un total annuel de 104.10 Hm³, pour un débit moyen annuel de 3.30 m³/s.

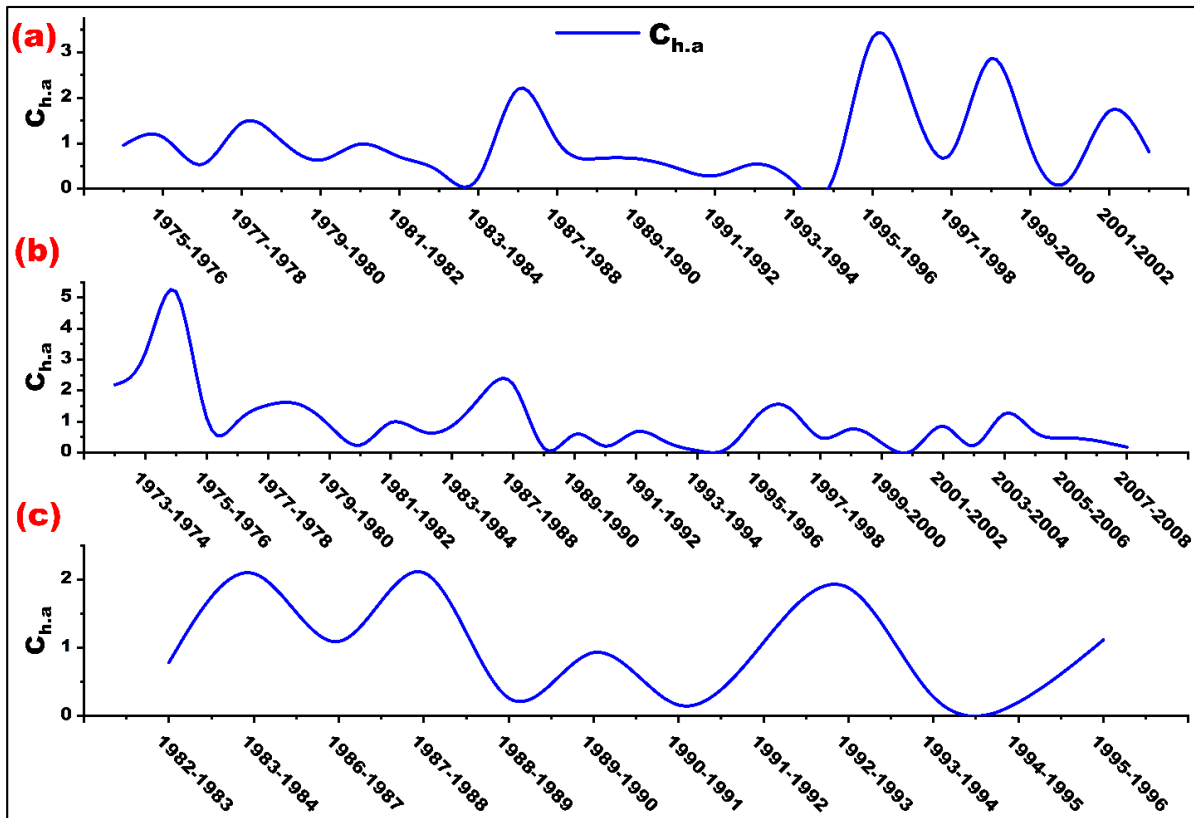


Figure V.5 Représentation graphique des coefficients d'hydraulicité annuels : (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda

V.3 Analyse des données pluviométriques

V.3.1 Problème des lacunes

Le traitement des données incomplètes est un problème réel et souvent déroutant. Bien que cela puisse ne pas avoir de répercussions significatives lorsqu'il y a une grande quantité de données disponibles, cela peut rendre l'étude peu intéressante si le nombre de données restantes est trop faible.

Dans le bassin du Cheliff, plusieurs stations pluviométriques en service peuvent être utilisées, mais elles présentent souvent des lacunes correspondant à des années et des mois d'observations incomplètes, qu'il est difficile de combler. Les stations avec plus de 10 % de lacunes sont directement éliminées. Dans les petits bassins, en raison du manque de données, une seule station peut être utilisée, ce qui rend l'étude spécifique au sous-bassin versant.

Pour cette étude, cinq stations pluviométriques ont été sélectionnées parmi les stations disponibles, une pour chaque sous-bassin : Oued Lili, El Ababsa, Arib Abda, Bir Ouled Tahar et Ponteba Barrage (Figure V.6). Les données de séries temporelles quotidiennes et mensuelles de ces stations ne dépassent pas 10 % de lacunes.

Le comblement des lacunes d'observation a été effectué par deux méthodes : la méthode de la moyenne mensuelle et la méthode de la régression linéaire.

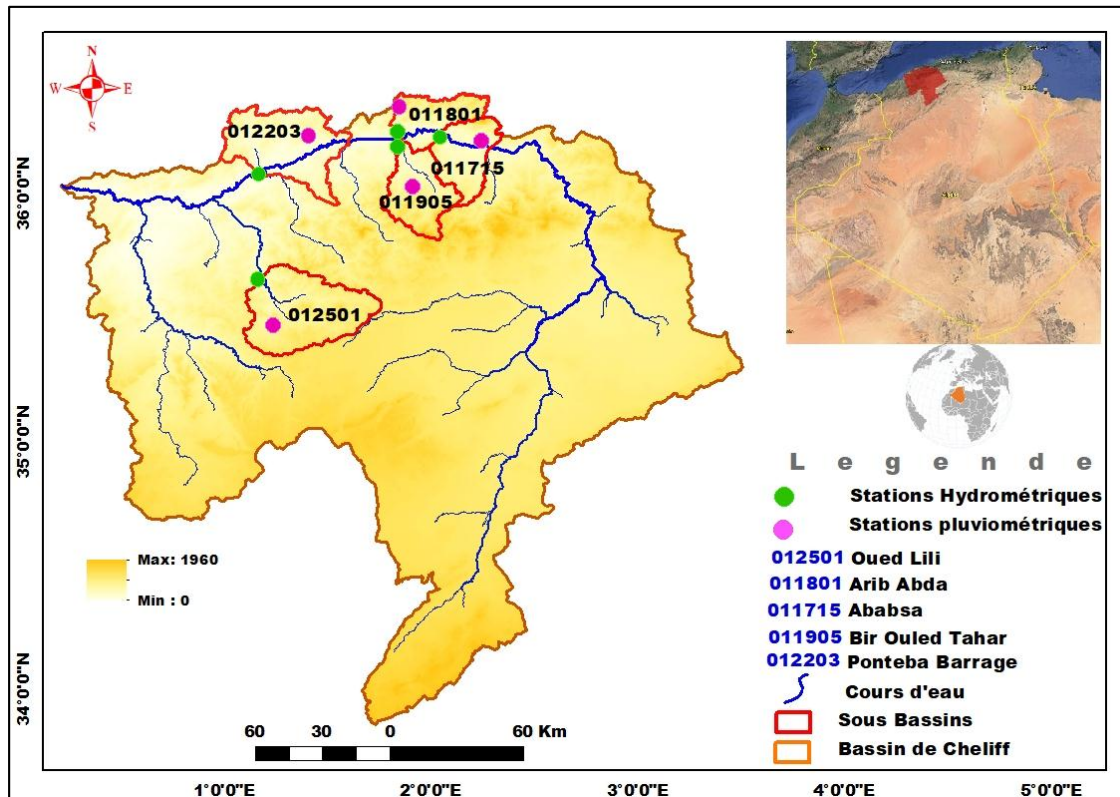


Figure V.6 Localisation des stations pluviométriques et hydrométriques dans le bassin de Cheliff

V.4 Séries chronologiques et leur utilisation

Une série chronologique est une séquence de points de données qui se produisent dans un ordre successif sur une certaine période de temps. On peut l'opposer aux données transversales, qui capturent un point dans le temps. Le temps est défini comme variable discrète et les données observées peuvent être quotidiennes, hebdomadaires, mensuelles, L'utilisation de ces séries peut avoir plusieurs objectifs, Parmi ces objectifs la prédiction et la prévision future à partir de celles qui ont été observées précédentes.

V.5 Différentes phases de la modélisation

La prédiction de séries chronologiques consiste à estimer ou anticiper les valeurs futures d'une variable sur une période de temps donnée. Elle implique la création d'un modèle à partir de données historiques afin de projeter des informations et d'orienter les décisions stratégiques à venir.

Au cours des dernières décennies, de nombreuses recherches sur la prédiction des séries temporelles ont montré la nécessité de suivre certaines étapes essentielles afin d'obtenir des résultats fiables, qui peuvent être résumées en quatre axes principaux.

- 1- Identification et séparation de la série chronologique ;
- 2- Identification du type du modèle ;
- 3- Estimation des paramètres du modèle (apprentissage) ;
- 4- Tests d'adéquation des modèles (test).

Dans le but de procéder à la modélisation d'une série temporelle, il convient de séparer la série en deux sections : la première pour l'apprentissage du modèle sélectionné ainsi que la seconde pour le test du modèle lui-même. L'apprentissage d'un modèle est un problème en général, dans la plupart des cas pratiques nous avons un modèle pour un phénomène physique, à savoir un modèle qui comprend plusieurs paramètres inconnus. Il convient de déterminer ces paramètres à partir des observations connues, cette détermination est connue sous le nom de "training" du modèle. Pour cela, il faudrait répartir les données selon un ensemble d'apprentissage et un ensemble de test. Nous formons d'abord notre modèle sur l'ensemble d'apprentissage, puis nous utilisons les données de l'ensemble de test pour évaluer la précision du modèle obtenu.

De nombreuses études soulignent la pertinence de l'utilisation de 70% des données pour la l'apprentissage et de 30% pour les tests, ainsi que la démonstration statistique.

La répartition des données journalières (débit et précipitations) utilisées dans cette étude est présentée dans **la figure V.7** reliant les stations Oued Lili, El Ababsa et Arib Abda, respectivement, et **la figure V.8** reliant les stations Bir Ouled Tahar et Ponteba Barrage, également, et dans **la figure V.9** la répartition des ensembles de données mensuelles des stations Oued Lili, El Ababsa et Arib Abda.

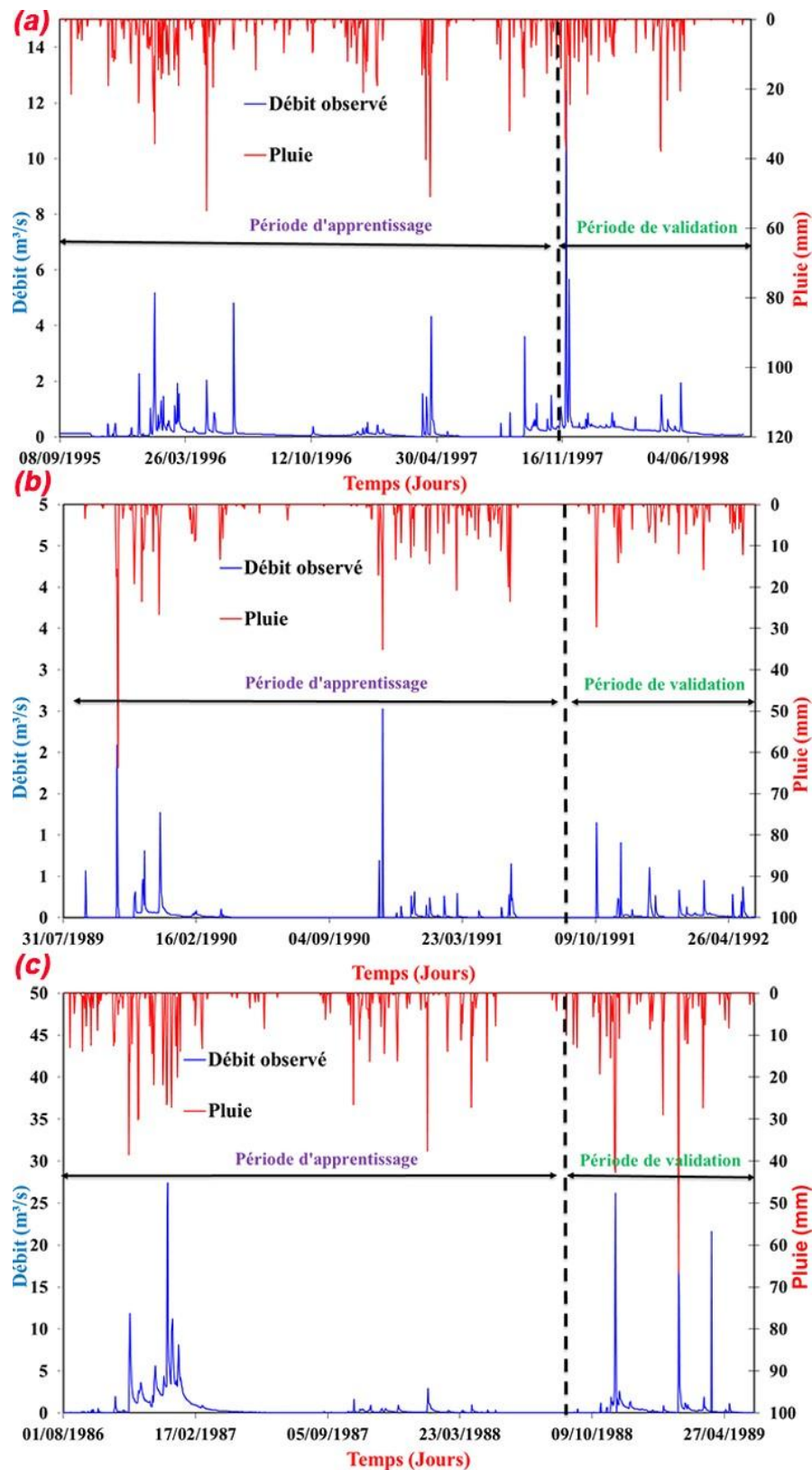


Figure V.7 Séparation de la série des débits journalières des stations (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda

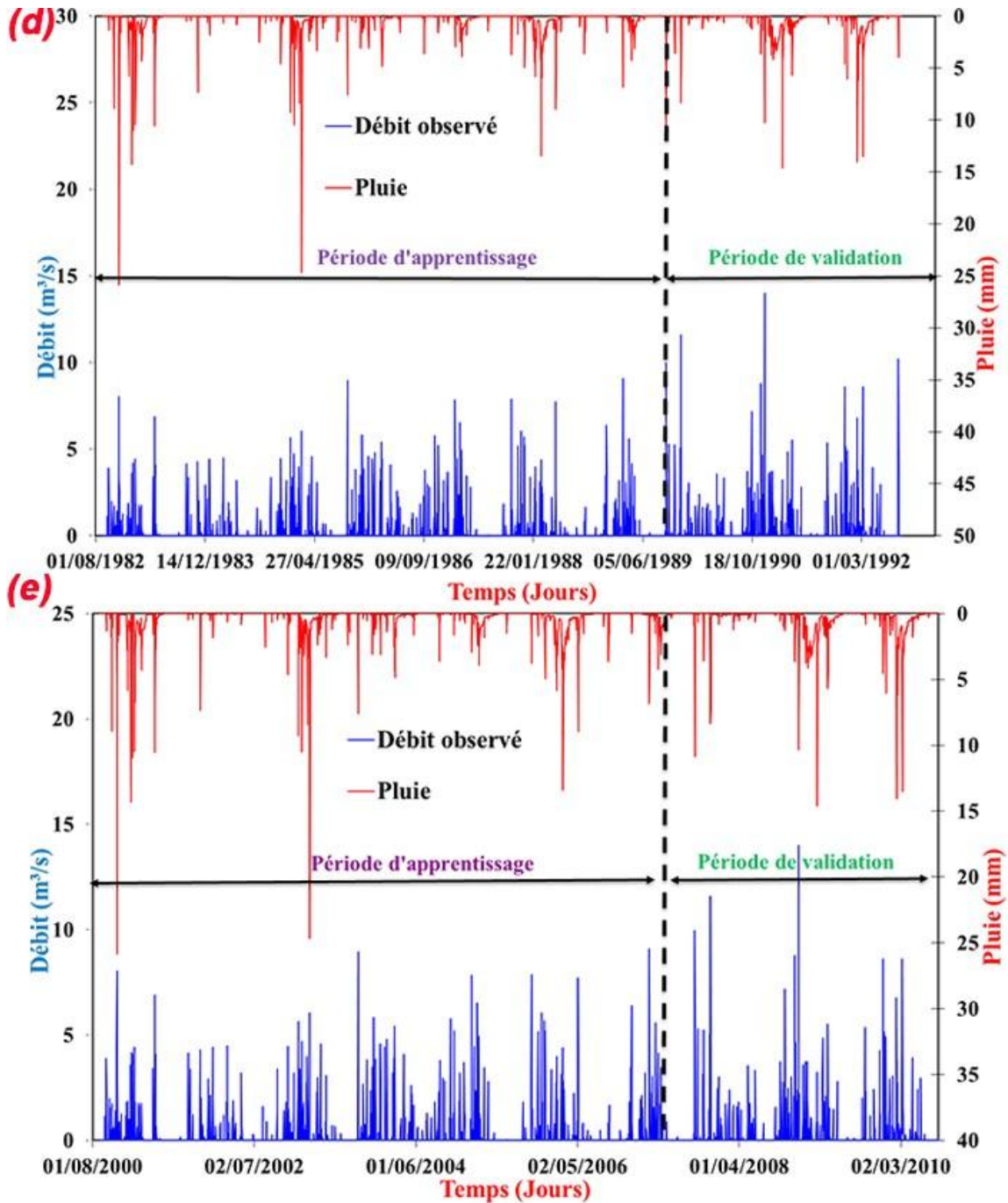


Figure V.8 Séparation de la série des débits journalières des stations (d) Bir Ouled Tahar et (e) Ponteba Barrage

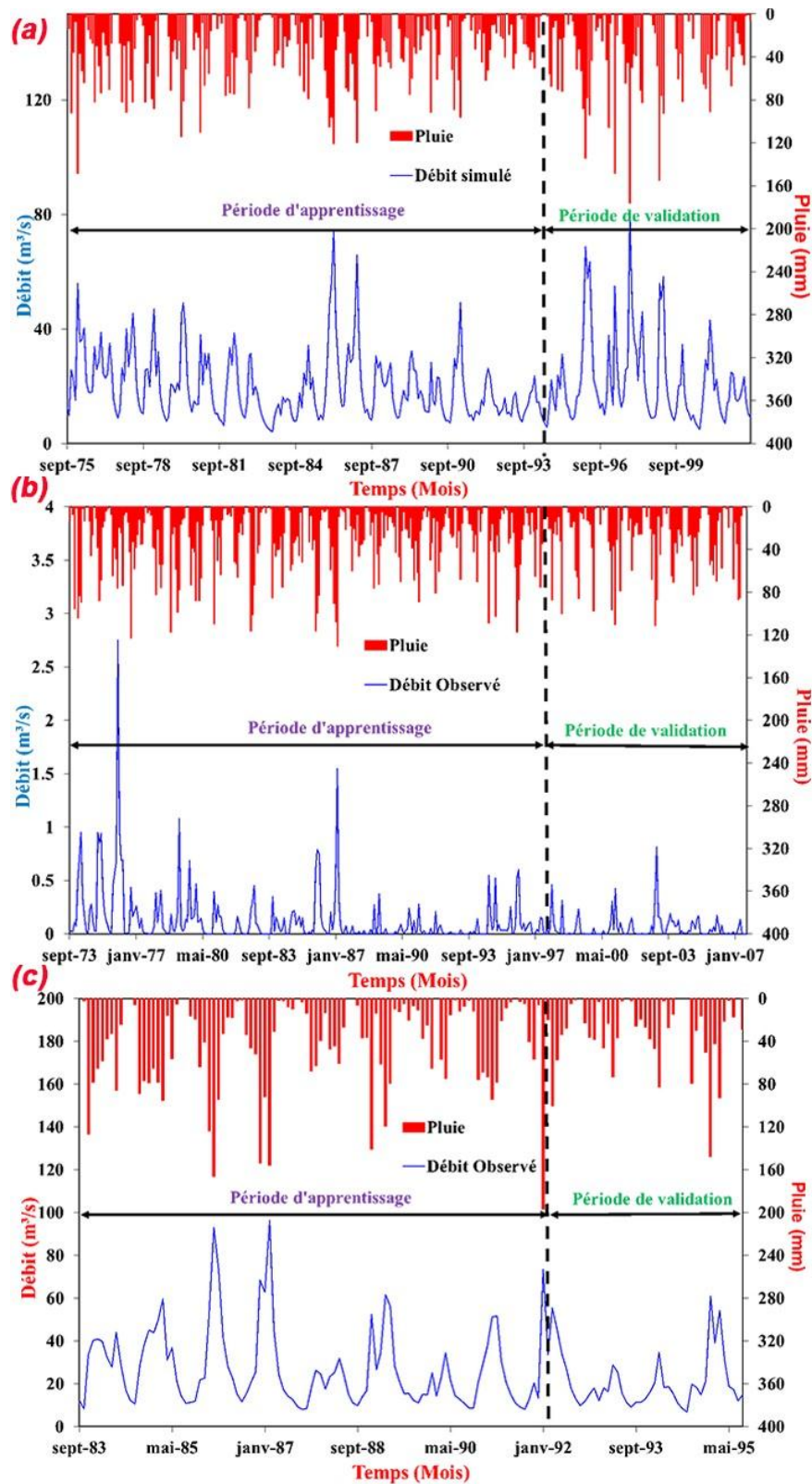


Figure V.9 Séparation de la série des débits mensuels des stations (a) Oued Lili, (b) El Ababsa, et (c) Arib Abda

Les paramètres statistiques de l'ensemble de données utilisé dans cette recherche pour le pas journalier et mensuelle sont résumés dans le **tableau V.5, V.6 et le tableau V.7**, nous avons calculé les paramètres statistiques pour les différentes variables climatiques au pas de temps journalier et mensuelle

- La valeur maximale (X_{\max}) qui est la valeur la plus élevée de la série et qui correspond à la borne supérieure de la série ;
- La valeur moyenne (X_{moy}) qui est égale à la somme des valeurs divisée par leur nombre ;
- La valeur minimale (X_{\min}) qui est la valeur la plus faible de la série et qui constitue la borne inférieure de la série ;
- L'écart type (S_x) qui représente la caractéristique de dispersion la plus usuelle ;
- Le coefficient de corrélation linéaire (R) permet d'illustrer la relation entre P et Q ;
- Le coefficient de variation (C_v) est le rapport de l'écart-type à la moyenne. Plus la valeur de coefficient de variation est élevée plus la dispersion autour de la moyenne est grande.

Tableau V.5 Caractéristiques statistiques des parties : Apprentissage, test et totale des données pour les stations Oued Lili, El Ababsa, et Arib Abda à l'échelle journalière

Variables	Subset	Unité	X_{max}	X_{moy}	X_{min}	S_x	C_v	R
<i>Oued Lili (station ID : 012501)</i>								
Q	Apprentissage	m ³ /s	15,61	0,51	0,00	1,25	2,46	1,00
	Validation		37,42	1,04	0,21	2,32	2,23	1,00
	Total		37,42	0,67	0,00	1,67	2,49	1,00
P	Apprentissage	mm	55,00	1,67	0,00	5,20	3,12	0,34
	Validation		38,00	1,64	0,00	5,23	3,20	0,39
	Total		55,00	1,66	0,00	5,20	3,14	0,34
<i>El Ababsa (station ID : 011715)</i>								
Q	Apprentissage	m ³ /s	7,61	0,09	0,00	0,43	4,97	1,00
	Validation		3,46	0,11	0,00	0,31	2,98	1,00
	Total		7,61	0,09	0,00	0,40	4,33	1,00
P	Apprentissage	mm	63,70	1,00	0,00	3,99	4,00	0,40
	Validation		29,60	0,93	0,00	3,04	3,27	0,64
	Total		63,70	0,98	0,00	3,73	3,81	0,44
<i>Arib Abda (station ID : 011801)</i>								
Q	Apprentissage	m ³ /s	82,45	1,72	0,00	5,42	3,15	1,00
	Validation		78,87	1,64	0,00	6,89	4,21	1,00
	Total		82,45	1,70	0,00	5,90	3,48	1,00
P	Apprentissage	mm	38,50	1,28	0,00	4,36	3,41	0,26
	Validation		66,70	1,58	0,00	6,15	3,88	0,49
	Total		66,70	1,37	0,00	4,96	3,62	0,36

Tableau V.6 Caractéristiques statistiques des parties : Apprentissage, test et totale des données pour les stations Bir Ouled Tahar et Ponteba Barrage à l'échelle journalière

Variables	Subset	Unité	X_{max}	X_{moy}	X_{min}	S_x	C_v	R
<i>Bir Ouled Tahar (station ID: 011905)</i>								
<i>Q</i>	Apprentissage	m ³ /s	25.84	0.33	0.00	1.22	3.72	1.00
	Validation		14.62	0.44	0.00	1.22	2.80	1.00
	Total		25.84	0.36	0.00	1.22	3.39	1.00
<i>P</i>	Apprentissage	mm	27.30	0.80	0.00	2.68	3.33	0.30
	Validation		42.10	1.00	0.00	3.54	3.55	0.30
	Total		42.10	0.87	0.00	2.97	3.40	0.29
<i>Ponteba Barrage (station ID : 012203)</i>								
<i>Q</i>	Apprentissage	m ³ /s	193.13	6.62	0.00	18.70	2.83	1.00
	Validation		221.34	4.74	0.00	16.62	3.50	1.00
	Total		221.34	6.05	0.00	17.66	2.99	1.00
<i>P</i>	Apprentissage	mm	47.20	1.03	0.00	3.71	3.62	0.13
	Validation		27.20	0.85	0.00	2.85	3.37	0.07
	Total		47.20	0.97	0.00	3.48	3.59	0.12

Tableau V.7 Caractéristiques statistiques des parties : Apprentissage, test et totale des données pour les stations d'étude à l'échelle mensuelle

Variables	Subset	Unité	X_{max}	X_{moy}	X_{min}	S_x	C_v	R
<i>Oued Lili (station ID : 012501)</i>								
Q	Apprentissage	m ³ /s	3.792	0.311	0.000	0.510	1.638	1.000
	Validation		3.437	0.530	0.000	0.596	1.124	1.000
	Total		3.792	0.373	0.000	0.545	1.460	1.000
P	Apprentissage	mm	148.400	31.309	0.000	30.522	0.975	0.541
	Validation		176.100	34.377	0.000	38.887	1.131	0.664
	Total		176.100	32.263	0.000	33.089	1.026	0.584
<i>El Ababsa (station ID : 011715)</i>								
Q	Apprentissage	m ³ /s	2.750	0.142	0.000	0.266	1.874	1.000
	Validation		0.815	0.053	0.000	0.106	1.980	1.000
	Total		2.750	0.117	0.000	0.234	2.008	1.000
P	Apprentissage	mm	130.700	30.268	0.000	30.751	1.016	0.559
	Validation		111.000	26.262	0.000	29.684	1.130	0.661
	Total		130.700	29.132	0.000	30.460	1.046	0.548
<i>Arib Abda (station ID : 011801)</i>								
Q	Apprentissage	m ³ /s	19.379	1.842	0.000	2.940	1.596	1.000
	Validation		8.553	1.031	0.000	1.957	1.898	1.000
	Total		19.379	1.692	0.000	2.810	1.661	1.000
P	Apprentissage	mm	166.100	41.462	0.000	39.771	0.959	0.706
	Validation		147.500	29.651	0.000	31.045	1.047	0.657
	Total		196.800	39.692	0.000	40.428	1.019	0.721

V.6 Sélection des variables d'entrée pour les modèles.

V.6.1 Fonction d'autocorrélation (ACF) et les d'autocorrélation partiels (PACF) et fonction de corrélation croisée (XCF)

L'autocorrélation est la corrélation entre deux valeurs d'une série temporelle. En d'autres termes, les données de la série temporelle sont corrélées dans leur ensemble, ainsi que le nom l'indique. Les corrélations sont désignées par le terme "décalages". Ce terme est utilisé pour décrire ces corrélations. Les analystes enregistrent des données de séries temporelles en mesurant une caractéristique à intervalles réguliers, par exemple quotidiennement, mensuellement ou annuellement. Le nombre d'intervalles entre les deux observations est le décalage. Le décalage entre l'observation actuelle et l'observation précédente, par exemple, est égal à 1. Si vous reculez d'un intervalle supplémentaire, le décalage est de 2 et ainsi de suite.

En termes mathématiques, les observations à y_t et y_{t-k} sont séparées par k unités de temps. K est le décalage. Ce décalage peut être de quelques jours, de quelques trimestres ou de quelques années, selon la nature des données. Lorsque $k=1$, les observations adjacentes sont évaluées. Pour chaque décalage, il existe une corrélation.

L'ACF est utilisé pour déterminer le degré de corrélation linéaire entre les observations d'une série temporelle qui sont séparées par un décalage temporel k . L'idée principale de l'ACF est de calculer le coefficient de corrélation entre la série et la série elle-même avec un décalage k . Une corrélation élevée est sélectionnée comme variable d'entrée, ce qui signifie que les variables qui n'ont pas un coefficient de corrélation significatif sont éliminées du vecteur d'entrée.

Le PACF est utilisé pour identifier le nombre de termes autorégressifs requis pour mettre en évidence la saisonnalité de la série, ou pour faire apparaître la tendance de la moyenne ou de la variance de la série.

La corrélation croisée (XCF) est une mesure qui suit les mouvements de deux ou plusieurs ensembles de données de séries temporelles les uns par rapport aux autres. Elle est utilisée pour comparer plusieurs séries temporelles et déterminer objectivement dans quelle mesure elles correspondent l'une à l'autre et, en particulier, à quel moment la meilleure correspondance se produit.

La corrélation croisée est généralement utilisée pour mesurer les informations entre deux séries temporelles différentes. Le coefficient de corrélation des données de la série temporelle peut varier de $-1,0$ à $+1,0$. Plus la valeur de la corrélation croisée est proche de 1, plus les séries sont identiques.

La représentation graphique de la fonction d'autocorrélation (ACF) est appelée corrélogramme. Sur ce dernier, chaque barre représente la taille et la direction de la corrélation. Les barres qui dépassent la ligne sont statistiquement significatives.

Les figures V.10, V.11 et V.12 présentent les tracés des ACF, PACF et XCF des séries de débit journalière et mensuel des stations d'étude.

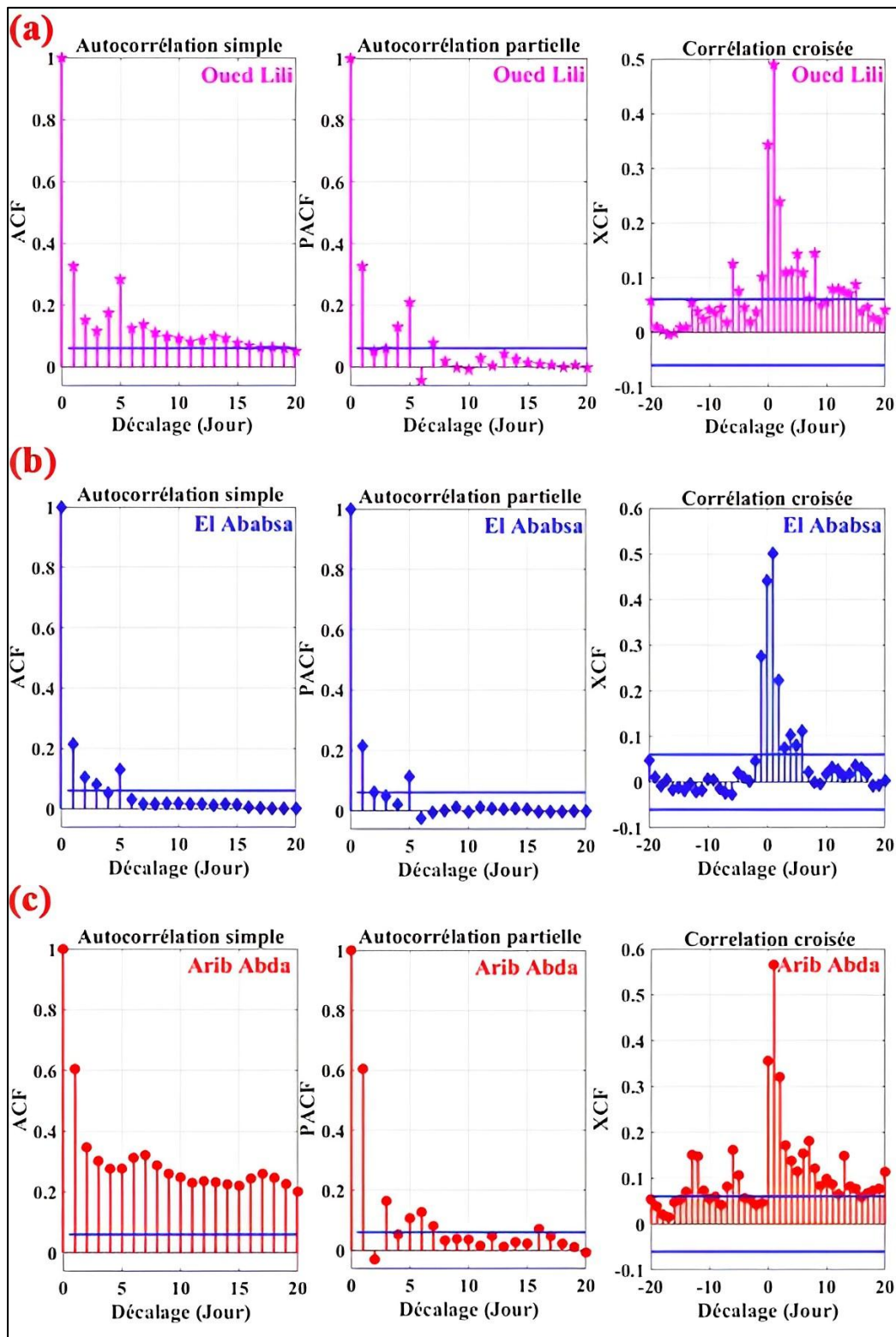


Figure V.10 Autocorrélation (ACF), fonction d'autocorrélation partielle (PACF) et fonction de corrélation croisée (XCF) des séries de débit journalier pour les stations (a) Oued Lili, (b) El Ababsa et (c) Arib Abda

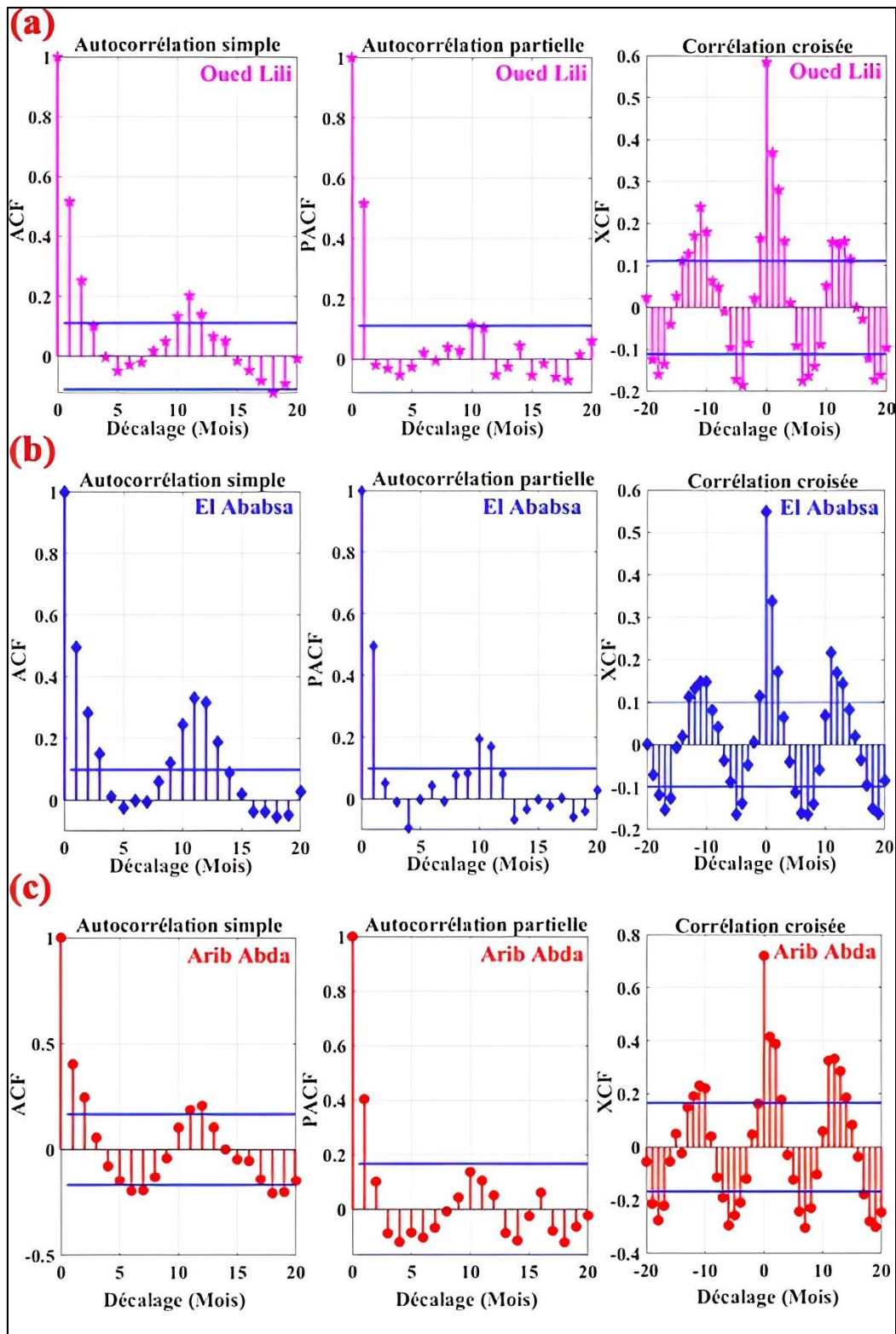


Figure V.11 Autocorrélation (ACF), fonction d'autocorrélation partielle (PACF) et fonction de corrélation croisée (XCF) des séries de débit mensuelles pour les stations (a) Oued Lili, (b) El Ababsa et (c) Arib abda

Le débit journalier et mensuel (Q) à l'instant (t) Pour Les trois stations Oued Lili (code 012501), El Ababsa (code 011715) et Arib Abda (code 011801) a été prédit en utilisant plusieurs temps de décalage de Q et P, et afin de déterminer le meilleur temps de décalage pertinent, l'ACF, le PACF, et le XCF pour les trois stations ont été calculés, comme indiqué dans la figure IV.10 et IV.11. En outre, dans la présente étude, nous avons adopté une approche pour prédire le débit en utilisant uniquement P et Q, étant donné que seules les données P sont disponibles simultanément sans enregistrement complet avec le débit. Comme la montre la figure IV.10 et IV.11, trois décalages temporels ont été sélectionnés et utilisés comme variables d'entrée, à savoir le débit journalière et mensuel mesuré à (t - 1), (t - 2), (t - 3) et les précipitations journalières et mensuelles mesurées à (t - 1), (t - 2), (t - 3) ; tandis que la variable de sortie commence à partir du temps (t), c'est-à-dire Q au temps (t). Ainsi, six combinaisons d'entrées sont adoptées et utilisées pour la présente étude (tableau V.8).

Tableau V.8 Combinaisons des variables d'entrées de différents modèles pour la station d'Oued Lili, El Ababsa et Arib abda dans le pas journalier et mensuelle

N Modèle	Combinaison des variables d'entré	Sortie
1	$Q(t-3), Q(t-2), Q(t-1), P(t-3), P(t-2), P(t-1), P(t)$	$Q(t)$
2	$P(t-3), P(t-2), P(t-1), P(t)$	$Q(t)$
3	$Q(t-3), Q(t-2), Q(t-1), P(t)$	$Q(t)$
4	$P(t-3), P(t-2), P(t-1)$	$Q(t)$
5	$Q(t-3), Q(t-2), Q(t-1)$	$Q(t)$
6	$P(t)$	$Q(t)$

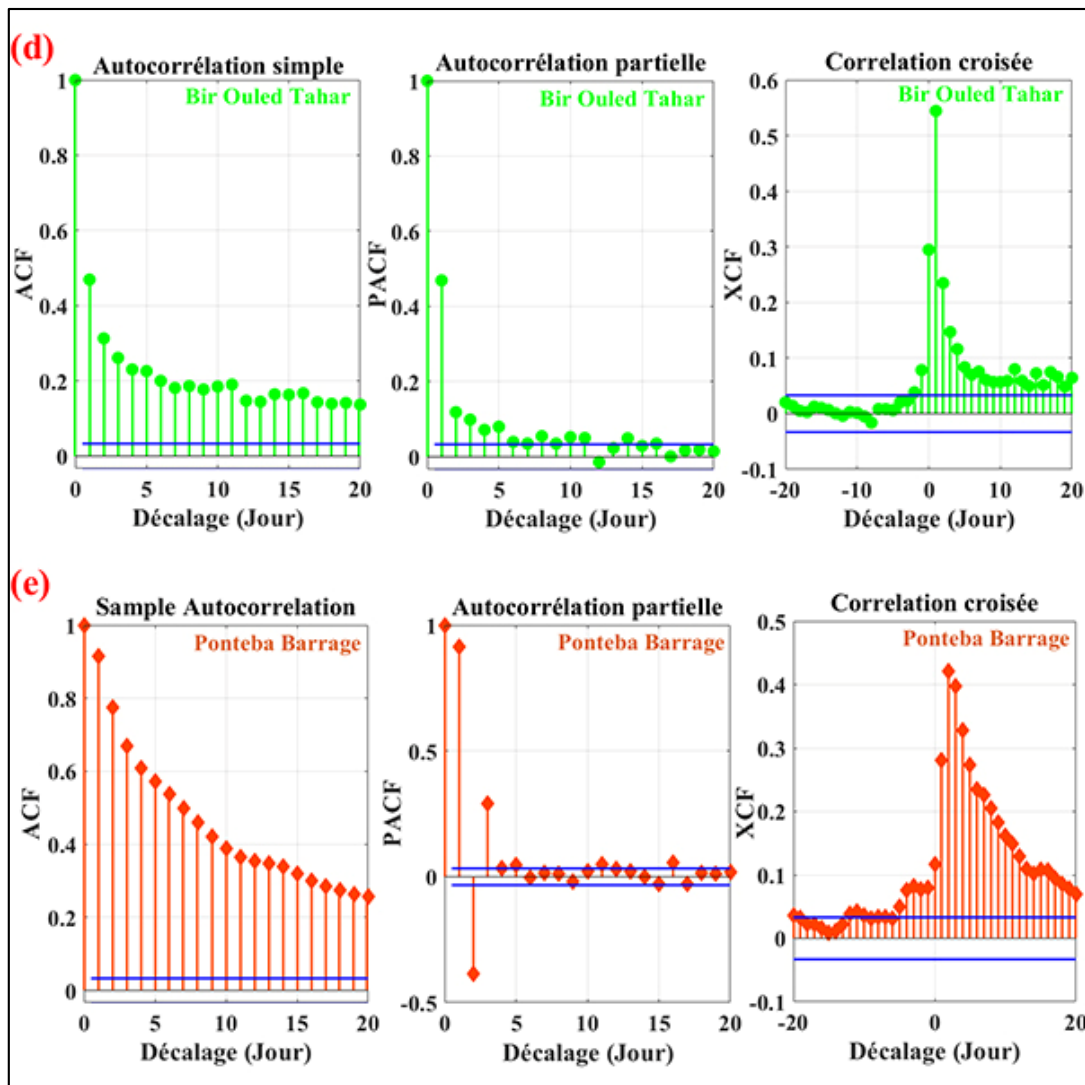


Figure V.12 Autocorrélation (ACF), fonction d'autocorrélation partielle (PACF) et fonction de corrélation croisée (XCF) des séries de débit mensuelles pour les stations (d) Bir Ouled Tahar et (e) Ponteba Barrage

Pour prédire le débit journalier des stations Bir Ouled Tahar et Ponteba Barrage en utilisant uniquement P et Q, étant donné que seules les données P sont disponibles en continu sans enregistrement complet. Conformément à la figure V.12, des décalages de deux temps ont été sélectionnés et utilisés comme variables d'entrée, c'est-à-dire le débit journalier mesuré à $(t-1)$, $(t-2)$, les précipitations journalières mesurées à $(t-1)$, $(t-2)$; tandis que la variable de sortie commence au temps (t) , c'est-à-dire Q au temps (t) . Par conséquent, quatre combinaisons d'entrées sont adoptées et utilisées pour la présente étude (tableau V.9).

Tableau V.9 Combinaisons des variables d'entrées de différents modèles pour la station de Bir Ouled Tahar et Ponteba Barrage

N Modèle	Combinaison des variables d'entrée	Sortie
1	$Q(t-2), Q(t-1), P(t-2), P(t-1), P(t)$	$Q(t)$
2	$P(t-2), P(t-1), P(t)$	$Q(t)$
3	$Q(t-2), Q(t-1), P(t)$	$Q(t)$
4	$P(t-2), P(t-1)$	$Q(t)$
5	$Q(t-2), Q(t-1)$	$Q(t)$
6	$P(t)$	$Q(t)$

Conclusion

Dans ce chapitre, nous avons examiné les données journalières et mensuelles des stations pluviométriques et hydrométriques pour les 5 sous-bassins versants étudiés. La qualité des résultats des prédictions peut être affectée par un nombre important de variables d'entrée. Par conséquent, le choix des bonnes données d'entrée est essentiel pour obtenir des performances robustes, ce qui constitue potentiellement une étape très importante dans la modélisation ou la prévision hydrologique. En utilisant les fonctions d'autocorrélation, d'autocorrélation partielle et de corrélation croisée (ACF, PACF et XCF), Nous identifions les variables d'entrée les plus pertinentes pour la prévision des débits journaliers et mensuels. La base de données sera utilisée dans les chapitres suivants, afin de développer les différents modèles proposés.

CHAPITRE VI Prévision des débits journaliers et mensuels par des techniques d'apprentissage automatique

Introduction

L'estimation précise et immédiate du débit des cours d'eau est essentielle pour améliorer notre compréhension de ses variations dans le temps et l'espace, ce qui est utile pour les études hydrologiques. Les mesures directes du débit des cours d'eau à l'aide de stations hydrométriques fournissent des données réelles et instantanées, mais ces stations ne sont pas disponibles en nombre suffisant et dans certains cas, les données sont incomplètes. Bien que la prévision du débit à différents pas de temps soit bien documentée, il est nécessaire d'améliorer la précision des prévisions avec de nouvelles méthodes, car le débit est souvent non linéaire et non stationnaire, ce qui a motivé le développement de nouveaux algorithmes de modélisation capables de surmonter les limites de certains modèles autonomes, en particulier en termes d'amélioration des capacités de généralisation et surtout d'estimation correcte des débits extrêmes.

La problématique de la prévision du débit des cours d'eau, à l'aide de l'apprentissage automatique basé sur les mesures du débit des cours d'eau en conjonction avec les variables météorologiques, a fait l'objet d'une vaste étude. Une approche courante pour estimer le débit consiste à le modéliser en fonction de ses valeurs antérieures mesurées avec différents décalages. Les méthodes ACF, PACF et XCF sont largement utilisées avec d'autres variables météorologiques telles que les précipitations, l'évapotranspiration et la température de l'air. Les algorithmes métaheuristiques peuvent être efficaces pour améliorer les performances des modèles d'apprentissage automatique dans la prédiction du débit des cours d'eau. Dans ce chapitre, nous présentons et discutons les résultats de la prévision des débits journaliers et mensuels pour trois sous-bassins (Ebda, Cheliff Harrezal et Tiguiguest) dans le grand bassin du Cheliff en Algérie, en utilisant six modèles d'intelligence artificielle (IA) : cinq modèles autonomes et un modèle hybride déjà défini dans le chapitre précédent (**Chapitre III**) sous les abréviations (MLPNN, GRNN, SVR, GPR, ELM, BAT-ELM).

VI.1 La mise en œuvre des modèles d'apprentissage automatique proposée dans l'étude

VI.1.1 Acquisition des données pour les modèles

La base de données utilisée pour le développement des modèles ML est constituée des valeurs journalières et mensuelles de pluie (P) et de débit (Q). Trois sous-bassins situés dans le grand bassin

du Cheliff, au nord de l'Algérie, sont proposés dans notre cas d'étude : Oued Ebda, Cheliff Harrezal et Tiguiguest (figures VI.1).

Le sous-bassin versant de l'oued de Tiguiguest est contrôlé par la station d'Oued Lili (code 012501) dans le bassin versant du Bas Chélif et de La Mina. L'Oued Harreza est contrôlé par la station d'El Ababsa (code 011715), ainsi que par la station d'Arib Abda (code 011801), qui contrôle l'Oued Ebda dans le bassin versant Haut et Moyen Cheliff. Les séries chronologiques ont été collectées à partir de la base de données disponible à l'ANRH en Algérie (Agence Nationale des Ressources Hydrauliques).

La base de données a été divisée en deux parties : une première partie dite d'apprentissage pour le développement et le calage des modèles, et une deuxième partie dite de validation. L'apprentissage est la propriété la plus intéressante des modèles d'apprentissage automatique. C'est la phase de développement durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. Une fois le modèle développé, il sera validé avec des données qui n'ont pas été utilisées lors de la phase d'apprentissage, à l'aide de données de validation. Dans notre étude, nous avons opté pour la répartition suivante : 70% des données pour l'apprentissage et 30% pour la validation.

La répartition des données journalières et mensuelles (débit et précipitations) utilisées dans cette partie est présentée dans le chapitre précédent (figures V.7 et V.9), de même que les paramètres statistiques de l'ensemble de données utilisé dans cette recherche à l'échelle journalière et mensuelle, qui sont résumés dans les tableaux V.5 et V.7.

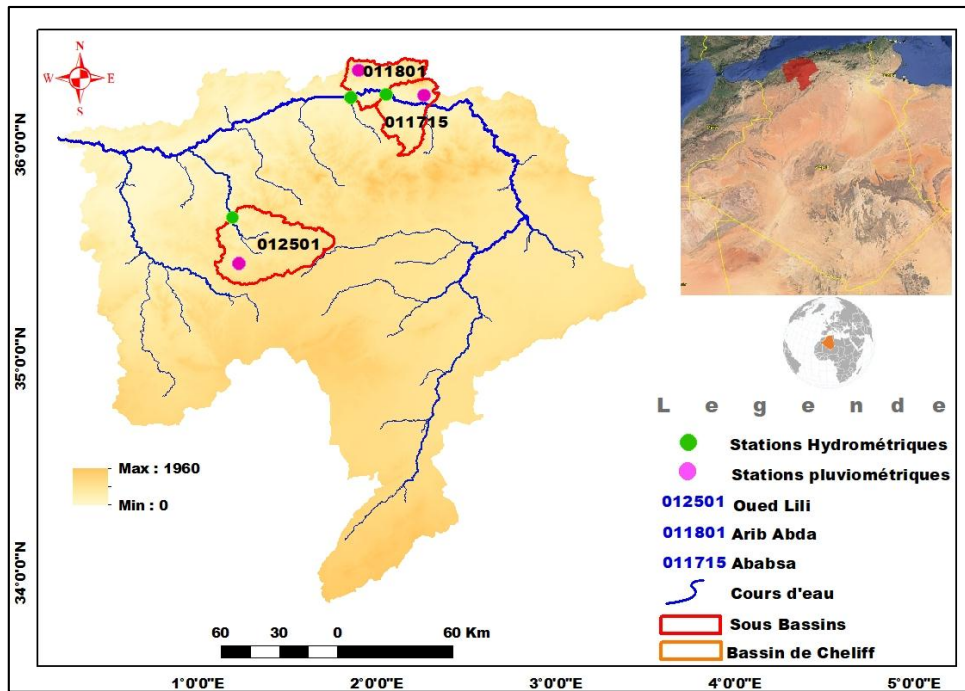


Figure VI.1 Localisation des stations pluviométriques et hydrométriques dans le bassin de Cheliff

VI.1.2 Sélection des variables d'entrée par les méthodes ACF, PAC et XCF

Le choix des variables d'entrée constitue une phase importante qui conditionne la performance des résultats de la modélisation par les méthodes de l'intelligence artificielle. Pour la sélection des variables d'entrée nous avons testé plusieurs modèles et plusieurs variantes en se basant sur les résultats obtenus à l'aide des techniques ACF, PACF et XCF, pour notre cas d'étude, nous avons élaboré six (6) combinaisons des variables climatiques et hydrologiques. Dans le chapitre précédent nous avons présenté les huit modèles (M1 jusqu'à M6). Le **tableau VI.1** c'est un rappel pour la sélection des entrées (inputs) pour le pas mensuel et journalier.

Tableau VI.1 Combinaisons des variables d'entrées de différents modèles pour la station d'Oued Lili, El Ababsa et Arib abda au pas de temps journalier et mensuelle

N Modèle	Combinaison des variables d'entré	Sortie
1	$Q(t-3), Q(t-2), Q(t-1), P(t-3), P(t-2), P(t-1), P(t)$	$Q(t)$
2	$P(t-3), P(t-2), P(t-1), P(t)$	$Q(t)$
3	$Q(t-3), Q(t-2), Q(t-1), P(t)$	$Q(t)$
4	$P(t-3), P(t-2), P(t-1)$	$Q(t)$
5	$Q(t-3), Q(t-2), Q(t-1)$	$Q(t)$
6	$P(t)$	$Q(t)$

VI.1.3 Critères de validation et de performances des modèles d'apprentissage automatique

L'évaluation de la performance des modèles de prévision implique une comparaison entre les valeurs prédites par les modèles et les observations réelles. Pour ce faire, différents indicateurs statistiques sont utilisés afin de quantifier la qualité des prédictions sur l'ensemble des données, à la fois pour les données d'apprentissage et de validation. Cette évaluation peut être qualitative, en examinant visuellement les données observées et prédites, ou quantitative, en utilisant des mesures numériques telles que l'erreur quadratique moyenne (RMSE), le coefficient de corrélation (R), l'erreur absolue moyenne (MAE) et le coefficient de Nash-Sutcliffe (NSE) ([chapitre I](#)). Outre ces mesures numériques, des représentations graphiques sont souvent utilisées, notamment des diagrammes de dispersion, des diagrammes en boîte et en violon, des diagrammes de Taylor, des diagrammes radar et des courbes de comparaison entre les valeurs observées et prédites.

VI.1.4 Développement des modèles d'apprentissage automatique

Pour le développement des modèles, le logiciel MATLAB 2020a a été employé comme environnement de programmation, offrant des fonctionnalités pour la mise en œuvre et l'exécution d'applications basées sur l'intelligence artificielle. Dans cette section de l'étude, nous avons évalué la performance de six modèles différents - MLPNN, GRNN, SVR, GPR, ELM et BAT-ELM - sur trois stations hydrologiques distinctes. Les modèles ont été entraînés en utilisant différentes combinaisons de variables d'entrée, comme indiqué dans le [tableau V.1](#). Chaque modèle nécessite l'optimisation de ses paramètres spécifiques tout au long du processus d'apprentissage.

VI.1.4.1 Développement des modèles ML autonome

Pour le modèle MLPNN, nous avons effectué une exploration du nombre de neurones cachés afin de déterminer la configuration optimale. Nos résultats ont montré que des performances optimales étaient atteintes avec un nombre de neurones cachés compris entre 10 et 13. En ce qui concerne le modèle GRNN, nous avons ajusté uniquement la valeur de l'écart, constatant que des valeurs dans la plage de 0,15 à 1,30 produisaient les meilleures performances selon les scénarios. Pour le modèle SVR, nous avons exploré une plage de valeurs pour le paramètre d'epsilon, variant de 0,05 à 0,7. Concernant le modèle GPR, plusieurs fonctions de noyau ont été testées, chacune nécessitant un

ensemble spécifique de paramètres. Enfin, pour le modèle ELM, nous avons ajusté le nombre de neurones cachés dans la plage de 10 à 20 pour obtenir les meilleures performances.

VI.1.4.2 Développement d'un modèle hybride BAT-ELM

Pour le BAT-ELM (**Figure VI.2**), nous avons plusieurs paramètres à optimiser : (i) la population de chauves-souris (N) égale à 25, (ii) l'intensité sonore (A_0) égale à 0,1, (iii) la fréquence d'impulsion (r_0) égale à 0.1, (iv) les fréquences minimale et maximale (f_{min} et f_{max}) avec des valeurs égales à (-1) et (+1) respectivement, (v) le nombre d'itérations égal à 500, (vi) le nombre de neurones cachés égal à 55, et (vii) pour la fonction d'activation, nous avons utilisé la sigmoïdale.

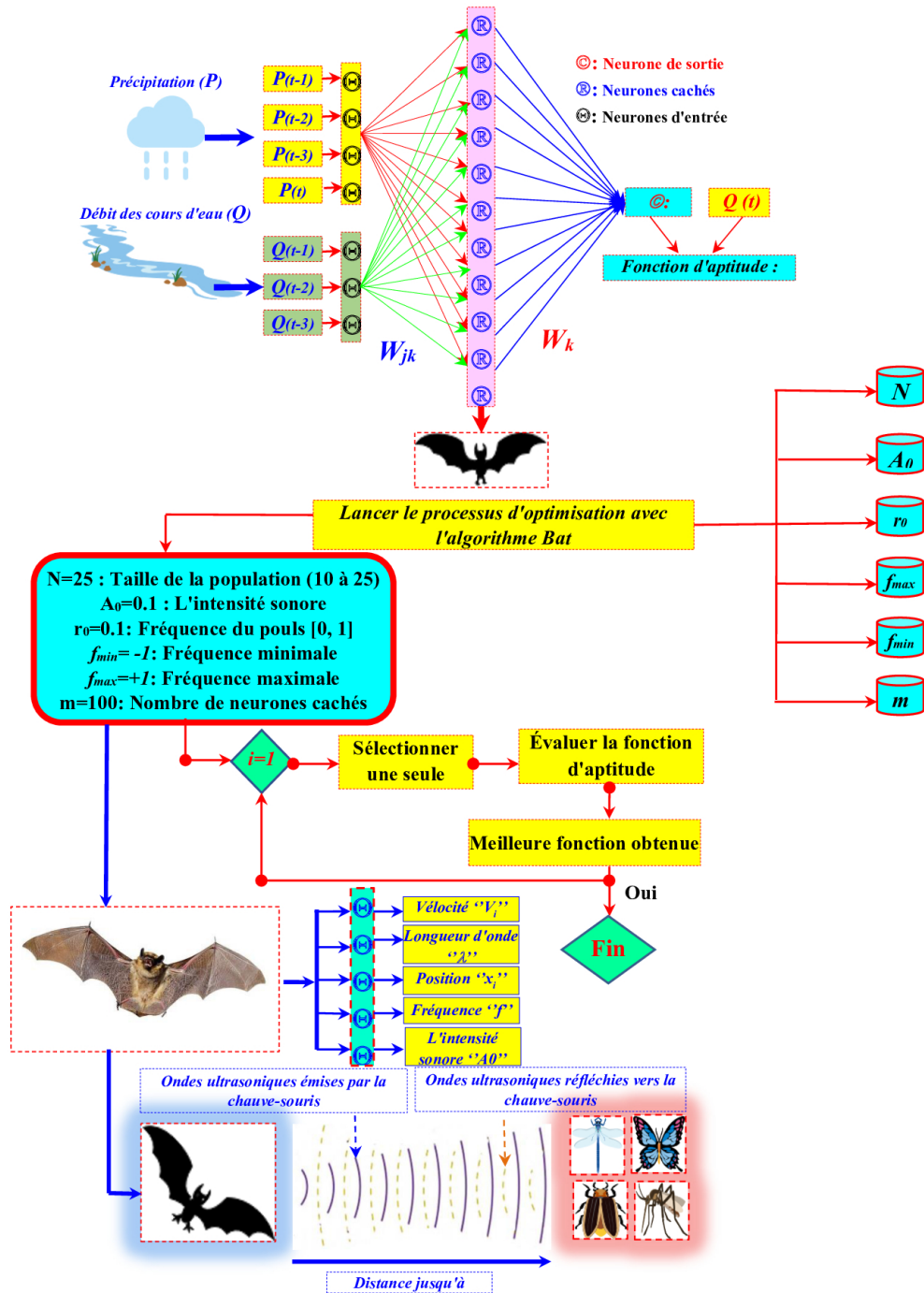


Figure VI.2 Organigramme des modèles hybrides BAT-ELM proposés

VI.2 Résultats de la modélisation et analyse comparative

Les indices R, NSE, RMSE et MAE sont utilisés pour la comparaison des modèles. Les résultats numériques sont présentés dans les tableaux VI.2 à VI.7, et les résultats significatifs de l'algorithme sont indiqués en gras.

VI.2.1 Résultats de la prédiction dans le pas journalier et évaluation des modèles

Dans la section suivante, nous présentons et discutons tous les résultats obtenus selon les six scénarios de modélisation pour la prédiction des débits journalière des cours d'eaux.

À la station d'Arib Abda, les performances de chaque scénario parmi les six proposés ont été évaluées pour déterminer leur efficacité. Les résultats de cette évaluation sont résumés dans la **Figure VI.3**, qui présente les valeurs moyennes des critères de performance pour chaque algorithme dans les différents scénarios lors de la phase de validation (**Tableau VI.2**).

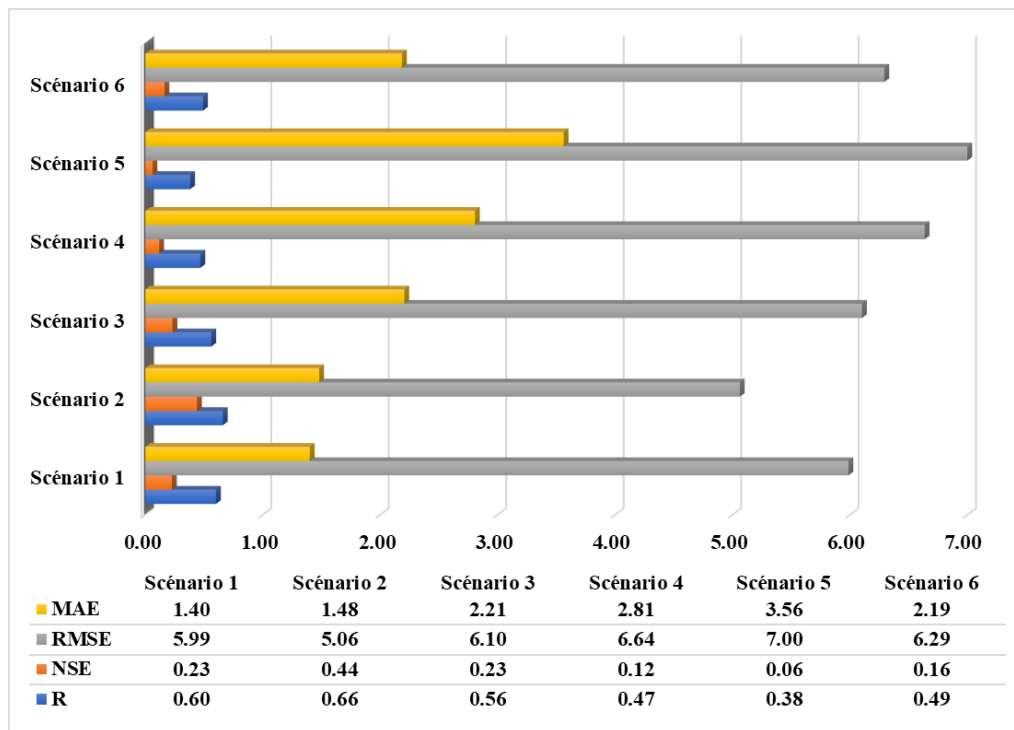


Figure VI.3 Histogramme de performance des algorithmes dans les six scénarios proposés pour la station d'Arib Abda au pas de temps journaliers (phase de validation)

Selon les données de la **figure VI.3**, le deuxième scénario présente des performances supérieures par rapport aux premier, troisième, quatrième, cinquième et sixième scénario, en termes de pourcentages de R et de NSE, avec des valeurs respectives d'environ 5,83 % et 21,33 %, 9,83 % et 20,83 %, 19,17 % et 32,17 %, 27,83 % et 37,83 %, et 16,83 % et 27,67 %.

Ces résultats indiquent que les deux premiers scénarios sont assez similaires et offrent les meilleures performances. Les troisième et sixième scénario présentent des résultats globalement similaires, mais sont moins efficaces que les deux premiers. Le quatrième scénario affiche des

résultats moins performants que les troisième et sixième scénario, tandis que le cinquième scénario est le moins performant parmi les six scénarios proposés.

Les résultats des simulations des algorithmes BAT-ELM, MLPNN, GRNN, SVR, GPR et ELM ont révélé que les meilleures précisions prédictives ont été obtenues avec les scénarios 1 et 2 (**tableau VI.2**). Le scénario 1 est "un ensemble de combinaisons de quatre variables d'entrée relatives aux précipitations (P) et de trois variables d'entrée relatives au débit (Q) comprenant (P(t), P (t - 1), P (t - 2), P (t - 3), Q (t - 1), Q (t - 2) et Q (t - 3)), dans lequel il a donné les meilleurs résultats avec les modèles BAT-ELM1 et SVR1. Le scénario 2 est "un ensemble de combinaisons de quatre variables d'entrée des précipitations (P) comprenant (P(t), P (t - 1), P (t - 2), P (t - 3)) dans lequel il a fourni les meilleurs résultats avec les modèles MLPNN2, GRNN2, GPR 2, ELM2.

Le modèle BAT-ELM1 a produit les valeurs R et NSE les plus élevées $\approx 0,84$ et $\approx 0,71$ et les valeurs RMSE et MAE les plus faibles $\approx 3,71 \text{ m}^3/\text{s}$ et $\approx 0,71 \text{ m}^3/\text{s}$, respectivement. Il convient de souligner que les meilleurs résultats obtenus par l'algorithme SVR utilisant le 1er scénario par rapport aux modèles proposés dans les différents scénarios montrent que le SVR1 est le moins efficace numériquement (R=0.41, NSE=0.15, RMSE=6.37 m^3/s et MAE=1.80 m^3/s) dans la prédiction des débits journaliers.

On peut également noter que le modèle BAT-ELM dans le deuxième scénario et troisième scénario a montré des valeurs de R et NSE les plus élevées et les valeurs RMSE et MAE les plus faibles, les modèles BAT-ELM2 (R ≈ 0.81 , NSE ≈ 0.65 , RMSE $\approx 4.11 \text{ m}^3/\text{s}$ et MAE $\approx 1.29 \text{ m}^3/\text{s}$) et le modèle BAT-ELM3 (R ≈ 0.77 , NSE ≈ 0.58 , RMSE $\approx 4.47 \text{ m}^3/\text{s}$ et MAE $\approx 1.56 \text{ m}^3/\text{s}$) ont surpassé les modèles MLPNN2, GRNN2, SVR1, GPR2 et ELM2. Le modèle BAT-ELM1 a considérablement amélioré la précision des prédictions par rapport aux autres algorithmes appliqués, les RMSE et MAE de MLPNN2, GRNN2, SVR1, GPR2 et ELM2 ayant été améliorés d'environ $\approx 30.39\%$ et $\approx 51.03\%$, $\approx 25.35\%$ et $\approx 47.01\%$, $\approx 41.76\%$ et $\approx 60.56\%$, $\approx 30.65\%$ et $\approx 59.20\%$, et $\approx 19.17\%$, et $\approx 56.17\%$, respectivement.

Ce modèle BAT-ELM1 a amélioré la précision de la prédiction par rapport à l'ensemble des algorithmes utilisés, avec un pourcentage moyen de RMSE $\approx 29.47\%$ et de MAE $\approx 54.79\%$. Ces observations conduisent à la conclusion que la fiabilité des prédictions de débits journaliers, basées sur le modèle BAT-ELM1, a été améliorée dans des proportions acceptables.

Tableau VI.2 Résultats de prévision des différents modèles au pas de temps journalier pour la station d'Arrib Abda

Modèles	Apprentissage				Validation			
	R	NSE	RMSE, m ³ /s	MAE, m ³ /s	R	NSE	RMSE, m ³ /s	MAE, m ³ /s
<i>Résultats obtenus à l'aide des modèles MLPNN</i>								
MLPNN1	0.67	0.44	4.05	1.69	0.59	0.16	6.33	1.22
MLPNN2	0.92	0.85	2.14	0.54	0.64	0.41	5.33	1.45
MLPNN3	0.89	0.78	2.53	0.56	0.52	0.23	6.66	1.66
MLPNN4	0.69	0.48	3.92	1.66	0.50	0.24	6.75	1.98
MLPNN5	0.86	0.73	2.83	0.57	0.40	0.10	6.77	2.11
MLPNN6	0.87	0.47	3.05	0.68	0.51	0.20	6.68	2.10
<i>Résultats obtenus à l'aide des modèles GRNN</i>								
GRNN1	0.89	0.78	2.53	1.29	0.58	0.05	6.75	1.69
GRNN2	0.98	0.96	1.04	0.54	0.70	0.48	4.97	1.34
GRNN3	0.89	0.78	2.53	0.87	0.59	0.34	5.62	1.62
GRNN4	0.79	0.61	3.40	1.49	0.48	0.21	6.27	1.95
GRNN5	0.83	0.68	3.06	0.90	0.42	0.14	6.57	2.05
GRNN6	0.87	0.77	2.60	0.88	0.57	0.31	5.73	1.63
<i>Résultats obtenus à l'aide des modèles SVR</i>								
SVR1	0.54	0.08	5.21	2.07	0.41	0.15	6.37	1.80
SVR2	0.54	0.08	5.21	2.05	0.40	0.14	6.42	2.02
SVR3	0.46	0.07	5.24	2.06	0.39	0.13	6.43	2.20
SVR4	0.35	0.04	5.31	2.13	0.31	0.01	6.87	2.33
SVR5	0.33	0.04	5.32	2.14	0.27	0.00	7.8	6.29
SVR6	0.46	0.07	5.24	2.06	0.40	0.14	6.42	2.23
<i>Résultats obtenus à l'aide des modèles GPR</i>								
GPR1	0.94	0.98	1.87	1.03	0.62	0.09	6.59	1.36
GPR2	0.99	0.99	0.04	0.02	0.66	0.40	5.35	1.74
GPR3	0.99	0.99	0.09	0.03	0.53	0.07	6.67	4.53
GPR4	0.95	0.91	1.64	0.85	0.47	0.03	6.88	4.56
GPR5	0.99	0.98	0.81	0.18	0.37	0.01	7.6	5.28
GPR6	0.52	0.27	4.64	2.04	0.50	0.04	6.79	4.54
<i>Résultats obtenus à l'aide des modèles ELM</i>								
ELM1	0.71	0.51	3.82	1.87	0.57	0.20	6.17	1.62
ELM2	0.90	0.80	2.40	0.85	0.75	0.56	4.59	1.62
ELM3	0.90	0.81	2.38	0.67	0.57	0.04	6.76	1.66
ELM4	0.72	0.52	3.78	1.73	0.55	0.02	6.85	1.59
ELM5	0.66	0.43	4.09	1.18	0.41	0.12	6.47	1.55
ELM6	0.90	0.81	2.38	0.67	0.55	0.02	6.83	1.68
<i>Résultats obtenus à l'aide des modèles BAT-ELM</i>								
BAT-ELM1	0.97	0.94	1.33	0.49	0.84	0.71	3.71	0.71
BAT-ELM2	0.92	0.86	2.07	1.17	0.81	0.65	4.11	1.29
BAT-ELM3	0.92	0.86	2.07	0.50	0.77	0.58	4.47	1.56
BAT-ELM4	0.74	0.55	3.64	1.68	0.50	0.20	6.19	1.88
BAT-ELM5	0.88	0.77	2.61	0.68	0.42	0.13	6.45	1.54
BAT-ELM6	0.92	0.86	2.07	0.50	0.67	0.41	5.30	1.07

Dans le cas de la station d'El Ababsa (Tableau VI.3, Figure. VI.4) on constate, que les meilleurs résultats ont été obtenus en utilisant les algorithmes présentés dans le premier scénario en comparaison avec les autres scénarios proposés.

Les résultats du premier scénario diffèrent significativement de ceux des deuxième, troisième, quatrième, cinquième et sixième scénario en termes de critères de performance R et NSE, avec des pourcentages de $\approx 4.17\%$ et $\approx 6.50\%$, $\approx 9.67\%$ et $\approx 13.67\%$, $\approx 20\%$ et $\approx 26.17\%$, $\approx 40.33\%$ et $\approx 43.33\%$, et $\approx 12.33\%$ et $\approx 22\%$, respectivement. Toutefois, on peut noter que le premier scénario est le plus favorable tandis que le cinquième est le plus défavorable.

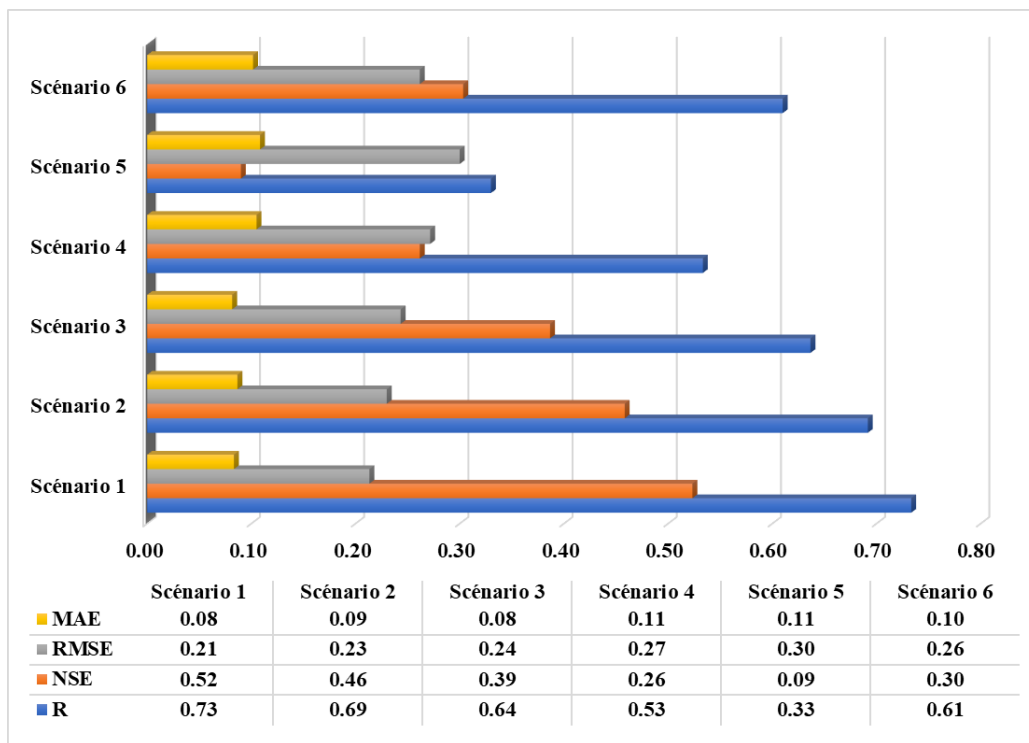


Figure VI.4 Histogramme de performance des algorithmes dans les six scénarios proposés pour la station d'El Ababsa au pas de temps journaliers (phase de validation)

Comme c'est le cas pour la station d'Arib Abda, le modèle BAT-ELM1 a également montré les meilleurs résultats, avec les valeurs R et NSE les plus élevées et les valeurs RMSE et MAE les plus faibles ou $R \approx 0.91$, $NSE \approx 0.83$, $RMSE \approx 0.13\text{m}^3/\text{s}$ et $MAE \approx 0.05\text{m}^3/\text{s}$, le modèle SVR1 a obtenu les résultats les moins significatifs avec les valeurs R et NSE les plus faibles et les valeurs RMSE et NSE les plus élevées ou $R \approx 0.51$, $NSE \approx 0.24$, $RMSE \approx 0.28\text{m}^3/\text{s}$ et $MAE \approx 0.11\text{m}^3/\text{s}$.

Le modèle BAT-ELM1 a considérablement amélioré la précision des prédictions par rapport aux autres algorithmes appliqués, les RMSE et MAE de MLPNN1, GRNN1, SVR1, GPR1 et ELM1 ayant été améliorés d'environ $\approx 38.10\%$ et $\approx 37.50\%$, $\approx 48.00\%$ et $\approx 44.44\%$, $\approx 53.57\%$ et $\approx 54.55\%$, $\approx 38.10\%$ et $\approx 28.57\%$, et $\approx 35.00\%$, et $\approx 50.00\%$, respectivement.

Ce modèle BAT-ELM1 a amélioré la précision de la prédiction par rapport aux algorithmes utilisés dans leur ensemble, avec un pourcentage moyen de RMSE $\approx 42.55\%$ et de MAE $\approx 43.01\%$.

Tableau VI.3 Résultats de prévision des différents modèles au pas de temps journalier pour la station d'El Ababsa

Modèles	Apprentissage				Validation			
	R	NSE	RMSE, m ³ /s	MAE, m ³ /s	R	NSE	RMSE, m ³ /s	MAE, m ³ /s
<i>Résultats obtenus à l'aide des modèles MLPNN</i>								
MLPNN1	0.80	0.63	0.26	0.06	0.77	0.55	0.21	0.08
MLPNN2	0.74	0.54	0.29	0.09	0.71	0.50	0.22	0.09
MLPNN3	0.69	0.48	0.31	0.07	0.68	0.46	0.23	0.07
MLPNN4	0.58	0.32	0.35	0.10	0.58	0.33	0.26	0.10
MLPNN5	0.34	0.11	0.41	0.10	0.27	0.05	0.31	0.10
MLPNN6	0.69	0.48	0.31	0.10	0.65	0.24	0.28	0.11
<i>Résultats obtenus à l'aide des modèles GRNN</i>								
GRNN1	0.98	0.95	0.09	0.04	0.70	0.39	0.25	0.09
GRNN2	0.97	0.92	0.12	0.05	0.70	0.38	0.25	0.09
GRNN3	0.76	0.56	0.28	0.08	0.67	0.31	0.26	0.10
GRNN4	0.72	0.49	0.31	0.07	0.54	0.29	0.27	0.10
GRNN5	0.31	0.08	0.41	0.10	0.30	0.09	0.30	0.11
GRNN6	0.68	0.44	0.32	0.10	0.63	0.28	0.27	0.11
<i>Résultats obtenus à l'aide des modèles SVR</i>								
SVR1	0.60	0.34	0.35	0.08	0.51	0.24	0.28	0.11
SVR2	0.56	0.29	0.36	0.08	0.49	0.22	0.28	0.13
SVR3	0.49	0.23	0.38	0.08	0.47	0.20	0.28	0.14
SVR4	0.39	0.15	0.40	0.08	0.38	0.12	0.30	0.16
SVR5	0.39	0.15	0.40	0.08	0.22	0.01	0.31	0.18
SVR6	0.48	0.22	0.38	0.08	0.46	0.19	0.28	0.15
<i>Résultats obtenus à l'aide des modèles GPR</i>								
GPR1	1.00	1.00	0.00	0.00	0.74	0.54	0.21	0.07
GPR2	0.99	0.99	0.05	0.03	0.71	0.50	0.22	0.09
GPR3	0.99	0.98	0.06	0.03	0.57	0.32	0.26	0.09
GPR4	0.87	0.76	0.21	0.07	0.43	0.12	0.30	0.15
GPR5	0.72	0.48	0.31	0.06	0.54	0.20	0.28	0.16
GPR6	0.99	0.99	0.05	0.01	0.56	0.18	0.29	0.17
<i>Résultats obtenus à l'aide des modèles ELM</i>								
ELM1	0.81	0.64	0.26	0.09	0.77	0.59	0.20	0.10
ELM2	0.79	0.63	0.26	0.09	0.77	0.56	0.21	0.12
ELM3	0.68	0.46	0.32	0.09	0.60	0.34	0.26	0.11
ELM4	0.60	0.36	0.34	0.10	0.56	0.22	0.28	0.12
ELM5	0.30	0.09	0.41	0.09	0.30	0.08	0.30	0.11
ELM6	0.65	0.43	0.33	0.09	0.57	0.30	0.26	0.11
<i>Résultats obtenus à l'aide des modèles BAT-ELM</i>								
BAT-ELM1	0.98	0.97	0.08	0.04	0.91	0.83	0.13	0.05
BAT-ELM2	0.92	0.84	0.17	0.05	0.77	0.59	0.20	0.09
BAT-ELM3	0.88	0.77	0.21	0.05	0.83	0.69	0.17	0.08
BAT-ELM4	0.81	0.65	0.25	0.09	0.71	0.49	0.22	0.11
BAT-ELM5	0.42	0.18	0.39	0.09	0.35	0.11	0.30	0.11
BAT-ELM6	0.87	0.75	0.21	0.05	0.79	0.63	0.19	0.07

Concernant la station d'Oued Lili, et d'après le tableau VI.4 et la figure VI.5, La performance du premier scénario diffère significativement des deuxième, troisième, quatrième, cinquième et sixième scénario en termes de critères de performance R et NSE, avec des pourcentages de $\approx 2.50\%$ et $\approx 3.17\%$, $\approx 4.83\%$ et $\approx 05\%$, $\approx 15.17\%$ et $\approx 14\%$, $\approx 17\%$ et $\approx 15\%$, et $\approx 10.67\%$ et 6.50% , respectivement. D'après ces résultats, on peut constater que le premier scénario est le plus approprié, tandis que le cinquième scénario est considéré comme le plus faible parmi les six scénarios proposés.

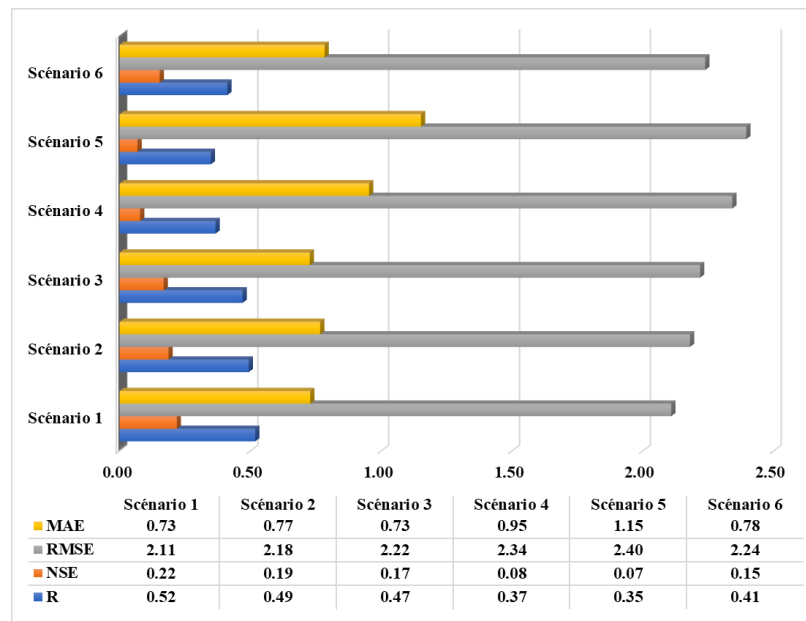


Figure VI.5 Histogramme de performance des algorithmes dans les six scénarios proposés pour la station d'Oued Lili au pas de temps journaliers (phase de validation)

Les résultats de la simulation pour les algorithmes BAT-ELM, MLPNN, GRNN, SVR, GPR et ELM ont révélé que la meilleure précision prédictive a été obtenue en utilisant les scénarios 1 et 3 (tableau VI.4), le scénario 1 donnant les meilleurs résultats avec les modèles BAT-ELM1, ELM1, GPR1 et SVR1. Le scénario 3 est "un ensemble de combinaisons de trois variables d'entrée des débits (Q) et une seule variable de précipitation (P) comprenant (P(t), Q (t - 1), Q (t - 2), Q (t - 3)) qui a fourni les meilleurs résultats avec les modèles MLPNN3 et GRNN3. Les performances des modèles BAT-ELM en termes de RMSE et de MAE sont supérieures à celles des autres modèles, affichant des taux croissants d'ordre de $\approx 52.63\%$ et $\approx 33.96\%$ par rapport au MLPNN3 ; $\approx 56.58\%$ et $\approx 46.15\%$ par rapport au GRNN3 ; $\approx 65.98\%$, et $\approx 72.22\%$ par rapport au SVR1 ; $\approx 55.00\%$ et $\approx 54.55\%$ par rapport au GPR1 ; et enfin $\approx 52.17\%$ et $\approx 46.15\%$ par rapport à l'ELM1, respectivement.

Tableau VI.4 Résultats de prévision des différents modèles au pas de temps journalier pour la station d'Oued Lili

Modèles	Apprentissage				Validation			
	R	NSE	RMSE, m ³ /s	MAE, m ³ /s	R	NSE	RMSE, m ³ /s	MAE, m ³ /s
<i>Résultats obtenus à l'aide des modèles MLPNN</i>								
MLPNN1	0.67	0.44	0.94	0.36	0.38	0.13	2.17	0.67
MLPNN2	0.66	0.43	0.94	0.24	0.39	0.14	2.18	0.67
MLPNN3	0.68	0.44	0.93	0.40	0.47	0.19	2.09	0.53
MLPNN4	0.49	0.24	1.09	0.32	0.37	0.13	2.17	0.73
MLPNN5	0.41	0.17	1.14	0.45	0.31	0.10	2.21	0.75
MLPNN6	0.63	0.39	0.97	0.40	0.36	0.09	2.22	0.73
<i>Résultats obtenus à l'aide des modèles GRNN</i>								
GRNN1	0.61	0.36	1.00	0.32	0.42	0.02	2.31	0.67
GRNN2	0.84	0.69	0.70	0.23	0.42	0.01	2.36	0.66
GRNN3	0.85	0.71	0.67	0.26	0.48	0.04	2.28	0.65
GRNN4	0.75	0.54	0.85	0.33	0.41	0.00	2.32	0.78
GRNN5	0.75	0.53	0.86	0.33	0.41	0.00	2.33	0.86
GRNN6	0.61	0.36	1.00	0.32	0.41	0.17	2.12	0.69
<i>Résultats obtenus à l'aide des modèles SVR</i>								
SVR1	0.64	0.31	1.04	1.19	0.46	0.02	2.91	1.26
SVR2	0.62	0.29	1.05	1.29	0.42	0.02	2.93	1.27
SVR3	0.60	0.28	1.06	1.38	0.38	0.02	2.94	1.36
SVR4	0.57	0.25	1.08	2.30	0.36	0.01	2.98	2.27
SVR5	0.43	0.11	1.18	2.35	0.35	0.01	3.21	3.27
SVR6	0.58	0.25	1.08	1.40	0.35	0.02	2.96	1.38
<i>Résultats obtenus à l'aide des modèles GPR</i>								
GPR1	0.99	0.98	0.19	0.12	0.47	0.11	2.20	0.77
GPR2	0.99	0.97	0.20	0.13	0.43	0.05	2.27	0.78
GPR3	0.99	0.98	0.20	0.03	0.38	0.12	2.18	0.58
GPR4	0.99	0.97	0.23	0.02	0.27	0.02	2.31	0.75
GPR5	0.97	0.94	0.30	0.02	0.27	0.01	2.31	0.76
GPR6	0.77	0.59	0.81	0.34	0.34	0.08	2.23	0.75
<i>Résultats obtenus à l'aide des modèles ELM</i>								
ELM1	0.68	0.47	0.91	0.37	0.47	0.21	2.07	0.65
ELM2	0.68	0.46	0.92	0.34	0.45	0.17	2.12	0.67
ELM3	0.68	0.46	0.92	0.34	0.41	0.16	2.14	0.68
ELM4	0.36	0.13	1.17	0.35	0.29	0.08	2.23	0.69
ELM5	0.35	0.12	1.17	0.35	0.28	0.08	2.24	0.79
ELM6	0.68	0.46	0.92	0.34	0.32	0.10	2.20	0.69
<i>Résultats obtenus à l'aide des modèles BAT-ELM</i>								
BAT-ELM1	0.91	0.83	0.52	0.23	0.91	0.82	0.99	0.35
BAT-ELM2	0.87	0.76	0.62	0.29	0.85	0.73	1.22	0.55
BAT-ELM3	0.78	0.61	0.78	0.36	0.70	0.48	1.68	0.52
BAT-ELM4	0.45	0.21	1.12	0.45	0.50	0.23	2.04	0.50
BAT-ELM5	0.60	0.37	1.00	0.32	0.47	0.21	2.07	0.48
BAT-ELM6	0.70	0.49	0.89	0.25	0.69	0.46	1.70	0.46

Les nuages de points des valeurs de débit journalier prévues et observées pour les trois stations d'étude au cours de la période de test fournies par les six modèles sont présentés dans la [figure VI.6](#), [figure VI.7](#) et [figure VI.8](#). Comme cela est clairement observé, la tendance linéaire de BAT-ELM est la plus proche de la droite $y=x$ par rapport à celles de MLPNN, GRNN, SVR, GPR et ELM. Ceci peut également être compris dans la station d'El Ababsa à partir des équations de la droite ajustée dans lesquelles les coefficients de pente (0,9913) et de biais (0,0134) de BAT-ELM, respectivement, sont beaucoup plus proches de 1 et de 0 que ceux des autres modèles. Par rapport aux résultats obtenus au niveau des stations d'Arib Abda et Oued Lili, il apparaît clairement que les taux d'amélioration sont plus importants. De même, les séries chronologiques du débit prédit par BAT-ELM sont comparées à celles observées durant la période de validation. On constate une bonne adéquation et un accord satisfaisant entre le débit observé et le débit prédit par le modèle BAT-ELM.

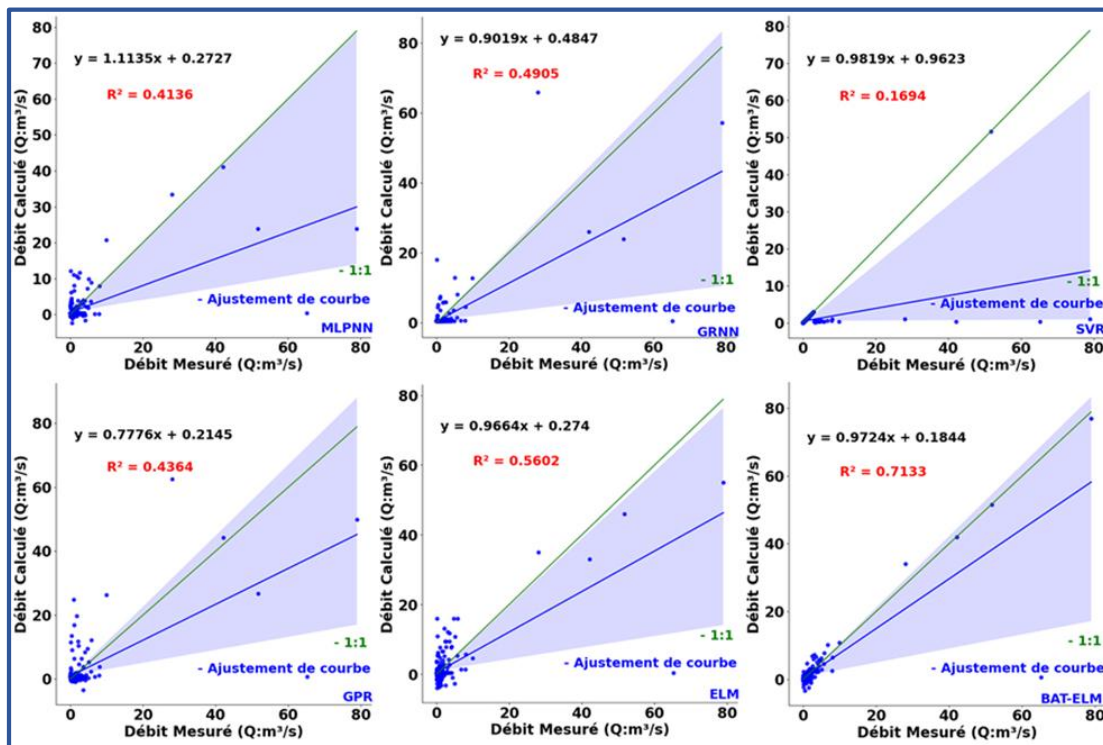


Figure VI.6 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Arib Abda pendant la phase de test (pas de temps journalier)

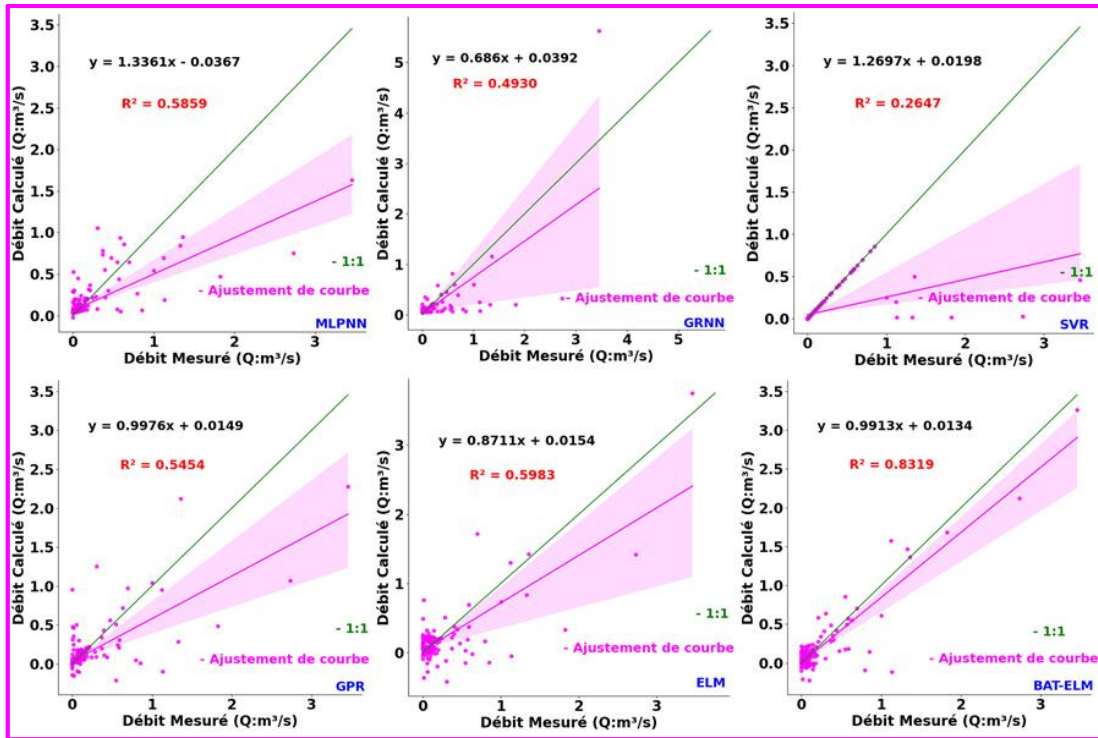


Figure VI.7 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'El Ababsa pendant la phase de test (pas de temps journalier)

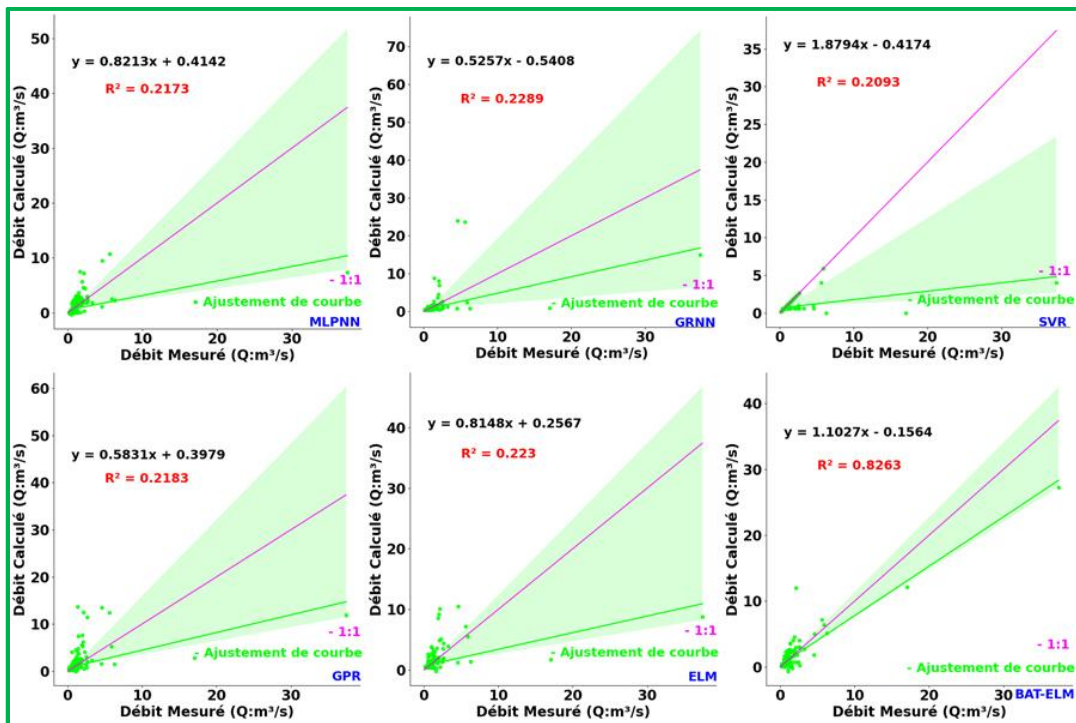


Figure VI.8 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Oued Lili pendant la phase de test (pas de temps journalier)

VI.2.2 Résultats de la prévision à l'échelle mensuelle et évaluation des modèles

Pour les trois stations, la meilleure précision de prévision mensuelles a été obtenue en utilisant le BAT-ELM1 avec sept variables d'entrée, c'est-à-dire $P(t)$, $P(t-1)$, $P(t-2)$, $P(t-3)$, $Q(t-1)$, $Q(t-2)$ et $Q(t-3)$, où les valeurs les plus élevées de R et NSE et les valeurs RMSE et MAE les plus faibles ont été obtenues.

Dans la station d'Arib Abda ([tableau VI.5](#) et [figure VI.9](#)), le BAT-ELM1 a donné les valeurs R et NSE les plus élevées ≈ 0.927 et ≈ 0.846 et les valeurs RMSE et MAE les plus faibles $\approx 0.749 \text{ m}^3/\text{s}$ et $\approx 0.472 \text{ m}^3/\text{s}$, respectivement. La performance la plus faible a été obtenue par le modèle MLPNN dans la prédiction des débits mensuels (R ≈ 0.838 , NSE ≈ 0.649 , RMSE $\approx 1.132 \text{ m}^3/\text{s}$ et MAE $\approx 0.67 \text{ m}^3/\text{s}$). Cependant, il apparaît que si BAT-ELM1 était le meilleur modèle, ce n'est pas le cas des autres algorithmes pour lesquels la meilleure précision a été obtenue pour différentes combinaisons d'entrées, à savoir MLPNN2, GRNN1, SVR2, GPR1 et ELM2. Compte tenu des indicateurs numériques, il est clair que le BAT-ELM1, par rapport aux MLPNN2, GRNN1, SVR2, GPR1 et ELM2, a réduit la RMSE et la MAE de $\approx 33.83\%$ et $\approx 30.17\%$, $\approx 25.02\%$ et $\approx 2.68\%$, $\approx 20.48\%$ et $\approx 33.70\%$, $\approx 31.22\%$ et $\approx 23.77\%$, $\approx 24.95\%$, et $\approx 34.71\%$, respectivement. Cette constatation permet de conclure que la précision de prévision, basée sur les BAT-ELM1, est significativement améliorée. D'après la [figure VI.9](#), la méthode hybride BAT-ELM permet de reproduire de manière adéquate les débits de pointe avec des données moins dispersées.

Tableau VI.5 Résultats de prédiction des différents modèles au pas de temps mensuelle pour la station d'Arib Abda

Modèles	Apprentissage				Validation			
	R	NSE	RMSE, m ³ /s	MAE, m ³ /s	R	NSE	RMSE, m ³ /s	MAE, m ³ /s
Résultats obtenus à l'aide des modèles MLPNN								
MLPNN1	0.888	0.786	1.340	0.672	0.815	0.648	1.133	0.719
MLPNN2	0.921	0.846	1.135	0.607	0.838	0.649	1.132	0.676
MLPNN3	0.882	0.770	1.387	0.670	0.796	0.630	1.161	0.661
MLPNN4	0.727	0.524	1.996	1.090	0.611	0.358	1.530	0.877
MLPNN5	0.579	0.334	2.363	1.280	0.438	0.187	1.722	0.950
MLPNN6	0.852	0.726	1.517	0.886	0.599	0.325	1.569	0.787
Résultats obtenus à l'aide des modèles GRNN								
GRNN1	0.891	0.782	1.351	0.494	0.891	0.726	0.999	0.485
GRNN2	0.879	0.724	1.521	0.672	0.873	0.714	1.022	0.508
GRNN3	0.871	0.649	1.716	0.848	0.855	0.548	1.284	0.594
GRNN4	0.632	0.314	2.398	1.263	0.687	0.354	1.535	0.822
GRNN5	0.521	0.220	2.557	1.405	0.444	0.187	1.722	0.910
GRNN6	0.753	0.416	2.212	1.213	0.644	0.362	1.525	0.781
Résultats obtenus à l'aide des modèles SVR								
SVR1	0.891	0.762	1.412	0.991	0.884	0.740	0.973	0.699
SVR2	0.903	0.809	1.267	0.867	0.900	0.756	0.942	0.712
SVR3	0.751	0.488	2.072	0.755	0.607	0.308	1.588	0.725
SVR4	0.688	0.460	2.127	1.296	0.809	0.652	1.127	0.666
SVR5	0.616	0.237	2.529	1.003	0.324	0.047	1.864	0.799
SVR6	0.771	0.429	2.187	0.965	0.518	0.137	1.774	0.822
Résultats obtenus à l'aide des modèles GPR								
GPR1	0.930	0.859	1.087	0.497	0.822	0.675	1.089	0.616
GPR2	0.943	0.885	0.981	0.477	0.789	0.600	1.207	0.605
GPR3	0.999	0.999	0.019	0.008	0.701	0.452	1.414	0.715
GPR4	0.514	0.264	2.484	1.286	0.514	0.263	1.640	0.889
GPR5	0.727	0.493	2.062	1.111	0.361	0.122	1.789	0.915
GPR6	0.810	0.655	1.700	0.974	0.593	0.334	1.558	0.815
Résultats obtenus à l'aide des modèles ELM								
ELM1	0.828	0.686	1.622	1.068	0.869	0.752	0.951	0.725
ELM2	0.903	0.815	1.245	0.765	0.877	0.727	0.998	0.723
ELM3	0.874	0.763	1.408	0.662	0.848	0.717	1.016	0.596
ELM4	0.715	0.512	2.023	1.272	0.762	0.545	1.288	0.843
ELM5	0.558	0.312	2.402	1.324	0.492	0.239	1.666	0.864
ELM6	0.755	0.569	1.900	1.091	0.659	0.400	1.479	0.818
Résultats obtenus à l'aide des modèles BAT-ELM								
BAT-ELM1	0.938	0.880	1.002	0.528	0.927	0.846	0.749	0.472
BAT-ELM2	0.964	0.930	0.765	0.570	0.918	0.818	0.815	0.648
BAT-ELM3	0.937	0.878	1.013	0.443	0.881	0.741	0.972	0.599
BAT-ELM4	0.882	0.779	1.362	0.669	0.773	0.532	1.306	0.862
BAT-ELM5	0.788	0.621	1.783	1.067	0.569	0.292	1.607	1.054
BAT-ELM6	0.813	0.661	1.687	0.893	0.680	0.445	1.422	0.704

Pour la station d'Ababsa (tableau VI.6 et figure VI.10), ils évident que BAT-ELM1 fournit les meilleurs indicateurs statistiques avec les valeurs R et NSE les plus élevées (0,944 et 0,883) et les valeurs RMSE et MAE les plus faibles (0,036 m³/s et 0,026 m³/s), respectivement. Plus précisément, les meilleures performances ont été obtenues avec le scénario 01. Le modèle BAT-ELM1 a considérablement amélioré la précision des prédictions par rapport aux autres algorithmes appliqués, les RMSE et MAE de MLPNN1, GRNN1, SVR1, GPR1 et ELM1 ayant été améliorés d'environ $\approx 46.27\%$ et $\approx 36.58\%$, $\approx 55.55\%$ et $\approx 38.09\%$, $\approx 56.62\%$ et $\approx 27.77\%$, $\approx 47.82\%$ et $\approx 33.33\%$, et $\approx 53.84\%$, et $\approx 48.00\%$, respectivement.

La **figure VI.10** montre que les prédictions garanties par BAT-ELM sont plus remarquables que celles des autres algorithmes. Ceci explique la valeur la plus élevée du coefficient R (0,9522) et le biais le plus faible (0,0047) pour BAT-ELM qui sont respectivement plus proches de 1 et 0 que ceux des autres modèles. Par rapport aux résultats obtenus à l'Arib Abda, il est clair que les taux d'amélioration sont plus importants.

Tableau VI.6 Résultats de prédiction des différents modèles au pas de temps mensuelle pour la station El Ababsa

Modèles	Apprentissage				Validation			
	R	NSE	RMSE, m ³ /s	MAE, m ³ /s	R	NSE	RMSE, m ³ /s	MAE, m ³ /s
Résultats obtenus à l'aide des modèles MLPNN								
MLPNN1	0.885	0.779	0.125	0.068	0.783	0.606	0.067	0.041
MLPNN2	0.709	0.503	0.188	0.095	0.585	0.329	0.087	0.051
MLPNN3	0.780	0.607	0.167	0.078	0.719	0.483	0.077	0.042
MLPNN4	0.421	0.175	0.242	0.135	0.341	0.107	0.101	0.063
MLPNN5	0.593	0.341	0.216	0.134	0.320	0.055	0.103	0.066
MLPNN6	0.647	0.410	0.204	0.099	0.614	0.371	0.084	0.045
Résultats obtenus à l'aide des modèles GRNN								
GRNN1	0.911	0.794	0.121	0.066	0.791	0.415	0.081	0.042
GRNN2	0.673	0.402	0.206	0.093	0.769	0.423	0.081	0.043
GRNN3	0.903	0.797	0.120	0.064	0.752	0.434	0.080	0.040
GRNN4	0.679	0.432	0.201	0.111	0.663	0.413	0.082	0.051
GRNN5	0.550	0.281	0.226	0.122	0.302	0.080	0.102	0.060
GRNN6	0.641	0.408	0.205	0.100	0.623	0.385	0.083	0.045
Résultats obtenus à l'aide des modèles SVR								
SVR1	0.737	0.436	0.200	0.073	0.700	0.390	0.083	0.036
SVR2	0.652	0.348	0.215	0.090	0.696	0.388	0.083	0.040
SVR3	0.752	0.111	0.251	0.193	0.620	0.011	0.106	0.082
SVR4	0.471	0.075	0.256	0.111	0.317	0.020	0.105	0.052
SVR5	0.596	0.150	0.245	0.105	0.315	0.014	0.106	0.053
SVR6	0.590	0.235	0.233	0.148	0.606	0.240	0.093	0.063
Résultats obtenus à l'aide des modèles GPR								
GPR1	0.968	0.933	0.069	0.038	0.759	0.574	0.069	0.039
GPR2	0.999	0.999	0.003	0.001	0.717	0.506	0.075	0.042
GPR3	0.999	0.999	0.007	0.004	0.710	0.498	0.075	0.041
GPR4	0.952	0.863	0.099	0.060	0.355	0.112	0.100	0.059
GPR5	0.812	0.650	0.158	0.095	0.280	0.014	0.106	0.060
GPR6	0.639	0.407	0.205	0.100	0.616	0.378	0.084	0.045
Résultats obtenus à l'aide des modèles ELM								
ELM1	0.873	0.762	0.130	0.086	0.731	0.469	0.078	0.050
ELM2	0.650	0.423	0.202	0.104	0.713	0.504	0.075	0.043
ELM3	0.740	0.548	0.179	0.091	0.702	0.490	0.076	0.045
ELM4	0.401	0.161	0.244	0.135	0.416	0.173	0.097	0.060
ELM5	0.528	0.279	0.226	0.123	0.368	0.110	0.100	0.058
ELM6	0.632	0.400	0.206	0.101	0.623	0.386	0.083	0.045
Résultats obtenus à l'aide des modèles BAT-ELM								
BAT-ELM1	0.978	0.956	0.056	0.040	0.944	0.883	0.036	0.026
BAT-ELM2	0.909	0.826	0.111	0.081	0.791	0.615	0.066	0.048
BAT-ELM3	0.981	0.963	0.051	0.037	0.875	0.754	0.053	0.037
BAT-ELM4	0.908	0.824	0.112	0.081	0.761	0.482	0.077	0.057
BAT-ELM5	0.877	0.769	0.128	0.082	0.680	0.390	0.083	0.056
BAT-ELM6	0.652	0.425	0.202	0.093	0.635	0.402	0.082	0.045

En ce qui concerne la station d'Oued Lili ([tableau VI.7](#) et [figure VI.12](#)), les performances des modèles BAT-ELM en termes de RMSE et de MAE sont supérieures aux autres modèles, affichant des taux croissants égaux à $\approx 66.10\%$ et $\approx 59.70\%$ par rapport au MLPNN3 ; $\approx 69.61\%$ et $\approx 62.96\%$ par rapport au GRNN3 ; $\approx 72.83\%$, et $\approx 72.90\%$ par rapport au SVR1 ; $\approx 63.37\%$ et $\approx 58.01\%$ par rapport au GPR1 ; et enfin $\approx 65.86\%$ et $\approx 61.53\%$ par rapport à l'ELM1, respectivement.

Comme pour les stations précédentes, la [figure VI.11](#) montre clairement la supériorité de BAT-ELM dans la prévision du débit mensuel. Dans toutes les plages des valeurs (faibles, moyennes et élevées), il a fourni des prévisions beaucoup moins dispersées qui se concentrent sur la ligne 1 :1 par rapport aux autres modèles d'apprentissage automatique. Sur la base des valeurs moyennes des quatre indices statistiques de performance, (RMSE, MAE, R et NSE) pour toutes les stations, la supériorité du BAT-ELM devient plus évidente par rapport aux autres algorithmes.

Tableau VI.7 Résultats de prédiction des différents modèles au pas de temps mensuelle pour la station d'Oued Lili

Modèles	Apprentissage				Validation			
	R	NSE	RMSE, m ³ /s	MAE, m ³ /s	R	NSE	RMSE, m ³ /s	MAE, m ³ /s
Résultats obtenus à l'aide des modèles MLPNN								
MLPNN1	0.871	0.712	0.274	0.157	0.755	0.436	0.447	0.264
MLPNN2	0.871	0.756	0.252	0.150	0.557	0.150	0.549	0.333
MLPNN3	0.855	0.730	0.265	0.144	0.763	0.513	0.416	0.253
MLPNN4	0.606	0.366	0.407	0.232	0.342	0.048	0.581	0.385
MLPNN5	0.573	0.329	0.419	0.232	0.444	0.154	0.548	0.357
MLPNN6	0.690	0.476	0.370	0.210	0.605	0.280	0.505	0.361
Résultats obtenus à l'aide des modèles GRNN								
GRNN1	0.935	0.864	0.188	0.110	0.669	0.408	0.458	0.294
GRNN2	0.831	0.674	0.292	0.159	0.657	0.400	0.461	0.300
GRNN3	0.768	0.547	0.344	0.174	0.695	0.394	0.464	0.297
GRNN4	0.599	0.321	0.421	0.235	0.335	0.107	0.563	0.364
GRNN5	0.590	0.327	0.419	0.231	0.289	0.035	0.585	0.393
GRNN6	0.638	0.324	0.420	0.230	0.692	0.406	0.459	0.333
Résultats obtenus à l'aide des modèles SVR								
SVR1	0.759	0.295	0.429	0.300	0.712	0.241	0.519	0.406
SVR2	0.689	0.347	0.413	0.196	0.688	0.358	0.477	0.313
SVR3	0.766	0.358	0.409	0.258	0.612	0.225	0.524	0.387
SVR4	0.546	0.147	0.472	0.281	0.336	0.077	0.572	0.408
SVR5	0.632	0.246	0.444	0.212	0.404	0.122	0.558	0.367
SVR6	0.611	0.258	0.440	0.222	0.627	0.308	0.495	0.325
Résultats obtenus à l'aide des modèles GPR								
GPR1	0.979	0.956	0.107	0.065	0.764	0.582	0.385	0.262
GPR2	0.999	0.997	0.027	0.014	0.607	0.331	0.487	0.335
GPR3	0.961	0.920	0.144	0.084	0.704	0.465	0.436	0.268
GPR4	0.543	0.289	0.431	0.244	0.364	0.125	0.557	0.364
GPR5	0.806	0.626	0.312	0.177	0.401	0.134	0.554	0.343
GPR6	0.665	0.442	0.382	0.210	0.656	0.390	0.465	0.333
Résultats obtenus à l'aide des modèles ELM								
ELM1	0.842	0.710	0.275	0.172	0.755	0.520	0.413	0.286
ELM2	0.664	0.441	0.382	0.223	0.694	0.481	0.429	0.298
ELM3	0.729	0.532	0.350	0.201	0.679	0.451	0.441	0.291
ELM4	0.512	0.262	0.439	0.252	0.422	0.174	0.541	0.354
ELM5	0.576	0.332	0.418	0.232	0.531	0.268	0.510	0.329
ELM6	0.648	0.420	0.389	0.214	0.665	0.422	0.453	0.332
Résultats obtenus à l'aide des modèles BAT-ELM								
BAT-ELM1	0.988	0.975	0.080	0.062	0.973	0.944	0.141	0.110
BAT-ELM2	0.987	0.974	0.082	0.063	0.864	0.744	0.302	0.226
BAT-ELM3	0.868	0.754	0.253	0.140	0.839	0.703	0.325	0.204
BAT-ELM4	0.682	0.465	0.374	0.229	0.495	0.232	0.522	0.317
BAT-ELM5	0.624	0.389	0.399	0.222	0.700	0.490	0.425	0.268
BAT-ELM6	0.670	0.449	0.379	0.209	0.604	0.339	0.484	0.353

À la station d'Arib Abda, la valeur R moyenne du BAT-ELM était supérieure de ≈ 0.791 à ≈ 0.751 , ≈ 0.751 , ≈ 0.630 , ≈ 0.674 , ≈ 0.732 , et ≈ 0.683 obtenue par les modèles ELM, GPR, SVR, GRNN et MLPNN, respectivement. En outre, la valeur NSE moyenne du modèle proposé était de 0.612 , supérieure aux valeurs de ≈ 0.563 , ≈ 0.408 , ≈ 0.440 , ≈ 0.482 , et ≈ 0.466 obtenues par les modèles ELM, GPR, SVR, GRNN et MLPNN, respectivement. De même, à la station d'El Ababsa, le R moyen du BAT-ELM était de 0.781 , supérieur à ≈ 0.592 , ≈ 0.573 , ≈ 0.542 , ≈ 0.650 , et ≈ 0.560 obtenus par les modèles ELM, GPR, SVR, GRNN et MLPNN, respectivement. En outre, la valeur NSE moyenne du modèle hybride était supérieure de ≈ 0.588 aux valeurs de ≈ 0.355 , ≈ 0.347 , ≈ 0.177 , ≈ 0.358 , et ≈ 0.325 obtenues par les modèles ELM, GPR, SVR, GRNN et MLPNN, respectivement. Enfin, à la station d'Oued Lili, les valeurs moyennes R et NSE de BAT-ELM étaient respectivement de ≈ 0.74 et ≈ 0.575 supérieures aux valeurs de ≈ 0.624 et ≈ 0.386 , ≈ 0.583 et ≈ 0.338 , ≈ 0.563 et ≈ 0.222 , ≈ 0.556 et ≈ 0.292 , ≈ 0.578 et ≈ 0.264 , obtenues par les modèles ELM, GPR, SVR, GRNN et MLPNN, respectivement. La comparaison des performances des modèles au-delà de la première combinaison d'entrées rend difficile la conclusion globale, car les résultats varient d'une station à l'autre. Dans la station d'Arib Abda, entre la première et la deuxième combinaison d'entrées, les performances du BAT-ELM1 par rapport au BAT-ELM2 ont diminué de $\approx 1,00\%$, $\approx 2,8\%$, $\approx 9\%$ et $\approx 27,16\%$ en termes de R, NSE, RMSE et MAE, respectivement ; à la station d'El Ababsa, les performances du BAT-ELM1 par rapport au BAT-ELM2 ont diminué de $\approx 15,30\%$, $\approx 26,80\%$, $\approx 45,45\%$ et $\approx 45,83\%$ en termes de R, NSE, RMSE et MAE, respectivement, Tandis qu'à la station Oued Lili, les différences entre BAT-ELM1 et BAT-ELM2 étaient plus évidentes, les valeurs R et NSE ayant diminué de $\approx 10,90\%$ et $\approx 20,00\%$, et les valeurs RMSE et MAE ayant augmenté de $\approx 53,31\%$ et $\approx 51,32\%$, respectivement.

Pour toutes les stations, il apparaît clairement que les performances prédictives les plus mauvaises sont obtenues à l'aide des scénarios quatre et cinq (**Tableau VI.1**), représentés par de très mauvais indices de performance, c'est-à-dire les valeurs RMSE et MAE les plus élevées et les valeurs R et NSE les plus faibles. En outre, même le modèle BAT-ELM a démontré son incapacité à surmonter les limites des modèles indépendants. De plus, si l'on exclut le BAT-ELM de la comparaison, on peut conclure qu'il n'y a pas de supériorité évidente d'un modèle par rapport à un autre, et que les performances globales varient de modérées à faibles (**Tableaux VI.5 à VI.7**).

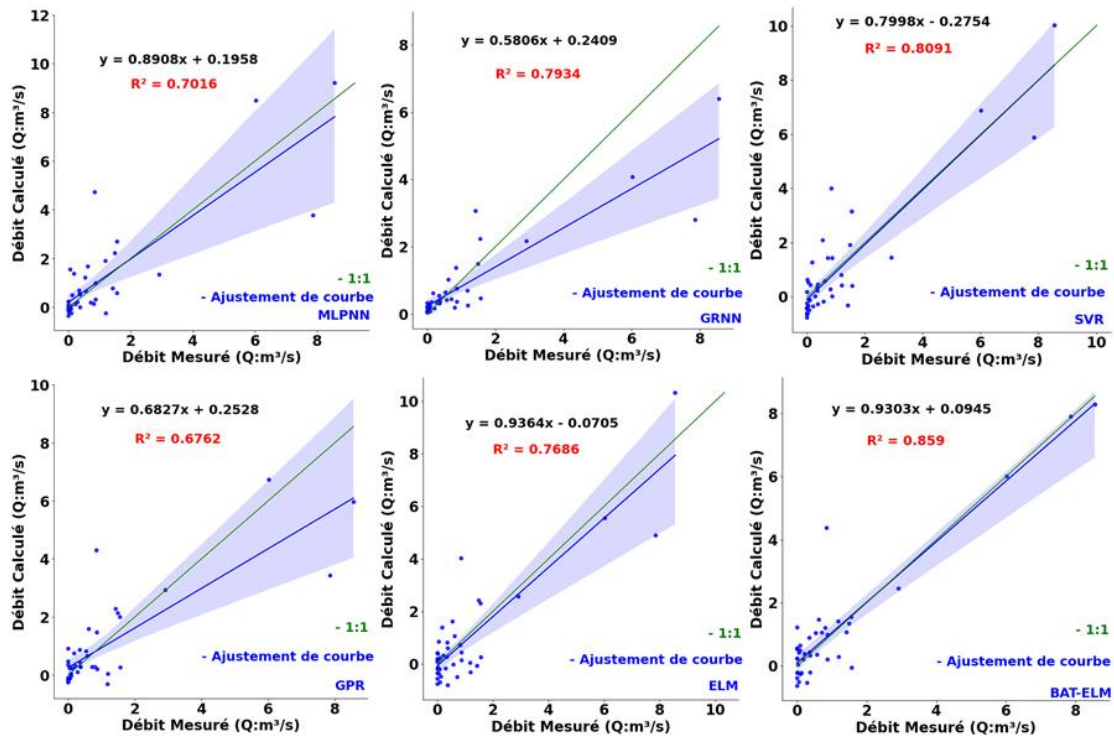


Figure VI.9 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Arib Abda pendant la phase de test (pas de temps mensuelle)

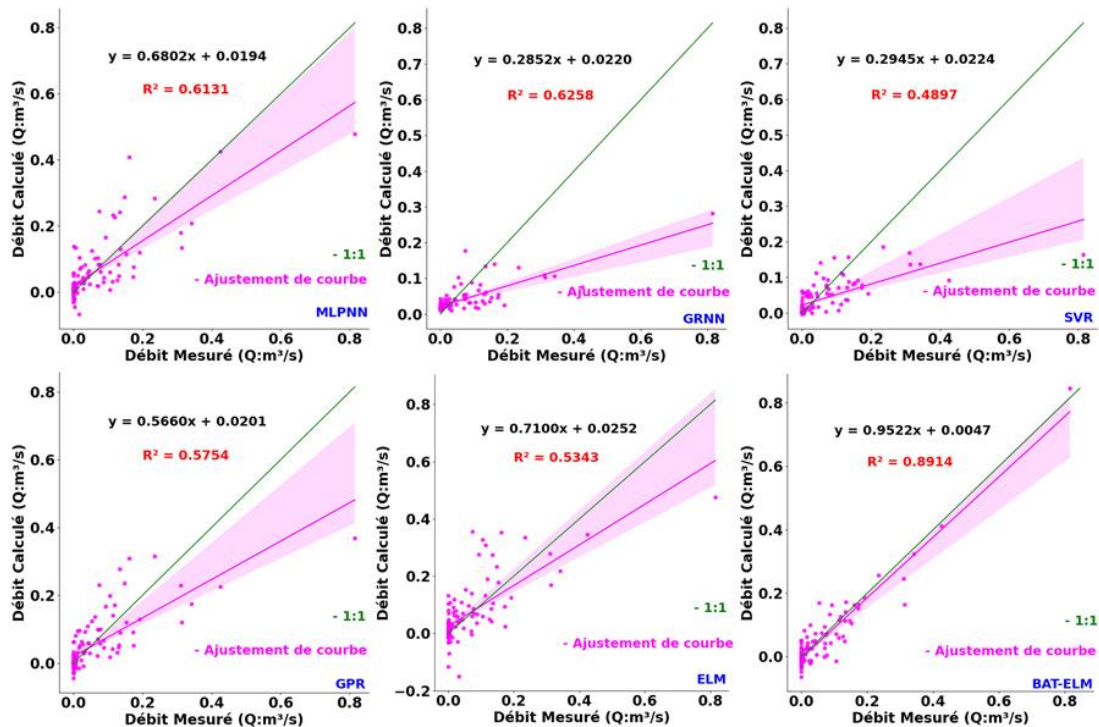


Figure VI.10 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'El Ababsa pendant la phase de test (pas de temps mensuelle)

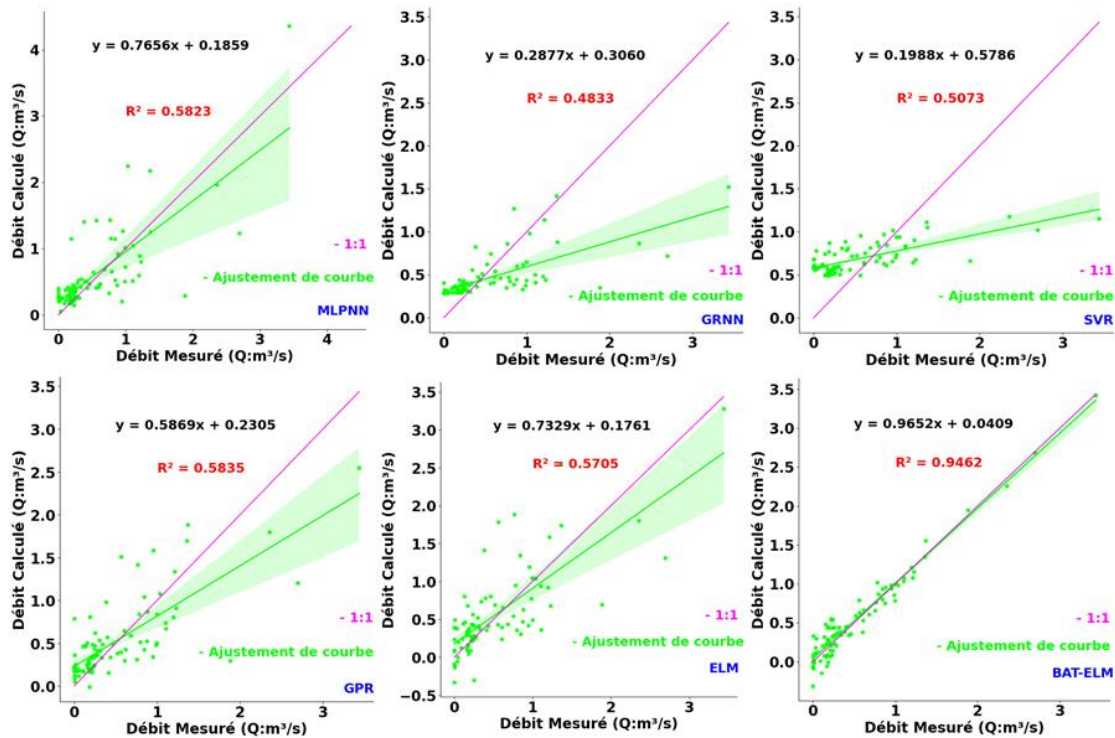


Figure VI.11 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé à la station d'Oued Lili pendant la phase de test (pas de temps mensuelle)

VI.2.3 Evaluations de performance des modèles de prédiction des débits dans le pas journalier et mensuelle par des critères des performances graphiques

Pour une comparaison plus approfondie des performances du modèle aux trois stations au pas de temps journalier et mensuelle, nous présentons dans les figures VI.12 à VI.15 les graphiques récapitulatifs comprenant le diagramme en boîte, le diagramme en violon, le diagramme radar et le diagramme de Taylor. Il semble de plus en plus évident que le modèle BAT-ELM est très performant et que sa supériorité ressort clairement de ces graphiques.

Les diagrammes en boîte et les diagrammes en violon sont des descriptions graphiques utile des principales caractéristiques d'un ensemble d'observations, Tandis qu'un diagramme en boîte présente des statistiques sommaires telles que la moyenne et l'intervalle interquartile, le diagramme en violon montre la distribution complète des données. Le diagramme en violon peut être utilisé pour les données multimodales (plus d'un pic). Dans ce cas, le diagramme du violon montre la présence de différents pics, leur position et leur amplitude relative. Les diagrammes en boîte et les diagrammes en violon (Figure VI.12 et Figure VI.13) montrent que les prévisions par le modèle

BAT-ELM dans le pas journalier et mensuelle sont les plus proches des valeurs de débit mesurées que les autres modèles.

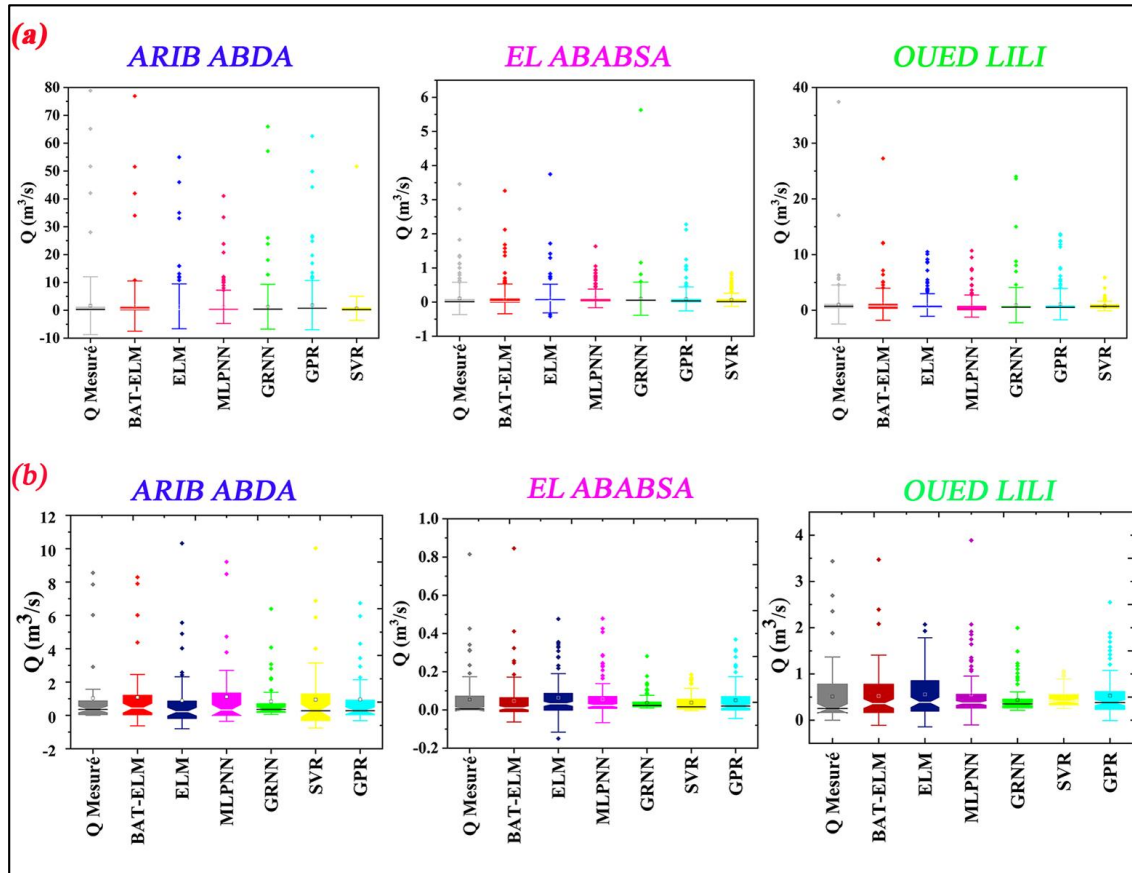


Figure VI.12 Diagramme en boîte (Box plot) des débits prévus par les meilleurs modèles par rapport aux débits observés pendant la phase de validation dans les trois stations d'étude : (a) Pas journalier, (b) Pas mensuelle

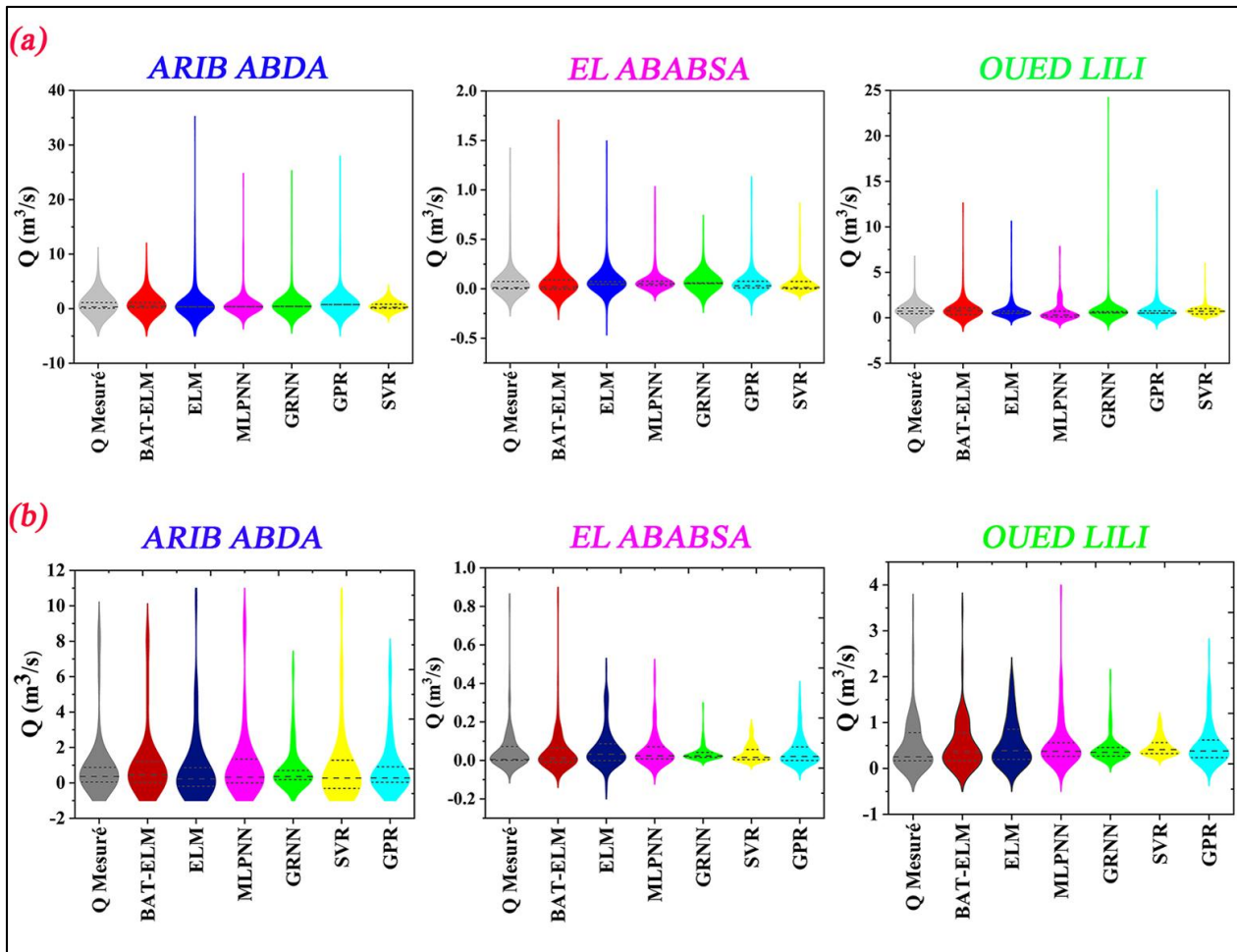


Figure VI.13 Diagramme en violon (Violin plot) des débits prévus par les meilleurs modèles par rapport aux débits observés pendant la phase de validation dans les trois stations d'étude : (a) Pas journalier, (b) Pas mensuelle

En utilisant les diagrammes de Taylor (Figure VI.14) pour la comparaison des modèles qui est principalement basée sur la représentation de distribution spatiale des valeurs estimées (c'est-à-dire, le test) par rapport au champ observé (champ de référence) en compilant le RMSE, l'écart type et le coefficient de corrélation dans le système polaire. Il ressort de ces chiffres là que le modèle BAT-ELM est plus proche du champ observé (référence) que les autres modèles dans les trois stations, à l'échelle journalière et mensuelle. De plus, le modèle BAT-ELM a la plus faible valeur de RMSE et de l'écart type et une corrélation plus élevée par rapport aux autres modèles.

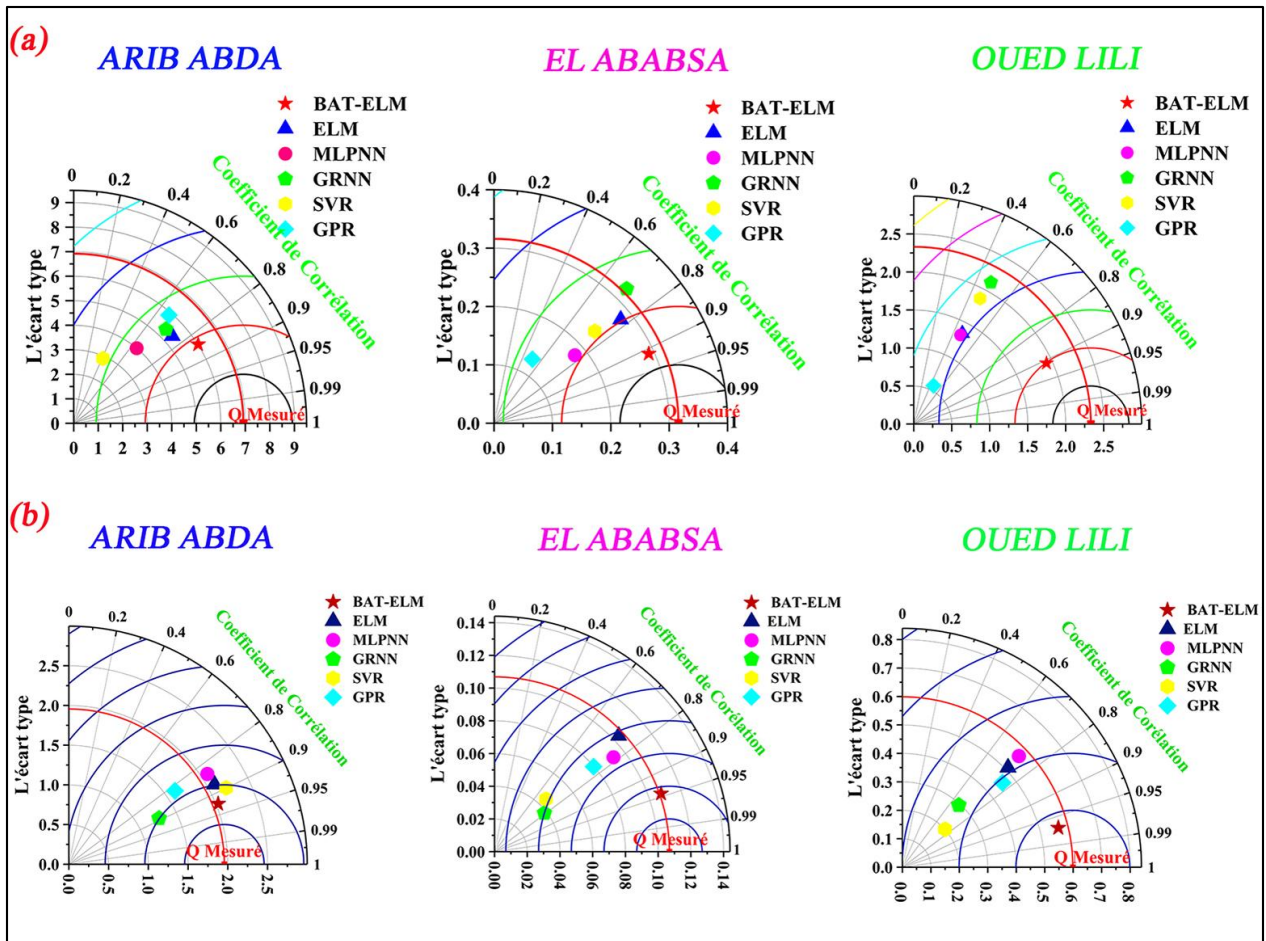


Figure VI.14 Diagramme de Taylor des débits prévus par les meilleurs modèles par rapport aux débits observés pendant la phase de validation dans les trois stations d'étude :(a) Pas journalier, (b) Pas mensuelle

Les diagrammes radar (Figure VI.15) justifient visuellement les statistiques RMSE, MAE, NSE et R montrant que le BAT-ELM est plus performant que les autres modèles pour la prédiction des débits journaliers et mensuels.

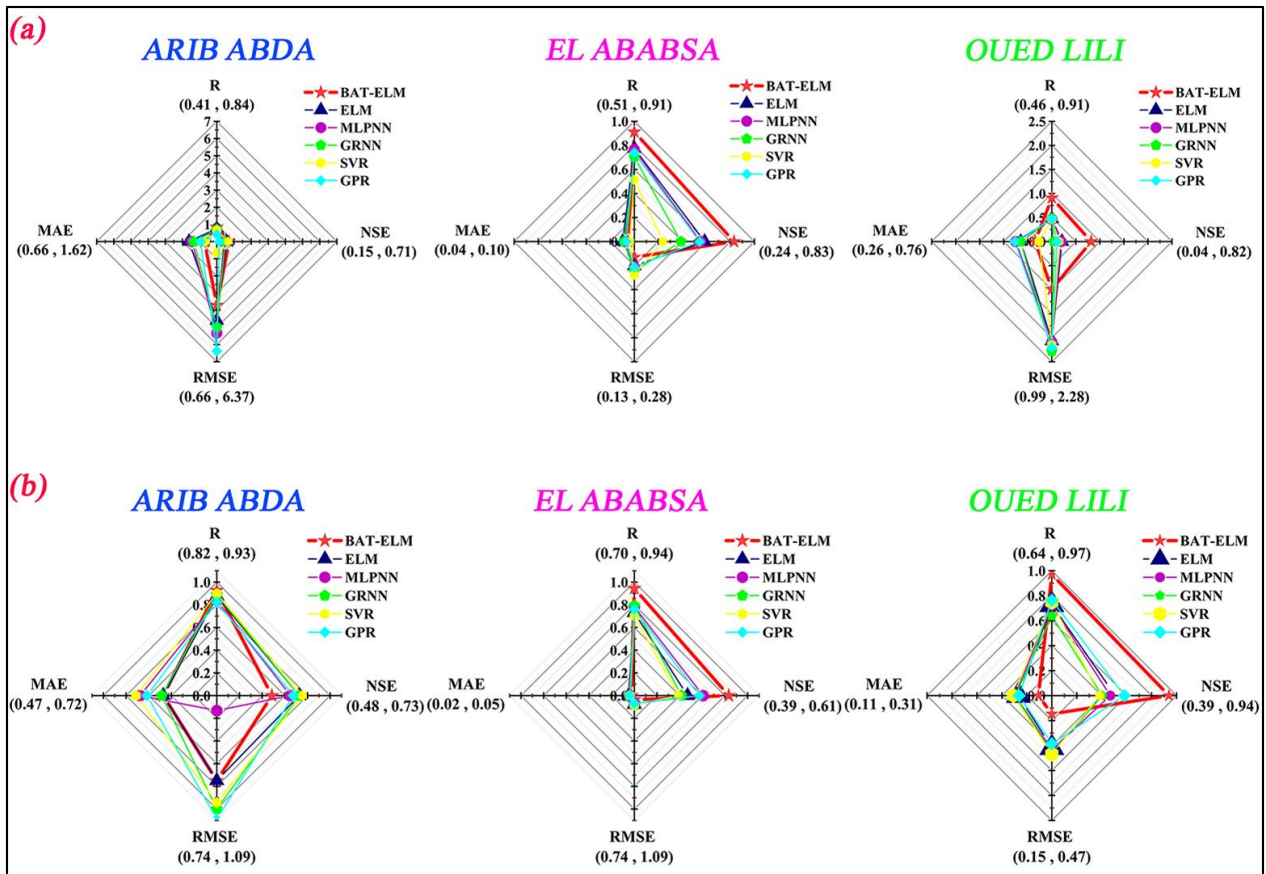


Figure VI.15 Diagramme radar (Radar chart) de meilleurs modèles pendant la phase de validation dans les trois stations d'étude :(a) Pas journalier, (b) Pas mensuelle

Enfin, la **figure VI.16** met en évidence les capacités des modèles BAT-ELM à capturer les fortes non-linéarités de l'ensemble de données et la manière dont la prévision des valeurs maximales a été réalisée avec succès.

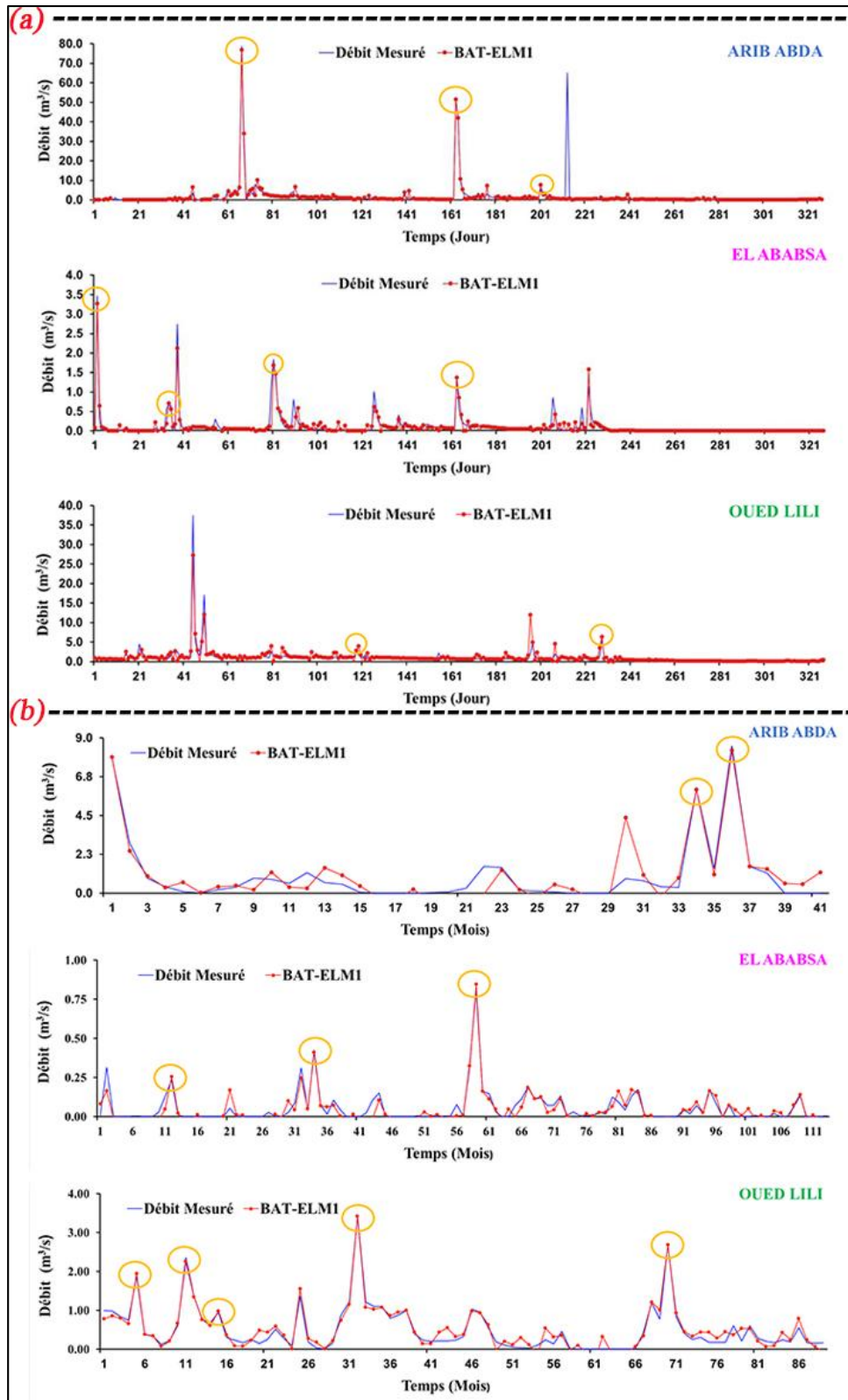


Figure VI.16 Variation temporelle des débits observés et calculé dans la phase de validation par le modèle (BAT-ELM) pour les trois stations d'étude : (a) Pas journalier, (b) Pas mensuelle

VI.3 Discussion des résultats

Cette étude a examiné le potentiel de prévision des débits journaliers et mensuels d'un nouveau modèle d'intelligence artificielle optimisé, appelé BAT-ELM. L'approche proposée BAT-ELM est inspirée de la combinaison de l'algorithme Bat et de la machine d'apprentissage extrême autonome. Les prévisions ont été effectuées pour trois stations situées dans le bassin du Cheliff, au nord de l'Algérie. Les résultats montrent que l'hybridation de l'ELM à l'aide de l'algorithme Bat a permis d'améliorer les performances de l'ELM et d'obtenir une grande précision de prévision. En outre, il a été constaté que le BAT-ELM a produit des performances numériques élevées avec les erreurs de prévision les plus faibles. À la connaissance des auteurs, aucune étude antérieure n'a été menée sur l'utilisation de l'algorithme BAT pour la prévision des débits journaliers et mensuels. On a introduit une nouvelle approche BAT-ELM dans la présente étude, et une comparaison de nos résultats avec ceux rapportés dans des études précédentes est nécessaire pour mieux mettre en évidence la robustesse du modèle proposé.

Pour le pas journalier, (Gomaa et al., 2023) ont examiné l'application d'un modèle hybride de perceptron multicouche optimisé par essaimage de particules (MLPNN-PSO) dans l'analyse de la relation entre les précipitations et le ruissellement et dans la prévision du débit de ruissellement. Les modèles de réseau neuronal à fonction de base radiale gaussienne (GRNN) et de régression par processus gaussien (GPR) sont utilisés à titre de procédure de comparaison pour évaluer la performance et la supériorité du modèle MLPNN-PSO. Les résultats montrent que le modèle MLPNN-PSO est plus performant et a donné des valeurs de R et NSE presque égaux à ≈ 0.971 et ≈ 0.986 , respectivement, en dessus des valeurs de $\approx 0,84$ à $\approx 0,91$, et $\approx 0,71$ à $\approx 0,83$ obtenues dans notre étude. (Khosravi et al., 2022) La présente étude démontre les premières applications d'un nouvel algorithme d'apprentissage profond optimisé d'un réseau neuronal convolutif (CNN) utilisant l'algorithme métaheuristique BAT (CNN-BAT). En utilisant les pouvoirs de prévision de quatre algorithmes bien connus comme références - perceptron multicouche (MLPNN-BAT), système d'inférence neuro-floue adaptatif (ANFIS-BAT), régression vectorielle de soutien (SVR-BAT) et forêt aléatoire (RF-BAT), les modèles CNN-BAT, MLPNN-BAT, ANFIS-BAT, SVR-BAT et RF-BAT est testé pour la prévision du débit journalier dans le bassin versant de Korkorsar, au nord de l'Iran et ont rapporté des précisions de prédiction basées sur R $\approx 0,91$, $\approx 0,79$, $\approx 0,87$, $\approx 0,82$ et $\approx 0,80$, respectivement. Ces résultats montrent que CNN-BAT a donné des résultats égaux à ceux obtenus aux stations d'El-Ababsa (R $\approx 0,91$) et d'Oued Lili (R $\approx 0,91$). En outre, ces résultats

révèlent que le modèle BAT-ELM est supérieur aux modèles MLP-BAT, ANFIS-BAT, SVR-BAT, RF-BAT, étant donné que leurs valeurs de $R \approx 79$ à 0.87 sont inférieures à celles obtenues dans les deux stations susmentionnées. (Kilinc, 2022) a proposé une approche hybride intégrant les réseaux de mémoire à long terme (LSTM) et l'algorithme PSO pour la prévision du débit quotidiens des cours d'eau. À cette fin, L'auteur a employé trois stations hydrologiques dans l'étude le long du bassin de l'Oronte, respectivement Karasu, Demirköprü et Samandağ. Les résultats ont démontré l'efficacité et la supériorité du modèle hybride LSTM-PSO proposé pour la prévision des débits de cours d'eau quotidiens et ont rapporté qu'une bonne précision a été obtenue en utilisant LSTM-PSO avec $R^2 \approx 0.927$ à 0.974 sont supérieures à celles obtenues dans les trois stations de la région d'étude ($R^2 \approx 0.713$ à 0.831). (Kilinc & Haznedar, 2022) Ont proposé une approche hybride, intégrant des réseaux de mémoire à long et court terme (LSTM) et un algorithme génétique (GA), pour prédire le débit quotidien de l'Euphrate en Turquie. La régression linéaire a été utilisée dans le cadre du processus de comparaison pour évaluer les performances de la méthode proposée et démontrer sa capacité prédictive supérieure. La comparaison des résultats de la prévision des débits journaliers a révélé l'efficacité et les performances efficaces du modèle LSTM-GA. Les résultats ont démontré la supériorité du modèle hybride LSTM-GA avec $R^2 \approx 0.866$ à 0.998 . (Adnan et al., 2022) ont intégré le modèle SVM avec l'algorithme SA et l'algorithme d'optimisation de la mouche de mai (MOA) pour prédire le débit dans le bassin de la rivière Helium. Dodangeh et al. (2021) ont proposé l'hybridation du modèle ANFIS en utilisant les algorithmes GWO, WOA et PSO pour la prévision du débit journalier de la rivière Taleghan, en Iran. Le modèle ANFIS-WOA s'est avéré le plus efficace et a donné une valeur NSE d'environ $\approx 0,93$. (Malik et al., 2020) ont optimisé SVR par six algorithmes méta-heuristiques, à savoir l'optimisation Ant Lion (SVR-ALO), l'optimiseur Multi-Verse (SVR-MVO), l'optimiseur Spotted Hyena (SVR-SHO), l'optimisation Harris Hawks (SVR-HHO), l'optimisation Particle Swarm (SVR-PSO) et l'optimisation bayésienne (SVR-BO) pour prédire le débit journalier dans le bassin versant de Naula, dans l'État d'Uttarakhand, en Inde. Les résultats montrent que le nouvel algorithme HHO a surpassé les autres algorithmes méta-heuristiques en termes de précision de prévision avec $R^2 \approx 0.77$ inférieures à ceux celles obtenus aux stations d'El Ababsa ($R^2 \approx 0.826$) et Oued Lili ($R^2 \approx 0.831$). Hadi et al. (2019) ont appliqué le modèle ELM couplé à l'extrême gradient boosting (XGB) pour prédire le débit journalier à Goksu-Himmeti, en Turquie, et ils ont rapporté une valeur NSE d'environ ≈ 0.704 . (Adnan et al., 2019) ont examiné la capacité de prédiction et d'estimation d'une nouvelle méthode heuristique, le modèle de

machine d'apprentissage extrême élaguée de manière optimale (OP-ELM), pour les débits journaliers des stations de Fujiangqiao et de Shehang sur la rivière Fujiang. La précision de la prévision de la méthode OP-ELM est comparée à d'autres modèles d'informatique douce, à savoir le système d'inférence neuro-floue adaptatif - optimisation par essais de particules (ANFIS-PSO), MARS et M5Tree. Les résultats globaux indiquent que la méthode OP-ELM peut être utilisée avec succès pour prédire et estimer le débit journalier.

Pour le pas mensuel, Suite à une étude précédente, [Tikhamarine et al. \(2019\)](#) ont utilisé quatre modèles hybrides d'apprentissage automatique, à savoir la régression à vecteur de support basée sur l'optimiseur du loup gris (GWO-SVR), la SVR basée sur l'optimisation par essaim particulière (PSO-SVR), la SVR basée sur l'optimiseur multiverse (MVO-SVR) et la SVR basée sur l'algorithme d'évolution complexe mélangée (SCE-UA-SVR) pour prédire le débit mensuel à deux stations à l'est de l'Algérie. Ils ont constaté que des précisions de prévision très faibles ont été obtenues avec R et NSE presque égaux à $\approx 0,558$ et $\approx 0,331$, respectivement, en dessous des valeurs de $\approx 0,927$ à $\approx 0,973$, et $\approx 0,846$ à $\approx 0,944$ obtenues dans notre étude. Cependant, ils ont démontré qu'une amélioration supplémentaire des performances du modèle a été obtenue en combinant la décomposition en ondelettes avec les modèles hybrides et les valeurs R et NSE ont été améliorées pour atteindre les valeurs de $\approx 0,98$ et $\approx 0,96$, respectivement. Dans une autre étude, [Tikhamarine et al. \(2020\)](#) ont comparé GWO-SVR, GWO-MLPNN et GWO-MLR pour la prévision du débit mensuel du fleuve Nil, en Égypte. Leurs résultats ont révélé que le R et le NSE étaient inférieurs aux valeurs $\approx 0,930$ et $\approx 0,861$, inférieures aux valeurs obtenues dans notre étude (R $\approx 0,927$ à $\approx 0,973$) et (NSE $\approx 0,846$ à $\approx 0,944$), respectivement. [Guo et al. \(2011\)](#) ont rapporté qu'une excellente précision de prévision a été obtenue en utilisant un algorithme PSO amélioré combiné avec le modèle SVR (IPSO-SVR) pour lequel R $\approx 0,913$ et NSE $\approx 0,834$ ont été obtenues et étaient inférieures à celles obtenues dans notre étude. [Zhao et al. \(2021\)](#) ont hybridé le GRU (gated recurrent unit deep learning) en utilisant un GWO amélioré pour la prévision des débits mensuels, et ils ont renforcé et complété le modèle en utilisant un ensemble complet de décomposition de mode empirique avec bruit adaptatif (CEEDAN) pour obtenir une valeur de NSE comprise entre $\approx 0,36$ et $\approx 0,84$, bien inférieure aux valeurs obtenues dans notre étude (NSE $\approx 0,846$ à $\approx 0,944$). Dans une étude récente menée par [Adnan et al. \(2021b\)](#), un système d'inférence floue neuronale dynamique et évolutive (DENFIS) a été appliqué pour prédire le débit mensuel au Pakistan. Les résultats obtenus ont montré que la valeur du NSE était de l'ordre de $\approx 0,915$, supérieure à la valeur

obtenue à la station d'Arib Abda ($NSE \approx 0.846$) et à la station d'El Ababsa ($NSE \approx 0.883$), et inférieure à la valeur obtenue à la station d'Oued Lili ($NSE \approx 0.944$). Enfin, dans l'étude de [Jiang et al. \(2020\)](#), l'ELM a été hybridé en utilisant l'optimisation par essaim particulière améliorée, PSO et GA, et les valeurs R et NSE étaient comprises entre $\approx 0,76$ et $\approx 0,78$, et entre $\approx 0,88$ et $\approx 0,89$, respectivement, soit des valeurs inférieures à celles obtenues par le BAT-ELM dans le cadre de notre étude. Dans une autre étude, [Afan et al. \(2020\)](#) ont utilisé le réseau neuronal à fonction de base radiale hybridé avec GA (RBFNN-GA) pour la prédiction des débits mensuels et ont obtenu des précisions élevées avec des valeurs R^2 allant de ≈ 0.876 à ≈ 0.955 pour les douze mois de l'année. [Al-Sudani et al. \(2019\)](#) ont utilisé la spline de régression adaptative multivariée intégrée à l'évolution différentielle (MARS-DE) pour prédire le débit mensuel du fleuve Tigre, en Irak, et ont rapporté une meilleure performance du modèle avec une valeur NSE de ≈ 0.893 , Enfin [Yaseen et al. \(2019\)](#) ont utilisé ANFIS hybridé avec PSO, GA et DE et ont rapporté qu'une grande précision a été obtenue en utilisant ANFIS-PSO avec $R^2 \approx 0.990$.

Comme indiqué à la fin de l'introduction, l'une des limites de l'approche proposée dans la présente étude est son application à un ensemble de données limité, à savoir trois stations situées dans la même région climatique ; une analyse plus approfondie et un plus grand nombre de stations situées dans des régions climatiques différentes peuvent contribuer à fournir des conclusions plus solides. En plus, d'après les résultats numériques obtenus, il est clair que certaines limitations de l'application du modèle en termes d'entrées de données doivent être soulignées, plus précisément, aux trois stations, le modèle BAT-ELM a complètement échoué si : (i) seules les précipitations sont utilisées comme variables d'entrée au même moment, c'est-à-dire $P(t)$, (ii) l'utilisation des seules précipitations aux trois moments précédents (c'est-à-dire $P(t-1)$, $P(t-2)$, $P(t-3)$) ne permet pas d'obtenir une grande précision de prévision, et (iii) l'utilisation des seuls débits aux trois moments précédents (c'est-à-dire $Q(t-1)$, $Q(t-2)$, $Q(t-3)$) n'est pas non plus appropriée. Par conséquent, nous pouvons conclure que la sélection de la meilleure combinaison d'entrées est l'une des conditions du succès du modèle BAT-ELM. En conclusion, sur la base de la comparaison entre les six scénarios, il est clair que le modèle BAT-ELM a mieux fonctionné en utilisant la combinaison des données de précipitations et de débit, et si seules les précipitations sont utilisées comme variable d'entrée, il est obligatoire d'augmenter le nombre de décalages jusqu'à trois pour obtenir des performances numériques élevées.

Le diagramme en boîte, le diagramme en violon, le diagramme de Taylor et le diagramme radar présentés dans les figures VI.12, VI.13, VI.14, et VI.15 sont présentés pour une meilleure comparaison entre les performances des modèles à l'échelle journalier et mensuelle. Le diagramme en boîte montre clairement que BAT-ELM est le meilleur modèle pour les trois stations. Certaines informations importantes sont disponibles : (i) les valeurs médianes sont représentées par les lignes médianes, (ii) les valeurs comprises entre les valeurs inférieures et supérieures sont représentées par les moustaches, (iii) la partie importante, à savoir la boîte, est utilisée pour représenter les premier et troisième quartiles, et (iv) les valeurs aberrantes sont les points distincts représentés au-dessus de la boîte et sont très importantes pour évaluer la capacité des modèles à prédire les valeurs maximales du débit d'eau. Le diagramme en violon est une sorte de diagramme en boîte auquel on a ajouté la densité du noyau. Le diagramme radar est un simple tracé des performances des modèles avec les quatre indices statistiques montrant la supériorité de l'un par rapport à l'autre. Enfin, le diagramme de Taylor est un outil graphique puissant pour la comparaison des modèles en fonction de deux facteurs principaux : (i) les capacités d'ajustement entre les données mesurées et calculées et (ii) la variabilité mesurée par l'écart-type normalisé. Le diagramme de Taylor pourrait être structuré autour de trois axes : deux pour l'écart type normalisé et le coefficient de corrélation dans une distance polaire. Il est clair que le BAT-ELM est le modèle le plus significatif, représenté par un point proche du modèle mesuré.

Conclusion

Dans la présente étude, une nouvelle approche hybride appelée BAT-ELM a été étudiée en combinant un système d'apprentissage extrême autonome avec l'algorithme Bat afin d'obtenir une prévision robuste et précise des débits journaliers et mensuels. Les résultats des prévisions et l'analyse comparative avec les MLPNN, GRNN, SVR, GPR et ELM permettent de tirer les conclusions suivantes.

1. La performance du modèle hybride BAT-ELM s'est considérablement améliorée par rapport à celle des MLPNN, GRNN, SVR, GPR et du modèle ELM autonome.
2. La comparaison des cinq modèles autonomes : MLPNN, GRNN, SVR, GPR et ELM, a révélé qu'ils étaient presque égaux et qu'il existait de légères différences entre eux.

3. Parmi les six combinaisons d'entrées testées dans la présente étude, il a été constaté que les plus mauvaises précisions de prévision ont été obtenues en utilisant les quatrième, cinquième et sixième combinaisons d'entrées dans lesquelles seuls P ou Q à plusieurs moments de décalage antérieurs ont été utilisés comme variables d'entrée, en outre, les précisions les plus faibles ont été obtenues en utilisant la cinquième combinaison pour laquelle seul le débit a été utilisé, même le BAT-ELM n'a pas réussi à améliorer les précisions de prévision pour ces combinaisons.
4. Dans la présente étude, il est clair que seul le modèle ELM a été optimisé à l'aide de l'algorithme BAT et que les cinq autres modèles autonomes n'ont en aucun cas été optimisés. Par conséquent, les améliorations possibles des modèles MLPNN, GRNN, SVR et GPR seront également explorées dans une étude future. La comparaison de l'algorithme bat avec d'autres algorithmes métaheuristiques peut être effectuée dans le cadre d'une étude future.

Sur la base de cette étude, il reste quelques recommandations à prendre en compte pour les recherches futures. Tout d'abord, la principale préoccupation concerne l'inclusion d'autres variables d'entrée qui peuvent contribuer à améliorer la précision des prévisions, et l'évapotranspiration et la température peuvent être les candidats les plus appropriés. Deuxièmement, outre le fait que l'algorithme d'optimisation a contribué à l'amélioration du modèle ELM, un algorithme de sélection plus efficace pour les variables d'entrée peut également contribuer au développement de modèles robustes en ne sélectionnant que les prédicteurs pertinents.

CHAPITRE VII Amélioration de la prévision des débits dans le pas journalier par des modèles hybrides ondelettes-intelligence artificielle

Introduction

Dans la prédiction des débits, l'utilisation de modèles linéaires et de modèles simples d'intelligence artificielle a montré ses limites, en particulier à l'échelle journalière pour les modèles SVR et ELM, comme clairement démontré dans les sections précédentes de ce travail. Cependant, les modèles MLPNN, GRNN et GPR ont fait des progrès remarquables dans la prédiction des débits journaliers des cours d'eau. Les progrès dans le domaine du traitement de signal et des transformées en ondelettes permettent de pallier de manière fiable les lacunes des modèles d'IA dans le traitement des données non stationnaires. L'analyse par ondelettes fournit une représentation à l'échelle temporelle du processus et de ses relations, ce qui permet de construire des modèles conjoints traçables ayant diverses applications dans les domaines de l'hydrologie et de l'optimisation, ce qui a motivé notre attention à utiliser cette technique dans le domaine de prédiction des débits des cours d'eau au pas de temps journaliers.

Dans ce chapitre, nous présentons et discutons les résultats de la prévision des débits journaliers pour deux sous-bassins (Rouina Zeddine et Ras Ouahrane) dans le grand bassin du Cheliff en Algérie, Nous intégrons une famille d'ondelettes sous l'abréviation MODWT déjà été définis dans le chapitre précédent (chapitre III) avec les modèles d'apprentissage automatique MLPNN, GRNN, GPR et le modèle LSTM.

VII.1 La mise en œuvre des modèles d'apprentissage automatique proposée dans l'étude

VII.1.1 Acquisition des données pour les modèles et sélection des variables d'entrée par les méthodes ACF, PAC et XCF

Cette étude a utilisé les données de l'Agence Nationale des Ressources Hydrauliques (ANRH) ; deux stations hydrométriques ont été sélectionnées pour l'étude de cas ([figure VII.1](#)) : (i) Bir Ouled Tahar (code 011905) et (ii) Ponteba Barrage (code 012203) et; toutes ces stations appartiennent au bassin du Cheliff, qui est situé dans le nord de l'Algérie, Le sous bassin versant d'oued Rouina Zeddine est contrôlé par la station de Bir Ouled Tahar dans le bassin versant Haut moyen Cheliff ainsi que la stations de Ponteba Barrage qui contrôle le sous bassin d'Oued de Ras Ouahrane dans le bassin versant Haut moyen Cheliff.

Deux stations hydrométriques ont été sélectionnées pour l'étude de cas (figure VII.1) : (i) Bir Ouled Tahar (code 011905) et (ii) Ponteba Barrage (code 012203). Les sous bassins versants de l'oued Rouina Zeddine et de l'Oued de Ras Ouahrane dans le bassin versant Haut et Moyen Cheliff sont contrôlés respectivement par la station de Bir Ouled Tahar et la station de Ponteba Barrage.

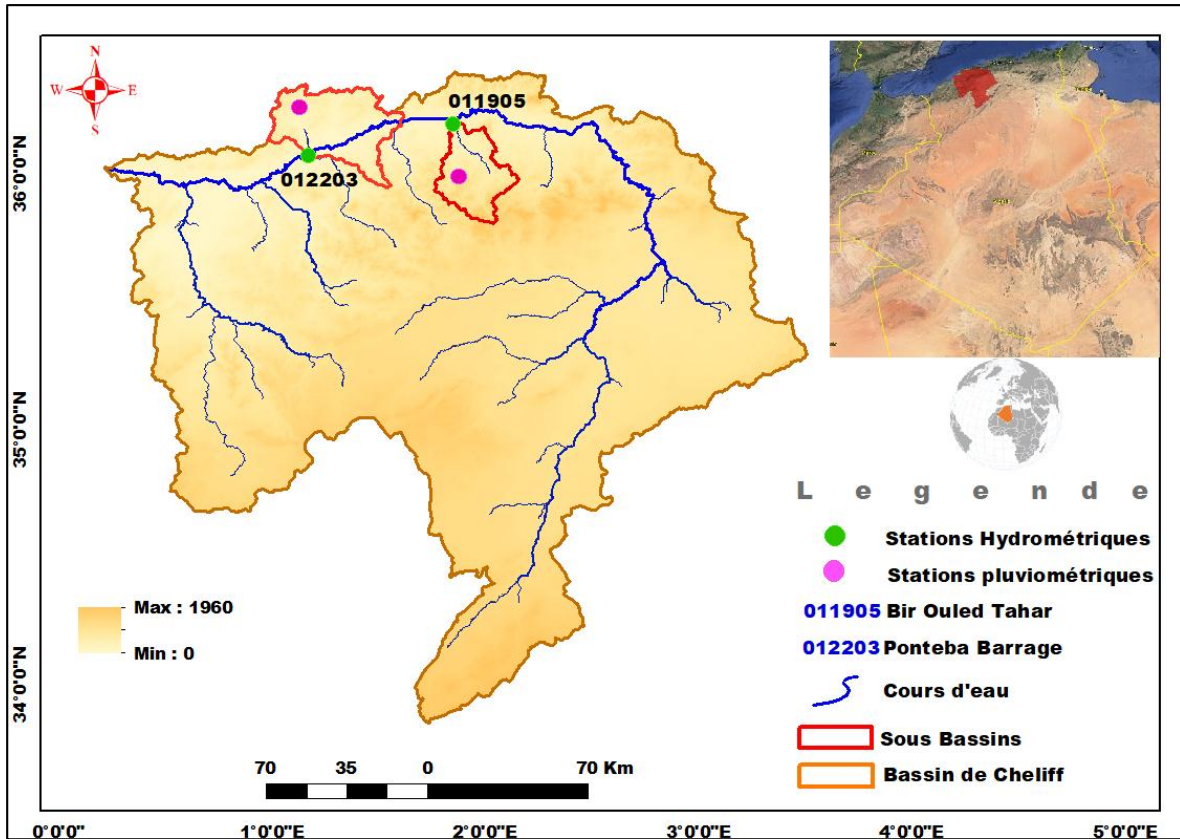


Figure VII.1 Localisation des stations pluviométriques et hydrométriques dans le bassin de Cheliff

Les données sur le débit des cours d'eau (Q) et les précipitations (P) sont disponibles à une échelle journalière. Les données ont été divisées en ensembles de formation (d'entraînement) (70 %) et de validation (30 %) pour chaque station. Les sous-ensembles d'entraînement et de validation étaient donc respectivement de 2555 et 1094 pour les stations (i) et (ii). Dans le [chapitre V](#) le [tableau V.6](#), en ce qui concerne la description statistique de (Q) et (P), nous indiquons la moyenne, le maximum, le minimum, l'écart type et le coefficient de variation (R). D'après les résultats des paramètres statistiques du [tableau V.6](#), en particulier le coefficient de corrélation R, nous pouvons identifier que des valeurs R très faibles ont été détectées dans les deux stations et qu'aucune corrélation significative entre (P) et (Q) n'a été trouvée, ce qui conduira à une étude de modélisation

intéressante, étant donné que la régression linéaire simple n'est pas appropriée pour une évaluation directe et une estimation efficace de (Q).

Afin de prédire le débit journalier (Q) à l'instant (t), plusieurs retards de (Q) et (P) ont été sélectionnés. En outre, pour le meilleur retard approprié, la fonction d'autocorrélation (ACF), la fonction d'autocorrélation partielle (PACF) et la fonction de corrélation croisée (XCF) pour chaque station ont été calculées, comme le montre la **figure V.12**. Par conséquent, dans la présente étude, nous avons appliqué une approche pour prédire le débit journalier en utilisant uniquement P et Q, étant donné que seules les données P sont disponibles en continu sans enregistrement complet. Conformément à la **figure V.12**, des décalages de deux temps ont été sélectionnés et utilisés comme variables d'entrée, c'est-à-dire le débit journalier mesuré à (t-1), (t-2), les précipitations journalières mesurées à (t-1), (t-2) ; tandis que la variable de sortie commence au temps (t), c'est-à-dire Q au temps (t). Deux scénarios ont été étudiés dans ce chapitre : le premier est une stratégie de modélisation autonome pour laquelle des modèles ML ont été développés en utilisant plusieurs temps de décalage de Q et P sans prétraitement, tandis que le second scénario utilise cinq techniques de décomposition du signal soutenues par la transformée en ondelettes à recouvrement maximal (MODWT) pour chaque ondelette mère. Six (6) combinaisons des variables climatiques ont été développé. Comme le montre **le tableau VII.1** qui représente les variables d'entrée après transformation on ondelettes au pas de temps journalier.

Tableau VII.1 Combinaisons des variables d'entrées de différents modèles pour la station de Ponteba et Bir Ouled Tahar dans le pas journalier

N Modèle	Variables d'entrée pour les Modèles autonomes	Variables d'entrée pour les modèles hybrides	Sortie
1	$Q(t-2), Q(t-1), P(t-2), P(t-1), P(t)$	$Q_{t-2}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $Q_{t-1}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_{t-2}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_{t-1}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_t(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$	$Q(t)$
2	$P(t-2), P(t-1), P(t)$	$P_{t-2}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_{t-1}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_t(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$	$Q(t)$
3	$Q(t-2), Q(t-1), P(t)$	$Q_{t-2}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $Q_{t-1}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_t(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$	$Q(t)$
4	$P(t-2), P(t-1)$	$P_{t-2}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $P_{t-1}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$	$Q(t)$
5	$Q(t-2), Q(t-1)$	$Q_{t-2}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$ $Q_{t-1}(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$	$Q(t)$
6	$P(t)$	$P_t(MRA1, MRA2, MRA3, MRA4, MRA5, MRA6, MRA7, RES),$	$Q(t)$

VII.1.2 Développement des modèles d'apprentissage automatique

Dans la présente étude, quatre modèles ML ont été comparés et leurs performances ont été évaluées en termes de prédiction de débits journaliers (Q). Les quatre modèles étaient respectivement le GPR, LSTM, GRNN et MLPNN. Les performances des modèles ont été évaluées selon deux

scénarios : (i) modèles simples sans décomposition (autonome), c'est-à-dire GPR, LSTM, GRNN, MLPNN, (ii) modèles hybrides avec algorithme MODWT, c'est-à-dire MODWT-GPR, MODWT-LSTM, MODWT-GRNN, MODWT-MLPNN et l'ensemble de ces modèles sont développés dans Matlab R2020a.

VII.1.2.1 Développement des modèles ML autonome

Pour les modèles MLPNN, GPR et GRNN, les mêmes paramètres ont été utilisés pour le développement du chapitre précédent et qui a conduit aux meilleures performances, tandis que pour le modèle LSTM, le nombre de nœuds de la couche cachée est l'un des paramètres à déterminer dans le modèle. D'après les expériences, le modèle avec 10 nœuds offre les meilleures performances. L'optimiseur, la taille du lot et les époques sont également des paramètres qui influencent les performances du modèle. Le choix de l'optimiseur influence la façon dont la fonction de perte est minimisée, et donc la façon dont le modèle aboutit au résultat final. Les choix standard incluent Momentum, Adagrad, RMSProp, Adam, etc. Par expérience, l'optimiseur Momentum est choisi. La taille des lots affecte la quantité de données traitées à la fois. Grâce aux lots, le modèle est mis à jour plusieurs fois avant de traiter l'ensemble des données, ce qui affecte la dynamique du processus. Étant donné qu'une petite taille de lot ralentit considérablement la vitesse d'apprentissage et qu'une grande taille de lot entraîne un sur-ajustement, la taille de lot est fixée à 50 dans cette expérience. Les époques sont les périodes pendant lesquelles le modèle s'exécute sur l'ensemble des données. Lorsque les époques sont d'environ 40, la perte de l'ensemble de test est la plus faible. Les périodes sont donc fixées à 40 dans cette expérience.

VII.1.2.2 Développement des modèles hybride ML avec algorithme MODWT

Les modèles GPR, LSTM et GRNN et MLPNN ont été intégrés à la MODWT dans cette étude pour former des modèles d'ondelettes hybrides. MODWT décompose les signaux d'entrée originaux (données quotidiennes sur le débit des cours d'eau et les précipitations) en sous-signaux (composantes de détail et d'approximation). Les sous-signaux sont ensuite utilisés comme entrées dans les modèles simples, GPR, LSTM, GRNN et MLPNN.

Comme il est mentionné dans le [chapitre III](#), Les approches de modélisation de l'état des cours d'eau basées sur le MODWT consistent en un algorithme en trois étapes, (i) Décomposition des signaux d'entrée originaux en sous-signaux (composantes de détail et d'approximation) à l'aide de

MODWT, (ii) Sélection des entrées efficaces parmi les sous-signaux, (iii) Apprentissage et test de modèles uniques, GPR, LSTM, GRNN et MLPNN, à l'aide des entrées effectives. Dans l'étape (i), Le MODWT peut être utilisé pour effectuer une analyse multi résolution (MRA) d'un signal. Pour décomposer les signaux d'entrée à l'aide de MODWT, le niveau de décomposition 7 a été déterminé sur la base de l'équation (III.38) et sélectionner le meilleur ordre N de chaque famille d'ondelettes (coif1, db3, sym4, haar ou db1et fk8) a été sélectionné pour les obtenir les meilleures résultats décompositions des séries pluies-débits (figure VII.2).

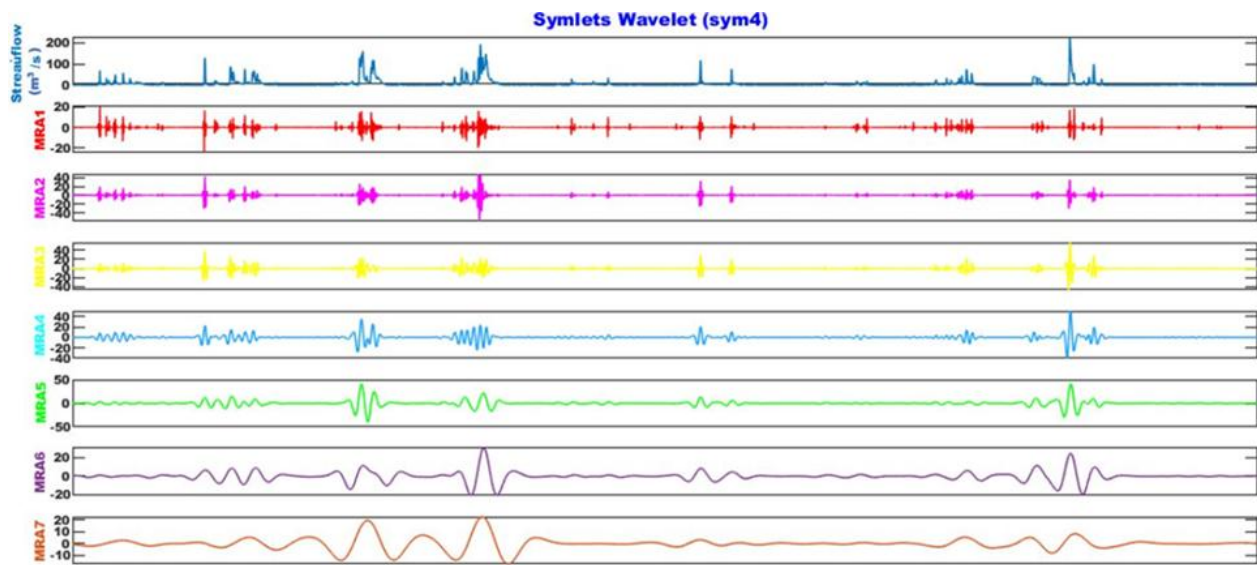


Figure VII.2 Exemple de décomposition en ondelettes MODWT (sym4) de la série de débit d'ordre 7 pour le pas journalier dans la station de Bir Ould Taher

À l'étape (ii), les entrées efficaces des sous-signaux (débit et pluie) ont été choisies, comme indiqué dans le tableau VII.1. Dans l'étape (iii), l'apprentissage et la validation ont été faits à partir des modèles ML autonomes (GPR, LSTM, GRNN et MLPNN) avec les sous signaux des cinq familles d'ondelette MODWT. Les étapes globales et l'Organigramme des modèles hybrides ML et MODWT sont illustrés dans la figure VII.3.

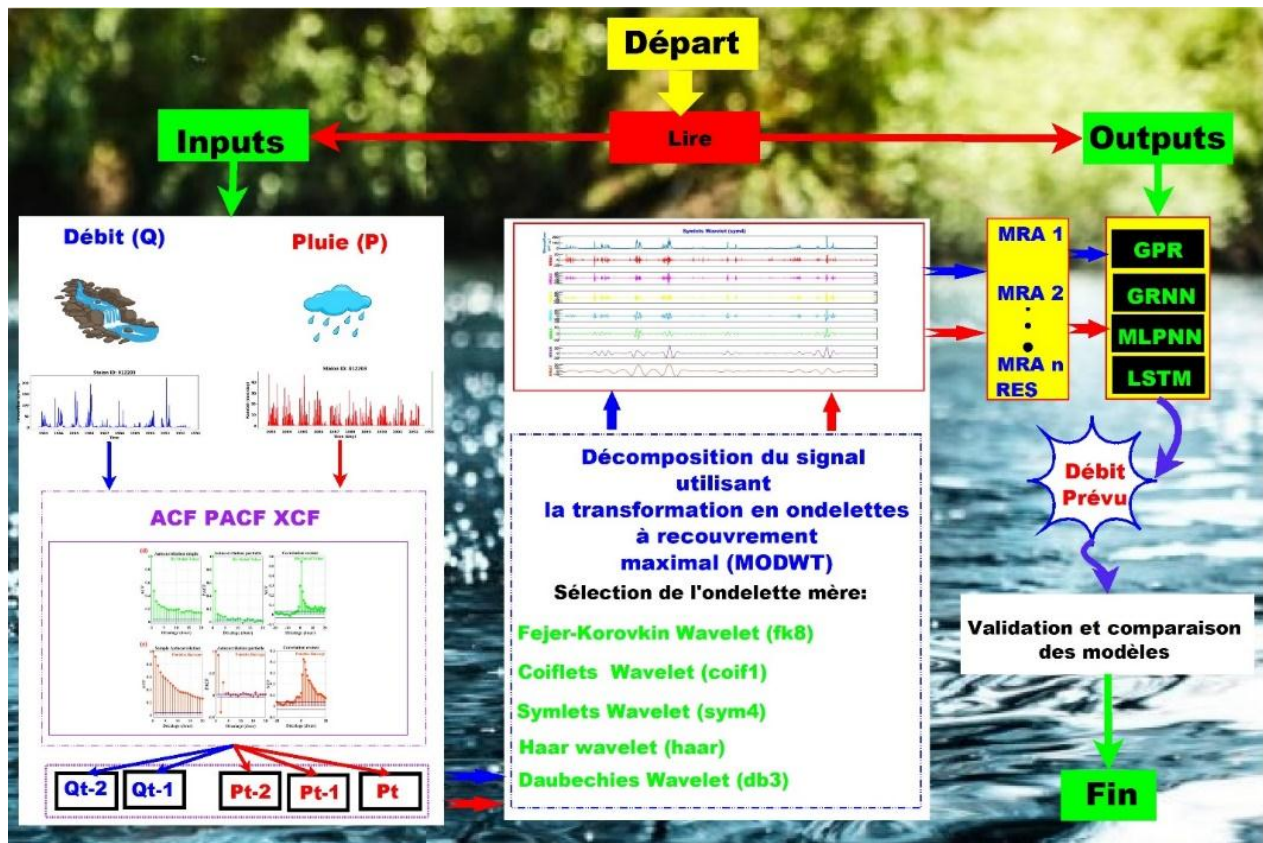


Figure VII.3 Organigramme des modèles hybrides ML-MODWT proposés

VII.2 Résultats de la modélisation et analyse comparative

L'utilité des modèles ML autonomes et hybrides anticipés a été évaluée en comparant les valeurs de débit prévues et observées à chaque station. L'efficacité des modèles pour prédire le débit journalier a été examinée à l'aide des critères diverses mesures statistiques (RMSE, MAE, R et NSE) et d'une évaluation graphique des données via le diagramme de dispersion, diagramme en boîte, diagramme en violon, diagramme de Taylor, diagramme radar et courbes de comparaison entre les valeurs mesurées et les valeurs calculées ont été produites.

VII.2.1 Résultats de prédiction des débits journalières dans les deux stations d'étude

Dans la section suivante, nous présentons et discutons tous les résultats obtenus selon les deux scénarios de modélisation pour la prévision des débits journaliers des cours d'eau en utilisant les modèles ML autonomes et hybrides. Les résultats numériques sont présentés dans les **tableaux VII.2** au **tableau VII.11**, et les résultats de l'algorithme supérieur dans deux phases (apprentissage et validation) sont indiqués en gras.

VII.2.1.1 Station de Bir Ouled Taher

Dans la station de Bir Ould Taher, comme le montre la **figure VII.4 (a)** qui résume les valeurs optimales de R, NSES, RMSE et MAE pour les meilleurs modèles ML dans la prévision des débits journalier, qui sont indiqués en gras dans **le tableau VII.2**, pendant la phase d'apprentissage, les résultats indiquent que le modèle GPR3 est le plus performant que les modèles LSTM1, GRNN1 et MLPNN1. Le modèle GPR3 a donné des valeurs R et NSE les plus élevées et les valeurs RMSE et MAE plus faibles ≈ 0.993 et ≈ 0.986 , $\approx 0.143 \text{ m}^3/\text{s}$ et $\approx 0.032 \text{ m}^3/\text{s}$, respectivement. Si l'on examine le modèle GPR3, LSTM1, GRNN1 et MLPNN1, qui ont donné les meilleurs résultats avec la troisième et la première combinaison des variables d'entrées (**tableau VII.1**), où la première combinaison assemblée les entrées de précipitations (P), de précipitations du jour précédent (Pt-1), de précipitations deux jours avant (Pt-2) et de débit du jour précédent (Qt-1) et de débit deux jours avant (Qt-2) et la troisième combinaison assemblée les entrées de précipitations (P), et de débit du jour précédent (Qt-1) et de débit deux jours avant (Qt-2). L'analyse des corrélations de Pearson (R) et le critère de Nash-Sutcliffe (NSE) indique que les précipitations sont le principal facteur prédictif du débit dans le bassin versant d'oued Rouina Zeddine. Cela ressort également des résultats de R et NSE pendant de la formation pour les modèles GPR4, LSTM4, GRNN4 et MLPNN4 lors de l'utilisation d'entrées contenant le débit un jour avant (Qt-1) et deux jours avant (Qt-2), alors qu'en l'absence de l'entrée de précipitation, nous constatons une diminution des performances. Au cours des phases de validation (**tableau VII.2, figure VII.4 (b)**). Le modèle LSTM1 a produit le plus haut niveau de précision en termes de $R \approx 0.805$, $NSE \approx 0.642$, $RMSE \approx 0.729 \text{ m}^3/\text{s}$ et $MAE \approx 0.225 \text{ m}^3/\text{s}$ pendant la période de validation, suivi par le modèle MLPNN1. De même, le modèle GPR1 donne des valeurs minimales de RMSE et MAE et des valeurs plus élevées de R et de NSE pour la période de validation. En outre, les résultats révèlent la faible performance du modèle GRNN1 dans la prédiction du débit journalier dans le bassin d'oued Rouina Zeddine. Toutefois, les quatre ML autonomes proposés dans cette étude n'ont pas donné de résultats satisfaisants au cours de la phase de validation.

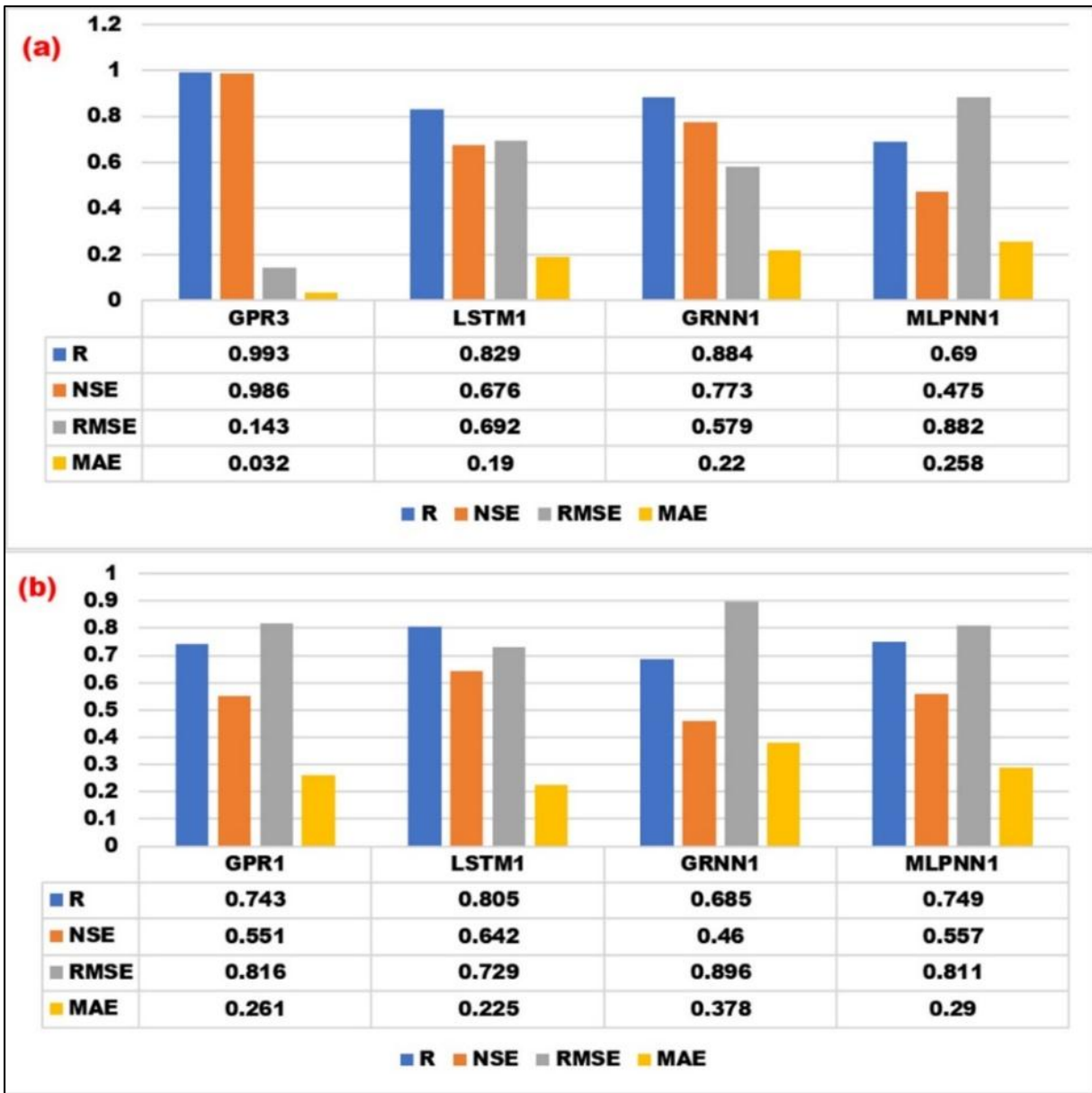


Figure VII.4 Graphique à barres des statistiques des critères de performance pour les meilleurs modèles ML autonomes pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation)

Tableau VII.2 Résultats de prévision des différents modèles ML autonomes dans le pas journalier pour la station de Bir Ouled Tahar

Modèles	Apprentissage				Validation			
	R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>GPR</i>								
GPR1	0.990	0.979	0.175	0.043	0.743	0.551	0.816	0.261
GPR2	0.959	0.920	0.345	0.160	0.571	0.317	1.007	0.575
GPR3	0.993	0.986	0.143	0.032	0.551	0.293	1.025	0.487
GPR4	0.902	0.811	0.529	0.224	0.539	0.267	1.043	0.614
GPR5	0.922	0.849	0.473	0.126	0.363	0.080	1.266	0.861
GPR6	0.645	0.413	0.932	0.378	0.125	0.190	1.329	0.557
<i>LSTM</i>								
LSTM1	0.829	0.676	0.692	0.190	0.805	0.642	0.729	0.225
LSTM2	0.764	0.580	0.788	0.248	0.741	0.485	0.875	0.474
LSTM3	0.784	0.610	0.759	0.229	0.714	0.507	0.856	0.311
LSTM4	0.705	0.496	0.864	0.329	0.695	0.477	0.881	0.411
LSTM5	0.603	0.363	0.971	0.269	0.615	0.378	0.961	0.304
LSTM6	0.689	0.468	0.887	0.265	0.729	0.502	0.860	0.333
<i>GRNN</i>								
GRNN1	0.884	0.773	0.579	0.220	0.685	0.460	0.896	0.378
GRNN2	0.719	0.505	0.856	0.341	0.580	0.313	1.010	0.470
GRNN3	0.706	0.470	0.886	0.292	0.583	0.323	1.003	0.410
GRNN4	0.641	0.400	0.942	0.370	0.592	0.307	1.014	0.480
GRNN5	0.567	0.297	1.020	0.331	0.486	0.233	1.067	0.432
GRNN6	0.291	0.082	1.166	0.438	0.310	0.094	1.159	0.559
<i>MLPNN</i>								
MLPNN1	0.690	0.475	0.882	0.258	0.749	0.557	0.811	0.290
MLPNN2	0.624	0.389	0.951	0.376	0.666	0.441	0.911	0.464
MLPNN3	0.611	0.371	0.965	0.269	0.640	0.409	0.937	0.282
MLPNN4	0.589	0.346	0.984	0.384	0.637	0.393	0.949	0.470
MLPNN5	0.559	0.312	1.009	0.278	0.616	0.378	0.960	0.301
MLPNN6	0.319	0.101	1.153	0.444	0.309	0.094	1.159	0.575

Dans le deuxième scénario de cette étude, on a tenté d'améliorer la performance du modèle à l'aide de décompositions de signaux de différentes familles de l'ondelette MODWT : coif1, db3, sym4, haar et fk8, (tableau VII.3, VII.4, VII.5 et VII.6). Il est clair que pour tous les modèles hybrides en particulier lors de l'utilisation de MODWT (sym4), les précisions prédictives semblent être excellentes, pour lesquelles les mesures d'erreurs, c'est-à-dire le RMSE et le MAE, ont été considérablement et les indices d'ajustement, c'est-à-dire le R et le NSE, ont été notablement améliorés.

Comme nous l'avons mentionné précédemment, le modèle GPR3 est le meilleur modèle en termes de performances numérique par rapport aux autres modèles GPR en phase d'apprentissage, où ses performances en termes de RMSE et MAE étaient respectivement $\approx 0.43 \text{ m}^3/\text{s}$ et $\approx 0.16 \text{ m}^3/\text{s}$. De plus, ce modèle est le meilleur parmi les modèles LSTM1, GRNN1 et MLPNN1. D'après la figure

VII.5 (a) qui résume les résultats du modèle hybride MODWT-GPR le plus performant (tableau VII.3) et les résultats du meilleur modèle GPR autonome (tableau VII.2), il ressort que les valeurs de RMSE, MAE obtenues par les modèles GPR3 ont été améliorées en utilisant les modèles basés sur MODWT-GPR, cette amélioration a entraîné une réduction maximale de RMSE et MAE de GPR dans la phase d'apprentissage à environ $\approx 0,004 \text{ m}^3/\text{s}$ et $\approx 0,001 \text{ m}^3/\text{s}$, respectivement, Ces valeurs ont été obtenues en utilisant le modèle MODWT-GPR3(haar). Bien qu'un modèle MODWT-GPR3(haar) ait été le meilleur lors de la phase de d'apprentissage, Dans la phase de validation le modèle MODWT-GPR3 (sym4) avec un RMSE $\approx 0.171 \text{ m}^3/\text{s}$ et MAE $\approx 0.117 \text{ m}^3/\text{s}$ a montré une supériorité dans le processus d'amélioration de la performance des modèles autonomes GPR (figure VII.5 (b)), Qui sont représentés par le modèle GPR3 comme le meilleur modèle (RMSE $\approx 0.816 \text{ m}^3/\text{s}$, MAE $\approx 0.261 \text{ m}^3/\text{s}$), De plus le modèle MODWT-GPR3 (sym4) surpasse le modèle LSTM1 (RMSE $\approx 0.729 \text{ m}^3/\text{s}$, MAE $\approx 0.225 \text{ m}^3/\text{s}$) ce qui représente les quatre modèles ML autonomes comme le meilleur modèle pendant la phase validation. Le modèle MODWT-GPR (sym4) est considéré comme l'un des meilleurs modèles hybrides en phase d'apprentissage et de validation car il a maintenu ses performances dans les deux phases de simulation des débits journaliers dans la station de Bir Ouled Tahar par les valeurs minimales de RMSE et MAE et de valeurs plus élevées de R et de NSE.

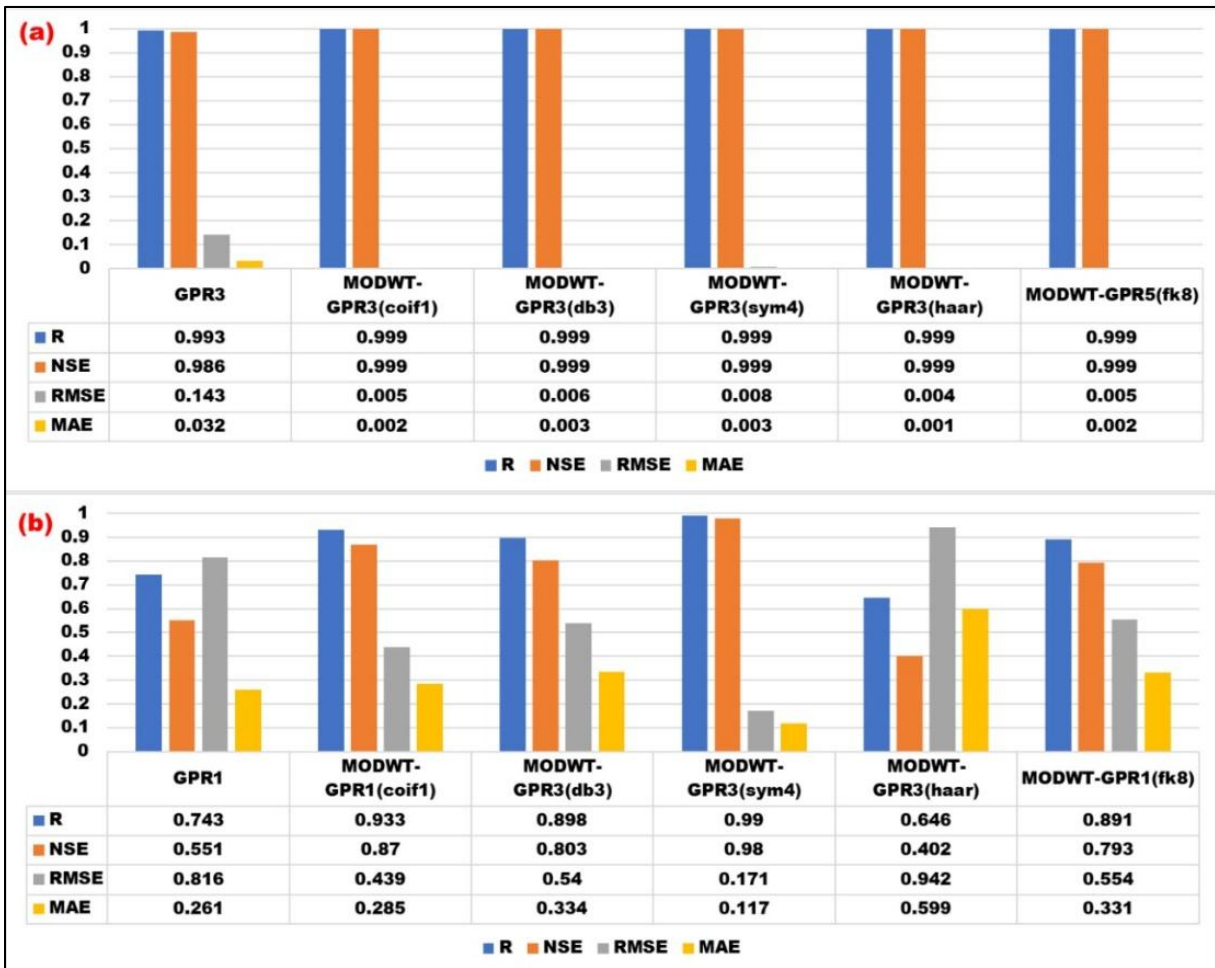


Figure VII.5 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GPR par rapport au modèle GPR optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation)

Tableau VII.3 Résultats de prévision des modèles hybrides GPR-MODWT dans le pas journalier pour la station de Bir Ouled Tahar

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT_GPR (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-GPR1	0.999	0.999	0.005	0.002	0.933	0.870	0.439	0.285
	MODWT-GPR2	0.997	0.993	0.099	0.040	0.549	0.286	1.029	0.560
	MODWT-GPR3	0.999	0.999	0.005	0.002	0.926	0.854	0.465	0.326
	MODWT-GPR4	0.996	0.993	0.103	0.042	0.516	0.247	1.057	0.570
	MODWT-GPR5	0.999	0.999	0.004	0.002	0.883	0.767	0.588	0.404
	MODWT-GPR6	0.997	0.994	0.095	0.035	0.426	0.143	1.128	0.596
<i>MODWT_GPR (db3)</i>									
Daubechies Wavelet (db3)	MODWT-GPR1	0.999	0.999	0.008	0.003	0.886	0.781	0.570	0.384
	MODWT-GPR2	0.997	0.994	0.093	0.035	0.518	0.226	1.072	0.590
	MODWT-GPR3	0.999	0.999	0.006	0.003	0.898	0.803	0.540	0.334
	MODWT-GPR4	0.997	0.994	0.094	0.034	0.530	0.242	1.061	0.602
	MODWT-GPR5	0.999	0.999	0.005	0.002	0.812	0.649	0.722	0.454
	MODWT-GPR6	0.997	0.994	0.097	0.038	0.471	0.140	1.130	0.671
<i>MODWT_GPR (sym4) s</i>									
Symlets Wavelet (sym4)	MODWT-GPR1	0.999	0.999	0.009	0.004	0.990	0.980	0.174	0.118
	MODWT-GPR2	0.997	0.994	0.098	0.036	0.593	0.334	0.994	0.479
	MODWT-GPR3	0.999	0.999	0.008	0.003	0.990	0.980	0.171	0.117
	MODWT-GPR4	0.997	0.993	0.101	0.037	0.578	0.311	1.011	0.488
	MODWT-GPR5	0.999	0.999	0.007	0.003	0.988	0.976	0.187	0.118
	MODWT-GPR6	0.997	0.994	0.096	0.034	0.390	0.076	1.171	0.607
<i>MODWT_GPR (haar)</i>									
Haar wavelet (haar)	MODWT-GPR1	0.999	0.999	0.004	0.001	0.605	0.313	1.010	0.645
	MODWT-GPR2	0.999	0.999	0.022	0.006	0.173	-0.027	1.235	0.640
	MODWT-GPR3	0.999	0.999	0.004	0.001	0.646	0.402	0.942	0.599
	MODWT-GPR4	0.999	0.999	0.028	0.008	0.154	-0.053	1.250	0.644
	MODWT-GPR5	0.999	0.999	0.004	0.001	0.623	0.362	0.973	0.643
	MODWT-GPR6	0.999	0.999	0.040	0.010	0.140	-0.009	1.224	0.633
<i>MODWT_GPR (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-GPR1	0.998	0.995	0.086	0.030	0.891	0.793	0.554	0.331
	MODWT-GPR2	0.996	0.991	0.114	0.039	0.556	0.297	1.021	0.485
	MODWT-GPR3	0.999	0.999	0.011	0.003	0.792	0.627	0.744	0.373
	MODWT-GPR4	0.995	0.991	0.117	0.037	0.535	0.274	1.038	0.497
	MODWT-GPR5	0.999	0.999	0.005	0.002	0.718	0.515	0.848	0.457
	MODWT-GPR6	0.996	0.993	0.105	0.037	0.329	0.019	1.207	0.596

Dans l'hybridation des modèles LSTM avec les différentes familles MODWT, généralement selon le [tableau VII.4](#), les modèles hybrides MODWT-LSTM montrent une amélioration en termes de critères de performance numérique R, NSE, RMSE et MAE dans les phases d'apprentissage et de validation. Le modèle MODWT-LSTM3 (sym4) de la [figure VII.6](#) montre une supériorité dans les deux phases en termes de RMSE et de MAE. Dans la phase d'apprentissage ([figure VII.6 \(a\)](#)), ces valeurs sont respectivement de $\approx 0,356 \text{ m}^3/\text{s}$ et $\approx 0,113 \text{ m}^3/\text{s}$ et dans la phase de validation, de $\approx 0,347 \text{ m}^3/\text{s}$ et $\approx 0,381 \text{ m}^3/\text{s}$. Ces valeurs de RMSE et de MAE du modèle MODWT-LSTM3 (sym4)

comparées à celles données par le modèle autonome LSTM1 dans la phase d'apprentissage (RMSE $\approx 0,692 \text{ m}^3/\text{s}$, MAE $\approx 0,190 \text{ m}^3/\text{s}$) et dans la phase de validation (RMSE $\approx 0,729 \text{ m}^3/\text{s}$, MAE $\approx 0,325 \text{ m}^3/\text{s}$) montrent une réduction considérable des valeurs RMSE et MAE, et il est utile de rappeler que le modèle LSTM1 était le meilleur modèle autonome en termes de performance numérique dans la phase de validation (figure VII.6 (b)).

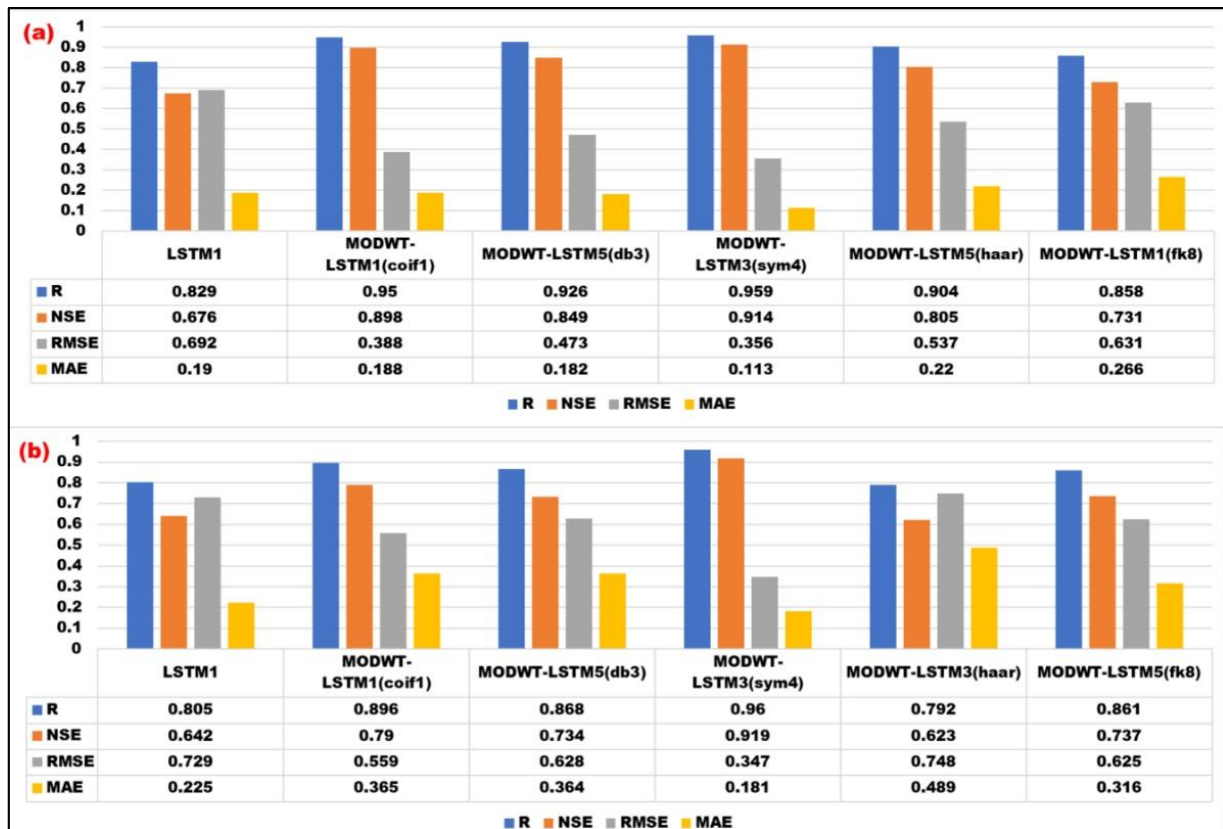


Figure VII.6 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-LSTM par rapport au modèle LSTM optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation)

Tableau VII.4 Résultats de prévision des modèles hybrides LSTM-MODWT dans le pas journalier pour la station de Bir Ouled Tahar

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT LSTM (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-LSTM1	0.950	0.898	0.388	0.188	0.896	0.790	0.559	0.365
	MODWT-LSTM2	0.817	0.660	0.710	0.279	0.650	0.415	0.932	0.451
	MODWT-LSTM3	0.932	0.861	0.453	0.179	0.852	0.725	0.638	0.361
	MODWT-LSTM4	0.780	0.603	0.767	0.307	0.622	0.379	0.960	0.440
	MODWT-LSTM5	0.945	0.885	0.412	0.170	0.848	0.717	0.648	0.374
	MODWT-LSTM6	0.750	0.556	0.811	0.331	0.630	0.378	0.961	0.459
<i>MODWT LSTM (db3)</i>									
Daubechies Wavelet (db3)	MODWT-LSTM1	0.885	0.772	0.581	0.231	0.825	0.676	0.694	0.418
	MODWT-LSTM2	0.737	0.537	0.828	0.327	0.546	0.290	1.027	0.535
	MODWT-LSTM3	0.910	0.821	0.515	0.209	0.837	0.699	0.669	0.392
	MODWT-LSTM4	0.758	0.570	0.798	0.317	0.363	0.082	1.267	0.642
	MODWT-LSTM5	0.926	0.849	0.473	0.182	0.868	0.734	0.628	0.364
	MODWT-LSTM6	0.782	0.607	0.762	0.303	0.600	0.348	0.984	0.461
<i>MODWT LSTM (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-LSTM1	0.949	0.892	0.399	0.137	0.951	0.901	0.383	0.220
	MODWT-LSTM2	0.837	0.696	0.671	0.284	0.658	0.367	0.969	0.497
	MODWT-LSTM3	0.959	0.914	0.356	0.113	0.960	0.919	0.347	0.181
	MODWT-LSTM4	0.822	0.672	0.697	0.290	0.758	0.575	0.795	0.389
	MODWT-LSTM5	0.946	0.883	0.417	0.111	0.966	0.929	0.325	0.182
	MODWT-LSTM6	0.748	0.554	0.812	0.325	0.688	0.473	0.885	0.443
<i>MODWT LSTM (haar)</i>									
Haar wavelet (haar)	MODWT-LSTM1	0.888	0.780	0.570	0.259	0.784	0.596	0.774	0.479
	MODWT-LSTM2	0.672	0.451	0.902	0.362	0.233	-0.154	1.309	0.747
	MODWT-LSTM3	0.898	0.800	0.544	0.264	0.792	0.623	0.748	0.489
	MODWT-LSTM4	0.589	0.346	0.984	0.385	0.363	0.120	1.143	0.581
	MODWT-LSTM5	0.904	0.805	0.537	0.220	0.724	0.412	0.934	0.567
	MODWT-LSTM6	0.633	0.399	0.943	0.355	0.247	0.037	1.196	0.575
<i>MODWT LSTM (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-LSTM1	0.858	0.731	0.631	0.266	0.757	0.571	0.798	0.403
	MODWT-LSTM2	0.755	0.564	0.803	0.304	0.644	0.410	0.936	0.462
	MODWT-LSTM3	0.852	0.719	0.645	0.255	0.810	0.654	0.717	0.367
	MODWT-LSTM4	0.753	0.563	0.805	0.324	0.537	0.271	1.040	0.486
	MODWT-LSTM5	0.858	0.727	0.635	0.227	0.861	0.737	0.625	0.316
	MODWT-LSTM6	0.726	0.526	0.838	0.333	0.621	0.383	0.957	0.482

Afin d'obtenir des résultats d'apprentissage et de validation satisfaisants dans le processus de prévision du débit journalier dans la station de Bir Ouled Tahar, le processus d'amélioration des performances du modèle GRNN indépendant a été réalisé en le combinant avec le MODWT (tableau VII.5, figure VII.7). Les modèles hybrides GRNN-MODWT donnent des bons résultats en termes de critères de performance numérique R, NSE, RMSE et MAE dans les phases d'apprentissage, Notez que les résultats du modèle GPR3 obtenus au cours du processus d'apprentissage étaient optimaux en termes de paramètres de performance numériques suivants

$R \approx 0.993$, $NSE \approx 0.986$, $RMSE \approx 0.143$ m³/s et $MAE \approx 0.143$ m³/s proposés au cours de cette étude. Cependant, quatre des cinq meilleurs modèles hybrides MODWT-GRNN1 (coif1), MODWT-GRNN1 (db3), MODWT-GRNN1 (sym4) et MODWT-GRNN1 (fk8) ont montré des résultats qui dépassaient les performances du modèle indépendant, tandis que le modèle hybride restant MODWT-GRNN (sym4) a donné des résultats sont beaucoup généralement plus proches ou égaux à ceux fournis par le modèle GPR3 autonome. Dans la phase de validation, trois des cinq meilleurs modèles hybrides MODWT-GRNN3 (coif1), MODWT-GRNN3 (db3) et MODWT-GRNN5 (sym4) ont montré une supériorité en termes de R, NSE, RMSE et MAE par rapport au modèle GRNN3. De plus, le modèle MODWT-GRNN5 (sym4) avec $R \approx 0.92$, $NSE \approx 0.82$, $RMSE \approx 0,52$ m³/s, $MAE \approx 0.27$ m³/s a surpassé le modèle autonome LSTM1 en termes des quatre critères numériques proposé dans cette étude.

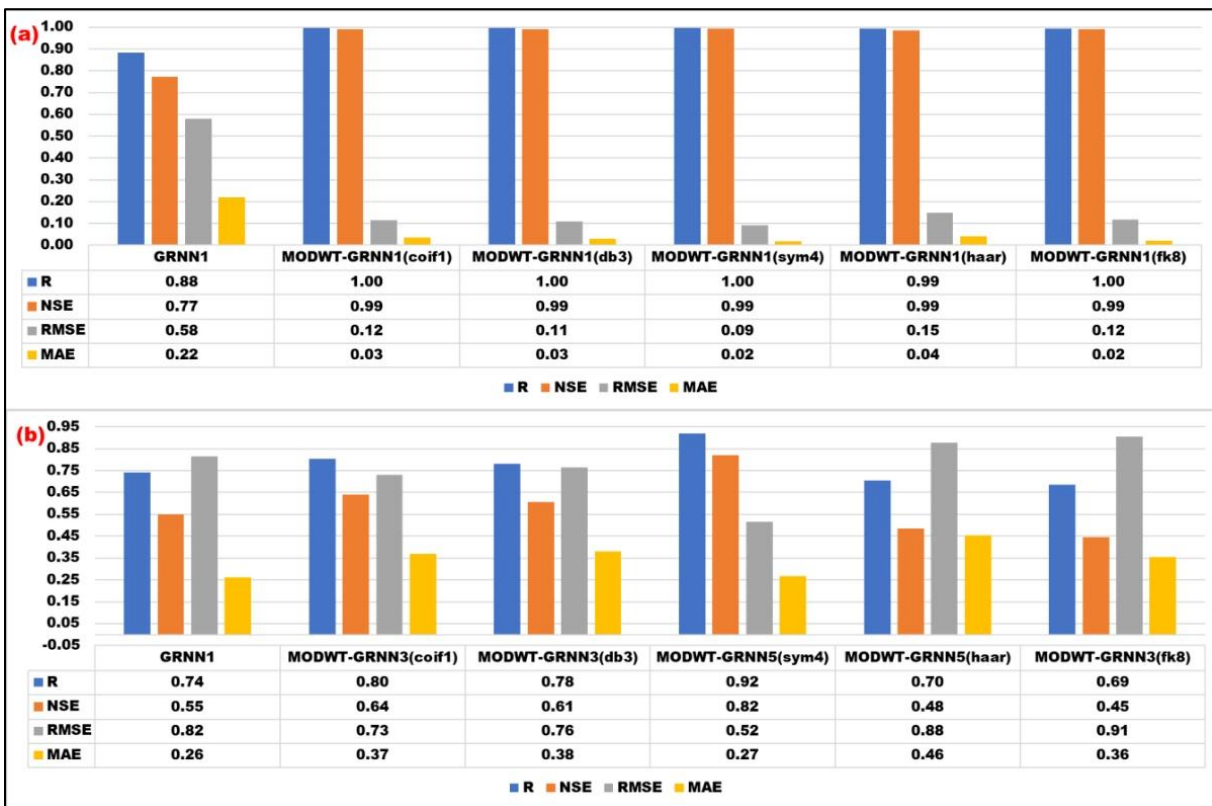


Figure VII.7 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GRNN par rapport au modèle GRNN optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation)

Tableau VII.5 Résultats de prévision des modèles hybrides GRNN-MODWT dans le pas journalier pour la station de Bir Ouled Tahar

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT GRNN (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-GRNN1	0.996	0.991	0.115	0.034	0.767	0.583	0.787	0.377
	MODWT-GRNN2	0.982	0.964	0.229	0.086	0.382	0.121	1.142	0.464
	MODWT-GRNN3	0.988	0.975	0.191	0.071	0.803	0.640	0.731	0.368
	MODWT-GRNN4	0.966	0.929	0.324	0.135	0.406	0.155	1.120	0.460
	MODWT-GRNN5	0.964	0.927	0.329	0.141	0.803	0.635	0.736	0.388
	MODWT-GRNN6	0.796	0.606	0.764	0.280	0.391	0.107	1.151	0.489
<i>MODWT GRNN (db3)</i>									
Daubechies Wavelet (db3)	MODWT-GRNN1	0.996	0.992	0.109	0.030	0.678	0.446	0.907	0.421
	MODWT-GRNN2	0.984	0.967	0.222	0.088	0.295	-0.005	1.221	0.511
	MODWT-GRNN3	0.989	0.978	0.182	0.063	0.782	0.607	0.764	0.381
	MODWT-GRNN4	0.968	0.933	0.314	0.137	0.310	0.055	1.184	0.495
	MODWT-GRNN5	0.968	0.935	0.310	0.129	0.772	0.591	0.779	0.398
	MODWT-GRNN6	0.857	0.702	0.664	0.270	0.480	0.221	1.075	0.488
<i>MODWT GRNN (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-GRNN1	0.997	0.994	0.093	0.017	0.866	0.707	0.660	0.275
	MODWT-GRNN2	0.988	0.976	0.190	0.069	0.541	0.284	1.031	0.433
	MODWT-GRNN3	0.995	0.990	0.121	0.034	0.910	0.805	0.538	0.257
	MODWT-GRNN4	0.970	0.939	0.300	0.114	0.574	0.325	1.001	0.422
	MODWT-GRNN5	0.981	0.962	0.239	0.076	0.920	0.821	0.515	0.267
	MODWT-GRNN6	0.825	0.656	0.713	0.264	0.343	0.004	1.216	0.468
<i>MODWT GRNN (haar)</i>									
Haar wavelet (haar)	MODWT-GRNN1	0.993	0.985	0.148	0.041	0.664	0.433	0.918	0.442
	MODWT-GRNN2	0.934	0.859	0.456	0.127	0.282	0.068	1.176	0.500
	MODWT-GRNN3	0.971	0.941	0.295	0.103	0.641	0.404	0.941	0.454
	MODWT-GRNN4	0.867	0.716	0.649	0.203	0.331	0.097	1.158	0.496
	MODWT-GRNN5	0.921	0.844	0.481	0.189	0.704	0.484	0.876	0.455
	MODWT-GRNN6	0.471	0.139	1.129	0.406	0.104	0.010	1.212	0.562
<i>MODWT GRNN (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-GRNN1	0.995	0.991	0.118	0.022	0.647	0.406	0.939	0.357
	MODWT-GRNN2	0.983	0.965	0.229	0.071	0.425	0.156	1.119	0.437
	MODWT-GRNN3	0.989	0.977	0.183	0.048	0.685	0.446	0.907	0.355
	MODWT-GRNN4	0.948	0.892	0.400	0.133	0.481	0.195	1.093	0.434
	MODWT-GRNN5	0.946	0.891	0.401	0.120	0.633	0.375	0.963	0.378
	MODWT-GRNN6	0.833	0.631	0.739	0.257	0.500	0.197	1.092	0.451

En ce qui concerne les résultats du dernier cas d'amélioration du deuxième scénario (tableau VII.6, figure VII.8), il est très clair que ce processus de prévision des débits journaliers à la station de Bir Ouled Tahar a été bien mené par hybridation du modèle MLPNN avec les différents algorithmes MODWT, en particulier en utilisant le modèle hybride MODWT-MLPNN5 (sym4), ceci aussi bien dans la phase d'apprentissage que dans la phase de validation. Dans la phase d'apprentissage, ce modèle avec $R \approx 0.991$, $NSE \approx 0.98$, $RMSE \approx 0.171 \text{ m}^3/\text{s}$ et $MAE \approx 0.056 \text{ m}^3/\text{s}$ a été plus performant que le modèle MLPNN1 et il a montré des performances très proches du modèle GPR3, Dans la

phase de validation, ce modèle a été plus performant que les modèles MLPNN1 et LSTM1, présentant les valeurs les plus élevées pour $R \approx 0.985$ et $NSE \approx 0.968$ et les valeurs les plus faibles pour $RMSE \approx 0.218 \text{ m}^3/\text{s}$ et $MAE \approx 0.128 \text{ m}^3/\text{s}$.

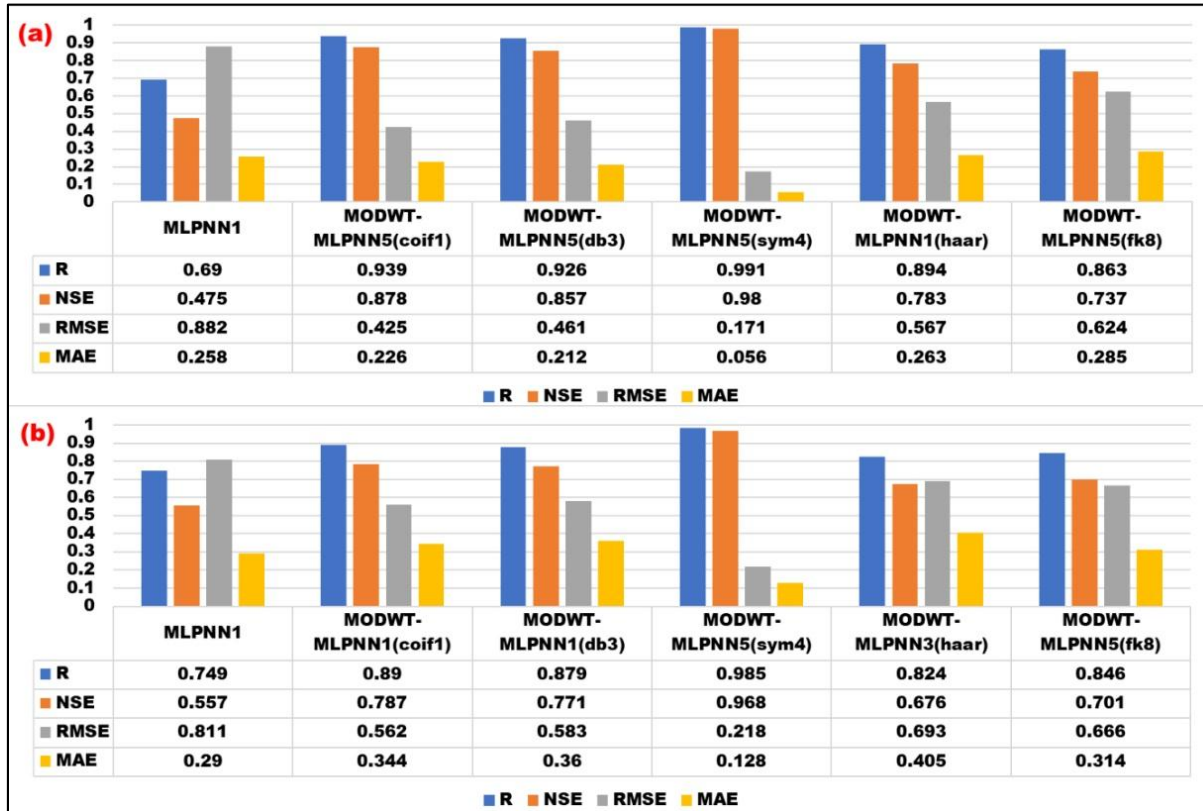


Figure VII.8 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-MLPNN par rapport au modèle MLPNN optimale pour la station de Bir Ouled Tahar (a : Apprentissage, b : Validation)

Tableau VII.6 Résultats de prévision des modèles hybrides MODWT-MLPNN dans le pas journalier pour la station de Bir Ouled Tahar

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT MLPNN (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-MLPNN1	0.923	0.817	0.521	0.235	0.890	0.787	0.562	0.344
	MODWT-MLPNN2	0.664	0.436	0.914	0.399	0.656	0.424	0.925	0.512
	MODWT-MLPNN3	0.931	0.851	0.469	0.220	0.892	0.794	0.553	0.354
	MODWT-MLPNN4	0.640	0.408	0.936	0.389	0.690	0.461	0.894	0.444
	MODWT-MLPNN5	0.939	0.878	0.425	0.226	0.887	0.786	0.563	0.348
	MODWT-MLPNN6	0.513	0.262	1.045	0.409	0.543	0.292	1.025	0.491
<i>MODWT MLPNN (db3)</i>									
Daubechies Wavelet (db3)	MODWT-MLPNN1	0.884	0.775	0.577	0.251	0.879	0.771	0.583	0.360
	MODWT-MLPNN2	0.636	0.401	0.942	0.402	0.597	0.355	0.978	0.524
	MODWT-MLPNN3	0.837	0.686	0.681	0.297	0.817	0.644	0.727	0.463
	MODWT-MLPNN4	0.653	0.423	0.924	0.379	0.511	0.254	1.052	0.525
	MODWT-MLPNN5	0.926	0.857	0.461	0.212	0.863	0.701	0.666	0.388
	MODWT-MLPNN6	0.515	0.238	1.062	0.392	0.531	0.251	1.054	0.494
<i>MODWT MLPNN (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-MLPNN1	0.984	0.960	0.245	0.068	0.984	0.967	0.221	0.138
	MODWT-MLPNN2	0.600	0.350	0.981	0.370	0.639	0.389	0.952	0.451
	MODWT-MLPNN3	0.952	0.904	0.377	0.100	0.972	0.945	0.285	0.147
	MODWT-MLPNN4	0.657	0.431	0.918	0.352	0.630	0.393	0.949	0.458
	MODWT-MLPNN5	0.991	0.980	0.171	0.056	0.985	0.968	0.218	0.128
	MODWT-MLPNN6	0.491	0.231	1.067	0.374	0.523	0.256	1.051	0.487
<i>MODWT MLPNN (haar)</i>									
Haar wavelet (haar)	MODWT-MLPNN1	0.894	0.783	0.567	0.263	0.794	0.628	0.743	0.444
	MODWT-MLPNN2	0.562	0.315	1.007	0.458	0.347	0.106	1.152	0.612
	MODWT-MLPNN3	0.888	0.776	0.576	0.268	0.824	0.676	0.693	0.405
	MODWT-MLPNN4	0.520	0.267	1.042	0.445	0.341	0.097	1.158	0.609
	MODWT-MLPNN5	0.873	0.750	0.608	0.267	0.821	0.674	0.695	0.388
	MODWT-MLPNN6	0.305	0.091	1.160	0.461	0.047	-0.056	1.252	0.617
<i>MODWT MLPNN (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-MLPNN1	0.810	0.639	0.731	0.272	0.836	0.690	0.679	0.330
	MODWT-MLPNN2	0.625	0.390	0.950	0.382	0.580	0.336	0.993	0.501
	MODWT-MLPNN3	0.775	0.598	0.772	0.273	0.827	0.677	0.692	0.339
	MODWT-MLPNN4	0.605	0.359	0.974	0.386	0.548	0.299	1.020	0.505
	MODWT-MLPNN5	0.863	0.737	0.624	0.285	0.846	0.701	0.666	0.314
	MODWT-MLPNN6	0.521	0.268	1.041	0.386	0.534	0.279	1.035	0.473

Les nuages de points pour les valeurs de débit journalier prévues et observées pour la station de Bir Ouled Taher au cours de la période de validation fournies par les meilleures modèles autonomes et hybrides sont présenté dans les figures VII.9, VII.10, respectivement. Selon ces illustrations, il est évident l'effet de l'opération de l'hybridation sur les modèles d'apprentissage automatiques autonomes car ces modèles avant l'hybridation sont plus dispersés, Comme il est clairement observé, la tendance linéaire des modèles hybrides MODWT-GPR3 (sym4), MODWT-LSTM3 (sym4), et MODWT-GRNN5 (sym4) et MODWT-MLPNN5 (sym4) est la plus proche de la ligne

$y=x$ par rapport à celles de GPR1, LSTM1, GRNN1 et MLPNN1. Ceci peut également être compris dans le modèle hybride MODWT-GPR3 (sym4) à partir des équations de la droite ajustée dans lesquelles les coefficients de pente (0,9974) et de biais (0,0018), respectivement, sont beaucoup plus proches de 1 et de 0 que ceux des autres modèles.

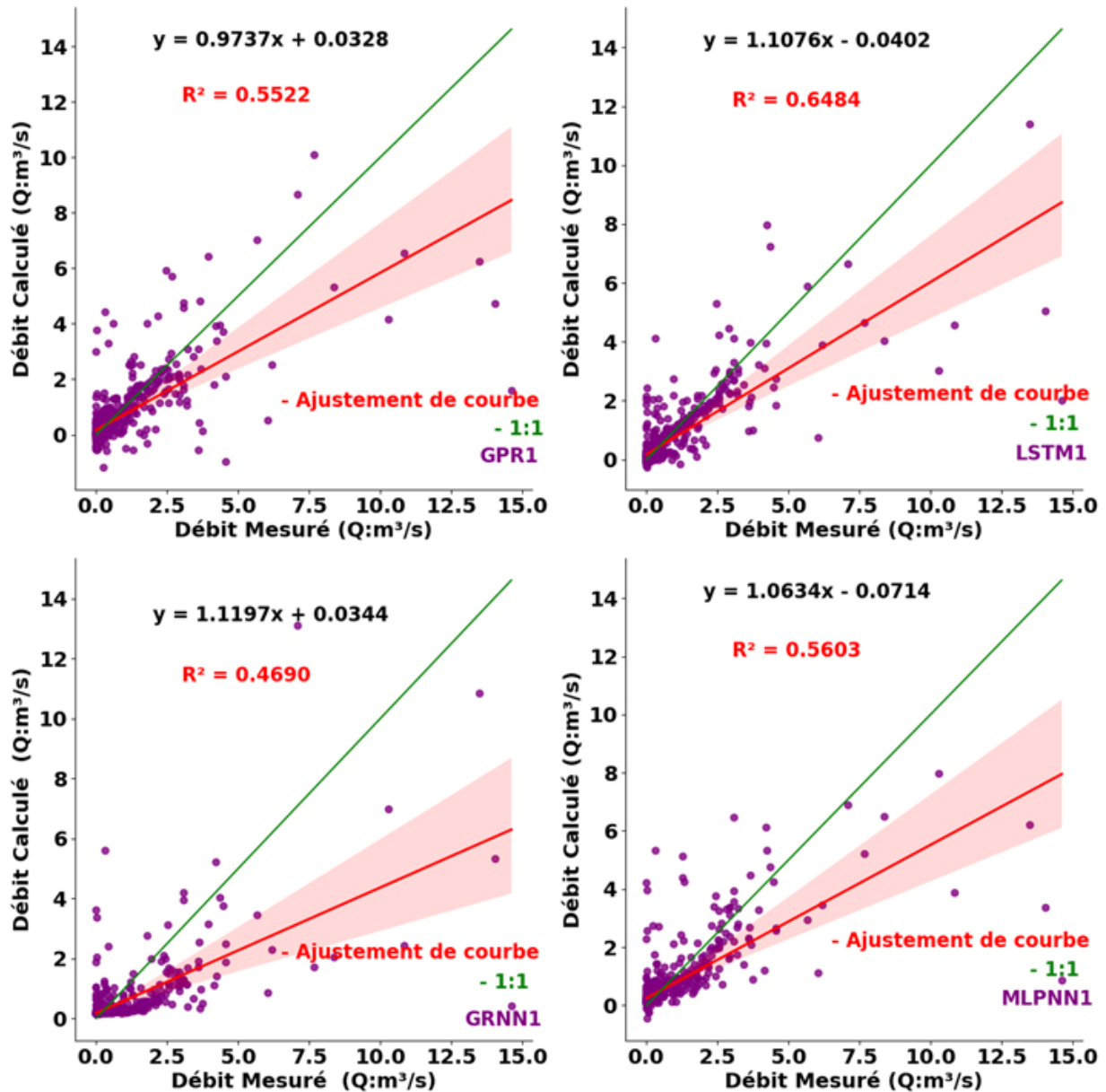


Figure VII.9 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles autonomes dans la station de Bir Ouled Tahar pendant la phase de validation

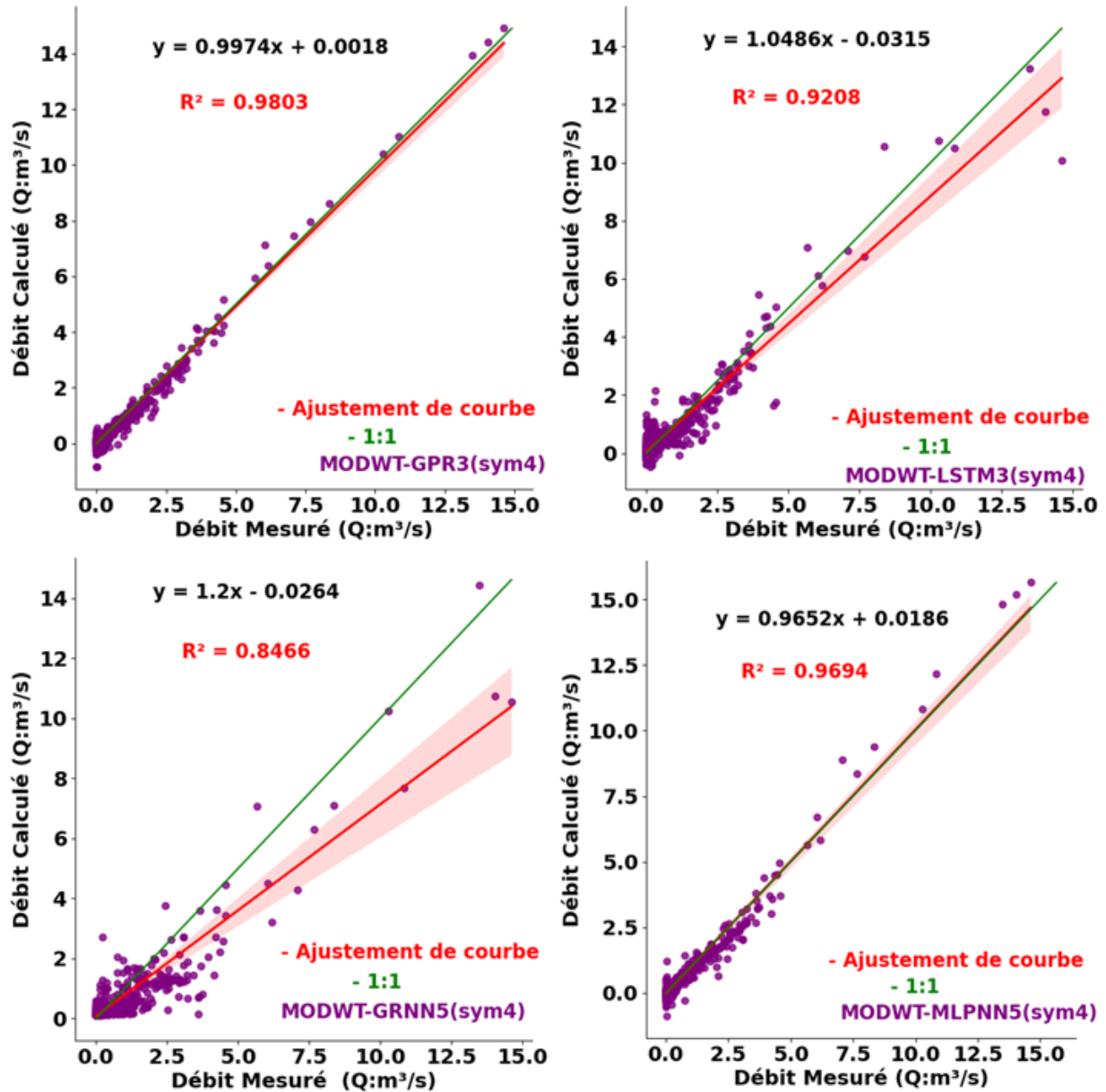


Figure VII.10 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles hybrides dans la station de Bir Ouled Tahar pendant la phase de validation

VII.2.1.2 Résultats récapitulatifs des différents modèles pour la station de Bir Ouled Tahar

- Dans les modèles ML autonomes les meilleures performances ([tableau VII.2](#)) dans les deux phases sont toujours obtenues avec les modèles utilisant la première et troisième combinaison des variables d'entrées ([tableau VII.1](#)). En l'absence de l'entrée de précipitation, nous constatons une diminution des performances. Au cours des phases de l'apprentissage et de validation, Cela ressort également selon les résultats obtenus des modèles GPR4, LSTM4, GRNN4 et MLPNN4.
- Le modèle GPR3 ($R \approx 0.993$, $NSE \approx 0.986$, $RMSE \approx 0.143 \text{ m}^3/\text{s}$ et $MAE \approx 0.032 \text{ m}^3/\text{s}$) a été le meilleur model autonomes dans la phase d'apprentissage, et le modèle LSTM1 ($R \approx 0.805$, $NSE \approx 0.642$, $RMSE \approx 0.729 \text{ m}^3/\text{s}$ et $MAE \approx 0.225 \text{ m}^3/\text{s}$) a été le meilleur model dans la phase de validation.
- Il ressort des tableaux [VII.3](#), [VII.4](#), [VII.5](#) et [VII.6](#) que les meilleurs résultats d'apprentissage hybridant les quatre algorithmes d'apprentissage profond et les cinq familles MODWT ont montré les meilleures performances dans le premier modèle, qui contient la décomposition du signal du débit $Qt-1$, $Qt-2$ avec la précipitation Pt , $Pt-1$, $Pt-2$. Le troisième modèle, contenant la décomposition des signaux $Qt-1$ et $Qt-2$ avec Pt , a également montré de bons résultats. Il est à noter que le cinquième modèle, contenant également la décomposition des signaux de débit $Qt-1$ et $Qt-2$, a donné des résultats satisfaisants dans la phase d'apprentissage.
- Dans l'hybridation des modèles GPR, LSTM, GRNN et MLPNN avec les différents familles de l'ondelette MODWT : coif1, db3, sym4, haar et fk8 au cours du processus de d'apprentissage, par rapport aux modèle optimal GPR3, les meilleurs résultats ont été fournis par les modèles hybrides MODWT-GPR3(coif1, db3, sym4, haar), MODWT-GPR5(fk8) et MODWT-GRNN1 (coif1, db3, sym4, fk8), Le modèle MODWT-GPR3(haar) était le meilleur en termes des valeurs de critères de performances suivantes $R \approx 0.999$, $NSE \approx 0.999$, $RMSE \approx 0.004 \text{ m}^3/\text{s}$ et $MAE \approx 0.001 \text{ m}^3/\text{s}$.
- Dans la phase de validation, l'hybridation des modèles ML autonomes avec le MODWT, par rapport aux modèle optimale LSTM1, les meilleurs résultats ont été fournis par les modèles hybrides MODWT-GPR3(sym4), MODWT-LSTM3 (sym4), MODWT-GRNN5 (sym4) et MODWT-MLPNN5 (sym4), Le modèle MODWT-GPR3(sym4) était meilleure

que tous les autres modèles, avec les R et NSE les plus élevées ($R \approx 0.99$, $NSE \approx 0.98$) et les RMSE et MAE les plus faibles ($RMSE \approx 0.171 \text{ m}^3/\text{s}$, $MAE \approx 0.117 \text{ m}^3/\text{s}$), suivie par le modèle MODWT-MLPNN5 (sym4), par le modèle MODWT-LSTM3 (sym4) et à la fin le modèle MODWT-GRNN5 (sym4).

- La **figure VII.11** présente des graphiques récapitulatifs des performances des meilleures modèles autonomes et hybrides pendant la phase de validation comprenant le diagramme en boîte, le diagramme en violon, le diagramme radar et le diagramme de Taylor, Il ressort de ces graphiques que le modèle MODWT-GPR3(sym4) est plus proche du champ observé (débit mesuré) que les autres modèles. De plus, le modèle MODWT-GPR3 (sym4) a la plus faible valeur de RMSE et de l'écart type et une corrélation plus élevée par rapport aux autres modèles.
- Dans la comparaison des courbes de variation temporelle des débits mesuré et calculé dans la phase de validation (**figure VII.12**), Le modèle MODWT-GPR3(sym4) montre une forte capacité de prédire et de capturer les fortes non-linéarités de l'ensemble de données et la manière dont la prévision des valeurs maximales a été réalisée avec succès.

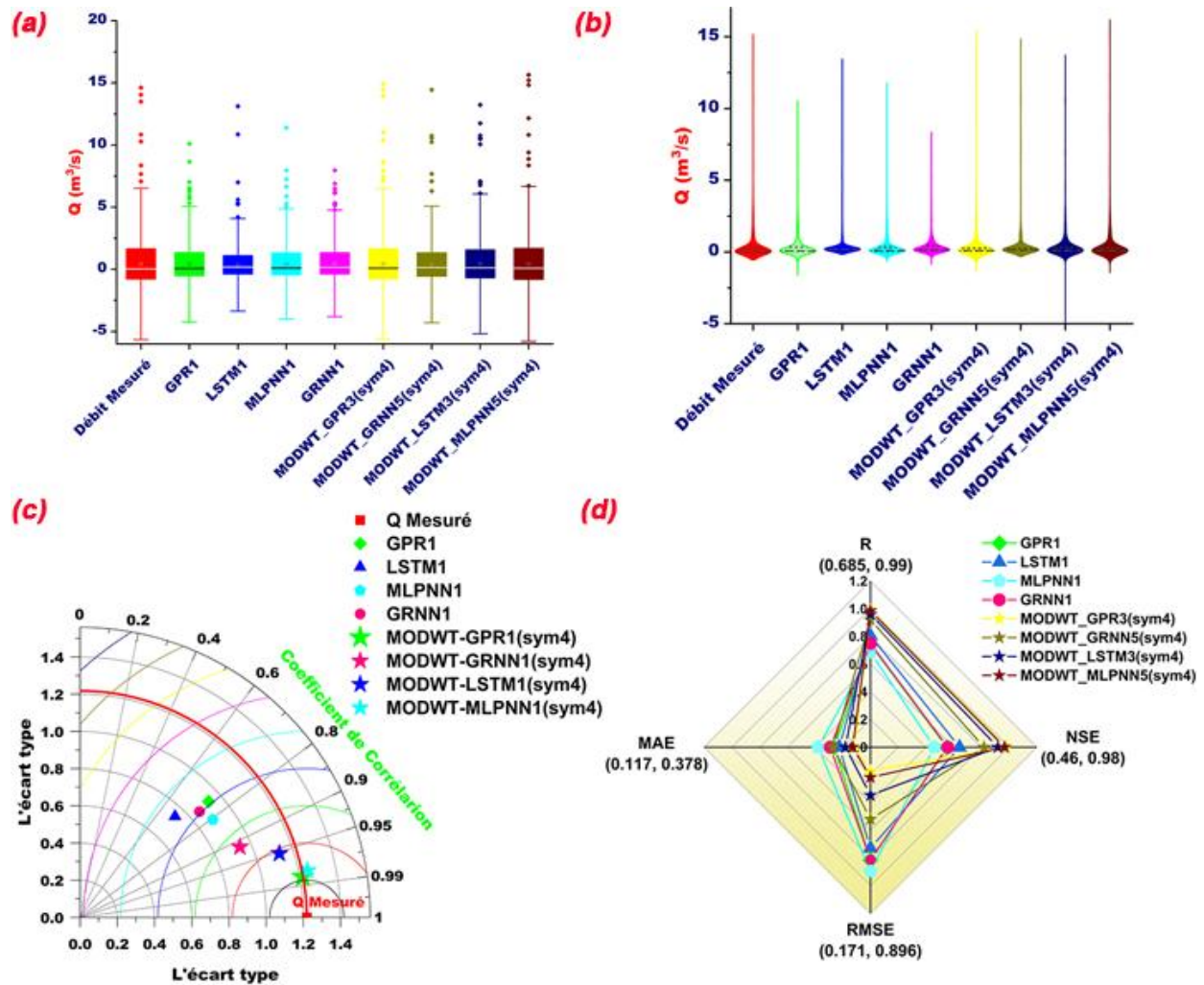


Figure VII.11 Résultats récapitulatifs des meilleurs modèles dans la station de Bir Ouled Taha pendant la phase de validation:(a) Diagramme en boîte, (b) Diagramme en violon, (c) Diagramme de Taylor et (c) Diagramme radar

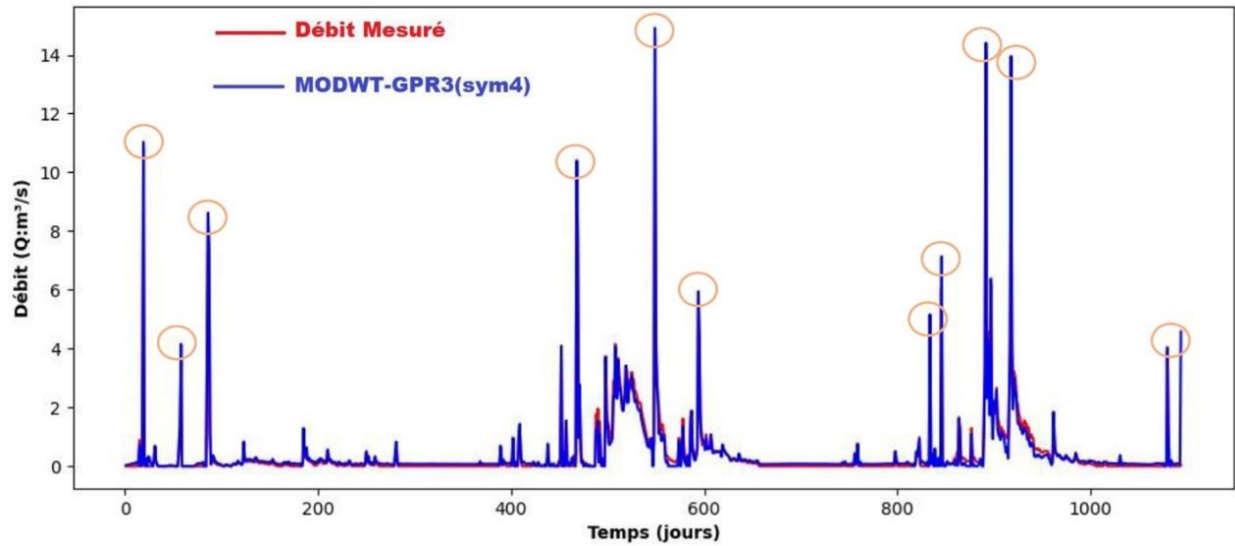


Figure VII.12 Variation temporelle des débits mesuré et calculé dans la phase de validation par le modèle MODWT-GPR3(sym4) pour la station de Bir Ouled Tahar

VII.2.1.3 Station de Ponteba Barrage

Dans la station de Ponteba Barrage, comme le montre le [tableau VII.7](#) et la [figure VII.13 \(a\)](#), et pendant la phase d'apprentissage, les résultats indiquent que le modèle GPR1 est le plus performant que les modèles LSTM1, GRNN1 et MLPNN1. Le modèle GPR3 a obtenu des valeurs RMSE et MAE plus faibles, $\approx 0.221 \text{ m}^3/\text{s}$ et $\approx 0.073 \text{ m}^3/\text{s}$, par rapport à $4.559 \text{ m}^3/\text{s}$, $1.654 \text{ m}^3/\text{s}$, $6.410 \text{ m}^3/\text{s}$, $3.170 \text{ m}^3/\text{s}$, $5.979 \text{ m}^3/\text{s}$ et $2.495 \text{ m}^3/\text{s}$. Ces valeurs ont été obtenues en utilisant respectivement les modèles LSTM1, GRNN1 et MLPNN1. Comme la station de Bir Ouled Taher, les quatre modèles ML autonomes ont montré les meilleures performances avec les valeurs R et NSE les plus élevées avec la 1^{ère} et 3^{-ème} combinaisons des variables d'entrée ([tableau VII.1](#)). L'analyse des corrélations de Pearson (R) et le critère de Nash-Sutcliffe (NSE) indique que les précipitations sont le principal facteur prédictif du débit dans le bassin versant d'oued Ras Ouahrane. Cela ressort également des résultats de R et NSE lors de la formation pour les modèles GPR4, LSTM4, GRNN4 et MLPNN4 lors de l'utilisation de 4^{-ème} combinaisons des variables d'entrées, nous constatons une diminution des performances. Au cours des phases de validation, La [figure VII.13 \(b\)](#) montre clairement que le modèle GPR1 a produit le plus haut niveau de précision en termes de $R \approx 0.951$, $NSE \approx 0.904$, $RMSE \approx 5.138 \text{ m}^3/\text{s}$ et $MAE \approx 2.284 \text{ m}^3/\text{s}$, suivi par le modèle MLPNN5. De même, le modèle LSTM3 avec des valeurs minimales RMSE et MAE et des valeurs plus élevées de R et

de NSE. En outre, les résultats révèlent la faible performance du modèle GRNN3 dans la prédiction du débit journalier dans le bassin d'oued Ras Ouahrane.

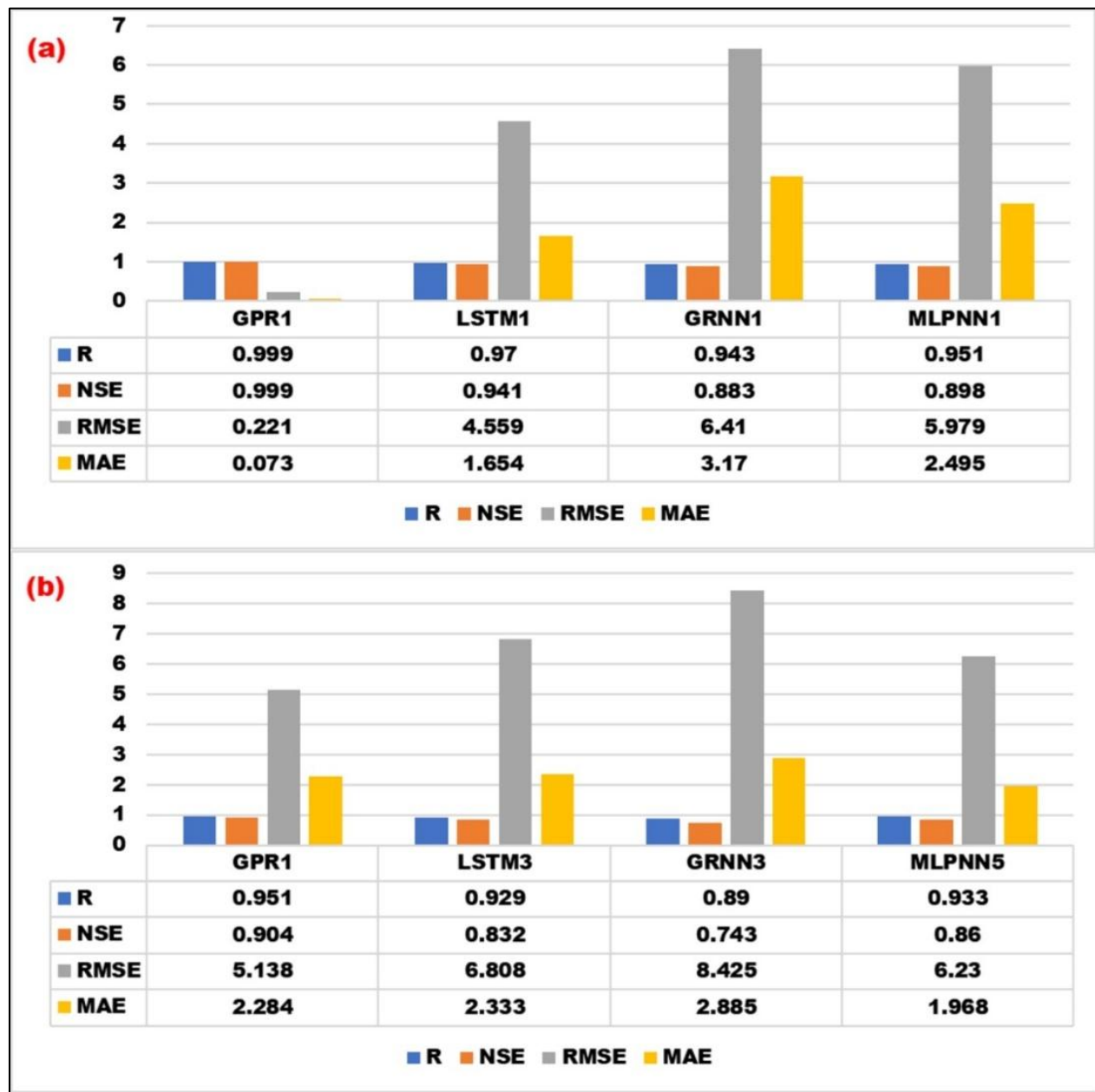


Figure VII.13 Graphique à barres des statistiques des critères de performance pour les meilleurs modèles ML autonomes pour la station de Ponteba Barrage (a : Apprentissage, b : Validation)

Tableau VII.7 Résultats de prévision des différents modèles dans le pas journalier pour la station de Ponteba Barrage

Modèles	Apprentissage				Validation			
	R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>GPR</i>								
GPR1	0.999	0.999	0.221	0.073	0.951	0.904	5.138	2.284
GPR2	0.842	0.698	10.270	4.329	0.385	0.136	15.450	6.703
GPR3	0.995	0.991	1.785	0.492	0.936	0.875	5.874	2.336
GPR4	0.798	0.630	11.378	4.932	0.388	0.140	15.420	6.836
GPR5	0.998	0.997	1.059	0.268	0.937	0.876	5.863	2.330
GPR6	0.431	0.174	17.000	8.077	0.010	0.142	17.769	11.680
<i>LSTM</i>								
LSTM1	0.970	0.941	4.559	1.654	0.914	0.815	7.156	2.259
LSTM2	0.903	0.815	8.040	3.423	0.644	0.337	13.539	6.022
LSTM3	0.959	0.919	5.313	1.940	0.929	0.832	6.808	2.333
LSTM4	0.874	0.764	9.080	3.723	0.713	0.460	12.211	5.488
LSTM5	0.939	0.881	6.448	2.105	0.911	0.820	7.056	2.028
LSTM6	0.862	0.742	9.499	4.095	0.662	0.420	12.664	5.986
<i>GRNN</i>								
GRNN1	0.943	0.883	6.410	3.170	0.876	0.730	8.645	2.811
GRNN2	0.528	0.270	15.975	6.874	0.379	0.139	15.428	5.699
GRNN3	0.873	0.752	9.323	4.060	0.890	0.743	8.425	2.885
GRNN4	0.494	0.237	16.336	7.129	0.399	0.155	15.279	5.767
GRNN5	0.867	0.741	9.510	4.165	0.887	0.741	8.468	2.926
GRNN6	0.149	0.020	18.514	8.705	0.057	0.000	16.623	6.792
<i>MLPNN</i>								
MLPNN1	0.951	0.898	5.979	2.495	0.901	0.785	7.711	2.324
MLPNN2	0.528	0.279	15.880	7.192	0.366	0.092	15.841	5.858
MLPNN3	0.928	0.852	7.190	3.407	0.933	0.837	6.720	2.960
MLPNN4	0.511	0.261	16.076	7.278	0.359	0.086	15.895	6.052
MLPNN5	0.931	0.867	6.825	2.119	0.933	0.860	6.230	1.968
MLPNN6	0.191	0.036	18.360	8.779	0.039	0.019	16.785	6.803

Les résultats de deuxième scénario (tableau VII.8, figure VII.14), en cours d'hybridation des modèles MODWT-GPR pour la station de Ponteba Barrage pendant la phase d'apprentissage ont montré la supériorité de quatre modèles hybrides : MODWT-GPR3 (coif1), MODWT-GPR1 (db3), MODWT-GPR3 (sym4) et MODWT-GPR5 (haar) en termes de critères de performance R, NSE, RMSE et MAE dans la comparaison avec le modèle autonome GPR1 (R ≈ 0.999, NSE ≈ 0.999, RMSE ≈ 0.221 m³/s et MAE ≈ 0.073 m³/s). Quant au modèle MODWT-GPR1 (fk8), son efficacité lors de la phase d'apprentissage a été inférieure que celle du modèle GPR1. Bien que les quatre modèles mentionnés ci-dessus aient montré des signes de supériorité dans la phase d'apprentissage, un seul de ces modèles hybrides a maintenu cette supériorité dans la phase de validation, MODWT-GPR3 (sym4) où R ≈ 0.957, NSE ≈ 0.915, RMSE ≈ 4.852 m³/s et MAE ≈ 1.941 m³/s, relativement aux quatre critères de performance numérique proposés dans cette étude,

ces statistiques témoignent de sa supériorité sur le modèle GPR1 ($R \approx 0,951$, $NSE \approx 0,904$, $RMSE \approx 5,138 \text{ m}^3/\text{s}$ et $MAE \approx 2,284 \text{ m}^3/\text{s}$).

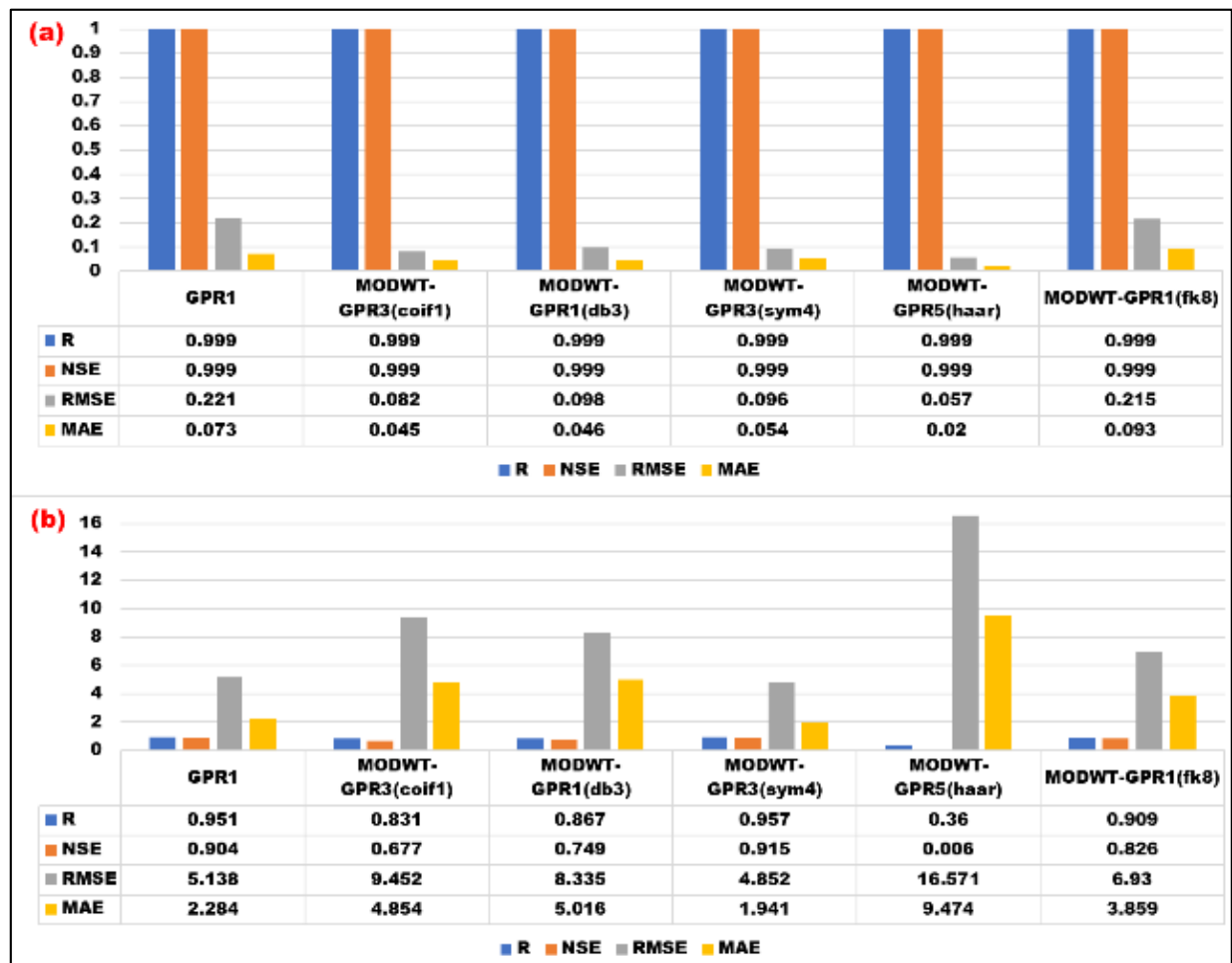


Figure VII.14 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GPR par rapport au modèle GPR optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation)

Tableau VII.8 Résultats de prévision des modèles hybrides GPR-MODWT dans le pas journalier pour la station de Ponteba Barrage

Mother wavelet	Models	Training				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT_GPR (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-GPR1	0.999	0.999	0.070	0.038	0.830	0.687	9.301	5.385
	MODWT-GPR2	0.999	0.999	0.198	0.056	0.227	-0.081	17.283	9.502
	MODWT-GPR3	0.999	0.999	0.082	0.045	0.831	0.677	9.452	4.854
	MODWT-GPR4	0.999	0.999	0.170	0.042	0.191	-0.150	17.831	9.999
	MODWT-GPR5	0.999	0.999	0.345	0.191	0.771	0.572	10.873	6.289
	MODWT-GPR6	0.999	0.999	0.176	0.043	0.099	-0.165	17.945	9.979
<i>MODWT_GPR (db3)</i>									
Daubechies Wavelet (db3)	MODWT-GPR1	0.999	0.999	0.098	0.046	0.867	0.749	8.335	5.016
	MODWT-GPR2	0.999	0.999	0.245	0.076	0.377	0.064	16.083	9.034
	MODWT-GPR3	0.999	0.999	0.096	0.049	0.853	0.722	8.773	5.023
	MODWT-GPR4	0.999	0.999	0.229	0.068	0.366	0.067	16.059	8.944
	MODWT-GPR5	0.999	0.999	0.081	0.041	0.775	0.574	10.855	7.276
	MODWT-GPR6	0.999	0.999	0.206	0.053	0.287	0.011	16.537	8.787
<i>MODWT_GPR (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-GPR1	0.999	0.999	0.106	0.058	0.957	0.915	4.856	3.348
	MODWT-GPR2	0.999	0.999	0.238	0.089	0.562	0.227	14.612	7.952
	MODWT-GPR3	0.999	0.999	0.096	0.054	0.957	0.915	4.852	1.941
	MODWT-GPR4	0.999	0.999	0.228	0.084	0.552	0.218	14.701	7.948
	MODWT-GPR5	0.999	0.999	0.104	0.056	0.957	0.914	4.864	3.350
	MODWT-GPR6	0.999	0.999	0.244	0.095	0.349	-0.105	17.472	8.809
<i>MODWT_GPR (haar)</i>									
Haar wavelet (haar)	MODWT-GPR1	0.999	0.999	0.023	0.010	0.309	-0.310	19.028	12.907
	MODWT-GPR2	0.999	0.999	0.045	0.011	0.147	-0.086	17.320	8.831
	MODWT-GPR3	0.999	0.999	0.028	0.012	0.317	-0.295	18.920	12.689
	MODWT-GPR4	0.999	0.999	0.041	0.010	0.135	-0.083	17.298	8.754
	MODWT-GPR5	0.999	0.999	0.057	0.020	0.360	0.006	16.571	9.474
	MODWT-GPR6	0.999	0.999	0.080	0.018	0.089	-0.050	17.035	7.871
<i>MODWT_GPR (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-GPR1	0.999	0.999	0.215	0.093	0.909	0.826	6.930	3.859
	MODWT-GPR2	0.999	0.999	0.278	0.111	0.533	0.183	15.031	7.971
	MODWT-GPR3	0.999	0.999	0.225	0.101	0.906	0.819	7.069	4.083
	MODWT-GPR4	0.999	0.999	0.287	0.114	0.500	0.109	15.696	8.176
	MODWT-GPR5	0.999	0.999	0.096	0.044	0.872	0.749	8.323	4.963
	MODWT-GPR6	0.999	0.999	0.299	0.128	0.464	0.130	15.510	7.946

Conformément aux dispositions de [tableau VII.9](#) et la [figure VII.15](#). Au cours de la phase d'apprentissage, les valeurs de R, NSE, RMSE et MAE fournies par les modèle autonome LSTM1 sont de l'ordre de $\approx 0,97$, $\approx 0,941$, $\approx 4,559 \text{ m}^3/\text{s}$ et $\approx 1,654 \text{ m}^3/\text{s}$, respectivement. Compte tenu ces statistiques pour les meilleures modèles hybrides MODWT-LSTM, seul le modèle hybride MODWT LSTM5 (sym4) ont amélioré le processus d'apprentissage où $R \approx 0,994$, $NSE \approx 0,987$, $RMSE \approx 2,097 \text{ m}^3/\text{s}$ et $MAE \approx 1,062 \text{ m}^3/\text{s}$. lors de la phase de validation en général, les modèles hybrides MODWT-LSTM n'ont apporté aucune amélioration dans la prédiction des débits

journalier par rapport aux valeurs optimale de $R \approx 0,929$, $NSE \approx 0,832$, $RMSE \approx 6,808 \text{ m}^3/\text{s}$ et $MAE \approx 2.233 \text{ m}^3/\text{s}$ fournies par les modèle autonome LSTM3.

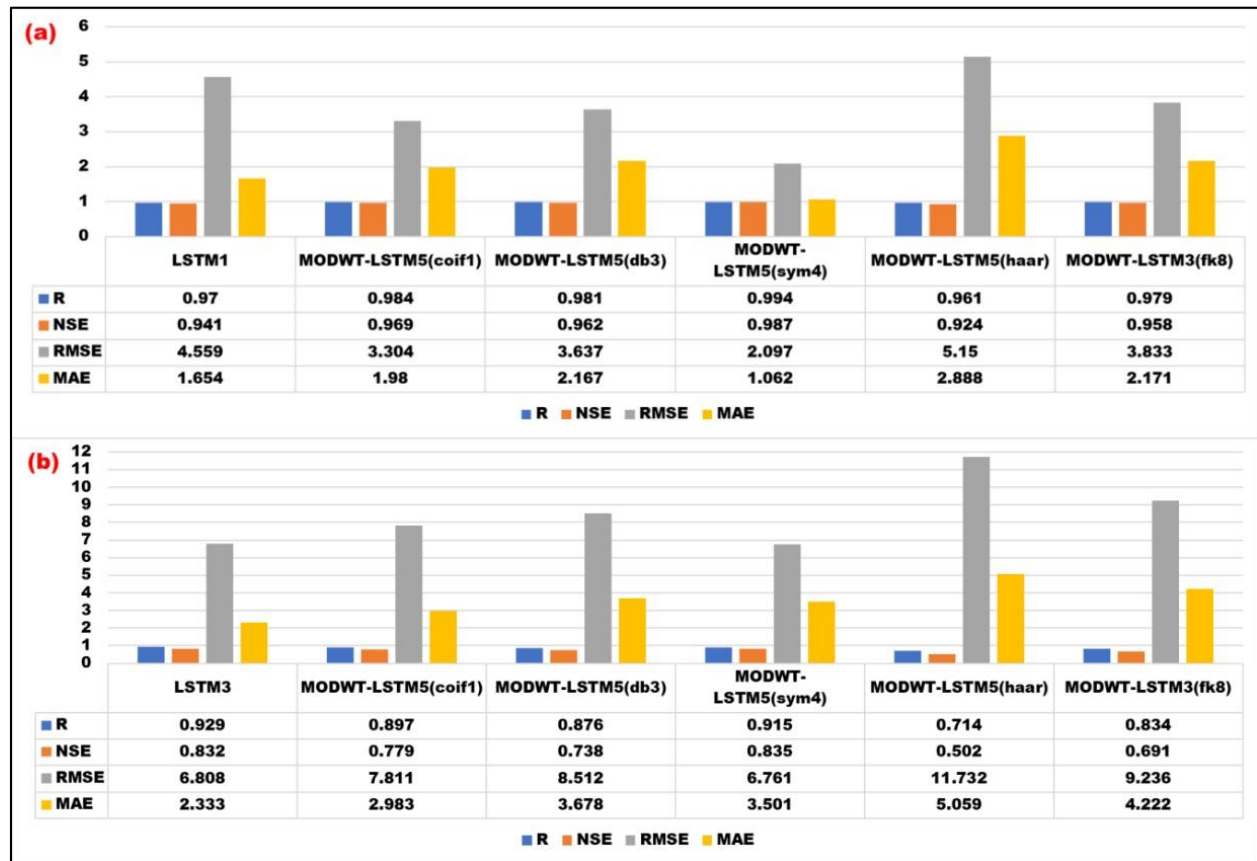


Figure VII.15 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-LSTM par rapport au modèle LSTM optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation)

Tableau VII.9 Résultats de prévision des modèles hybrides LSTM-MODWT dans le pas journalier pour la station de Ponteba Barrage

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT_LSTM (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-LSTM1	0.988	0.977	2.836	1.920	0.796	0.624	10.189	4.540
	MODWT-LSTM2	0.885	0.782	8.724	4.847	0.512	0.236	14.528	6.821
	MODWT-LSTM3	0.986	0.973	3.087	1.968	0.860	0.707	9.002	4.167
	MODWT-LSTM4	0.876	0.766	9.055	4.849	0.517	0.260	14.298	6.286
	MODWT-LSTM5	0.984	0.969	3.304	1.980	0.897	0.779	7.811	2.983
	MODWT-LSTM6	0.838	0.700	10.246	5.070	0.462	0.199	14.879	6.033
<i>MODWT_LSTM (db3)</i>									
Daubechies Wavelet (db3)	MODWT-LSTM1	0.981	0.962	3.624	2.275	0.807	0.645	9.912	4.719
	MODWT-LSTM2	0.894	0.797	8.437	4.551	0.398	0.091	15.851	6.521
	MODWT-LSTM3	0.980	0.961	3.687	2.323	0.811	0.648	9.867	4.543
	MODWT-LSTM4	0.930	0.864	6.887	3.834	0.486	0.210	14.774	6.167
	MODWT-LSTM5	0.981	0.962	3.637	2.167	0.876	0.738	8.512	3.678
	MODWT-LSTM6	0.837	0.700	10.244	5.623	0.522	0.262	14.285	6.731
<i>MODWT_LSTM (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-LSTM1	0.992	0.984	2.356	1.352	0.882	0.771	7.953	3.341
	MODWT-LSTM2	0.960	0.922	5.220	2.872	0.407	-0.588	20.948	9.298
	MODWT-LSTM3	0.996	0.992	1.695	0.967	0.905	0.818	7.083	3.389
	MODWT-LSTM4	0.960	0.922	5.230	2.742	0.496	0.095	15.816	6.949
	MODWT-LSTM5	0.994	0.987	2.097	1.062	0.915	0.835	6.761	3.501
	MODWT-LSTM6	0.934	0.872	6.690	3.394	0.406	-0.419	19.804	9.199
<i>MODWT_LSTM (haar)</i>									
Haar wavelet (haar)	MODWT-LSTM1	0.957	0.916	5.414	3.302	0.645	0.366	13.236	6.347
	MODWT-LSTM2	0.842	0.709	10.094	5.442	0.260	-0.041	16.959	7.590
	MODWT-LSTM3	0.953	0.909	5.645	3.276	0.711	0.475	12.051	5.821
	MODWT-LSTM4	0.872	0.760	9.167	4.656	0.256	-0.252	18.599	8.841
	MODWT-LSTM5	0.961	0.924	5.150	2.888	0.714	0.502	11.732	5.059
	MODWT-LSTM6	0.716	0.512	13.059	6.613	0.338	-0.235	18.474	9.675
<i>MODWT_LSTM (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-LSTM1	0.980	0.960	3.763	2.214	0.755	0.527	11.439	4.659
	MODWT-LSTM2	0.936	0.876	6.594	3.615	0.671	0.334	13.564	6.889
	MODWT-LSTM3	0.979	0.958	3.833	2.171	0.834	0.691	9.236	4.222
	MODWT-LSTM4	0.938	0.879	6.507	3.353	0.602	0.117	15.622	7.667
	MODWT-LSTM5	0.976	0.952	4.110	2.147	0.795	0.629	10.126	4.378
	MODWT-LSTM6	0.931	0.867	6.813	3.536	0.632	0.050	16.201	7.766

Dans la phase d'apprentissage (tableau VII.10, figure VII.16 (a)), Les modèles hybrides MODWT-GRNN à l'exception du modèle MODWT-GRNN1 (haar) ont montent des améliorations au niveau de la performance des modèles autonomes avec les valeurs de R et NSE les plus élevé, et les valeurs de RMSE et MAE les plus faibles, Cela apparaît lorsque l'on compare leurs performances avec le modèle GRNN1(R ≈0,994, NSE ≈0,987, RMSE ≈2,097 m³/s et MAE ≈1,062m³/s). D'après la figure VII.16 (b) et en phase de validation le meilleur modèle parmi les cinq modèles hybrides MODWT-GRNN était le modèle MODWT-GRNN5 (sym4) ou R ≈0,823, NSE ≈0,667, RMSE

$\approx 9,592 \text{ m}^3/\text{s}$ et $\text{MAE} \approx 3.272 \text{ m}^3/\text{s}$. Lorsque ces statistiques sont comparées à celles présentées par le modèle GRNN3 ($R \approx 0,823$, $\text{NSE} \approx 0,667$, $\text{RMSE} \approx 9,592 \text{ m}^3/\text{s}$ et $\text{MAE} \approx 3.272 \text{ m}^3/\text{s}$), en utilisant les modèles hybrides MODWT-GRNN, nous pouvons affirmer que aucun processus d'amélioration de la prédiction du débit journalier du cours d'eau au niveau de la station de Ponteba Barrage ne se produit, dans la mesure où les valeurs R et NSE sont les plus faibles, et les valeurs RMSE et MAE les plus élevées.

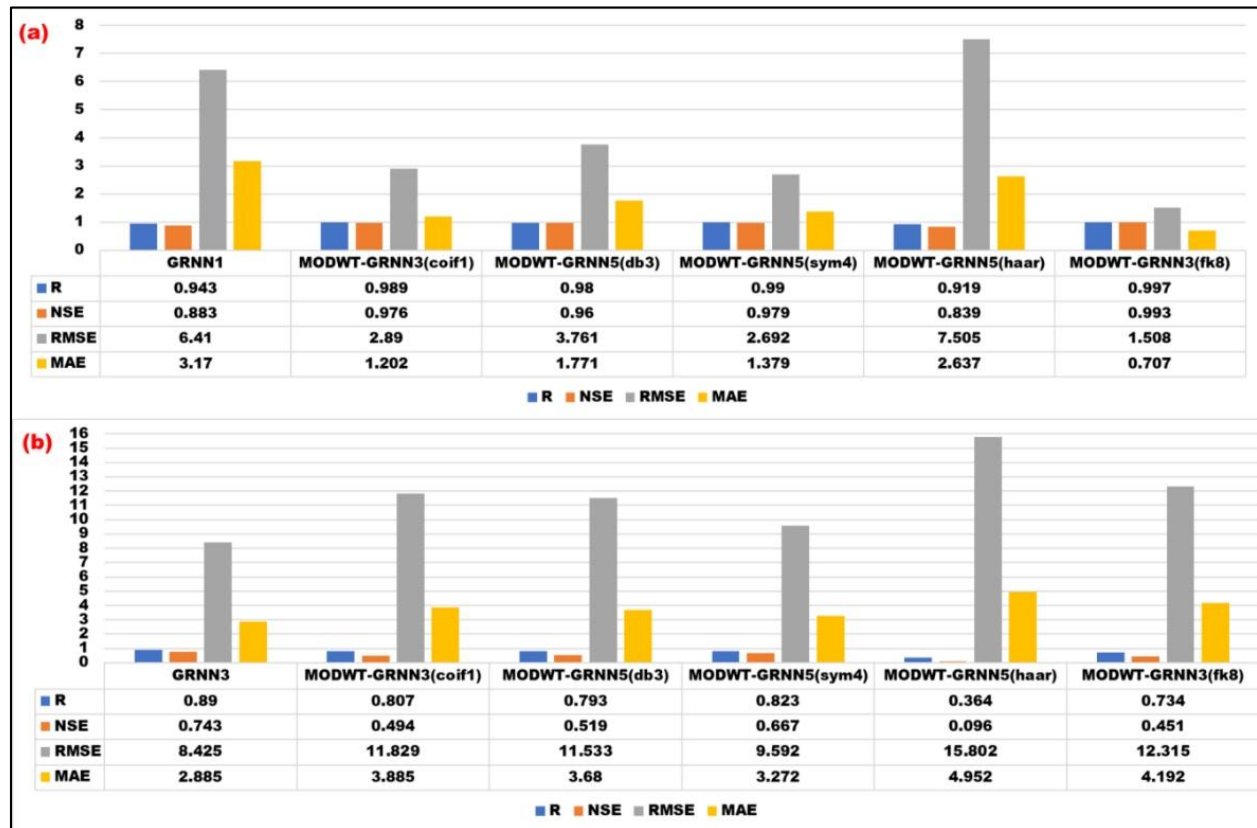


Figure VII.16 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-GRNN par rapport au modèle GRNN optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation)

Tableau VII.10 Résultats de prévision des modèles hybrides GRNN-MODWT dans le pas journalier pour la station de Ponteba Barrage

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT GRNN (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-GRNN1	0.997	0.994	1.443	0.513	0.748	0.445	12.385	4.155
	MODWT-GRNN2	0.934	0.865	6.872	2.219	0.216	-0.037	16.927	5.467
	MODWT-GRNN3	0.989	0.976	2.890	1.202	0.807	0.494	11.829	3.885
	MODWT-GRNN4	0.873	0.748	9.384	3.283	0.257	0.024	16.422	5.305
	MODWT-GRNN5	0.972	0.942	4.488	2.015	0.796	0.526	11.448	3.699
	MODWT-GRNN6	0.614	0.315	15.476	5.993	0.156	0.003	16.599	5.688
<i>MODWT GRNN (db3)</i>									
Daubechies Wavelet (db3)	MODWT-GRNN1	0.998	0.995	1.285	0.455	0.675	0.391	12.974	4.288
	MODWT-GRNN2	0.935	0.870	6.739	2.112	0.317	0.067	16.061	5.100
	MODWT-GRNN3	0.993	0.985	2.297	1.011	0.769	0.466	12.150	3.943
	MODWT-GRNN4	0.884	0.774	8.897	3.142	0.361	0.113	15.655	4.968
	MODWT-GRNN5	0.980	0.960	3.761	1.771	0.793	0.519	11.533	3.680
	MODWT-GRNN6	0.688	0.429	14.128	5.718	0.397	0.135	15.462	5.189
<i>MODWT GRNN (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-GRNN1	0.999	0.999	0.642	0.297	0.771	0.467	12.134	4.004
	MODWT-GRNN2	0.986	0.972	3.126	1.114	0.484	0.183	15.025	5.356
	MODWT-GRNN3	0.997	0.994	1.442	0.698	0.827	0.651	9.815	3.368
	MODWT-GRNN4	0.971	0.940	4.569	1.721	0.526	0.263	14.275	5.040
	MODWT-GRNN5	0.990	0.979	2.692	1.379	0.823	0.667	9.592	3.272
	MODWT-GRNN6	0.822	0.644	11.165	4.301	0.384	0.136	15.450	4.991
<i>MODWT GRNN (haar)</i>									
Haar wavelet (haar)	MODWT-GRNN1	0.975	0.950	4.185	0.924	0.364	0.098	15.787	5.270
	MODWT-GRNN2	0.837	0.670	10.751	3.527	0.024	-0.139	17.743	6.472
	MODWT-GRNN3	0.951	0.904	5.809	1.754	0.326	0.064	16.082	5.169
	MODWT-GRNN4	0.703	0.446	13.922	5.188	0.043	-0.068	17.181	6.267
	MODWT-GRNN5	0.919	0.839	7.505	2.637	0.364	0.096	15.802	4.952
	MODWT-GRNN6	0.491	0.188	16.849	7.448	0.029	-0.023	16.813	6.475
<i>MODWT GRNN (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-GRNN1	0.999	0.999	0.708	0.302	0.684	0.398	12.902	4.529
	MODWT-GRNN2	0.991	0.981	2.571	0.957	0.439	0.165	15.195	5.301
	MODWT-GRNN3	0.997	0.993	1.508	0.707	0.734	0.451	12.315	4.192
	MODWT-GRNN4	0.967	0.931	4.920	1.814	0.400	0.139	15.423	5.094
	MODWT-GRNN5	0.983	0.966	3.456	1.543	0.710	0.486	11.921	3.825
	MODWT-GRNN6	0.876	0.743	9.490	3.691	0.526	0.253	14.372	4.512

Dans le dernier cas du deuxième scénario (tableau VII.11, figure VII.17), le processus de prédiction des débits journaliers a été bien mené par l'hybridation MODWT-MLPNN, en particulier en utilisant le modèle hybride MODWT-MLPNN5 (sym4), ceci aussi bien dans la phase d'apprentissage que dans la phase de validation. Dans la phase d'apprentissage, le modèle MODWT-MLPNN5 (sym4) avec $R \approx 0.999$, $NSE \approx 0.999$, $RMSE \approx 0.454 \text{ m}^3/\text{s}$ et $MAE \approx 0.224 \text{ m}^3/\text{s}$ a été plus performant que le modèle MLPNN1 ou $R \approx 0.951$, $NSE \approx 0.898$, $RMSE \approx 5.979 \text{ m}^3/\text{s}$ et $MAE \approx 2.495 \text{ m}^3/\text{s}$, Dans la phase de validation, le modèle MODWT-MLPNN5 (sym4)

présentant les valeurs les plus élevées pour $R \approx 0.946$ et $NSE \approx 0.896$ et les valeurs les plus faibles pour $RMSE \approx 5.373 \text{ m}^3/\text{s}$ et $MAE \approx 1.951 \text{ m}^3/\text{s}$. ce modèle a été plus performant que le modèle MLPNN5 ($R \approx 0.933$, $NSE \approx 0.86$, $RMSE \approx 6.23 \text{ m}^3/\text{s}$ et $MAE \approx 1.968 \text{ m}^3/\text{s}$).

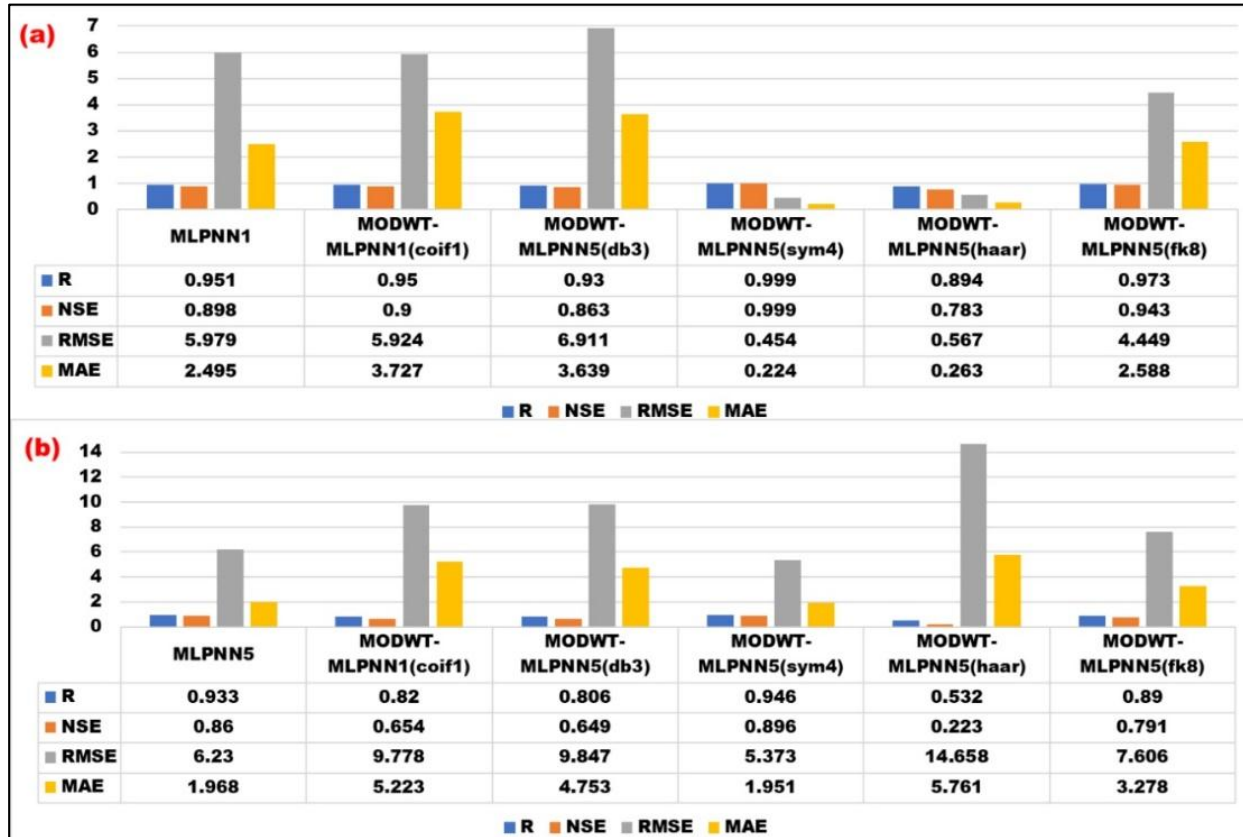


Figure VII.17 Graphique à barres des statistiques des critères de performance pour top modèles hybrides MODWT-MLPNN par rapport au modèle MLPNN optimale pour la station de Ponteba Barrage (a : Apprentissage, b : Validation)

Tableau VII.11 Résultats de prévision des modèles hybrides MODWT-MLPNN dans le pas journalier pour la station de Ponteba Barrage

L'ondelette mère	Modèles	Apprentissage				Validation			
		R	NSE	RMSE	MAE	R	NSE	RMSE	MAE
<i>MODWT MLPNN (coif1)</i>									
Coiflets Wavelet (coif1)	MODWT-MLPNN1	0.950	0.900	5.924	3.727	0.820	0.654	9.778	5.223
	MODWT-MLPNN2	0.752	0.564	12.346	6.564	0.431	0.124	15.558	7.973
	MODWT-MLPNN3	0.946	0.888	6.270	4.012	0.725	0.524	11.469	5.700
	MODWT-MLPNN4	0.707	0.499	13.234	6.956	0.350	0.044	16.253	8.110
	MODWT-MLPNN5	0.931	0.865	6.867	3.737	0.792	0.626	10.173	5.500
	MODWT-MLPNN6	0.529	0.263	16.053	8.848	0.193	-0.068	17.177	9.137
<i>MODWT MLPNN (db3)</i>									
Daubechies Wavelet (db3)	MODWT-MLPNN1	0.927	0.849	7.257	3.832	0.778	0.598	10.544	5.643
	MODWT-MLPNN2	0.726	0.521	12.945	6.823	0.423	0.142	15.400	7.837
	MODWT-MLPNN3	0.940	0.881	6.444	3.395	0.789	0.617	10.288	5.084
	MODWT-MLPNN4	0.724	0.520	12.964	6.792	0.563	0.286	14.050	7.344
	MODWT-MLPNN5	0.930	0.863	6.911	3.639	0.806	0.649	9.847	4.753
	MODWT-MLPNN6	0.498	0.241	16.295	8.229	0.396	0.141	15.412	7.623
<i>MODWT MLPNN (sym4)</i>									
Symlets Wavelet (sym4)	MODWT-MLPNN1	0.998	0.994	1.436	0.436	0.927	0.858	6.258	3.532
	MODWT-MLPNN2	0.846	0.697	10.301	4.890	0.421	0.083	15.922	7.067
	MODWT-MLPNN3	0.998	0.996	1.239	0.473	0.928	0.860	6.221	3.434
	MODWT-MLPNN4	0.895	0.800	8.364	3.936	0.622	0.369	13.209	6.276
	MODWT-MLPNN5	0.999	0.999	0.454	0.224	0.946	0.896	5.373	1.951
	MODWT-MLPNN6	0.635	0.380	14.723	6.314	0.325	0.058	16.137	6.665
<i>MODWT MLPNN (haar)</i>									
Haar wavelet (haar)	MODWT-MLPNN1	0.824	0.674	10.672	5.217	0.497	0.168	15.165	7.518
	MODWT-MLPNN2	0.533	0.281	15.861	8.479	0.140	-0.112	17.530	8.841
	MODWT-MLPNN3	0.821	0.672	10.710	5.004	0.523	0.217	14.713	6.725
	MODWT-MLPNN4	0.455	0.207	16.656	8.531	0.050	-0.119	17.585	8.442
	MODWT-MLPNN5	0.815	0.665	10.830	4.767	0.532	0.223	14.658	5.761
	MODWT-MLPNN6	0.311	0.096	17.781	8.543	0.026	-0.053	17.058	7.559
<i>MODWT MLPNN (fk8)</i>									
Fejer-Korovkin Wavelet (fk8)	MODWT-MLPNN1	0.953	0.887	6.300	4.008	0.868	0.722	8.760	5.079
	MODWT-MLPNN2	0.769	0.587	12.025	6.707	0.523	0.220	14.685	8.780
	MODWT-MLPNN3	0.962	0.920	5.281	2.795	0.849	0.714	8.894	4.264
	MODWT-MLPNN4	0.775	0.600	11.831	6.308	0.501	0.204	14.830	8.001
	MODWT-MLPNN5	0.973	0.943	4.449	2.588	0.890	0.791	7.606	3.278
	MODWT-MLPNN6	0.656	0.430	14.123	6.553	0.509	0.214	14.743	7.014

D'après les figures VII.18 et VII.19, le modèle hybride MODWT-GPR3 (sym4) permet de reproduire de manière adéquate les débits de pointe. Il a fourni des prédictions beaucoup moins dispersées qui se concentrent sur la ligne 1 :1 par rapport aux autres modèles d'apprentissage automatique autonomes et hybrides.

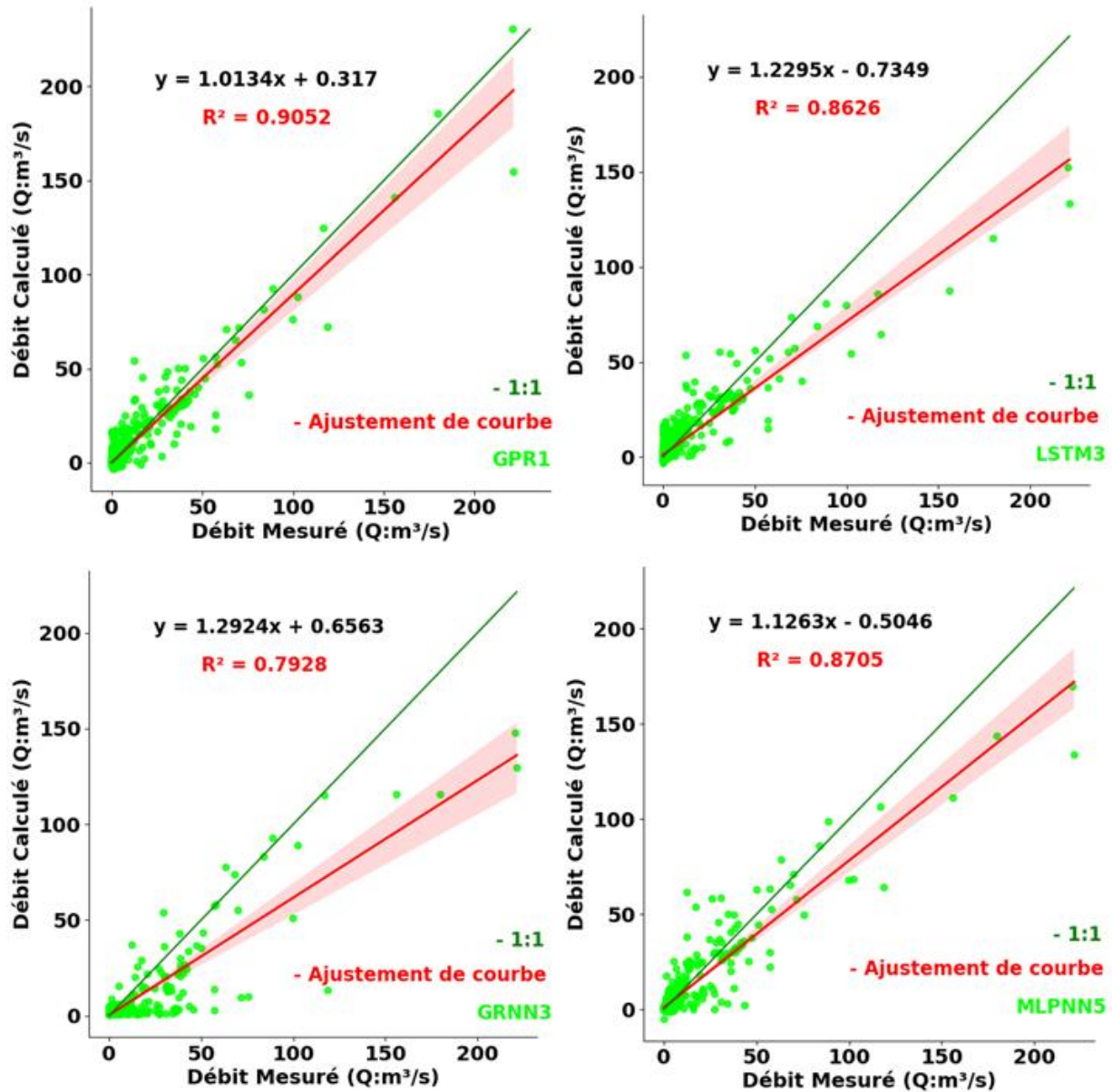


Figure VII.18 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles autonomes dans la station de Ponteba Barrage pendant la phase de validation

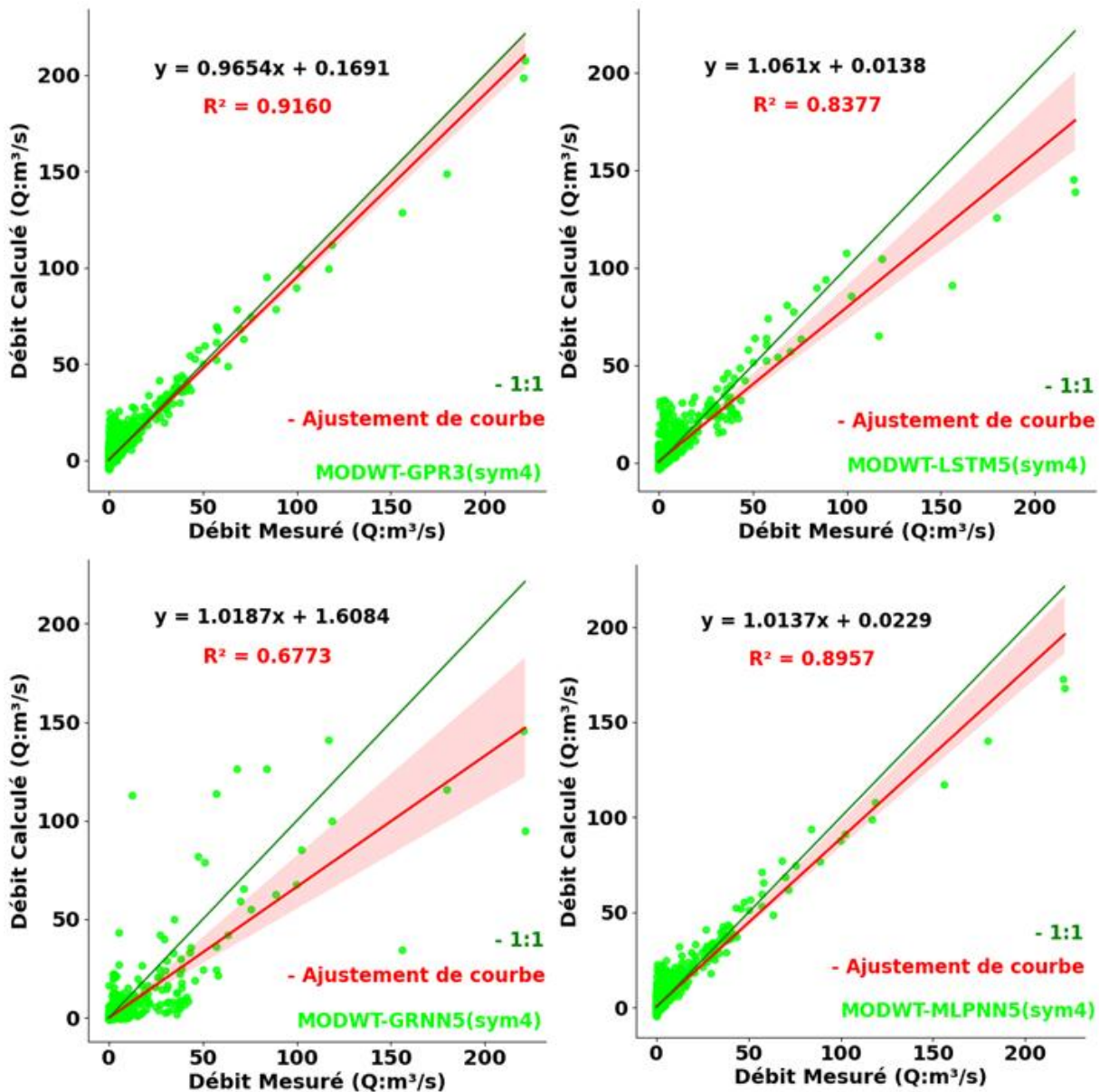


Figure VII.19 Comparaison des nuages de points pour le débit journalier observé par rapport au débit calculé par les meilleurs modèles hybrides dans la station de Ponteba Barrage pendant la phase de validation

VII.2.1.4 Résultats récapitulatifs des différents modèles pour la station de Ponteba Barrage

- Le modèle GPR1 a été le meilleur modèle dans les phases d'apprentissage et de validation parmi les quatre modèles ML indépendants proposés dans l'étude, où dans la phase d'apprentissage $R \approx 0.999$, $NSE \approx 0.999$, $RMSE \approx 0.221 \text{ m}^3/\text{s}$ et $MAE \approx 0.073 \text{ m}^3/\text{s}$, et dans la phase de validation $R \approx 0.951$, $NSE \approx 0.904$, $RMSE \approx 5.138 \text{ m}^3/\text{s}$, $MAE \approx 2.284 \text{ m}^3/\text{s}$.

- Les résultats des modèles d'apprentissage automatique autonomes dans les deux phases (**Tableau VII.9**) sont généralement plus élevés lors de l'utilisation des premiers et troisièmes ensembles de variables d'entrée (**Tableau VII.1**), comme c'est le cas dans la station de Bir Ouled Taher, En l'absence de l'entrée de précipitation, on constate une diminution des performances aux cours des phases de l'apprentissage et de validation, Cela ressort également des résultats obtenus lors de l'utilisation de modèles basés sur le quatrième ensemble de variables d'entrée.
- Dans l'hybridation des modèles GPR, LSTM, GRNN et MLPNN avec les différents familles de l'ondelette MODWT au cours du processus de d'apprentissage, par rapport au modèle optimal GPR1, les meilleurs résultats ont été fournis par les modèles hybrides MODWT-GPR1(db3, sym4 et fk8), MODWT-GPR1(coif1) et MODWT-GPR5(haar), Le modèle MODWT-GPR5(haar) était le meilleur en termes des valeurs de critères de performances suivantes $R \approx 0.999$, $NSE \approx 0.999$, $RMSE \approx 0.057 \text{ m}^3/\text{s}$ et $MAE \approx 0.02 \text{ m}^3/\text{s}$.
- Dans la phase de validation les résultats des modèles hybrides MODWT-LSTM, MODWT-GRNN, MODWT-MLPNN en termes de quatre critères de performances numériques proposées dans l'étude, n'ont montré aucune amélioration par rapport au modèle optimal GPR1.
- Les résultats des tableaux **VII.8**, **VII.9**, **VII.10** et **VII.11** montrent que les meilleurs résultats d'apprentissage du modèle ML hybride ont été obtenus en utilisant le premier modèle, qui contient la décomposition du signal de débit Q_{t-1} , Q_{t-2} avec la précipitation P_t , P_{t-1} , P_{t-2} et en utilisant le troisième modèle, qui contient la décomposition des signaux Q_{t-1} et Q_{t-2} avec P_t . Le cinquième modèle a également donné des résultats satisfaisants dans la phase d'apprentissage qui implique la décomposition des signaux de débit Q_{t-1} et Q_{t-2} .
- Dans le processus de validation, sur la base des résultats obtenus de la prédiction des débits journalier dans la station de Ponteba Barrage par les modèles hybrides MODWT-GPR, MODWT-LSTM, MODWT-GRNN et MODWT-MLPNN par rapport aux résultats du modèle optimal GPR1, les meilleurs résultats ont été fournis uniquement par le modèle hybride MODWT-GPR3(sym4) était supérieur au modèle GPR1 avec les valeurs de R et NSE les plus élevées ($R \approx 0.957$, $NSE \approx 0.915$) et les plus faibles RMSE et MAE ($RMSE \approx 4.852 \text{ m}^3/\text{s}$, $MAE \approx 1.941 \text{ m}^3/\text{s}$).

- Les graphiques des résultats récapitulatifs de la prédiction des débits journaliers dans la station de Ponteba Barrage à la **figure VII.20** pour les meilleures modèles autonomes et hybrides pendant la phase de validation montrent que le modèle MODWT-GPR3 (sym4) est plus fiable par rapport aux autres modèles. De plus, Ceci est également clair dans la comparaison des courbes de variation temporelle des débits mesuré et calculé par le modèle MODWT-GPR3 (sym4) en phase de validation (**figure VII.21**).

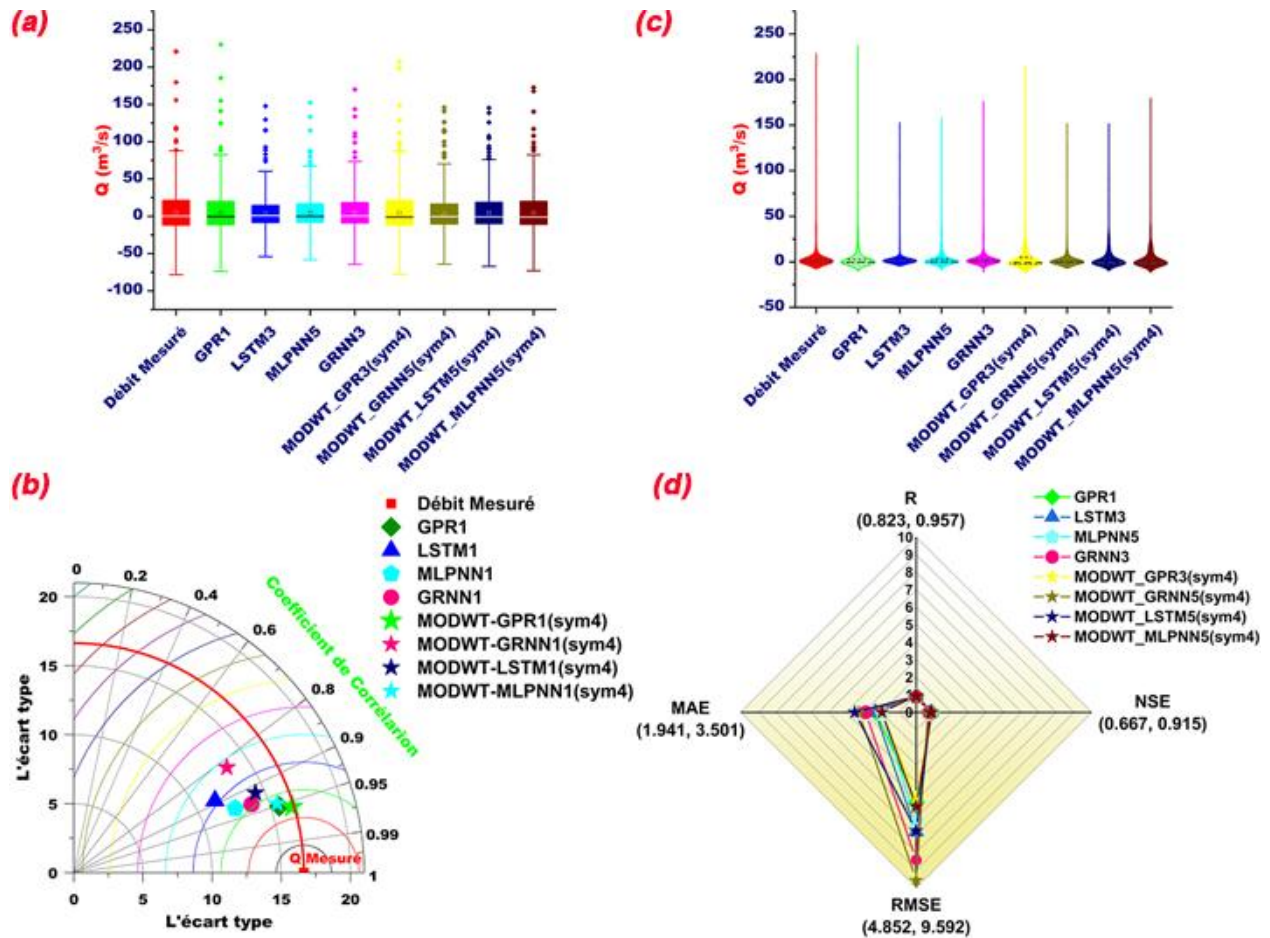


Figure VII.20 Résultats récapitulatifs des meilleurs modèles dans la station de Ponteba Barrage pendant la phase de validation:(a) Diagramme en boîte, (b) Diagramme en violon, (c) Diagramme de Taylor et (d) Diagramme radar

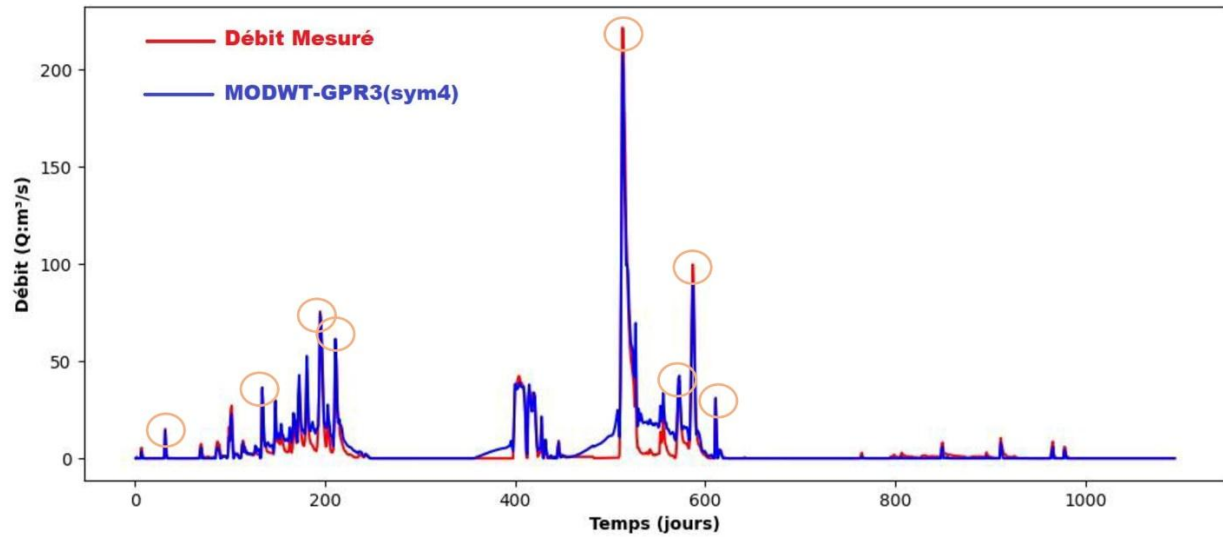


Figure VII.21 Variation temporelle des débits mesuré et calculé dans la phase de validation par le modèle MODWT-GPR3(sym4) pour la station de Ponteba Barrage

VII.3 Discussion des résultats

La présente étude porte sur la prédiction des débits journaliers via l'hybridation de modèles d'apprentissage automatique autonomes (GPR, LSTM, GRNN et MLPNN) avec cinq familles de Transformation en ondelettes discrètes à recouvrement maximal (MODWT) : Coiflets, Daubechies, Symlets, Haar et Fejer-Korovkin. Ces prédictions ont été effectuées dans deux stations du bassin de Cheliff, au nord de l'Algérie. Les résultats récapitulatifs selon les critères de performances numériques : R, NSE, RMSE et MAE et les critères de performance graphiques montrent que l'hybridation MODWT-GPR (sym) permettait d'obtenir une grande précision de prédiction. Les figures de Taylor diagramme VII.11 (c) et VII.20 (c) pour les stations de Bir Ouled Taher et Ponteba Barrage, respectivement, montrent la supériorité des modèles MODWT-GPR (sym4) avec les valeurs de RMSE les plus faibles dans les deux stations d'étude. En outre, Singh et al. (2004) et Moriasi et al. (2007) déclarent que les modèles hydrologiques avec des $RMSE < \frac{1}{2} SD$ (l'écart type) des données mesurées sont très performants. Dans cette étude, l'écart-type du débit dans la phase de validation est de 1,22 pour la station de Bir Ouled Taher et 16.62 pour la station de Ponteba Barrage, sur la base de cette mesure, l'algorithme MODWT-GPR (sym4) a obtenu de très bons résultats. Dans les études hydrologiques précédentes, notamment celles utilisant des modèles d'IA autonomes et hybrides, les auteurs n'ont pas mentionné les recherches sur l'hybridation des modèles MODWT-GPR. Nous présentons ce modèle pour la première fois afin de prédire les débits

journaliers, et pour confirmer sa robustesse et sa supériorité, il convient de le comparer aux études précédentes.

(Gomaa et al., 2023) dans leur étude, ils ont évalué l'application de l'analyse de mode empirique-transformation de Hilbert-Huang (EMD-HHT) en combinaison avec les modèles GPR et MLPNN-PSO pour analyser la relation entre les précipitations et le ruissellement, et dans prédire le taux de ruissellement. Les résultats obtenus prouvent que cette combinaison donne des résultats prometteurs dans la prédiction du ruissellement de surface, car le modèle MLP-PSO-EMD montre, à travers les valeurs $R=0.982$ et $NSE=0.961$, une précision supérieure à celle du modèle GPR-EMD ou $R=0.914$ et $NSE=0.834$. Le modèle hybride MODWT-GPR (sym4) proposé dans notre étude, plus précisément dans la station de Bir Ouled Taher, a surpassé les modèles EMD-MLP-PSO et EMD-GPR avec des valeurs de R et NSE atteignant respectivement 0.99 et 0.98 . (Chakraborty & Biswas, 2023) ont appliqué des modèles de prédiction du débit basés sur ANN et LSTM à trois stations hydrologiques, Teesta Bazaar, Domohani et Mekhliganj, dans le bassin de la rivière Teesta, en Inde, pour différents délais. La transformée en ondelettes a été combinée aux modèles ANN et LSTM pour développer deux modèles hybrides : le réseau neuronal artificiel basé sur les ondelettes (WANN) et le modèle WLSTM. Les résultats obtenus ont prouvé que les modèles hybrides proposés sont largement efficaces pour la prédiction des flux. Dans les trois stations d'étude, les valeurs de NSE et $RMSE$ étaient respectivement $=0.9985$ et $=24.32 \text{ m}^3/\text{s}$, $=0.9844$ et $91.13 \text{ m}^3/\text{s}$ et $=0.9985$ et $=28.75 \text{ m}^3/\text{s}$. Dans le cadre d'une autre nouvelle étude, (Shabbir et al., 2023) ont proposé une méthode hybride pour modéliser le débit quotidien de différents affluents du bassin de l'Indus (IRB), au Pakistan. La méthode hybride consiste en une analyse mixte et trois modèles basés sur les données. La décomposition hybride (HD) est basée sur la décomposition moyenne locale et la décomposition en mode empirique qui décompose la série de débits fluviaux en différentes composantes. Ensuite, les composants HD sont utilisés comme entrées pour prendre en charge les modèles de régression vectorielle (SVR), de voisin le plus proche (KNN) et de moyenne mobile intégrale automatique (ARIMA). Les prédictions des modèles HD-SVR, HD-KNN et HD-ARIMA sont combinées. La prédiction finale est la moyenne des prédictions des modèles HD-SVR, HD-KNN et HD-ARIMA (collectivement appelés HD-SKA). Les résultats obtenus ont montré que l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE) et l'erreur quadratique moyenne relative (RRSE) du modèle HD-SKA dans les données de test sont respectivement de $7,9314 \text{ m}^3/\text{s}$, $3,5315 \text{ m}^3/\text{s}$ et $0,2676 \text{ m}^3/\text{s}$, ce qu'est plus petit que tous les modèles concurrents de

l'étude. La conclusion de cette étude est que le modèle HD-SKA a une capacité supérieure à capturer le caractère aléatoire du débit fluvial. (Li et al., 2021) dans leur article, ont développé un modèle hybride d'analyse adaptative en mode variable (VMD) et de mémoire bidirectionnelle à long-court terme (Bi-LSTM) basé sur l'entropie énergétique pour prédire les débits journaliers. Le modèle a été analysé à l'aide de MAE, RMSE, NSE et du coefficient de corrélation (r). Par le VMD-BiLSTM proposé avec des modèles hybrides. À titre de comparaison, BP, GRNN, ELMAN, SVR, LSTM, Bi-LSTM, EMD-LSTM et VMD-LSTM sont adoptés pour l'évaluation et l'analyse. Ils ont constaté que le modèle proposé, avec MAE = 38,965, RMSE = 64,783 $r = 95\%$ et NSE = 95,7 %, était supérieur aux autres modèles. Par conséquent, les auteurs déclarent que le modèle hybride est robuste et efficace pour prédire un écoulement très instable et non linéaire. Il peut être utilisé comme un outil privilégié basé sur les données pour la prévision des flux quotidiens. Pour améliorer la précision des prévisions, Dans un autre article, (Wang et al., 2021) proposent un modèle hybride basé sur la décomposition des caractéristiques, la prédiction des composantes et la reconstruction des résultats, appelé VMD-LSTM-PSO. Le modèle est appliqué à trois stations hydrologiques du cours principal du fleuve Jaune en Chine et comparé aux modèles CEEMD-LSTM-PSO, WT-LSTM-PSO et LSTM-PSO. Les auteurs affirment la grande précision prédictive et la stabilité du nouveau modèle hybride VMD-LSTM PSO. (Hu et al., 2020) ont proposé un nouveau modèle basé sur les données qui combine l'analyse en mode variable (VMD) et des modèles de prévision incluent la moyenne mobile autorégressive (ARMA), l'arbre de régression boosté par gradient (GBRT), la régression vectorielle de support (SVR) et le réseau neuronal de rétro-propagation (BPNN) pour prédire le débit quotidien à partir des séries historiques de débits journaliers enregistrées aux stations hydrologiques de Wushan et Weijiabao en Chine. Les résultats de leur étude montrent que les modèles proposés basés sur VMD peuvent toujours atteindre de bonnes performances pendant la phase de test et avoir une stabilité et une représentativité relativement bonnes. Plus précisément, le modèle VMD-BPNN avec NSE = 0.9566, = 0.9865 et RMSE=3.79 m³/s et = 16.803.79 m³/s) pour les deux stations d'étude Wushan et Weijiabao, respectivement, a pris en compte la précision des prévisions et l'efficacité des calculs. (He et al., 2019) ont proposé dans leur article un modèle hybride basé sur l'analyse en mode variable (VMD) et les réseaux de neurones profonds (DNN), appelé VMD-DNN, pour effectuer des prévisions quotidiennes du ruissellement. Le modèle proposé est démontré à l'aide de données de séries de ruissellements quotidiens de la station hydrologique de Zhangjiashan dans la rivière Jing, en Chine. Les résultats

les plus importants obtenus ont révélé que le modèle hybride VMD-DNN proposé produit les meilleures performances basées sur l'efficacité de Nash-Sutcliffe ($NSE = 0,95$), l'erreur quadratique moyenne ($RMSE = 9,92$) et l'erreur absolue moyenne ($MAE = 3,82$). Ils ont déclaré dans le résumé de leur étude que le modèle hybride VMD-DNN proposé constitue une nouvelle méthode prometteuse pour prédire le débit quotidien. (Xie et al., 2019) dans leur article, ils utilisent un nouveau modèle hybride, appelé VMD-DBN-IPSO, pour améliorer la précision de la prévision du ruissellement aux stations hydrologiques de Yangxian et d'Ankang dans le bassin du fleuve Han, en Chine. L'analyse en mode variable (VMD) est utilisée pour analyser la série de ruissellement quotidien d'origine, puis, grâce au modèle hybride combinant l'algorithme d'optimisation améliorée des essaims de particules (IPSO) et le réseau de croyance profonde (DBN), le ruissellement est prédit. Les résultats indiquent que le modèle VMD-DBN-IPSO peut toujours atteindre les meilleures performances dans la phase de formation et de test, qu'il présente une bonne stabilité et une bonne représentation, que le coefficient NSE reste supérieur à 0,8 et que l'erreur de prédiction du débit de pointe est inférieure à 20 %.

Le débit journalier estimé à l'aide de modèles hybrides semble être influencé de manière significative par les algorithmes d'analyse du signal. Notre étude confirme que la précision de prédiction était presque significativement meilleure que la performance obtenue en utilisant des modèles autonomes. Le modèle MODWT-GPR (sym4) a montré bien une meilleure performance en termes d'amélioration de la précision de prédiction à la première station (Bir Oulad Taher) qu'à la deuxième station (Ponteba Barrage). Cela peut en effet imposer certaines limites à notre étude, car le développement de modèles hybrides de prévision des débits peut varier d'une station à l'autre en raison des différentes caractéristiques des non-linéarités dans les séries de précipitations et d'hydrométrie. D'autre part, la meilleure performance du modèle MODWT-GPR (sym) pour les deux stations était dans la troisième combinaison de variables d'entrée contenant les précipitations (P_t) et le débit pour le jour précédent (Q_{t-1}) et les deux jours précédents (Q_{t-2}), alors que la performance du modèle n'était pas la même pour les deux stations. Notre compréhension scientifique est que lors de l'utilisation de ACF, PACF et XCF pour vérifier les fortes corrélations entre les précipitations et les débits transportés et pour sélectionner les meilleures variables d'entrée pour les modèles d'apprentissage automatique autonomes et hybrides, le temps de concentration de chaque bassin versant doit être pris en compte, Parce qu'il s'agit d'un facteur important pour

connaître le comportement et l'étendue du ruissellement de surface de chaque bassin hydrographique.

Conclusion

Cette étude vise à analyser la relation entre les précipitations et le débit des cours d'eau, ainsi qu'à prédire ce débit à l'échelle journalière dans le bassin du Cheliff, au nord de l'Algérie, en utilisant des modèles d'apprentissage automatique tels que la Régression par Processus Gaussien (GPR), la Mémoire à Long Terme (LSTM), le Réseau Neuronal à Fonction de Base Radiale Gaussienne (GRNN) et le Perceptron Multicouche (MLPNN).

Dans le premier scénario de cette étude, l'efficacité de chaque modèle a été évaluée sur la base de diverses mesures statistiques, et leurs forces et limites ont été discutées. Les résultats indiquent que le modèle LSTM a surpassé les modèles GPR, MLPNN et GRNN dans la station de Bir Ouled Taher. Le modèle GPR a également montré des performances compétitives, en particulier dans la station de Ponteba Barrage, où il a surpassé les modèles LSTM, MLPNN et GRNN. Cependant, le modèle GRNN a généralement présenté des valeurs RMSE plus élevées par rapport aux autres modèles pour les deux stations. Malgré la supériorité des modèles GPR et LSTM dans le premier scénario de cette étude, leur performance n'est pas considérée comme satisfaisante pour prédire le débit journalier dans la zone d'étude.

Dans le second scénario, une analyse plus approfondie a exploré l'utilisation de la transformée en ondelettes à décomposition par recouvrement maximal (MODWT) avec les cinq mères algorithmes (Coiflets, Daubechies, Symlets, Haar et Fejer-Korovkin) en conjonction avec les quatre modèles ML pour la prédiction du débit journalier. Le modèle MODWT-GPR (sym4) a montré des résultats prometteurs tant dans la phase d'entraînement que dans la phase de validation, avec une excellente précision et une meilleure performance pour les deux stations. Ce modèle est apparu comme le plus performant, obtenant systématiquement les valeurs RMSE les plus faibles pour différentes combinaisons d'entrées. Il a démontré une amélioration significative par rapport aux modèles hybrides et aux modèles autonomes, soulignant son efficacité à capturer la relation complexe entre les précipitations et le ruissellement, et à prédire avec précision le débit. L'intégration des transformées en ondelettes de la décomposition par chevauchement maximal (sym4) au modèle GPR a joué un rôle crucial dans l'amélioration de la précision de la prédiction.

En outre, l'étude a examiné l'impact de l'incorporation de différents composants sur la précision de la prédiction du débit des cours d'eau. Les résultats ont révélé que la troisième combinaison de variables d'entrée, qui incorporait des composantes et des variables spécifiques, atteignait systématiquement les valeurs RMSE les plus basses, tant dans la phase d'apprentissage que dans la phase de test. Cette combinaison des variables d'entrée a démontré une précision supérieure dans la prédiction du débit et s'est avérée être le choix le plus favorable parmi les scénarios présentés.

Dans l'ensemble, cette étude contribue à la compréhension des relations entre les précipitations et le ruissellement, et fournit des informations précieuses sur les performances de différents modèles pour la prévision du débit des cours d'eau. Le modèle MODWT-GPR, en particulier lorsqu'il utilise l'ondelette mère ou la fonction Symlets d'ordre 4, apparaît comme une approche robuste et précise pour la prévision du débit journalier, avec des applications potentielles dans la gestion des ressources en eau, la prévision des inondations et la modélisation hydrologique. Cependant, des recherches supplémentaires sont nécessaires pour étudier la performance de ces modèles sur divers ensembles de données et dans des conditions hydrologiques variées.

CONCLUSION GENERALE

Conclusion générale

L'objectif principal de cette étude est d'améliorer les prévisions hydrologiques des débits des cours d'eau aux pas de temps journaliers et mensuels dans un milieu semi-aride. La recherche sur les prévisions hydrologiques favorise l'innovation en matière de technologie, de techniques d'analyse des données et d'outils informatiques. Elle encourage le développement d'algorithmes sophistiqués, de méthodes d'apprentissage automatique et d'applications de télédétection pour améliorer la collecte, le traitement et l'interprétation des données. Pour cela, nous avons cherché à développer des modèles d'intelligence artificielle et d'apprentissage automatique autonomes et hybrides pour améliorer les prévisions des débits des cours d'eau aux pas de temps journaliers et mensuels dans cinq sous-bassins versants du Cheliff, au nord de l'Algérie : (i) Ebda, (ii) Cheliff Harreza, (iii) Tiguiguest, (iv) Rouina Zeddine et (v) Ras Ouahrane. Le développement de modèles d'IA prédictifs repose sur les données des séries de précipitations (P) et hydrométriques (Q) fournies par l'ANRH. Atteindre les principaux objectifs de cette étude a nécessité de passer par plusieurs étapes, chacune ayant un but primordial essentiel pour atteindre l'étape suivante.

La première contribution de cette thèse consiste tout d'abord à tester la faisabilité des modèles d'intelligence artificielle autonomes pour prédire le débit journalier en utilisant MLPNN, GRNN, SVR, GPR et ELM avec différentes configurations d'entrée. La sélection des ensembles de données d'entrée des modèles était constituée des précipitations et des débits antérieurs, effectuée à l'aide des fonctions d'autocorrélation (ACF), d'autocorrélation partielle (PACF) et de corrélation croisée (XCF). Après le test de faisabilité sur la performance des modèles autonomes sur la prévision des débits journaliers dans les bassins versants (i), (ii) et (iii), un modèle hybride a été développé pour améliorer la prévision des débits journaliers dans les trois bassins proposés dans l'étude. Ce modèle combine la machine d'apprentissage extrême (ELM) avec l'algorithme métaheuristique BAT. De plus, les modèles prédictifs autonomes MLPNN, GRNN, SVR, GPR et ELM, ainsi que le modèle hybride BAT-EM, ont été implémentés dans la prévision des débits au pas de temps mensuel en retenant les mêmes scénarios que ceux proposés pour la prédiction des débits journaliers. Les résultats révélés sont :

1. Les performances du modèle hybride BAT-ELM se sont considérablement améliorées aux pas de temps journaliers et mensuels par rapport aux performances des MLPNN, GRNN, SVR, GPR et du modèle ELM autonome.
2. La comparaison des cinq modèles autonomes MLPNN, GRNN, SVR, GPR et ELM a révélé qu'ils étaient presque égaux et qu'il existait de légères différences entre eux.
3. Parmi les six combinaisons d'entrées testées dans la présente étude, les précisions de prévision les plus faibles ont été obtenues aux pas de temps journaliers et mensuels pour la quatrième, cinquième et sixième combinaison d'entrées. De plus, le BAT-ELM n'a pas réussi à améliorer la précision de prévision pour la cinquième combinaison dans laquelle seul le débit a été utilisé.
4. L'algorithme BAT est très efficace pour optimiser les paramètres du modèle ELM et améliorer ses performances de prédiction aux pas de temps journaliers et mensuels.

La deuxième contribution de cette thèse consiste à établir et développer une méthode plus fiable qui donne des résultats de prévision robustes aux pas de temps journaliers. Cette méthode est basée sur l'hybridation de modèles d'apprentissage automatique autonomes (GPR, LSTM, GRNN et MLPNN) avec cinq familles de transformations en ondelettes discrètes à recouvrement maximal (MODWT) : Coiflets, Daubechies, Symlets, Haar et Fejer-Korovkin. Les prédictions ont été effectuées dans les bassins versants (iv) et (v) de Cheliff. Les résultats obtenus sont :

1. Les modèles développés MODWT-GPR (sym4) ont montré une efficacité importante dans la prévision des débits des cours d'eau.
2. Les modèles développés MODWT-GPR (sym4) ont fourni de meilleures performances par rapport aux modèles d'apprentissage automatique autonomes aux deux stations étudiées pour les prévisions à l'échelle journalière.
3. Le MODWT-GPR (sym4) a donné de meilleures performances par rapport aux modèles hybrides MODWT-GPR (coif1), MODWT-GPR (db3), MODWT-GPR (haar) et MODWT-GPR (fk8) aux deux stations étudiées pour les prévisions journalières.
4. Le MODWT-GPR (sym4) présenté a donné de meilleures performances par rapport aux modèles hybrides MODWT-LSTM (coif1), MODWT- LSTM (db3), MODWT- LSTM (sym4), MODWT- LSTM (haar) et MODWT- LSTM (fk8) aux deux stations étudiées pour les prévisions journalières.

5. Le MODWT-GPR (sym4) a montré de meilleures performances par rapport aux modèles hybrides MODWT-GRNN (coif1), MODWT- GRNN (db3), MODWT- GRNN (sym4), MODWT- GRNN (haar) et MODWT- GRNN (fk8) aux deux stations étudiées pour les prévisions journalières.
6. Le MODWT-GPR (sym4) a donné de meilleures performances par rapport aux modèles hybrides MODWT-MLPNN (coif1), MODWT- MLPNN (db3), MODWT- MLPNN (sym4), MODWT- MLPNN (haar) et MODWT- MLPNN (fk8) aux deux stations étudiées pour les prévisions journalières.

Dans l'ensemble, cette étude contribue à la compréhension des relations entre les précipitations et le débit et fournit des informations précieuses sur les performances de différents modèles pour la prédiction du débit des cours d'eau aux pas de temps journaliers et mensuels. L'hybridation de l'ELM avec l'algorithme BAT a amélioré les performances de l'ELM et obtenu une précision prédictive élevée. De plus, il a été constaté que le BAT-ELM a produit des performances numériques élevées avec des erreurs de prédiction minimales, et que le modèle semblait représenter une approche robuste et précise pour la prévision des débits des cours d'eaux journaliers et mensuels. Le modèle MODWT-GPR, en particulier lorsqu'il utilise l'ondelette mère ou la fonction Symlets d'ordre 4, apparaît comme une méthode robuste et précise pour la prévision du débit journalier, avec des applications potentielles dans la gestion des ressources en eau et la prédiction des inondations, etc. Cependant, des recherches supplémentaires sont nécessaires pour étudier les performances de ces modèles sur divers ensembles de données et dans différentes conditions hydrologiques.

REFERENCES BIBLIOGRAPHIQUES

Références bibliographiques

Abda, Z., Zerouali, B., Chettih, M., Guimaraes Santos, C. A., de Farias, C. A. S., & Elbeltagi, A. (2022). Assessing machine learning models for streamflow estimation: a case study in Oued Sebaou watershed (Northern Algeria). *Hydrological Sciences Journal*, 67(9), 1328-1341. <https://doi.org/10.1080/02626667.2022.2083511>

Abdel-Nasser, M., & Mahmoud, K. (2019). Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural computing and applications*, 31, 2727-2740. <https://doi.org/10.1007/s00521-017-3225-z>

Abraham, T. H. (2002). (Physio) logical circuits: The intellectual origins of the McCulloch–Pitts neural networks. *Journal of the History of the Behavioral Sciences*, 38(1), 3-25. <https://doi.org/10.1002/jhbs.1094>

Adamowski, J., Chan, H. F., Prasher, S. O., & Sharda, V. N. (2012). Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. *Journal of hydroinformatics*, 14(3), 731-744. <https://doi.org/10.2166/hydro.2011.044>

Adnan, R. M., Kisi, O., Mostafa, R. R., Ahmed, A. N., & El-Shafie, A. (2022). The potential of a novel support vector machine trained with modified mayfly optimization algorithm for streamflow prediction. *Hydrological Sciences Journal*, 67(2), 161-174. <https://doi.org/10.1080/02626667.2021.2012182>

Adnan, R. M., Liang, Z., Trajkovic, S., Zounemat-Kermani, M., Li, B., & Kisi, O. (2019). Daily streamflow prediction using optimally pruned extreme learning machine. *Journal of Hydrology*, 577, 123981. <https://doi.org/10.1016/j.jhydrol.2019.123981>

Adnan, R. M., Mostafa, R. R., Kisi, O., Yaseen, Z. M., Shahid, S., & Zounemat-Kermani, M. (2021). Improving streamflow prediction using a new hybrid ELM model combined with hybrid particle swarm optimization and grey wolf optimization. *Knowledge-Based Systems*, 230, 107379. <https://doi.org/10.1016/j.knosys.2021.107379>

Agarwal, S., Roy, P., Choudhury, P., & Debbarma, N. (2022). Comparative study on stream flow prediction using the GMNN and wavelet-based GMNN. *Journal of Water and Climate Change*, 13(9), 3323-3337. <https://doi.org/10.2166/wcc.2022.226>

Aguiar- Conraria, L., & Soares, M. J. (2014). The continuous wavelet transform: Moving beyond uni- and bivariate analysis. *Journal of Economic Surveys*, 28(2), 344-375. <https://doi.org/10.1111/joes.12012>

Akçakoca, F., & Apaydın, H. (2020). Modelling of Bektas Creek daily streamflow with generalized regression neural network method. *International Journal of Advances in Scientific Research and Engineering*, 6(2), 97-103. DOI: [10.31695/IJASRE.2020.33884](https://doi.org/10.31695/IJASRE.2020.33884)

- Akhavan-Amjadi, M. (2020). Fetal electrocardiogram modeling using hybrid evolutionary firefly algorithm and extreme learning machine. *Multidimensional Systems and Signal Processing*, 31, 117-133. DOI: [10.1007/s11045-019-00669-7](https://doi.org/10.1007/s11045-019-00669-7)
- Alshamiri, A. K., Singh, A., & Surampudi, B. R. (2016). Artificial bee colony algorithm for clustering: an extreme learning approach. *Soft Computing*, 20, 3163-3176. DOI: [10.1007/s00500-015-1747-0](https://doi.org/10.1007/s00500-015-1747-0)
- Ambroise, B. J. F., Modèles. H G* A*, Bucarest. (1998). La dynamique du cycle de l'eau dans un bassin versant-Processus.
- Andersson, E., Bauer, P., Beljaars, A., Chevallier, F., Hólm, E., Janisková, M., Kållberg, P., Kelly, G., Lopez, P., & McNally, A. (2005). Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system. *Bulletin of the American Meteorological Society*, 86(3), 387-402. DOI: [10.1175/BAMS-86-3-387](https://doi.org/10.1175/BAMS-86-3-387)
- Andréassian, V. (2002). Impact de l'évolution du couvert forestier sur le comportement hydrologique des bassins versants. Doctorat en Géosciences et ressources naturelles spécialité hydrologie.
- Aoulmi, Y., Marouf, N., & Amireche, M. (2022). Modélisation pluie-débit et débits de crues en milieu semi-aride par différentes méthodes statistiques et modèles hydrologiques.
- Aoulmi, Y., Marouf, N., Amireche, M., Kisi, O., Shubair, R. M., & Keshtegar, B. (2021). Highly accurate prediction model for daily runoff in semi-arid basin exploiting Metaheuristic learning algorithms. *IEEE Access*, 9, 92500-92515. DOI: [10.1109/ACCESS.2021.3092074](https://doi.org/10.1109/ACCESS.2021.3092074)
- Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006). Multi-time scale stream flow predictions: The support vector machines approach. *Journal of Hydrology*, 318(1-4), 7-16. DOI: [10.1016/j.jhydrol.2005.06.017](https://doi.org/10.1016/j.jhydrol.2005.06.017)
- Atiquzzaman, M., & Kandasamy, J. (2016). Prediction of hydrological time-series using extreme learning machine. *Journal of Hydroinformatics*, 18(2), 345-353. DOI: [10.2166/hydro.2015.176](https://doi.org/10.2166/hydro.2015.176)
- Bakhshi Ostadkalayeh, F., Moradi, S., Asadi, A., Moghaddam Nia, A., & Taheri, S. (2023). Performance improvement of LSTM-based deep learning model for streamflow forecasting using Kalman filtering. *Water resources management*, 1-17. DOI: [10.1007/s11269-023-03218-4](https://doi.org/10.1007/s11269-023-03218-4)
- Barzegar, R., Aalami, M. T., & Adamowski, J. (2021). Coupling a hybrid CNN-LSTM deep learning model with a boundary corrected maximal overlap discrete wavelet transform for multiscale lake water level forecasting. *Journal of Hydrology*, 598, 126196. DOI: [10.1016/j.jhydrol.2021.126196](https://doi.org/10.1016/j.jhydrol.2021.126196)
- Bebis, G., & Georgiopoulos, M. (1994). Feed-forward neural networks. *Ieee Potentials*, 13(4), 27-31. DOI: [10.1109/45.312853](https://doi.org/10.1109/45.312853)

Bennis, S., & Bruneau, P. J. C. J. o. C. E. (1993). Amélioration de méthodes d'estimation des débits journaliers. 20(3), 490-499.

Besaw, L. E., Rizzo, D. M., Bierman, P. R., & Hackett, W. R. (2010). Advances in ungauged streamflow prediction using artificial neural networks. *Journal of Hydrology*, 386(1-4), 27-37. DOI: [10.1016/j.jhydrol.2010.03.027](https://doi.org/10.1016/j.jhydrol.2010.03.027)

Beven*, K. (2001). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1), 1-12. DOI: [10.5194/hess-5-1-2001](https://doi.org/10.5194/hess-5-1-2001)

Beven, K. J. J. o. h. (1989). Changing ideas in hydrology—the case of physically-based models. 105(1-2), 157-172. DOI: [10.1007/BF01230135](https://doi.org/10.1007/BF01230135)

Bhasme, P., & Bhatia, U. (2024). Improving the interpretability and predictive power of hydrological models: Applications for daily streamflow in managed and unmanaged catchments. *Journal of Hydrology*, 628, 130421. DOI: [10.1016/j.jhydrol.2022.130421](https://doi.org/10.1016/j.jhydrol.2022.130421)

Biancamaria, S. (2020). Variabilité du cycle hydrologique des grands bassins fluviaux. Université Toulouse III-Paul Sabatier.

Block, P. J., Souza Filho, F. A., Sun, L., & Kwon, H. H. (2009). A streamflow forecasting framework using multiple climate and hydrological models. *Journal of the American Water Resources Association*, 45(4), 828-843. DOI: [10.1111/j.1752-1688.2009.00328.x](https://doi.org/10.1111/j.1752-1688.2009.00328.x)

Bouaziz, M. (2017). Réseaux de neurones récurrents pour la classification de séquences dans des flux audiovisuels parallèles.

Cameron, D., Beven, K. J., Tawn, J., Blazkova, S., & Naden, P. (1999). Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *Journal of Hydrology*, 219(3-4), 169-187. DOI: [10.1016/S0022-1694\(99\)00075-8](https://doi.org/10.1016/S0022-1694(99)00075-8)

Carpenter, G. A., & Grossberg, S. (1991). Pattern recognition by self-organizing neural networks. MIT Press. DOI: [10.7551/mitpress/9780262034118.001.0001](https://doi.org/10.7551/mitpress/9780262034118.001.0001)

Carriere, P., Mohaghegh, S., & Gaskari, R. (1996). Performance of a virtual runoff hydrograph system. *Journal of Water Resources Planning and Management*, 122(6), 421-427.

DOI: [10.1061/\(ASCE\)0733-9496\(1996\)122:6\(421\)](https://doi.org/10.1061/(ASCE)0733-9496(1996)122:6(421))

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific model development*, 7(3), 1247-1250. DOI: [10.5194/gmd-7-1247-2014](https://doi.org/10.5194/gmd-7-1247-2014)

Chakraborty, S., & Biswas, S. (2023). River discharge prediction using wavelet-based artificial neural network and long short-term memory models: a case study of Teesta River Basin, India. *Stochastic Environmental Research and Risk Assessment*, 1-22.

DOI: [10.1007/s00477-023-02023-3](https://doi.org/10.1007/s00477-023-02023-3)

Chamani, M., & Roushangar, K. (2020). Evaluation of CEEMD-GPR hybrid model in temporal and spatial daily stream flow forecasting. *Water and Irrigation Management*, 9(2), 277-289.

Charbonneau, P., & Tremblay, G. (2010). Création d'une banque référence pour l'identification des chauves-souris au Québec. *Le Naturaliste canadien*, 134, 50-61.

Chavan, M. S., Mastorakis, N., Chavan, M. N., & Gaikwad, M. (2011). Implementation of SYMLET wavelets to removal of Gaussian additive noise from speech signal. *Proceedings of Recent Researches in Communications, Automation, Signal Processing, Nanotechnology, Astronomy and Nuclear Physics: 10th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications (EHAC'11)*, Cambridge.

Chekroun, S. (2009). *Commande neuro-floue sans capteur de vitesse d'une machine asynchrone triphasée*. Mémoire de Magister, ENSET d'Oran.

Cheng, Y., Xu, C., Mashima, D., Thing, V. L., & Wu, Y. (2017). PowerLSTM: power demand forecasting using long short-term memory neural network. *Advanced Data Mining and Applications: 13th International Conference, ADMA 2017, Singapore, November 5–6, 2017, Proceedings 13*.

Chiang, Y.-M., Chang, L.-C., & Chang, F.-J. (2004). Comparison of static-feedforward and dynamic-feedback neural networks for rainfall–runoff modeling. *Journal of Hydrology*, 290(3-4), 297-311. DOI: [10.1016/j.jhydrol.2003.12.012](https://doi.org/10.1016/j.jhydrol.2003.12.012)

Chidepudi, S. K. R., Massei, N., Jardani, A., Henriot, A., Allier, D., & Baulon, L. (2023). A wavelet-assisted deep learning approach for simulating groundwater levels affected by low-frequency variability. *Science of The Total Environment*, 865, 161035.

DOI: [10.1016/j.scitotenv.2021.161035](https://doi.org/10.1016/j.scitotenv.2021.161035)

Chun, S., Liu, Y. Z., & Sung, H. J. (2007). Multi-resolution analysis of the large-scale coherent structure in a turbulent separation bubble affected by an unsteady wake. *Journal of fluids and structures*, 23(1), 85-100. DOI: [10.1016/j.jfluidstructs.2006.07.009](https://doi.org/10.1016/j.jfluidstructs.2006.07.009)

Cigizoglu, H. K. (2005a). Application of generalized regression neural networks to intermittent flow forecasting and estimation. *Journal of Hydrologic Engineering*, 10(4), 336-341.

DOI: [10.1061/\(ASCE\)1084-0699\(2005\)10:4\(336\)](https://doi.org/10.1061/(ASCE)1084-0699(2005)10:4(336))

Cigizoglu, H. K. (2005b). Generalized regression neural network in monthly flow forecasting. *Civil Engineering and Environmental Systems*, 22(2), 71-81. DOI: [10.1080/10286600500097321](https://doi.org/10.1080/10286600500097321)

Cigizoglu, H. K., & Alp, M. (2006). Generalized regression neural network in modelling river sediment yield. *Advances in Engineering Software*, 37(2), 63-68.

DOI: [10.1016/j.advengsoft.2005.02.001](https://doi.org/10.1016/j.advengsoft.2005.02.001)

Clarke, R. J. (1973). A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of Hydrology*, 19(1), 1-20.

DOI: [10.1016/0022-1694\(73\)90021-8](https://doi.org/10.1016/0022-1694(73)90021-8)

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20, 273-297.

DOI: [10.1023/A:1022627411411](https://doi.org/10.1023/A:1022627411411)

Corzo, G., & Solomatine, D. (2007). Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. *Hydrological Sciences Journal*, 52(3), 491-507.

DOI : [10.1623/hysj.52.3.491](https://doi.org/10.1623/hysj.52.3.491)

Coulibaly, P., Anctil, F., & Bobée, B. (1999). Prévision hydrologique par réseaux de neurones artificiels: état de l'art. *Canadian Journal of civil engineering*, 26(3), 293-304.

DOI: [10.1139/199-040](https://doi.org/10.1139/199-040)

Cunningham, P., Cord, M., & Delany, S. J. (2008). Supervised learning. In *Machine learning techniques for multimedia: case studies on organization and retrieval* (pp. 21-49). Springer.

Da Ros, D., & Borga, M. (1997). Adaptive use of a conceptual model for real-time flood forecasting. *Hydrological Research*, 28(3), 169-188. DOI: [10.2166/nh.1997.0010](https://doi.org/10.2166/nh.1997.0010)

Dai, R., Wang, W., Zhang, R., & Yu, L. (2024). Multimodal deep learning water level forecasting model for multiscale drought alert in Feiyun River basin. *Expert Systems with Applications*, 244, 122951. DOI: [10.1016/j.eswa.2021.122951](https://doi.org/10.1016/j.eswa.2021.122951)

Danandeh Mehr, A., Kahya, E., Şahin, A., & Nazemosadat, M. (2015). Successive-station monthly streamflow prediction using different artificial neural network algorithms. *International Journal of Environmental Science and Technology*, 12, 2191-2200.

DOI: [10.1007/s13762-014-0594-2](https://doi.org/10.1007/s13762-014-0594-2)

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41(7), 909-996. DOI: [10.1002/cpa.3160410705](https://doi.org/10.1002/cpa.3160410705).

Dawson, C., & Wilby, R. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1), 80-108. DOI: [10.1177/030913330102500105](https://doi.org/10.1177/030913330102500105)

Deo, R. C., & Şahin, M. (2016). An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. *Environmental Monitoring and Assessment*, 188(1), 1-24. DOI: [10.1007/s10661-015-5061-4](https://doi.org/10.1007/s10661-015-5061-4)

Diop, L., Bodian, A., Djaman, K., Yaseen, Z. M., Deo, R. C., El-Shafie, A., & Brown, L. C. (2018). The influence of climatic inputs on stream-flow pattern forecasting: case study of Upper Senegal River. *Environmental Earth Sciences*, 77, 1-13. DOI: [10.1007/s12665-018-7279-y](https://doi.org/10.1007/s12665-018-7279-y)

Dong, J., Wang, Z., Wu, J., Cui, X., & Pei, R. (2024). A Novel Runoff Prediction Model Based on Support Vector Machine and Gate Recurrent unit with Secondary Mode Decomposition. *Water Resources Management*, 1-20.

Doucoure, B., Agbossou, K., & Cardenas, A. (2016). Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data. *Renewable Energy*, 92, 202-211. DOI: [10.1016/j.renene.2016.01.014](https://doi.org/10.1016/j.renene.2016.01.014)

Elbeltagi, A., Salam, R., Pal, S. C., Zerouali, B., Shahid, S., Mallick, J., ... & Islam, A. R. M. T. (2022). Groundwater level estimation in the northern region of Bangladesh using hybrid locally weighted linear regression and Gaussian process regression modeling. *Theoretical and Applied Climatology*, 149(1-2), 131-151. DOI: [10.1007/s00704-020-03481-w](https://doi.org/10.1007/s00704-020-03481-w)

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179-211.

DOI: [10.1016/0364-0213\(90\)90002-E](https://doi.org/10.1016/0364-0213(90)90002-E)

Fang, Z., Wang, Y., Peng, L., & Hong, H. (2021). Predicting flood susceptibility using LSTM neural networks. *Journal of Hydrology*, 594, 125734. DOI: [10.1016/j.jhydrol.2020.125734](https://doi.org/10.1016/j.jhydrol.2020.125734)

Feng, R., Fan, G., Lin, J., Yao, B., & Guo, Q. (2021). Enhanced long short-term memory model for runoff prediction. *Journal of Hydrologic Engineering*, 26(2), 04020063. DOI: [10.1061/\(ASCE\)HE.1943-5584.0002034](https://doi.org/10.1061/(ASCE)HE.1943-5584.0002034)

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669. DOI: [10.1016/j.ejor.2018.04.017](https://doi.org/10.1016/j.ejor.2018.04.017)

Fortin, V., Ouarda, T., Rasmussen, P., & Bobée, B. (1997). Revue bibliographique des méthodes de prévision des débits. *Revue des sciences de l'eau*, 10(4), 461-487.

Frasconi, P., Gori, M., & Sperduti, A. (1998). A general framework for adaptive processing of data structures. *IEEE Transactions on Neural Networks*, 9(5), 768-786. DOI: [10.1109/72.712178](https://doi.org/10.1109/72.712178)

Gerard, L., & BERTHET, L. (2010). Sensibilité des performances d'un modèle de prévision des crues au critère de calage. *Mem. Master, Eau Environnement*. INP Toulouse.

Gers, F. A., & Schmidhuber, E. (2001). LSTM recurrent networks learn simple context-free and context-sensitive languages. *IEEE Transactions on Neural Networks*, 12(6), 1333-1340. DOI: [10.1109/72.963769](https://doi.org/10.1109/72.963769)

Ghasempour, R., Azamathulla, H. M., & Roushangar, K. (2021). EEMD- and VMD-based hybrid GPR models for river streamflow point and interval predictions. *Water Supply*, 21(7), 3960-3975. DOI: [10.2166/ws.2021.139](https://doi.org/10.2166/ws.2021.139)

Ghumman, A. R., Ghazaw, Y. M., Sohail, A. R., & Watanabe, K. (2011). Runoff forecasting by artificial neural network and conventional model. *Alexandria Engineering Journal*, 50(4), 345-350. DOI: [10.1016/j.aej.2011.10.004](https://doi.org/10.1016/j.aej.2011.10.004)

Gnouma, R. (2006). Aide à la calibration d'un modèle hydrologique distribué au moyen d'une analyse des processus hydrologiques: application au bassin versant de l'Yzeron. Thèse de doctorat, INSA Lyon.

Gomaa, E., Zerouali, B., Difi, S., El-Nagdy, K. A., Santos, C. A. G., Abda, Z., Ghoneim, S. S., Bailek, N., da Silva, R. M., & Rajput, J. (2023). Assessment of hybrid machine learning algorithms using TRMM rainfall data for daily inflow forecasting in Três Marias Reservoir, eastern Brazil. *Heliyon*, 9(8).

Grossmann, A., Kronland-Martinet, R., & Morlet, J. (1990). Reading and understanding continuous wavelet transforms. *Wavelets: Time-Frequency Methods and Phase Space Proceedings of the International Conference, Marseille, France, December 14–18, 1987*.

Guf, J.-S., & Jiang, W.-S. (1996). The Haar wavelets operational matrix of integration. *International Journal of Systems Science*, 27(7), 623-628. DOI: [10.1080/00207729608929109](https://doi.org/10.1080/00207729608929109)

Gupta, R. S. (2016). *Hydrology and hydraulic systems*. Waveland Press.

Halff, A. H., Halff, H. M., & Azmoodeh, M. (1993). Predicting runoff from rainfall using neural networks. *Engineering Hydrology*.

Han, H., & Morrison, R. R. (2022). Improved runoff forecasting performance through error predictions using a deep-learning approach. *Journal of Hydrology*, 608, 127653. DOI: [10.1016/j.jhydrol.2021.127653](https://doi.org/10.1016/j.jhydrol.2021.127653)

Han, L., Jing, H., Zhang, R., & Gao, Z. (2019). Wind power forecast based on improved Long Short Term Memory network. *Energy*, 189, 116300. DOI: [10.1016/j.energy.2019.116300](https://doi.org/10.1016/j.energy.2019.116300)

Hassan, K. S. M., Huang, Y. F., Koo, C. H., Weng, T. K., Ahmed, A. N., & Elshafie, A. H. K. A. (2022). Forecasting of Reservoir Inflow Using Machine Learning—Case Study: Klang Gate Dam Reservoir. *Proceedings of International Conference on Emerging Technologies and Intelligent Systems: ICETIS 2021 Volume 2*.

He, F., Zhang, H., Wan, Q., Chen, S., & Yang, Y. (2023). Medium Term Streamflow Prediction Based on Bayesian Model Averaging Using Multiple Machine Learning Models. *Water*, 15(8), 1548. DOI: [10.3390/w15081548](https://doi.org/10.3390/w15081548)

He, X., Luo, J., Zuo, G., & Xie, J. (2019). Daily runoff forecasting using a hybrid model based on variational mode decomposition and deep neural networks. *Water Resources Management*, 33, 1571-1590. DOI: [10.1007/s11269-019-02225-9](https://doi.org/10.1007/s11269-019-02225-9)

- Hidayat, R., Bejo, A., Sumaryono, S., & Winursito, A. (2018). Denoising speech for MFCC feature extraction using wavelet transformation in speech recognition system. 2018 10th international conference on information technology and electrical engineering (ICITEE).
- Hingray, B., Picouet, C., & Musy, A. (2009). Hydrologie: Une science pour l'ingénieur (Vol. 21). PPUR presses polytechniques.
- Hingray, B., Picouet, C., & Musy, A. (2009). Hydrologie 2.
- Hinton, G. E. (1992). How neural networks learn from experience. *Scientific American*, 267(3), 144-151.
- Hipel, K. W., & McLeod, A. I. (1994). Time series modelling of water resources and environmental systems. Elsevier.
- Ho, K., Prokopiw, W., & Chan, Y. (1995). Modulation identification by the wavelet transform. Proceedings of MILCOM'95.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Hosseinzadeh, P., Nassar, A., Boubrahimi, S. F., & Hamdi, S. M. (2023). ML-Based Streamflow Prediction in the Upper Colorado River Basin Using Climate Variables Time Series Data. *Hydrology*, 10(2), 29. DOI: [10.3390/hydrology10020029](https://doi.org/10.3390/hydrology10020029)
- Hsu, K.-l., Gupta, H. V., & Sorooshian, S. (1997). Application of a recurrent neural network to rainfall-runoff modeling. *Aesthetics in the Constructed Environment*.
- Hu, C., Wu, Q., Li, H., Jian, S., Li, N., & Lou, Z. (2018). Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. *Water*, 10(11), 1543.
DOI: [10.3390/w10111543](https://doi.org/10.3390/w10111543)
- Hu, H., Zhang, J., & Li, T. (2020). A comparative study of VMD-based hybrid forecasting model for nonstationary daily streamflow time series. *Complexity*, 2020, 1-21.
DOI: [10.1155/2020/5795103](https://doi.org/10.1155/2020/5795103)
- Hu, Y., Yan, L., Hang, T., & Feng, J. (2020). Stream-flow forecasting of small rivers based on LSTM. arXiv preprint arXiv:2001.05681. DOI: [10.1109/ICWICN51207.2020.9206751](https://doi.org/10.1109/ICWICN51207.2020.9206751)
- Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., & Zhang, H. (2019). Deep learning with long short-term memory for time series prediction. *IEEE Communications Magazine*, 57(6), 114-119. DOI: [10.1109/MCOM.001.1800247](https://doi.org/10.1109/MCOM.001.1800247)
- Huang, G.-B., Chen, L., & Siew, C. K. (2006). Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Trans. Neural Networks*, 17(4), 879-892. DOI: [10.1109/TNN.2006.875971](https://doi.org/10.1109/TNN.2006.875971)

- Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3), 489-501. DOI: [10.1016/j.neucom.2005.12.126](https://doi.org/10.1016/j.neucom.2005.12.126)
- Iamampai, S., Talaluxmana, Y., Kanasut, J., & Rangsiwanichpong, P. (2024). Enhancing rainfall-runoff model accuracy with machine learning models by using soil water index to reflect runoff characteristics. *Water Science & Technology*, 89(2), 368-381. DOI: [10.2166/wst.2022.454](https://doi.org/10.2166/wst.2022.454)
- Ikram, R. M. A., Goliatt, L., Kisi, O., Trajkovic, S., & Shahid, S. (2022). Covariance matrix adaptation evolution strategy for improving machine learning approaches in streamflow prediction. *Mathematics*, 10(16), 2971. DOI: [10.3390/math10162971](https://doi.org/10.3390/math10162971)
- Islam, M., Liong, S., Phoon, K., & Liaw, C. (2001). Forecasting of river flow data with a general regression neural network. *Iahs Publication*, 285-290.
- Jiang, Y., Bao, X., Hao, S., Zhao, H., Li, X., & Wu, X. (2020). Monthly streamflow forecasting using ELM-IPSO based on phase space reconstruction. *Water resources management*, 34, 3515-3531. DOI: [10.1007/s11269-020-02637-y](https://doi.org/10.1007/s11269-020-02637-y)
- Kalra, A., Miller, W. P., Lamb, K. W., Ahmad, S., & Piechota, T. (2013). Using large-scale climatic patterns for improving long lead time streamflow forecasts for Gunnison and San Juan River Basins. *Hydrological Processes*, 27(11), 1543-1559. DOI: [10.1002/hyp.9302](https://doi.org/10.1002/hyp.9302)
- Kalteh, A. M. (2015). Wavelet genetic algorithm-support vector regression (wavelet GA-SVR) for monthly flow forecasting. *Water resources management*, 29, 1283-1293. DOI: [10.1007/s11269-014-0880-2](https://doi.org/10.1007/s11269-014-0880-2)
- Karbasi, M. (2018). Forecasting of multi-step ahead reference evapotranspiration using wavelet-Gaussian process regression model. *Water resources management*, 32, 1035-1052. DOI: [10.1007/s11269-017-1833-1](https://doi.org/10.1007/s11269-017-1833-1)
- Karunanithi, N., Grenney, W. J., Whitley, D., & Bovee, K. (1994). Neural networks for river flow prediction. *Journal of computing in civil engineering*, 8(2), 201-220.
- Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 173, 24-49. DOI: [10.1016/j.isprsjprs.2021.12.001](https://doi.org/10.1016/j.isprsjprs.2021.12.001)
- Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95-international conference on neural networks*.
- Khosravi, K., Golkarian, A., & Tiefenbacher, J. P. (2022). Using optimized deep learning to predict daily streamflow: a comparison to common machine learning algorithms. *Water resources management*, 36(2), 699-716. DOI: [10.1007/s11269-022-03155-5](https://doi.org/10.1007/s11269-022-03155-5)
- Kilinc, H. C. (2022). Daily streamflow forecasting based on the hybrid particle swarm optimization and long short-term memory model in the Orontes Basin. *Water*, 14(3), 490. DOI: [10.3390/w14030490](https://doi.org/10.3390/w14030490)

Kilinc, H. C., & Haznedar, B. (2022). A hybrid model for streamflow forecasting in the Basin of Euphrates. *Water*, 14(1), 80. DOI: [10.3390/w14010080](https://doi.org/10.3390/w14010080)

Kim, R. J., Loucks, D. P., & Stedinger, J. R. (2012). Artificial neural network models of watershed nutrient loading. *Water resources management*, 26, 2781-2797. DOI: [10.1007/s11269-012-0045-5](https://doi.org/10.1007/s11269-012-0045-5)

Kim, T., Yang, T., Gao, S., Zhang, L., Ding, Z., Wen, X., Gourley, J. J., & Hong, Y. (2021). Can artificial intelligence and data-driven machine learning models match or even replace process-driven hydrologic models for streamflow simulation?: A case study of four watersheds with different hydro-climatic regions across the CONUS. *Journal of Hydrology*, 598, 126423. DOI: [10.1016/j.jhydrol.2021.126423](https://doi.org/10.1016/j.jhydrol.2021.126423)

Kirkpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598), 671-680. DOI: [10.1126/science.220.4598.671](https://doi.org/10.1126/science.220.4598.671)

Kişi, Ö. (2008a). River flow forecasting and estimation using different artificial neural network techniques. *Hydrology Research*, 39(1), 27-40. DOI: [10.2166/nh.2008.030](https://doi.org/10.2166/nh.2008.030)

Kişi, Ö. (2008b). Stream flow forecasting using neuro- wavelet technique. *Hydrological Processes: An International Journal*, 22(20), 4142-4152. DOI: [10.1002/hyp.7065](https://doi.org/10.1002/hyp.7065)

Kişi, Ö. (2011). A combined generalized regression neural network wavelet model for monthly streamflow prediction. *KSCE Journal of Civil Engineering*, 15, 1469-1479. DOI: [10.1007/s12205-011-0157-9](https://doi.org/10.1007/s12205-011-0157-9)

Kisi, O., & Kerem Cigizoglu, H. (2007). Comparison of different ANN techniques in river flow prediction. *Civil Engineering and Environmental Systems*, 24(3), 211-231. DOI: [10.1080/10286600701336516](https://doi.org/10.1080/10286600701336516)

Kisi, O., Heddami, S., Keshtegar, B., Piri, J., & Adnan, R. M. (2022). Predicting daily streamflow in a cold climate using a novel data mining technique: Radial M5 Model Tree. *Water*, 14(9), 1449. DOI: [10.3390/w14101449](https://doi.org/10.3390/w14101449)

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22(11), 6005-6022. DOI: [10.5194/hess-22-6005-2018](https://doi.org/10.5194/hess-22-6005-2018)

Kronland-Martinet, R. (1988). The wavelet transform for analysis, synthesis, and processing of speech and music sounds. *Computer Music Journal*, 12(4), 11-20. DOI: [10.2307/3679804](https://doi.org/10.2307/3679804)

Kumar, M., Elbeltagi, A., Srivastava, A., Kumari, A., Ali, R., Pande, C., Bajirao, T. S., Islam, A. R. M. T., & Kushwaha, D. P. (2021). Prediction of Daily Streamflow Using Various Kernel Function Based Regression: A Case Study in India.

- Labbas, M. r. (2015). Modélisation hydrologique de bassins versants périurbains et influence de l'occupation du sol et de la gestion des eaux pluviales: Application au bassin de l'Yzeron (130km²) Université Grenoble Alpes].
- Laqrichi, S. (2015). Approach to build realistic models for estimating project effort/cost in an uncertain environment : application to the software development field
- Lark, R., & Webster, R. (2001). Changes in variance and correlation of soil properties with scale and location: analysis using an adapted maximal overlap discrete wavelet transform. *European journal of soil science*, 52(4), 547-562. DOI: [10.1046/j.1365-2389.2001.00401.x](https://doi.org/10.1046/j.1365-2389.2001.00401.x)
- Legates, D. R., & McCabe Jr, G. J. (1999). Evaluating the use of “goodness- of- fit” measures in hydrologic and hydroclimatic model validation. *Water resources research*, 35(1), 233-241. DOI: [10.1029/1998WR900018](https://doi.org/10.1029/1998WR900018)
- Leonard, J., & Kramer, M. A. (1990). Improvement of the backpropagation algorithm for training neural networks. *Computers & Chemical Engineering*, 14(3), 337-341. DOI: [10.1016/0098-1354\(90\)85032-S](https://doi.org/10.1016/0098-1354(90)85032-S)
- Lettenmaier, D. P., & Wood, E. F. J. H. o. h. (1993). Hydrologic forecasting. In (Vol. 26): McGraw-Hill New York.
- Li, B., & Cheng, C. (2014). Monthly discharge forecasting using wavelet neural networks with extreme learning machine. *Science China Technological Sciences*, 57, 2441-2452. DOI: [10.1007/s11431-014-5720-2](https://doi.org/10.1007/s11431-014-5720-2)
- Li, F., Ma, G., Chen, S., & Huang, W. (2021). An ensemble modeling approach to forecast daily reservoir inflow using bidirectional long-and short-term memory (Bi-LSTM), variational mode decomposition (VMD), and energy entropy method. *Water resources management*, 35, 2941-2963.
- Li, R., Wang, X., Song, Y., & Lei, L. (2021). Hierarchical extreme learning machine with L21-norm loss and regularization. *International Journal of Machine Learning and Cybernetics*, 12(5), 1297-1310. DOI: [10.1007/s13042-020-01189-4](https://doi.org/10.1007/s13042-020-01189-4)
- Liang, Z., Li, Y., Hu, Y., Li, B., & Wang, J. (2018). A data-driven SVR model for long-term runoff prediction and uncertainty analysis based on the Bayesian framework. *Theoretical and applied climatology*, 133, 137-149. DOI: [10.1007/s00704-017-2096-x](https://doi.org/10.1007/s00704-017-2096-x)
- Lima, A. R., Cannon, A. J., & Hsieh, W. W. (2016). Forecasting daily streamflow using online sequential extreme learning machines. *Journal of Hydrology*, 537, 431-443. DOI: [10.1016/j.jhydrol.2016.03.020](https://doi.org/10.1016/j.jhydrol.2016.03.020)
- Lin, L.-J. (1992). Reinforcement learning for robots using neural networks. Carnegie Mellon University.

Lin, Y., Wang, D., Wang, G., Qiu, J., Long, K., Du, Y., Xie, H., Wei, Z., Shangguan, W., & Dai, Y. (2021). A hybrid deep learning algorithm and its application to streamflow prediction. *Journal of Hydrology*, 601, 126636. DOI: [10.1016/j.jhydrol.2021.126636](https://doi.org/10.1016/j.jhydrol.2021.126636)

Lina, J.-M., & Mayrand, M. (1995). Complex daubechies wavelets. *Applied and Computational Harmonic Analysis*, 2(3), 219-229. DOI: [10.1006/acha.1995.1016](https://doi.org/10.1006/acha.1995.1016)

Liu, D., Jiang, W., Mu, L., & Wang, S. (2020). Streamflow prediction using deep learning neural network: case study of Yangtze River. *IEEE Access*, 8, 90069-90086. DOI: [10.1109/ACCESS.2020.2998391](https://doi.org/10.1109/ACCESS.2020.2998391)

Luo, X., Yuan, X., Zhu, S., Xu, Z., Meng, L., & Peng, J. (2019). A hybrid support vector regression framework for streamflow forecast. *Journal of Hydrology*, 568, 184-193. DOI: [10.1016/j.jhydrol.2018.11.014](https://doi.org/10.1016/j.jhydrol.2018.11.014)

Ma, X., Yao, Y., Zhang, B., & Du, Z. (2022). FY-3A/MERSI precipitable water vapor reconstruction and calibration using multi-source observation data based on a generalized regression neural network. *Atmospheric Research*, 265, 105893.

DOI: [10.1016/j.atmosres.2021.105893](https://doi.org/10.1016/j.atmosres.2021.105893)

MAFTEI, C. P. (2002). Etudes concernant les écoulements superficiels modélisation spatialisée de l'écoulement sur le bassin versant de Voinești-Roumanie Université de Montpellier II].

Magand, C. (2014). Influence de la représentation des processus nivaux sur l'hydrologie de la Durance et sa réponse au changement climatique (Doctoral dissertation, Université Pierre et Marie Curie-Paris VI).

Malik, A., Tikhamarine, Y., Souag-Gamane, D., Kisi, O., & Pham, Q. B. (2020). Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction. *Stochastic Environmental Research and Risk Assessment*, 34, 1755-1773.

DOI: [10.1007/s00477-020-01837-0](https://doi.org/10.1007/s00477-020-01837-0)

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7), 674-693. DOI: [10.1109/34.192463](https://doi.org/10.1109/34.192463)

Manimaran, P., Panigrahi, P. K., & Parikh, J. C. (2005). Wavelet analysis and scaling properties of time series. *Physical Review E*, 72(4), 046120. DOI: [10.1103/PhysRevE.72.046120](https://doi.org/10.1103/PhysRevE.72.046120)

Mroczkowski, M., Raper, P. G., & Kuczera, G. J. W. R. R. (1997). The quest for more powerful validation of conceptual catchment models. 33(10), 2325-2335.

Muhammad Adnan, R., Yuan, X., Kisi, O., Yuan, Y., Tayyab, M., & Lei, X. (2019). Application of soft computing models in streamflow forecasting. *Proceedings of the institution of civil engineers-water management*,

- Mutlu, E., Chaubey, I., Hexmoor, H., & Bajwa, S. (2008). Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed. *Hydrological Processes: An International Journal*, 22(26), 5097-5106.
- Nascimento, N. D. O. (1995). Appréciation à l'aide d'un modèle empirique des effets d'actions anthropiques sur la relation pluie-débit à l'échelle d'un bassin versant.
- Ng, W., Panu, U., & Lennox, W. (2009). Comparative studies in problems of missing extreme daily streamflow records. *Journal of Hydrologic Engineering*, 14(1), 91-100.
- Ni, L., Wang, D., Singh, V. P., Wu, J., Wang, Y., Tao, Y., & Zhang, J. (2020). Streamflow and rainfall forecasting by two long short-term memory-based models. *Journal of Hydrology*, 583, 124296.
- Niu, W.-j., & Feng, Z.-k. (2021). Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management. *Sustainable Cities and Society*, 64, 102562.
- Nourani, V., Davanlou Tajbakhsh, A., Molajou, A., & Gokcekus, H. (2019). Hybrid wavelet-M5 model tree for rainfall-runoff modeling. *Journal of Hydrologic Engineering*, 24(5), 04019012. DOI: [10.1061/\(ASCE\)HE.1943-5584.0001775](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001775)
- Osisanwo, F., Akinsola, J., Awodele, O., Hinmikaiye, J., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, 48(3), 128-138.
- Oudin, L. (2004). Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global ENGREF (AgroParisTech)].
- Palmer, A., Montano, J. J., & Sesé, A. (2006). Designing an artificial neural network for forecasting tourism time series. *Tourism management*, 27(5), 781-790. DOI: [10.1016/j.tourman.2005.02.017](https://doi.org/10.1016/j.tourman.2005.02.017)
- Panahi, F., Ehteram, M., Ahmed, A. N., Huang, Y. F., Mosavi, A., & El-Shafie, A. (2021). Streamflow prediction with large climate indices using several hybrid multilayer perceptrons and copula Bayesian model averaging. *Ecological Indicators*, 133, 108285.
- Parisouj, P., Mohebzadeh, H., & Lee, T. (2020). Employing machine learning algorithms for streamflow prediction: a case study of four river basins with different climatic zones in the United States. *Water resources management*, 34, 4113-4131.
- Parizeau, M. (2004). Réseaux de neurones. GIF-21140 et GIF-64326, 124.
- Park, Y.-S., & Lek, S. (2016). Artificial neural networks: Multilayer perceptron for ecological modeling. In *Developments in environmental modelling* (Vol. 28, pp. 123-140). Elsevier.
- Pedrycz, W. (1993). Fuzzy neural networks and neurocomputations. *Fuzzy Sets and Systems*, 56(1), 1-28.

- Percival, D. B., & Walden, A. T. (2000). Wavelet methods for time series analysis (Vol. 4). Cambridge university press.
- Perrin, C. (2002). Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. *La Houille Blanche-Revue internationale de l'eau*(6-7), 84-91.
- Peters, H. (2005). Perceptrons. *Artificial Neural Networks: An Introduction to ANN Theory and Practice*, 67-81.
- Pham, Q. B., Afan, H. A., Mohammadi, B., Ahmed, A. N., Linh, N. T. T., Vo, N. D., Moazenzadeh, R., Yu, P.-S., & El-Shafie, A. (2020). Hybrid model to improve the river streamflow forecasting utilizing multi-layer perceptron-based intelligent water drop optimization algorithm. *Soft Computing*, 24, 18039-18056. DOI: [10.1007/s00500-020-04825-w](https://doi.org/10.1007/s00500-020-04825-w)
- Pratiwi, H., Windarto, A. P., Susliansyah, S., Aria, R. R., Susilowati, S., Rahayu, L. K., Fitriani, Y., Merdekawati, A., & Rahadjeng, I. R. (2020). Sigmoid activation function in selecting the best model of artificial neural networks. *Journal of Physics: Conference Series*, DOI: [10.1088/1742-6596/1469/1/012034](https://doi.org/10.1088/1742-6596/1469/1/012034)
- Rafik, B. (2020). These Baouche-pages-3-9,31-47,49-154
- Rahimzad, M., Moghaddam Nia, A., Zolfonoon, H., Soltani, J., Danandeh Mehr, A., & Kwon, H.-H. (2021). Performance comparison of an LSTM-based deep learning model versus conventional machine learning algorithms for streamflow forecasting. *Water resources management*, 35(12), 4167-4187. DOI: [10.1007/s11269-021-02954-5](https://doi.org/10.1007/s11269-021-02954-5)
- Rasmussen, C. E., & Williams, C. (2006). *Gaussian processes for machine learning* the mit press. Cambridge, MA, 32, 68.
- Refsgaard, J. C., & Abbott, M. B. (1990). The role of distributed hydrological modelling in water resources management. In *Distributed hydrological modelling* (pp. 1-16). Springer.
- Ren, K., Wang, X., Shi, X., Qu, J., & Fang, W. (2021). Examination and comparison of binary metaheuristic wrapper-based input variable selection for local and global climate information-driven one-step monthly streamflow forecasting. *Journal of Hydrology*, 597, 126152. DOI: [10.1016/j.jhydrol.2021.126152](https://doi.org/10.1016/j.jhydrol.2021.126152)
- Rezaie-Balf, M., & Kisi, O. (2018). New formulation for forecasting streamflow: evolutionary polynomial regression vs. extreme learning machine. *Hydrology Research*, 49(3), 939-953.
- Rojas, R., & Rojas, R. (1996). The backpropagation algorithm. *Neural networks: a systematic introduction*, 149-182.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning internal representations by error propagation. In: Institute for Cognitive Science, University of California, San Diego La

- Sahoo, A., Samantaray, S., & Ghose, D. K. (2022). Multilayer perceptron and support vector machine trained with grey wolf optimiser for predicting floods in Barak river, India. *Journal of Earth System Science*, 131(2), 85. DOI: [10.1007/s12040-022-01978-w](https://doi.org/10.1007/s12040-022-01978-w)
- Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2017). Recent advances in recurrent neural networks. arXiv preprint arXiv:1801.01078. DOI: [arXiv:1801.01078](https://doi.org/10.1007/s12040-022-01978-w)
- Salhi, A. (2015). Contribution à l'optimisation de l'écoulement de puissance en utilisant la logique floue associée aux réseaux de neurones (Neuro-Flou) Thèse de doctorat, Université de Biskra].
- Samantaray, S., & Sahoo, A. (2023). Prediction of flow discharge in Mahanadi River Basin, India, based on novel hybrid SVM approaches. *Environment, Development and Sustainability*, 1-25. DOI: [10.1007/s10668-023-01906-8](https://doi.org/10.1007/s10668-023-01906-8)
- Sammen, S. S., Ehteram, M., Abba, S., Abdulkadir, R., Ahmed, A. N., & El-Shafie, A. (2021). A new soft computing model for daily streamflow forecasting. *Stochastic Environmental Research and Risk Assessment*, 35(12), 2479-2491.
- Saravanan, S., Reddy, N. M., Pham, Q. B., Alodah, A., Abdo, H. G., Almohamad, H., & Al Dughairi, A. A. (2023). Machine Learning Approaches for Streamflow Modeling in the Godavari Basin with CMIP6 Dataset. *Sustainability*, 15(16), 12295. DOI: [10.3390/su151612295](https://doi.org/10.3390/su151612295)
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- Sefton, C., & Howarth, S. J. J. o. h. (1998). Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. 211(1-4), 1-16.
- Seo, Y., Choi, Y., & Choi, J. (2017). River stage modeling by combining maximal overlap discrete wavelet transform, support vector machines and genetic algorithm. *Water*, 9(7), 525. DOI: [10.3390/w9070525](https://doi.org/10.3390/w9070525)
- Shabbir, M., Chand, S., & Iqbal, F. (2023). Prediction of river inflow of the major tributaries of Indus river basin using hybrids of EEMD and LMD methods. *Arabian Journal of Geosciences*, 16(4), 257.
- Shabri, A., & Suhartono. (2012). Streamflow forecasting using least-squares support vector machines. *Hydrological Sciences Journal*, 57(7), 1275-1293. DOI: [10.1080/02626667.2012.715640](https://doi.org/10.1080/02626667.2012.715640)
- Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons & Fractals*, 140, 110212. DOI: [10.1016/j.chaos.2020.110212](https://doi.org/10.1016/j.chaos.2020.110212)
- Shamseldin, A. Y., O'CONNOR, K. M., & Nasr, A. E. (2007). A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall—runoff models. *Hydrological Sciences Journal*, 52(5), 896-916. DOI: [10.1623/hysj.52.5.896](https://doi.org/10.1623/hysj.52.5.896)

- Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6(12), 310-316.
- Sharp, A. A., O'Neil, M. B., Abbott, L., & Marder, E. (1993). The dynamic clamp: artificial conductances in biological neurons. *Trends in neurosciences*, 16(10), 389-394.
- Shen, Y., Oki, T., Utsumi, N., Kanae, S., & Hanasaki, N. (2008). Projection of future world water resources under SRES scenarios: water withdrawal/Projection des ressources en eau mondiales futures selon les scénarios du RSSE: prélèvement d'eau. *Hydrological sciences journal*, 53(1), 11-33.
- Sheoran, K., Tomar, P., & Mishra, R. (2020). A novel quality prediction model for component based software system using ACO–NM optimized extreme learning machine. *Cognitive Neurodynamics*, 14(4), 509-522. DOI: [10.1007/s11571-020-09580-7](https://doi.org/10.1007/s11571-020-09580-7)
- Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306. DOI: [10.1016/j.physd.2020.132306](https://doi.org/10.1016/j.physd.2020.132306)
- Shi, X., Wood, A. W., & Lettenmaier, D. P. J. J. o. H. (2008). How essential is hydrologic model calibration to seasonal streamflow forecasting? , 9(6), 1350-1363.
- Shoaib, M., Shamseldin, A. Y., Melville, B. W., & Khan, M. M. (2016). A comparison between wavelet based static and dynamic neural network approaches for runoff prediction. *Journal of Hydrology*, 535, 211-225.
- Siddiqi, T. A., Ashraf, S., Khan, S. A., & Iqbal, M. J. (2021). Estimation of data-driven streamflow predicting models using machine learning methods. *Arabian Journal of Geosciences*, 14(11), 1058. DOI: [10.1007/s12517-020-06142-3](https://doi.org/10.1007/s12517-020-06142-3)
- Singh, V. P. (1995). *Computer models of watershed hydrology*. Water Resources Publications. DOI: [10.1007/978-94-015-9259-6](https://doi.org/10.1007/978-94-015-9259-6)
- Singh, V. P., & Woolhiser, D. A. J. J. o. h. e. (2002). Mathematical modeling of watershed hydrology. 7(4), 270-292. DOI: [10.1080/713610851](https://doi.org/10.1080/713610851)
- Singh, Y., & Chauhan, A. S. (2009). NEURAL NETWORKS IN DATA MINING. *Journal of Theoretical & Applied Information Technology*, 5(1).
- Sinha, S., Routh, P. S., Anno, P. D., & Castagna, J. P. (2005). Spectral decomposition of seismic data with continuous-wavelet transform. *Geophysics*, 70(6), P19-P25.
- Sinitsin, A., Plokhotnyuk, V., Pyrkin, D., Popov, S., & Babenko, A. (2020). Editable neural networks. arXiv preprint arXiv:2004.00345.

- Siqueira, H., Boccato, L., Attux, R., & Lyra, C. (2014). Unorganized machines for seasonal streamflow series forecasting. *International journal of neural systems*, 24(03), 1430009. DOI: [10.1142/S0129065714300094](https://doi.org/10.1142/S0129065714300094)
- Sivapragasam, C., & Liong, S.-Y. (2005). Flow categorization model for improving forecasting. *Hydrology Research*, 36(1), 37-48. DOI: [10.2166/nh.2005.0004](https://doi.org/10.2166/nh.2005.0004)
- Sivapragasam, C., Vanitha, S., Muttill, N., Suganya, K., Suji, S., Thamarai Selvi, M., Selvi, R., & Jeya Sudha, S. (2014). Monthly flow forecast for Mississippi River basin using artificial neural networks. *Neural computing and applications*, 24, 1785-1793. DOI: [10.1007/s00521-013-1455-y](https://doi.org/10.1007/s00521-013-1455-y)
- Smola, A. J. (1996). Regression estimation with support vector learning machines Master's thesis, Technische Universität München]. DOI: [10.5281/zenodo.4022418](https://doi.org/10.5281/zenodo.4022418)
- Solaiman, B., & Richard, L. (2003). Les réseaux de neurones artificiels et leurs applications en imagerie et en vision par ordinateur. Presse Universitaire du Québec.
- Song, C., Yao, L., Hua, C., & Ni, Q. (2021). Comprehensive water quality evaluation based on kernel extreme learning machine optimized with the sparrow search algorithm in Luoyang River Basin, China. *Environmental earth sciences*, 80(16), 521. DOI: [10.1007/s12665-021-09883-2](https://doi.org/10.1007/s12665-021-09883-2)
- Sulugodu, B., & Deka, P. C. (2019). Evaluating the performance of CHIRPS satellite rainfall data for streamflow forecasting. *Water resources management*, 33, 3913-3927. DOI: [10.1007/s11269-019-02461-1](https://doi.org/10.1007/s11269-019-02461-1)
- Sun, A. Y., Wang, D., & Xu, X. (2014). Monthly streamflow forecasting using Gaussian process regression. *Journal of Hydrology*, 511, 72-81. DOI: [10.1016/j.jhydrol.2014.01.024](https://doi.org/10.1016/j.jhydrol.2014.01.024)
- Sun, N., Zhang, S., Peng, T., Zhou, J., & Sun, X. (2020). A composite uncertainty forecasting model for unstable time series: Application of wind speed and streamflow forecasting. *IEEE Access*, 8, 209251-209266. DOI: [10.1109/ACCESS.2020.3033841](https://doi.org/10.1109/ACCESS.2020.3033841)
- Sun, P., & Yang, L. (2022). Generalized eigenvalue extreme learning machine for classification. *Applied Intelligence*, 1-30. DOI: [10.1007/s10489-022-03240-3](https://doi.org/10.1007/s10489-022-03240-3)
- Svozil, D., Kvasnicka, V., & Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and intelligent laboratory systems*, 39(1), 43-62.
- Tan, K.-H., & Lim, B. P. (2018). The artificial intelligence renaissance: deep learning and the road to human-level machine intelligence. *APSIPA Transactions on Signal and Information Processing*, 7, e6. DOI: [10.1017/ATSIP.2018.1](https://doi.org/10.1017/ATSIP.2018.1)
- Tan, Q. F., Lei, X. H., Wang, X., Wang, H., Wen, X., Ji, Y., & Kang, A. Q. (2018). An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach. *Journal of hydrology*, 567, 767-780. DOI: [10.1016/j.jhydrol.2018.10.031](https://doi.org/10.1016/j.jhydrol.2018.10.031)

- Tareke, K. A., & Awoke, A. G. (2023). Hydrological drought forecasting and monitoring system development using artificial neural network (ANN) in Ethiopia. *Heliyon*, 9(2).
- Tayyab, M., Zhou, J., Adnan, R., Meng, C., & Zahra, A. (2016). Streamflow prediction by applying generalized regression network with time series decomposition method. *Indonesian Journal of Electrical Engineering and Computer Science*, 4(3), 611-616. DOI: 10.11591/ijeecs.v4.i3.pp611-616
- Thapa, S., Zhao, Z., Li, B., Lu, L., Fu, D., Shi, X., Tang, B., & Qi, H. (2020). Snowmelt-driven streamflow prediction using machine learning techniques (LSTM, NARX, GPR, and SVR). *Water*, 12(6), 1734. DOI: 10.3390/w12061734
- Tikhmarine, Y., Souag-Gamane, D., & Kisi, O. (2019). A new intelligent method for monthly streamflow prediction: hybrid wavelet support vector regression based on grey wolf optimizer (WSR-GWO). *Arabian Journal of Geosciences*, 12, 1-20.
- Tokar, A. S., & Johnson, P. A. (1999). Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering*, 4(3), 232-239. DOI: 10.1061/(ASCE)1084-0699(1999)4:3(232)
- Tongal, H., & Booij, M. J. (2018). Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. *Journal of Hydrology*, 564, 266-282. DOI: 10.1016/j.jhydrol.2018.07.040
- Tounsi, A., Abdelkader, M., & Temimi, M. (2023). Assessing the simulation of streamflow with the LSTM model across the continental United States using the MOPEX dataset. *Neural computing and applications*, 1-18.
- Uhlenbrook, S., Seibert, J., Leibundgut, C., & Rodhe, A. J. H. S. J. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. 44(5), 779-797. DOI: 10.1002/(SICI)1099-1085(19991115)13:14/15<779::AID-HYP847>3.0.CO;2-Q
- Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review*, 53, 5929-5955. DOI: 10.1007/s10462-020-09944-8
- Vatanchi, S. M., Etemadfar, H., Maghrebi, M. F., & Shad, R. (2023). A Comparative Study on Forecasting of Long-term Daily Streamflow using ANN, ANFIS, BiLSTM and CNN-GRU-LSTM. *Water resources management*, 1-17.
- Vázquez, R. F., & Feyen, J. (2010). Rainfall-runoff modelling of a rocky catchment with limited data availability: Defining prediction limits. *Journal of Hydrology*, 387(1-2), 128-140. DOI: 10.1016/j.jhydrol.2010.04.018
- Vidal, J.-P. (2005). Assistance au calage de modèles numériques en hydraulique fluviale-Appports de l'intelligence artificielle.

- Vu, D.-Q., Mai, S. T., & Dang, T. D. (2023). Streamflow prediction in the Mekong River Basin using deep neural networks. *IEEE Access*. DOI: [10.1109/ACCESS.2023.3144986](https://doi.org/10.1109/ACCESS.2023.3144986)
- Wang, L., Li, Z., Ye, F., & Liu, T. (2023). A Probability Model for Short-Term Streamflow Prediction Based on Multi-Resolution Data. *Water resources management*, 1-18.
- Wang, S., Gong, J., Gao, H., Liu, W., & Feng, Z. (2023). Gaussian Process Regression and Cooperation Search Algorithm for Forecasting Nonstationary Runoff Time Series. *Water*, 15(11), 2111. DOI: [10.3390/w15112111](https://doi.org/10.3390/w15112111)
- Wang, W.-C., Chau, K.-W., Cheng, C.-T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374(3-4), 294-306. DOI: [10.1016/j.jhydrol.2009.06.005](https://doi.org/10.1016/j.jhydrol.2009.06.005)
- Wang, X., Wang, Y., Yuan, P., Wang, L., & Cheng, D. (2021). An adaptive daily runoff forecast model using VMD-LSTM-PSO hybrid approach. *Hydrological Sciences Journal*, 66(9), 1488-1502. DOI: [10.1080/02626667.2021.1960180](https://doi.org/10.1080/02626667.2021.1960180)
- Wang, Z., Xin, J., Wang, Z., Gu, H., Zhao, Y., & Qian, W. (2021). Computer-aided dementia diagnosis based on hierarchical extreme learning machine. *Cognitive Computation*, 13, 34-48. DOI: [10.1007/s12559-020-09756-2](https://doi.org/10.1007/s12559-020-09756-2)
- Waqas, M., Shoaib, M., Saifullah, M., Naseem, A., Hashim, S., Ehsan, F., Ali, I., & Khan, A. (2021). Assessment of advanced artificial intelligence techniques for streamflow forecasting in Jhelum River Basin. *Pakistan Journal of Agricultural Research*, 34(3), 580. DOI: [10.17582/journal.pjar/2021/34.3.580.584](https://doi.org/10.17582/journal.pjar/2021/34.3.580.584)
- Westall, F., & Brack, A. (2018). The importance of water for life. *Space Science Reviews*, 214, 1-23. DOI: [10.1007/s11214-017-0440-7](https://doi.org/10.1007/s11214-017-0440-7)
- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, 30(1), 79-82. DOI: [10.3354/cr030079](https://doi.org/10.3354/cr030079)
- Wu, J. S., Han, J., Annambhotla, S., & Bryant, S. (2005). Artificial neural networks for forecasting watershed runoff and stream flows. *Journal of Hydrologic Engineering*, 10(3), 216-222. DOI: [10.1061/\(ASCE\)1084-0699\(2005\)10:3\(216\)](https://doi.org/10.1061/(ASCE)1084-0699(2005)10:3(216))
- Wurbs, R. A. (1998). Dissemination of generalized water resources models in the United States. *Water International*, 23(3), 190-198. DOI: [10.1080/02508069808692010](https://doi.org/10.1080/02508069808692010)
- Xie, T., Zhang, G., Hou, J., Xie, J., Lv, M., & Liu, F. (2019). Hybrid forecasting model for non-stationary daily runoff series: a case study in the Han River Basin, China. *Journal of Hydrology*, 577, 123915. DOI: [10.1016/j.jhydrol.2019.123915](https://doi.org/10.1016/j.jhydrol.2019.123915)

- Xu, C.-Y., & Vandewiele, G. (1995). Parsimonious monthly rainfall-runoff models for humid basins with different input requirements. *Water Resources*, 18(1), 39-48. DOI: [10.1080/02508069508686304](https://doi.org/10.1080/02508069508686304)
- Xu, Y., Zhao, S., Liang, J., Xie, B., Li, W., Lu, Z., Li, R., & Bal, H. (2023). Using Adaptive Chaotic Grey Wolf Optimization for the daily streamflow prediction. *Expert Systems with Applications*, 121113. DOI: [10.1016/j.eswa.2023.121113](https://doi.org/10.1016/j.eswa.2023.121113)
- Yadav, B., Ch, S., Mathur, S., & Adamowski, J. (2016). Discharge forecasting using an online sequential extreme learning machine (OS-ELM) model: a case study in Neckar River, Germany. *Measurement*, 92, 433-445. DOI: [10.1016/j.measurement.2016.06.020](https://doi.org/10.1016/j.measurement.2016.06.020)
- Yang, X., & Michel, C. (2000). Flood forecasting with a watershed model: a new method of parameter updating. *Hydrological Sciences Journal*, 45(4), 537-546. DOI: [10.1080/02626660009492383](https://doi.org/10.1080/02626660009492383)
- Yang, X., Parent, E., Michel, C., & Roche, P.-A. (1995). Comparison of real-time reservoir-operation techniques. *Journal of Water Resources Planning and Management*, 121(5), 345-351. DOI: [10.1061/\(ASCE\)0733-9496\(1995\)121:5\(345\)](https://doi.org/10.1061/(ASCE)0733-9496(1995)121:5(345))
- Yang, X.-S. (2010). A new metaheuristic bat-inspired algorithm. *Nature inspired cooperative strategies for optimization (NICSO 2010)*, 65-74. DOI: [10.1007/978-3-642-12538-6_6](https://doi.org/10.1007/978-3-642-12538-6_6)
- Yang, X.-S., & He, X. (2013). Bat algorithm: literature review and applications. *International Journal of Bio-inspired computation*, 5(3), 141-149. DOI: [10.1504/IJBIC.2013.055801](https://doi.org/10.1504/IJBIC.2013.055801)
- Yaseen, Z. M., Faris, H., & Al-Ansari, N. (2020). Hybridized extreme learning machine model with salp swarm algorithm: a novel predictive model for hydrological application. *Complexity*, 2020, 1-14. DOI: [10.1155/2020/5078954](https://doi.org/10.1155/2020/5078954)
- Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J., & El-Shafie, A. (2016). Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. *Journal of Hydrology*, 542, 603-614. DOI: [10.1016/j.jhydrol.2016.09.015](https://doi.org/10.1016/j.jhydrol.2016.09.015)
- Yin, S., Tang, D., Jin, X., Chen, W., & Pu, N. (2016). A combined rotated general regression neural network method for river flow forecasting. *Hydrological Sciences Journal*, 61(4), 669-682. DOI: [10.1080/02626667.2015.1008108](https://doi.org/10.1080/02626667.2015.1008108)
- Yin, Z., Feng, Q., Wen, X., Deo, R. C., Yang, L., Si, J., & He, Z. (2018). Design and evaluation of SVR, MARS and M5Tree models for 1, 2 and 3-day lead time forecasting of river flow data in a semiarid mountainous catchment. *Stochastic Environmental Research and Risk Assessment*, 32, 2457-2476. DOI: [10.1007/s00477-018-1606-1](https://doi.org/10.1007/s00477-018-1606-1)
- Yonaba, H., Anctil, F., & Fortin, V. (2010). Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting. *Journal of Hydrologic Engineering*, 15(4), 275-283. DOI: [10.1061/\(ASCE\)HE.1943-5584.0000159](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000159)

- Yu, P.-S., Yang, T.-C., Chen, S.-Y., Kuo, C.-M., & Tseng, H.-W. (2017). Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *Journal of Hydrology*, 552, 92-104. DOI: [10.1016/j.jhydrol.2017.06.008](https://doi.org/10.1016/j.jhydrol.2017.06.008)
- Yuan, Z., Xie, C., & Wang, J. (2020). Deconvolutional artificial neural network models for large eddy simulation of turbulence. *Physics of Fluids*, 32(11). DOI: [10.1063/5.0030622](https://doi.org/10.1063/5.0030622)
- Zaini, N., Malek, M., Yusoff, M., Mardi, N., & Norhisham, S. (2018). Daily river flow forecasting with hybrid support vector machine–particle swarm optimization. *IOP Conference Series: Earth and Environmental Science*. DOI: [10.1088/1755-1315/146/1/012052](https://doi.org/10.1088/1755-1315/146/1/012052)
- Zakhrouf, M., Hamid, B., Kim, S., & Madani, S. (2023). Novel insights for streamflow forecasting based on deep learning models combined the evolutionary optimization algorithm. *Physical Geography*, 44(1), 31-54. DOI: [10.1080/02723646.2022.2031631](https://doi.org/10.1080/02723646.2022.2031631)
- Zehai, G., Cunbao, M., Jianfeng, Z., & Weijun, X. (2021). Remaining useful life prediction of integrated modular avionics using ensemble enhanced online sequential parallel extreme learning machine. *International Journal of Machine Learning and Cybernetics*, 12(7), 1893-1911. DOI: [10.1007/s13042-021-01379-0](https://doi.org/10.1007/s13042-021-01379-0)
- Zerouali, B., Elbeltagi, A., Al-Ansari, N., Abda, Z., Chettih, M., Santos, C. A. G., ... & Araibia, A. S. (2022). Improving the visualization of rainfall trends using various innovative trend methodologies with time–frequency-based methods. *Applied water science*, 12(9), 207. DOI: [10.1007/s13201-021-01512-0](https://doi.org/10.1007/s13201-021-01512-0)
- Zhang, B., Tan, R., & Lin, C.-J. (2021). Forecasting of e-commerce transaction volume using a hybrid of extreme learning machine and improved moth-flame optimization algorithm. *Applied Intelligence*, 51, 952-965. DOI: [10.1007/s10489-020-02177-8](https://doi.org/10.1007/s10489-020-02177-8)
- Zhang, X., Liu, P., Cheng, L., Zhou, L., Liu, W., & Luo, X. (2024). Time-varying parameters from the same period in previous years to improve runoff forecasting. *Journal of Hydrology*, 631, 130685. DOI: [10.1016/j.jhydrol.2021.130685](https://doi.org/10.1016/j.jhydrol.2021.130685)
- Zhang, D., & Zhang, D. (2019). Wavelet transforms. *Fundamentals of image data mining: Analysis, Features, Classification and Retrieval*, 35-44. DOI: [10.1007/978-3-319-95077-3_3](https://doi.org/10.1007/978-3-319-95077-3_3)
- Zhang, X., Liu, P., Cheng, L., Zhou, L., Liu, W., & Luo, X. (2024). Time-varying parameters from the same period in previous years to improve runoff forecasting. *Journal of Hydrology*, 631, 130685. DOI: [10.1016/j.jhydrol.2021.130685](https://doi.org/10.1016/j.jhydrol.2021.130685)
- Zhu, S., Luo, X., Xu, Z., & Ye, L. (2019). Seasonal streamflow forecasts using mixture-kernel GPR and advanced methods of input variable selection. *Hydrology Research*, 50(1), 200-214. DOI: [10.2166/nh.2018.021](https://doi.org/10.2166/nh.2018.021)

Zhu, S., Zhou, J., Ye, L., & Meng, C. (2016). Streamflow estimation by support vector machine coupled with different methods of time series decomposition in the upper reaches of Yangtze River, China. *Environmental earth sciences*, 75, 1-12. DOI: [10.1007/s12665-015-5081-9](https://doi.org/10.1007/s12665-015-5081-9)

Zounemat-Kermani, M., Mahdavi-Meymand, A., & Hinkelmann, R. (2021). A comprehensive survey on conventional and modern neural networks: application to river flow forecasting. *Earth Science Informatics*, 14, 893-911. DOI: [10.1007/s12145-021-00616-8](https://doi.org/10.1007/s12145-021-00616-8)

Zounemat-Kermani, M., Stephan, D., Barjenbruch, M., & Hinkelmann, R. (2020). Ensemble data mining modeling in corrosion of concrete sewer: A comparative study of network-based (MLPNN & RBFNN) and tree-based (RF, CHAID, & CART) models. *Advanced Engineering Informatics*, 43, 101030. DOI: [10.1016/j.aei.2019.101030](https://doi.org/10.1016/j.aei.2019.101030)

Zuo, H., Yan, G., Lu, R., Li, R., Xiao, S., & Pang, Y. (2024). A Multi-Task Learning Based Runoff Forecasting Model for Multi-Scale Chaotic Hydrological Time Series. *Water Resources Management*, 38(2), 481-503.

