



**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**

**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**

**Faculté des Sciences et Sciences Appliquées**

**Université Larbi Ben M'Hidi. Oum-El-Bouaghi**

**Département : Hydraulique**

**Thèse présentée en vue de l'obtention du diplôme de doctorat en 3<sup>ème</sup> cycle LMD**

**Option : Sciences Hydrauliques et Développement Durable**

**N° d'ordre : .....**

**Sujet :**

**Etude et analyse statistique des paramètres des écoulements naturels dans certains oueds en climat semi-aride. Cas de l'Algérie.**

**Par : Yahia Takai Eddine**

**Présentée et soutenue publiquement le 13/12/2023**

**DEVANT LES MEMBRES DU JURY :**

<b><u>Nom et Prénom</u></b>	<b><u>Grade</u></b>	<b><u>Qualité</u></b>	<b><u>Université</u></b>
<b>BOUDAUD Zeineddine</b>	<b>Pr</b>	<b>Président</b>	<b>Univ. Oum El Bouaghi</b>
<b>MAROUF Nadir</b>	<b>Pr</b>	<b>Rapporteur</b>	<b>Univ. Oum El Bouaghi</b>
<b>SEHTAL Sabah</b>	<b>MCA</b>	<b>Co-Rapporteur</b>	<b>Univ. Oum El Bouaghi</b>
<b>HOUICHI Larbi</b>	<b>Pr</b>	<b>Examineur</b>	<b>Univ. Batna II</b>
<b>LOUKAM Imed</b>	<b>MCA</b>	<b>Examineur</b>	<b>Univ. Souk-Ahras</b>
<b>MOKHTARI Elhadj</b>	<b>MCA</b>	<b>Examineur</b>	<b>Univ. Msila</b>

## **Remerciements**

*Je tiens à exprimer ma profonde gratitude envers toutes les personnes qui ont contribué à la réalisation de cette thèse de doctorat. Ce travail n'aurait pas été possible sans leurs soutiens, leurs conseils et leurs encouragements.*

*Je tiens tout d'abord à remercier mon encadreur Pr Marouf Nadir, pour sa guidance inestimable tout au long de ce parcours académique et de recherche. Ses connaissances, sa patience et sa passion pour la recherche m'ont inspiré et m'ont aidé à devenir un chercheur plus compétent.*

*Je remercie également ma Co-encadrante Dr Sehtal Sabah, pour son expertise et ses conseils précieux. Sa contribution à ce projet de recherche a été essentielle, et je lui suis reconnaissant pour son engagement.*

*Je remercie aussi les membres du jury pour avoir pris le temps de lire et d'évaluer ce travail, ainsi que pour leurs commentaires constructifs ce qui pourra grandement améliorer cette thèse.*

*Je suis reconnaissant envers ma famille pour leur soutien constant. Leur encouragement et leur compréhension ont été une source de force tout au long de cette aventure.*

*Enfin, je tiens à remercier toutes les personnes qui ont participé de près ou de loin à cette recherche, ainsi que toutes les institutions qui m'ont aidé à l'élaboration de cette thèse de doctorat.*

*Ce travail de thèse représente le résultat de nombreuses années de travail acharnés, d'apprentissage et de dévouement. Je suis profondément reconnaissant envers toutes les personnes qui ont fait partie de ce document de recherche.*

*Merci du fond du cœur.*

## **Dédicace**

*À l'attention de mes proches,*

*Je souhaite prendre un moment pour adresser mes remerciements les plus sincères à toutes les personnes qui ont joué un rôle crucial dans mon parcours académique et personnel.*

*Tout d'abord, à mes parents, qui ont été une source constante de vie, d'amour et d'affection. Votre soutien inconditionnel a été la pierre angulaire sur laquelle j'ai bâti ma quête de connaissances et de réussite.*

*À mon cher frère, dont la présence a été synonyme de joie et de bonheur tout au long de ce travail. Ta camaraderie et ton encouragement ont été inestimables.*

*À ma famille élargie, vous avez été une source d'espoir et de motivation à chaque étape de mes études. Vos conseils avisés et votre soutien moral m'ont poussé à persévérer, même lorsque les défis semblaient insurmontables.*

*À tous mes amis, qui ont partagé avec moi des moments précieux et qui m'ont soutenu dans les hauts et les bas de la vie étudiante. Votre amitié est une boussole qui m'a guidé dans ma vie personnelle et académique.*

*À chacun de mes enseignants, depuis les premières années de ma scolarité jusqu'à mes études supérieures, je vous remercie pour votre dévouement à l'enseignement. Vos enseignements ont façonné mon chemin intellectuel et m'ont poussé à aller toujours plus loin.*

*Enfin, à vous, cher lecteur, qui prenez le temps de parcourir ces pages, je vous remercie pour l'intérêt que vous portez à mes travaux de recherche.*

## ملخص

تُعتبر محاكاة تدفق جريان المياه، خاصةً في حالة الأحواض الهيدرولوجية الغير المُقاسة، تحديًا مُعقدًا في ميدان هيدرولوجيا المياه السطحية. في الجزائر، يُحطّط للعديد من المشاريع الهيدروليكية في مناطق حيث تكون قِيَّاسَاتُ التدفقات محدودة أو غير مُتاحة. في مثل هذه الحالات، يُصنِّح تَقْدِيرُ المُعلَّمَاتِ الهيدرولوجية، والمعدلات القصوى بشكلٍ خاص.

حاليًا، لا تُوجدُ طَرِيقَةٌ تحليليةٌ لحسب التدفق في المجاري الطبيعية وفي هذا السياق، نُقدِّمُ هذه الدراسة التي تستند إلى بعض طُرُقِ الحسابة، مثل نظرية النموذج الخشن (RMM)، وطريقة Chiu، بالإضافة إلى طُرُقِ الذكاء الاصطناعي، وذلك بهدف التقدير المناسب لتدفقات السائلة في الأنهار الطبيعية. تمت الدراسة على خمس محطات هيدرومترية في بعض احواض وسط شمال الجزائر.

تعتمد نظرية MMR على معايير قابلة للقياس للمجرى المائي. أما بالنسبة لطريقة Chiu، فإن تطبيقها يُمكن من إنشاء نموذج خطي بين متوسط السرعة والسرعة القصوى لمقطع الجريان. بالإضافة إلى ذلك، تحدد هذه الطريقة معدلات التدفق باستخدام قيم معاملات الإنتروبيا.

تم تعزيز هذه الدراسة من خلال تطبيق ثمانية نماذج قائمة على استعمال الذكاء الاصطناعي وذلك لتحسين توقعات التدفقات المحسوبة. تشمل النماذج التي تمت دراستها ANN و ANFIS و Random Forest و Rep Tree و M5p و GPR و SVM و LibSVM. وقد تم كذلك إنشاء نموذج كلي للتعلم الآلي لكل بيانات منطقة الدراسة.

في الختام، نقترح نهج تصميم هجين يدمج نموذج آلية الدعم (SVM) مع العديد من تقنيات مجموعة التحسين، وهي Dagging، Bagging، Rotation Forest، و Random subspace، بهدف تحسين توقعات هذه التدفقات بشكل أفضل.

تم التحقق من هذه النماذج من خلال استخدام معايير إحصائية وتفسيرات رسوم بيانية. تظهر النتائج التي تم الحصول عليها باستعمال نظرية MMR أن هذه الأخير تتنبأ بتقدير التدفقات السائلة بطريقة مثالية، كما أثبتت كذلك فعاليتها في تقدير التدفقات في المجاري المائية الغير مقاسة أو السيئة القياس، وبالتالي تعمل على تعويض نقص البيانات في محطات القياس مقارنة بطريقة Chiu. تم أيضا تقدير مثالي للتدفقات باستخدام ثمانية نماذج قائمة على الذكاء الاصطناعي. وعلاوة على ذلك، تم اقتراح بنجاح أيضا أساليب هجينة تظهر نماذج عالية الجودة لتدفقات السطحية للمنطقة المقترحة للدراسة.

**كلمات مفتاحية:** مجرى مائي طبيعي، التدفق، MMR، Chiu، ANN، ANFIS، RF، Rep Tree، M5p، SVM، GPR، LibSVM، SVM-Hybrids.

## **Résumé**

La simulation d'écoulement, en particulier dans les bassins versants non jaugés, présente un défi complexe dans le domaine de l'hydrologie des eaux de surface. En Algérie, des nombreux projets hydrauliques sont prévus dans des zones où les mesures des débits sont limitées ou inexistantes. Dans de telles situations, l'estimation des paramètres hydrologiques, en particulier les débits extrêmes. Actuellement, il n'y a pas de méthode analytique pour le calcul du débit dans les canaux naturels. C'est dans ce contexte que nous proposons la présente étude, basée sur certaine méthode de calcul, La Méthode du Modèle Rugueux (MMR), la méthode de Chiu et les méthodes de l'intelligence artificielle, pour estimer les débits liquides dans des cours d'eau naturels, L'étude a été menée sur cinq stations hydrométriques dans quelque bassins versants du centre-nord Algérien.

La méthode MMR se base sur des paramètres mesurables du cours d'eau. Pour la méthode de Chiu, son application permet de définir un modèle linéaire entre la vitesse moyenne et la vitesse maximale d'une section d'écoulement. De plus, cette méthode détermine les débits en utilisant les valeurs des paramètres d'entropie.

Cette étude a été renforcée par l'application des huit modèles basés sur l'intelligence artificielle pour améliorer la prédiction de ces débits calculés. Les modèles étudiés comprennent ANN, ANFIS, Random Forest, Rep Tree, M5p, GPR, SVM, LibSVM. Un modèle global d'apprentissage automatique a été ainsi établi pour l'ensemble des données de la zone d'étude. Finalement, nous proposons une approche de modélisation hybridé qui intègre une machine à vecteurs de support (SVM) avec plusieurs techniques d'ensemble, à savoir, Bagging, Dagging, Random subspace et Rotation Forest, dans le but d'améliorer encore la prédiction de ces débits.

Ces modèles ont été validé à l'aide des critères statistiques et d'interprétations graphiques. Les résultats obtenus de la méthode MMR montrent que cette dernière prédire les débits liquides d'une façon optimale, elle s'avère également efficace pour estimer les débits dans des cours d'eau non ou mal jaugés, compensant ainsi le manque de données par rapport à la méthode alternative de Chiu. Ces résultats ont été ainsi bien améliorés par l'utilisation de huit modèles basés sur l'intelligence artificielle, Par ailleurs, des approches hybrides montrant des modèles globaux ont également été proposées avec succès pour les écoulements à surface libre dans de notre cas d'étude.

**Mot clé :** Cours d'eau naturel, débit, MMR, Chiu, ANN, ANFIS, RF, Rep Tree, M5p, GPR, SVM, LibSVM, SVM-Hybridés.

## **Abstract**

Flow simulation, especially in ungauged watersheds, presents a complex challenge in the field of surface water hydrology. In Algeria, many hydraulic projects are planned in areas where flow measurements are limited or non-existent. In such situations, the estimation of hydrological parameters, in particular flows, becomes particularly difficult. Currently, there is no analytical method for calculating flow in natural channels. It is in this context that we propose the present study, based on certain calculation method, The Rough Model Method (RMM), Chiu's method and artificial intelligence methods to estimate liquid flow rates in water courses. natural water, the study was conducted on five hydrometric stations in northern Algeria.

The MMR method does not take into account the Chézy and Manning coefficients. For the calculation of flow in natural channels, this method is based on measurable parameters of the watercourse. For Chiu's method, its application makes it possible to define a linear model between the average speed and the maximum speed of a flow section. Additionally, this method determines flow rates using entropy parameter values.

This study was enhanced by the application of eight models based on artificial intelligence to improve the prediction of these calculated flows. The models studied include ANN, ANFIS, Random Forest, Rep Tree, M5p, GPR, SVM, LibSVM. A global machine learning model was thus established for all the data from the study area. Finally, we propose a hybrid modeling approach that integrates a support vector machine (SVM) with several ensemble techniques, namely, Bagging, Dagging, Random subspace and Rotation Forest, with the aim of further improving the prediction of these flow rates.

These models were validated using statistical criteria and graphical interpretations. The results obtained from the MMR method show that the latter predicts liquid flow rates in an optimal manner, it also proves effective for estimating flow rates in ungauged or poorly gauged watercourses, thus compensating for the lack of data in relation to Chiu's alternative method. These results were thus much improved by the use of eight models based on artificial intelligence. Furthermore, hybrid approaches showing global models were also successfully proposed for the surface flows of our case study.

**Keyword:** Natural Streamwater, Flow, MMR, ANN, ANFIS, RF, Rep Tree, M5p, GPR, SVM, LibSVM, SVM-Hybrids.

## **Sommaire**

Remerciements .....	I
Dédicace.....	II
ملخص .....	III
Résumé.....	II
Abstract.....	V
Sommaire .....	VI
Liste des abréviations .....	XII
Liste des figures .....	XV
Liste des tableaux .....	XIX
Introduction générale.....	1

### **Chapitre I : Recherche bibliographique sur les écoulements dans les cours d'eau naturels**

I.1. Introduction .....	4
I.2. Les différents types des canaux.....	5
I.2.1 Définition.....	5
I.2.2 Les canaux artificiels.....	5
I.2.3 Les canaux naturels.....	6
I.2.4 Géométrie des canaux .....	7
I.3. Cours d'eau.....	8
I.3.1 Définition.....	8
I.3.2 Dynamique du cours d'eau .....	8
I.3.2.1. Débit liquide.....	8
I.3.2.2. Débit solide .....	9
I.3.2.3. Érosion Hydrique .....	9
I.3.3 Morphologie .....	9
I.3.3.1. Lit du cours d'eau.....	10

I.3.4	Les principaux styles géomorphologiques des cours d'eau .....	12
I.3.5	Vue longitudinale.....	14
I.4.	Les écoulements dans les cours d'eau .....	15
I.4.1	Classification d'écoulement .....	15
I.4.2	Régime d'écoulement.....	16
I.5.	Distribution de la vitesse dans les écoulements à surface libre.....	17
I.5.1	Distribution des vitesses dans un écoulement uniforme .....	17
I.5.2	Équation de Keulegan .....	19
I.6.	Mesure de l'écoulement dans les cours d'eau .....	21
I.6.1	Mesure instantanée.....	22
I.6.2	Jaugeage par exploration du champ de vitesse.....	23
I.6.3	Le jaugeage au moulinet.....	23
I.6.4	Mesure en continu.....	24
I.6.5	Jaugeage par radar mobile .....	26
I.7.	Conclusion.....	28

## **Chapitre II : Etude hydrologique des bassins versants**

II.1.	Introduction .....	29
II.2.	Cours d'eau étudiés.....	29
II.2.1	Bassin versant d'Oued Chiffa.....	30
II.2.2	Bassin versant d'Oued Sebaou .....	32
II.2.3	Bassin versant d'Oued Chellif.....	33
II.2.4	Bassin versant d'Oued Mazafran.....	35
II.2.5	Bassin versant d'Oued Bellah.....	36
II.3.	Analyse de l'écoulement.....	38
II.3.1	Variation de L'écoulement.....	39
II.3.2	Débit moyen et max annuel.....	46

II.4. Conclusion.....	49
-----------------------	----

### **Chapitre III : Traitement statistique et analyse des données**

III.1. Introduction .....	51
III.2. Collecte et contrôle des données.....	51
A. Les erreurs aléatoires (accidentelles).....	52
B. Les erreurs systématiques.....	52
III.3. Caractéristiques des stations.....	52
III.4. Choix des stations .....	52
III.5. Test d'homogénéité.....	56
III.6. Choix du modèle d'ajustement.....	56
III.7. Test de Wilcoxon.....	60
III.8. Conclusion.....	65

### **Chapitre IV : Méthodes MMR et CHIU**

IV.1. Introduction .....	66
IV.2. Méthode du modèle rugueux de référence (MMR) .....	66
IV.2.1 Historique .....	66
IV.2.2 Modèle rugueux de référence .....	66
IV.2.3 Relation régissant l'écoulement dans le modèle rugueux.....	68
IV.2.4 Calcul de l'écoulement turbulent par le modèle rugueux de référence.....	69
IV.2.4.1. Expression du nombre de Reynolds.....	69
IV.2.4.2. Formule modifiée de Colebrook-White.....	71
IV.2.4.3. Expression du coefficient de frottement.....	72
IV.2.4.4. Expression du débit volume Q.....	72
IV.3. Méthode CHIU .....	73
IV.3.1 Historique de la théorie de l'information (l'entropie).....	73
IV.3.2 Définition de la théorie de l'information.....	74

IV.3.3	Dérivations des équations analytiques pour une distribution de vitesse bidimensionnelle dans un canal ouvert.....	75
IV.3.4	Mesures de l'homogénéité de distribution de vitesse.....	79
IV.3.5	Relation entre la vitesse moyenne et maximale.....	80
IV.3.6	Détermination de la surface de la coupe transversale .....	80
IV.3.7	Expression du débit Chiu.....	81
IV.4.	Conclusion.....	81

### **Chapitre V : Modèles d'intelligence artificielle**

V.1.	Introduction .....	83
V.2.	Historique d'intelligence artificielle .....	84
V.2.1	Apprentissage automatique (Machine Learning).....	85
V.2.2	Apprentissage profond (Deep Learning).....	86
V.3.	Contexte de la théorie des modèles.....	87
V.3.1	Réseaux de neurones artificiels (Artificial neural networks) .....	87
V.3.1.1.	Neurone artificiel.....	87
V.3.2	Système d'inférence neuro floue adaptative (Adaptive neural-based fuzzy inference system).....	90
V.3.2.1.	Architecture de l'ANFIS.....	91
V.3.3	Machine à vecteurs de support (Support vector machine) .....	92
V.3.3.1.	SVM (principe de fonctionnement général) .....	93
V.3.4	Bibliothèque pour les machines à vecteurs de support (Library for Support Vector Machines).....	96
V.3.5	Régression de processus gaussien (gaussian process regression).....	97
V.3.6	Forêt aléatoire (Random Forest).....	99
V.3.7	Arbre d'élagage à erreur (Reduced Error Pruning Tree).....	101
V.3.8	Arbre M5P (M5p tree).....	102
V.4.	Conclusion.....	104

**Chapitre VI : Méthodes Hybridés du Modèle SVM**

VI.1. Introduction .....	105
VI.2. Apprentissage hybride (Hybrid Learning) .....	105
VI.3. Apprentissage d'ensemble (Ensemble Learning).....	106
VI.4. MACHINE A VECTEURS DE SUPPORT (SUPPORT VECTOR MACHINE) .....	107
VI.5. Dagging .....	108
VI.6. Bagging .....	109
VI.7. Sous-espace aléatoire (Random Subspace).....	110
VI.8. Forêt de rotation (Rotation Forest) .....	111
VI.9. Évaluation et comparaison des modèles .....	113
VI.10. Conclusion.....	115

**Chapitre VII : Résultat et Discussion**

VII.1. Introduction .....	116
VII.2. Méthode MMR et CHIU .....	117
VII.2.1. MMR .....	117
VII.2.2. Méthode de Chiu.....	122
VII.2.3. Comparaison entre les méthodes MMR et CHIU .....	125
VII.2.4. Le critère de Nash et Sutcliffe .....	127
VII.3. Exemple d'application (Méthode MMR et CHIU).....	127
VII.4. Modèle d'intelligence artificielle.....	130
VII.4.1. Modélisation des débits par l'intégration des données hydrométriques et l'algorithmes d'apprentissage automatique .....	136
VII.4.2. Caractéristiques des modèles d'algorithmes d'apprentissage pour la prédiction de débit dans le Nord de l'Algérie.....	139
VII.5. Modèles SVM-Hybridés.....	141
VII.6. Conclusion.....	145

Conclusion générale .....	148
Références bibliographiques .....	151

## Liste des abréviations

$\overline{Q_{obs}}$  = Débit moyen mesuré.

$y_n$  = Le débit est défini par la profondeur normale.

$\psi_Q$  = Facteur de correction des débits.

$(\Delta Q/Q)$  = Relation entre l'écart relative du débit (Débit mesuré et Débit calculé).

$A$  = Section mouillée.

$a_1, b_1$  et  $c_1$  = Coefficients.

$A_c$  = Surface calculée.

**ACP** = Analyse en composante principale.

**ANFIS** = Adaptive neural-based fuzzy inference system.

**ANN** = Artificial neural network.

**ANRH** = Agence Nationale des Ressources Hydrauliques.

$A_{obs}$  = Surface observée.

$D$  = Profondeur d'eau sur l'axe y.

$D_h$  = Diamètre hydraulique.

$f$  = Coefficient de frottement.

$G$  = Hauteur du niveau de l'eau à l'échelle.

$g$  = Accélération de la gravité.

**GEV** = Generalized Extreme Value.

**GPD** = Generalized Pareto Distribution

**GPR** = Gaussian process regression.

$H$  = la profondeur sur l'axe y.

$h$  = Profondeur réelle de  $U_{max}$ .

$i$  = Pente géométrique.

**IA** = Intelligence artificielle.

$K$  = Coefficient de Strickler.

$L$  = largeur du canal.

**LibSVM** = Library for Support Vector Machines.

$M$  = Paramètres d'entropie.

**M5P** = M5P tree.

**MAE** = Erreur moyenne absolue.

**ML** = apprentissage automatique (Machine Learning).

**MLP** = Perceptron multicouche.

**MMR** = Méthode du Modélé Ruguex

**N** = Nombre total de pas de temps de la période de simulation.

**NSE** = Critère de Nash Sutcliffe.

**P** = Périmètre mouillé.

**PCA** = Principal Component Analysis.

**Q** = Débit volumique.

**Q<sub>cal</sub>** = Débit Calculé.

**Q<sub>mes</sub>** = Débit mesuré.

**Q<sub>mes,i</sub>** = Débit mesuré au pas de temps i.

**Q<sub>mes,i</sub>** = Débit simulé au pas de temps i.

**R** = Coefficient de corrélation de Pearson.

**R** et  $\bar{R}$  = Nombre de Reynolds

**R<sup>2</sup>** = Coefficient de détermination.

**RBF** = Fonction de base radiale.

**Rep Tree** =Reduced Error PruningTree.

**RF** = Random Forest.

**RF\*** = Rotation Forest.

**R<sub>h</sub>** = Rayon hydraulique.

**RMSE** = Erreur moyenne quadratique.

**RNA** = Réseau de neurones artificiels.

**RSS** = Random Subspace.

**SVM** = Support Vector Machine.

**SVR** = Régression à vecteurs de support.

**U** = Vitesse.

**U<sub>max</sub>** = Vitesse maximale.

**y** = Distance vertical du lit du canal.

**α** = Dimension linéaire.

**ε** = Rugosité absolue.

**ε/D<sup>h</sup>** = Rugosité relative.

**η** = Le taux de remplissage du conduit.

**ξ** = Variables sans dimensions.

**ξ<sub>max</sub>** = Variable maximal.

**Φ(M)** = Fonction du parameter d'entropie M.

**M** = Parameter d'entropie.

**$\nu$**  = Viscosité cinématique.

**$K_s$**  = Coefficient de strickler

**DL** = Apprentissage profond (Deep Learning)

## Liste des figures

### CHAPITER I

<b>Figure. I. 1.</b> Types de canaux (Graf et Altinaka, 2000).....	5
<b>Figure. I. 2.</b> Éléments géométriques de la section (Graf et Altinaka, 2000). ....	7
<b>Figure. I. 3.</b> Coupe transversale d'un cours d'eau (Verniers,1995). ....	12
<b>Figure. I. 4.</b> Classification des styles cheneaux (Church, 2006 ; Nanson et Knighton,1996).13	
<b>Figure. I. 5.</b> Concept de corridor d'in cours d'eau (Verniers, G. 1995).....	14
<b>Figure. I. 6.</b> Schéma représentant les différents régimes d'écoulement (Träsch, 2019). ...	16
<b>Figure. I. 7.</b> Schéma de définition ayant servi de base à l'établissement de l'équation de Keulegan.....	20
<b>Figure. I. 8.</b> Distribution typique des vitesses dans une section transversale (Nicolas, 2012). .....	23
<b>Figure. I. 9.</b> Débits et champ de vitesse à travers une section (Ammari et Remini, 2014).24	
<b>Figure. I. 10.</b> Jaugeage au moulinet (Subramanya, 2008).....	24
<b>Figure. I. 11.</b> Station de jaugeage typique (Nicolas, 2012). ....	25
<b>Figure. I. 12.</b> Mesure par radar mobile (Dramais et al., 2014).....	27

### CHAPITER II

<b>Figure. II. 1.</b> Carte de la station hydrométrique des cinq cours d'eau étudiés dans le nord de l'Algérie. ....	30
<b>Figure. II. 2.</b> Coupe transversale du cours d'eau naturel "Chiffa"(ANRH). ....	30
<b>Figure. II. 3.</b> Station Amont des gorges "Oued Chiffa".....	31
<b>Figure. II. 4.</b> Coupe transversale du cours d'eau naturel "Sebaou" (ANRH). ....	32
<b>Figure. II. 5.</b> Station Baghlia "Oued Sebaou ".....	33
<b>Figure. II. 6.</b> Coupe transversale du cours d'eau naturel "Chellif"(ANRH).....	34
<b>Figure. II. 7.</b> Station Arib "Oued Chellif ".....	35
<b>Figure. II. 8.</b> Coupe transversale du cours d'eau naturel "Mazafran"(ANRH).....	35
<b>Figure. II. 9.</b> Station Fer à cheval "Oued Mazafran".....	36
<b>Figure. II. 10.</b> Coupe transversale du cours d'eau naturel "Bellah"(ANRH).....	37
<b>Figure. II. 11.</b> Station Pont RN 11 "Oued Bellah".....	38
<b>Figure. II. 12.</b> Histogramme de débits moyens mensuels. ....	45

<b>Figure. II. 13.</b> Variation annuelle des débits moyen et max ( $m^3/s$ ) d'écoulement d'Oued Chellif (Station Arib Chellif).....	46
<b>Figure. II. 14.</b> Variation annuelle des débits moyen et max ( $m^3/s$ ) d'écoulement d'Oued Mazafran (Station Fer à Cheval). ....	47
<b>Figure. II. 15.</b> Variation annuelle des débits moyen et max ( $m^3/s$ ) d'écoulement d'Oued Chiffa (Station Amont Des Gorges). ....	47
<b>Figure. II. 16.</b> Variation annuelle des débits moyen et max ( $m^3/s$ ) d'écoulement d'Oued Sebaou (Station Baghlia).....	48
<b>Figure. II. 17.</b> Variation annuelle des débits moyen et max ( $m^3/s$ ) d'écoulement Oued Bellah (Station Pont RN 11). ....	48
<b>Figure. II. 18.</b> Variation saisonnière des débits maxima mensuels. ....	49

### CHAPTER III

<b>Figure. III. 1.</b> Ajustement des débits moyens à la loi log normal d'Oued Chellif (Station Arib Chellif).....	58
<b>Figure. III.2.</b> Ajustement des débits moyens à la loi log normal d'Oued Mazafran (Station Fer à Cheval).....	58
<b>Figure. III.3.</b> Ajustement des débits moyens à Gumbel d'Oued Chiffa (Station Amont Des Gorges). ....	59
<b>Figure. III.4.</b> Ajustement des débits moyens à Gumbel d'Oued Sebaou (Station Baghlia). ..	59
<b>Figure. III.5.</b> Ajustement des débits moyens à la loi log normal d'Oued Bellah (Station Pont RN 11). ....	60

### CHAPTER IV

<b>Figure. IV. 1.</b> Schéma de la définition d'un canal de forme quelconque et de son modèle rugueux.....	66
<b>Figure. IV. 2.</b> Modèle de distribution de vitesse et système de coordonnées curvilignes (Chiu et Chen 1999). ....	76

### CHAPTER V

<b>Figure.V.1.</b> Relation entre l'intelligence artificielle, l'apprentissage automatique, les réseaux de neurones et l'apprentissage profond (Bunod et al., 2022). ....	86
<b>Figure.V. 2.</b> Conception de neurones biologiques et artificiels (Akgün et Demir, 2018).....	88

<b>Figure.V. 3.</b> Structure d'un MLP (Abroudi et al., 2013).....	90
<b>Figure. V. 4.</b> Structure d'un système d'inférence floue (FIS) (Voiron et al., 2013). .....	91
<b>Figure.V. 5.</b> Architecture équivalente d'ANFIS pour deux règles (Belhachat et al., 2007)...	92
<b>Figure.V. 6.</b> Représentation géométrique de la marge du SVM (Huang et al., 2018). .....	95
<b>Figure.V. 7.</b> Fonction noyau. Les données qui ne peuvent pas être séparées par SVM linéaire peuvent être transformées et séparées par une fonction noyau (Jakkula, 2006). .....	96
<b>Figure.V. 8.</b> Organisation de la bibliothèque dans LibSVM (Abdiansah et Wardoyo, 2015).	97
<b>Figure.V. 9.</b> Démonstration de la méthodologie RF ( Malekipirbazari & Aksakalli 2015 ).	100
<b>Figure.V. 10.</b> Organigramme du fonctionnement M5p tree.....	103
<b>Figure.V. 11.</b> Diagramme de méthodologie appliquée pour les modèles d'apprentissages utilisés.....	104

## CHAPITER VI

<b>Figure. VI. 1.</b> Combinaison d'un ensemble de classificateurs pour réduire les erreurs de classification et/ou la sélection de modèles (Polikar, 2012). .....	107
<b>Figure. VI. 2.</b> Organigramme du fonctionnement Bagging (Salcedo-Sanz et al.,2022). .....	110
<b>Figure. VI. 3.</b> Diagramme de méthodologie des méthodes SVM-Hybridés. ....	113

## CHAPITER VII

<b>Figure. VII. 1.</b> Coupe transversale d'un cours d'eau générique avec une profondeur d'écoulement $y$ , une largeur de canal $T$ et une largeur de surface $L$ . .....	117
<b>Figure. VII. 2.</b> Relation surface mesurée ( $A_{mes}$ ) et surface calculée ( $A_{cal}$ ) de la section transversale pour la méthode MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).....	119
<b>Figure. VII. 3.</b> Relation entre l'écart relative du débit ( $\Delta Q/Q$ ) et le rayon hydraulique ( $R_h$ ) obtenus par la méthode MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif). .....	120
<b>Figure. VII. 4.</b> Relation débits mesurés et débits calculés de la méthode MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).....	121
<b>Figure. VII. 5.</b> Relation entre la vitesse moyenne ( $U_m$ ) et la vitesse maximale ( $U_{max}$ ) obtenue par la méthode de Chiu (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).....	123

---

<b>Figure. VII. 6.</b> Relation entre la surface de la section transversale et le produit de la hauteur d'eau (H) et la largeur du canal (L) par la méthode de Chiu (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, .....	124
<b>Figure. VII. 7.</b> Relation débits observés et débits estimés pour les deux méthodes CHIU et MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).....	126
<b>Figure. VII. 8.</b> Relation débits observés et débits simulés (prévues) pour les cinq cours d'eau naturels étudiés (a- St. Sebaou, b- St. Mazafran, c- St. Chiffa, d- St. Chellif, e- St. Bellah).	134
<b>Figure. VII. 9.</b> Diagrammes de Taylor affichant une comparaison statistique des huit modèles pour les cours d'eau naturels des oueds (a) Sebaou, b) Mazafran, c) Chiffa, d) Chellif, e) Bellah). .....	135
<b>Figure. VII. 10.</b> Relation globale de débits observés et débits prédits pour l'ensemble des stations appliquées aux différents modèles d'intelligence artificielle .....	138
<b>Figure. VII. 11.</b> Diagrammes de Taylor : Comparaison statistique des modèles (toutes les stations). .....	138
<b>Figure. VII. 12.</b> Relation débits observés et débit prédits dans le cours d'eau "Chiffa" pour l'ensemble des modèles hybridés lors de la phase d'apprentissage : a) SVM ; b) SVM-Bagging ; c) SVM-Dagging ; e) SVM- Random subspace ; f) SVM- Rotation Forest. ....	143
<b>Figure. VII. 13.</b> Relation des débits observés et débits prédits dans le cours d'eau "Chiffa" pour l'ensemble des modèles hybridés lors de la phase de test : a) SVM ; b) SVM-Bagging ; c) SVM-Dagging ; e) SVM- Random subspace ; f) SVM- Rotation Forest. ....	144
<b>Figure. VII. 14.</b> Diagrammes de Taylor montrant la comparaison statistique des cinq modèles hybridés étudiés pendant la période d'essai. ....	145

## Liste des tableaux

### CHAPITER II

<b>Tableau. II. 1.</b> Caractéristiques de formes des sous bassins d'Oued chiffa (Station Amont des gorges).....	31
<b>Tableau. II. 2.</b> Caractéristiques de formes des sous bassins d'Oued Sebaou (Station Baghlia). .....	32
<b>Tableau. II. 3.</b> Caractéristiques de formes de sous bassin d'Oued Chellif (Station Arib Chellif). .....	34
<b>Tableau. II. 4.</b> Caractéristiques de formes des sous bassins d'Oued Mazafran (Station Fer à Cheval). .....	36
<b>Tableau. II. 5.</b> Caractéristiques de formes de sous bassin d'Oued Bellah (Station Pont RN 11). .....	37
<b>Tableau. II. 6.</b> Identification des Stations hydrométriques des cours d'eau étudiées. ....	39
<b>Tableau. II. 7.</b> Variation interannuelle de l'écoulement au droit de différentes stations. ....	40
<b>Tableau. II. 8.</b> Caractéristique de l'écoulement d'Oued Chellif (Station Arib Chellif). ....	41
<b>Tableau. II. 9.</b> Caractéristiques de l'écoulement d'Oued Mazafran (Station Fer à Cheval)...	41
<b>Tableau. II. 10.</b> Caractéristiques de l'écoulement d'Oued Chiffa (Station Amont Des Gorges). .....	42
<b>Tableau. II. 11.</b> Caractéristiques de l'écoulement d'Oued Sebaou (Station Baghlia). ....	43
<b>Tableau. II. 12.</b> Caractéristiques de l'écoulement d'Oued Bellah (Station Pont RN 11). ....	44

### CHAPITER III

<b>Tableau. III. 1.</b> Caractéristiques des données des stations hydrométriques. ....	53
<b>Tableau. III. 2.</b> Homogénéité et ajustement des données des stations hydrométriques étudiées. .....	62

### CHAPITER IV

<b>Tableau. IV. 1.</b> Caractéristiques géométriques et hydrauliques du MMR. ....	68
---	----

### CHAPITER VII

<b>Tableau. VII. 1.</b> Les valeurs de R des figures (VII.2, VII.3, et VII.4) pour les différentes stations étudiées.....	121
---	-----

<b>Tableau. VII. 2.</b> Valeurs de $\Phi$ , M et R pour chaque station étudiée.....	123
<b>Tableau. VII. 3.</b> Expressions des sections estimées pour l'ensemble des stations étudiées selon la méthode de Chiu.....	125
<b>Tableau. VII. 4.</b> Résultats du critère de NSE pour les modèles obtenus par les méthodes ..	127
<b>Tableau. VII. 5.</b> Représentation des données observées par l'ANRH et les résultats obtenus par l'application de la méthode MMR "Oued Chiffa". .....	128
<b>Tableau. VII. 6.</b> Valeurs de $\Phi$ à "Oued Chiffa".....	129
<b>Tableau. VII. 7.</b> Expression de section estimée par la méthode de Chiu à "Oued Chiffa"...	130
<b>Tableau. VII. 8.</b> Représentation des données observées par l'ANRH et les résultats obtenus par l'application de la méthode Chiu à "Oued Chiffa". .....	130
<b>Tableau. VII. 9.</b> Indicateurs de performance des modèles (ANN, ANFIS, SVM, LibSVM, Random Forest, GPR, Rep Tree, M5P) optimaux sur les périodes d'apprentissage et de test pour les cinq stations étudiées. ....	131
<b>Tableau. VII. 10.</b> Indicateurs de performance des modèles : ANN, ANFIS, SVM, LibSVM, Random Forest, GPR, Rep Tree, M5P, optimaux sur les périodes d'Apprentissage et de Test pour l'ensemble des cours d'eau étudiés. ....	137
<b>Tableau. VII. 11.</b> Caractéristiques des modèles d'algorithmes d'apprentissage automatique utilisés pour la prédiction de débit dans le Nord de l'Algérie. ....	139
<b>Tableau. VII. 12.</b> Performances du modèle au cours des phases de formation et de test. ....	141

# **INTRODUCTION GÉNÉRALE**

## **Introduction générale**

Les cours d'eau naturels forment l'une des principales sources d'eau douce exploitable pour répondre aux besoins domestiques, d'agricultures et industriels (Shiklomanov, 1990 ; Blöschl et al., 2013 ; Sivapalan et al, 2003) et font l'objet de mesures et d'observations en vue de préciser leurs régimes hydrologiques. Cependant, l'accroissement de ces besoins et le changement climatique amène les spécialistes d'approfondir ses travaux sur les régimes hydrologiques et le comportement des cours d'eau naturels (Ficke et al, 2007 ; Goswami et al., 2007 ; Zamoum and Souag., 2019).

Les débits des cours d'eau présentent d'importantes fluctuations saisonnières, dont les conséquences sur l'environnement sont multiples, en particulier en période de crue (Baker et al, 2004 ; Gibson et al, 2005 ; Marouf et Remini, 2019). Ces événements naturels peuvent engendrer des dommages importants et des modifications sur la qualité des eaux et sur les caractéristiques des oueds des bassins versants (Karr, 1991 ; Poff et al., 1997 ; Marouf et Remini, 2019).

Les différentes problématiques qui menacent nos bassins versants Algériens, tels que ; la qualité des eaux à améliorer, la perturbation physique et hydrologique du milieu aquatique, l'exploitation excessive des ressources en eau ont mis une réflexion a aboutir à une gestion globale des milieux physiques et aquatiques, en réalisant un état des lieux des cours d'eau, des oueds et de leurs fonctionnements (Touaibia et all, 2001 ; Ammari, 2012 ; Touazi, 2011 ; Marouf, 2012; Marouf et Remini, 2019).

En Algérie, l'ensemble des cours d'eau naturels présente un régime d'écoulement irrégulier et très variable (sec en été, torrentiel en hiver) avec de fortes charges sédimentaires (Marouf et Remini, 2011 ; Marouf, 2012 ; Taïbi, 1993), surtout au début de la saison des pluies. Ces conditions rendent très difficile la maintenance des sections de mesure, de plus, le manque des ressources humaines et matérielles a aggravé cette situation (Marouf, 2012 ; Marouf et Remini, 2011). Dans ces conditions, afin d'atteindre l'estimation optimale des débits avec un minimum de temps et d'effort, ce qui introduit l'application des formules empiriques pour le calcul des débits, non calées dans les régions d'étude où le coût d'exécution des projets est important.

Dans ce contexte, ce travail montre le suivi du fonctionnement hydrologique dans certains cours d'eau en Algérie, en se basant sur une analyse statistique de certains paramètres physiques et cinématiques des cours d'eau naturels. L'étude mettra en évidence les défis liés à la mesure des débits dans ces cours d'eau naturels, en particulier pendant les périodes de

crue, et l'exploration des solutions alternatives, telles que l'utilisation de techniques de l'apprentissage automatique et les méthodes du Modèle Rugueux MMR et Chiu.

Ce travail de recherche contribuera à une meilleure compréhension de fonctionnements de nos cours d'eau Algériens, à une gestion plus efficace de nos ressources en eau, une bonne maîtrise de la quantification des apports liquides et à une adaptation adéquate de nos infrastructures hydrauliques pour faire face aux défis futurs.

### **Objectif**

Cette étude se focalise sur divers domaines de recherche relatifs à l'évaluation appropriée des débits liquides dans certains cours d'eau naturels des bassins versants du nord Algérien.. Elle aborde notamment l'utilisation des deux méthodes Chiu et MMR, ainsi que le développement des modèles de l'intelligence artificielle simples et hybridés.

Ce travail montre l'application pour la première fois de la Méthode du Modèle Rugueux MMR dans les cours d'eau naturels dont l'objet de la détermination des débits pour des stations de mesure mal jaugées dans des cours d'eau situés dans un climat semi-aride et méditerranéen. L'application ainsi des modèles basés sur l'intelligence artificielle pour prédire ces débits liquide peuvent également à obtenir des meilleures performances de prévision.

### **Démarche et organisation de la thèse**

Le travail de thèse comporte sept chapitres, chacun ayant un objectif spécifique. Voici une présentation de la structure globale de la thèse et une brève description du contenu de chaque chapitre :

Le premier chapitre de la thèse consiste en une synthèse bibliographique portant sur les écoulements dans les cours d'eau naturels et les différentes méthodes de mesure de ces écoulements. Une attention particulière est accordée aux techniques de jaugeage utilisées dans ce contexte.

Le deuxième chapitre est dédié à la présentation des caractéristiques morphologiques et hydrographiques spécifiques des bassins étudiés. Ce chapitre vise à fournir une description détaillée des aspects physiques et géographiques des cours d'eau étudiés, ainsi que des données hydrologiques pertinentes pour cette recherche.

Le troisième chapitre se concentre sur l'analyse, le traitement et les caractéristiques statistiques des données utilisées. Ce chapitre vise à présenter la méthodologie appliquée pour

collecter, prétraiter et analyser les données, ainsi que les différentes techniques statistiques utilisées pour examiner les résultats obtenus.

Dans le quatrième chapitre, nous procéderons à l'application des deux méthodes MMR et Chiu. L'accent sera mis sur les démarches et les procédures impliquées dans l'application de ces modèles, ainsi que sur l'évaluation de leurs performances.

Dans le cinquième chapitre, nous avons examinées l'utilisation de plusieurs approches d'intelligence artificielle pour prédire les débits liquides. (Huit modèles et algorithmes ont été utilisés). Ce chapitre se concentrera sur l'évaluation de la performance des modèles d'intelligence artificielle appliquées dans la prédiction des débits, en comparant les résultats obtenus avec les méthodes traditionnelles. Nous explorerons également divers algorithmes utilisés dans ces modèles pour améliorer la précision des prédictions.

Dans le sixième chapitre, une approche de modélisation innovante a été proposée, qui intègre une machine à vecteurs de support (SVM) avec plusieurs techniques d'ensemble, notamment (Bagging, Dagging, Random subspace et Rotation Forest), dont le but de mieux encore prédire les écoulements naturels. Ce chapitre décrira en détail cette approche hybride, en expliquant comment chaque technique est utilisée et combinées pour améliorer la précision des prédictions des débits.

Le septième chapitre, présente les résultats obtenus à partir des méthodes et des modèles utilisés, accompagnés d'une discussion approfondie. Une analyse des différents critères de performance a été réalisée pour évaluer l'efficacité des méthodes et modèles appliqués. Ce chapitre permettra de mettre en lumière les forces et les limitations des approches utilisées, tout en soulignant les implications des résultats obtenus pour la gestion et la prédiction des apports liquides dans les cours d'eau naturels et bassins versants situés à un climat méditerranéen et semi-aride en Algérie.

Enfin, une conclusion générale résume de manière synthétique les démarches principales entreprises ainsi que les résultats essentiels obtenus au cours de cette thèse.

# **CHAPITRE I**

RECHERCHE BIBLIOGRAPHIQUE SUR LES  
ÉCOULEMENTS DANS LES COURS D'EAU  
NATURELS

## **I.1. Introduction**

L'objectif de ce chapitre est de se familiariser avec les termes appartenant au domaine de l'estimation des débits et apports en cours d'eau qui sont couramment utilisés. L'estimation des débits et apports est un maillon incontournable d'une bonne gestion des ressources en eau.

Dans ce fascinant chapitre, nous avons plongé au cœur des écoulements et des morphologies des cours d'eau, ainsi que des méthodes cruciales pour mesurer leurs débits. Au fil de cette exploration, nous avons découvert l'importance fondamentale de ces concepts dans la science des cours d'eau et leur rôle essentiel dans la gestion durable des ressources en eau et la préservation des écosystèmes fluviaux.

La disponibilité d'enregistrements du débit des rivières est essentielle pour développer notre compréhension du cycle hydrologique (Oki et Kanae, 2006). La surveillance du débit à des points discrets sur un système fluvial permet aux hydrologues de quantifier la sortie intégrée de tous les processus hydrologiques agissant sur un bassin versant, et sous-tend ainsi une gestion efficace de l'eau dans des domaines tels que l'estimation des risques d'inondation, la gestion des ressources en eau, l'évaluation hydro-écologique et production hydroélectrique (Hawng et al., 2016 ; Blöschl et al., 2013 ; Mishra et Coulibaly, 2009 ; Razavi et Coulibaly, 2013).

En acquérant une compréhension approfondie des différents types d'écoulements, tels que les écoulements laminaires et turbulents, nous sommes désormais mieux armés pour appréhender la complexité de la dynamique fluviale. L'analyse des facteurs qui influencent la morphologie des cours d'eau, notamment le régime hydraulique, l'érosion hydrique, nous a permis de saisir comment ces voies navigables évoluent naturellement et comment les activités humaines peuvent altérer leur forme et leur fonctionnement ( Gallart et al., 2016 , Prat et al., 2014 , Sánchez-Montoya et al., 2007 ).

Nous avons également plongé dans les méthodes sophistiquées de mesure des débits, des outils indispensables pour une gestion responsable des ressources en eau. Grâce aux jauges à moulinet, aux flotteurs et aux systèmes de jaugeage, nous pouvons désormais évaluer avec précision le débit des cours d'eau en différents points et moments. Cette capacité de surveillance nous offre une base solide pour prendre des décisions éclairées concernant l'utilisation de l'eau, qu'il s'agisse de satisfaire les besoins en eau potable, de soutenir l'agriculture par l'irrigation ou d'exploiter l'énergie hydroélectrique.

## I.2. Les différents types des canaux

### I.2.1 Définition

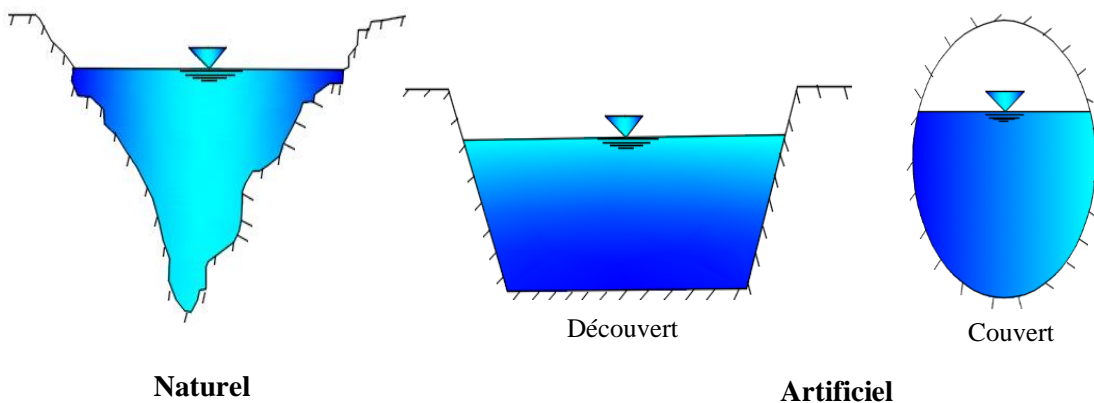
Les liquides sont transportés d'un endroit à un autre à l'aide de structures de transport naturelles ou construites. La section transversale de ces structures peut être ouverte ou fermée (Chaudhry, 2008).

Les structures dont le dessus est fermé sont appelées conduits fermés et celles dont le dessus est ouvert sont appelées canaux ouverts. Par exemple, les tunnels et les tuyaux sont des conduits fermés alors que les rivières, les ruisseaux, les estuaires, etc. sont des canaux ouverts. L'écoulement dans un canal ouvert ou dans un conduit fermé à surface libre est appelé écoulement à surface libre ou écoulement à canal ouvert.

Les termes écoulement en canal ouvert ou écoulement à surface libre. La surface libre est généralement soumise à la pression atmosphérique.

Les eaux souterraines ou les écoulements en charge. S'il n'y a pas de surface libre et que le conduit s'écoule à plein, le débit est appelé écoulement sous pression (Chaudhry, 2008).

- 1) Les canaux artificiels,
- 2) Les canaux naturels.



**Figure. I. 1.** Types de canaux (Graf et Altinaka, 2000).

### I.2.2 Les canaux artificiels

Il s'agit de cours d'eau créés par l'homme, peuvent être classés en deux catégories distinctes : les canaux découverts, construits au niveau du sol (tels que les canaux de navigation, d'adduction, d'évacuation, d'irrigation et de drainage), ainsi que les canaux

couverts où les liquides ne remplissent pas toute la section (comme les tunnels hydrauliques, les aqueducs, les drains et les égouts). Les canaux artificiels possèdent généralement des caractéristiques hydrauliques assez régulières. Lorsqu'on applique la théorie hydraulique à leur étude, les résultats obtenus sont souvent réalistes (Graf et Altinaka, 2000).

### **I.2.3 Les canaux naturels**

Les cours d'eau sont des flux d'eau continus ou temporaires qui existent naturellement sur ou sous terre (Graf et Altinaka, 2000).

Pour décrire les structures à travers lesquelles l'eau s'écoule, plusieurs termes sont couramment utilisés. Voici les principales définitions issues du dictionnaire Larousse (Verniers, 1995) :

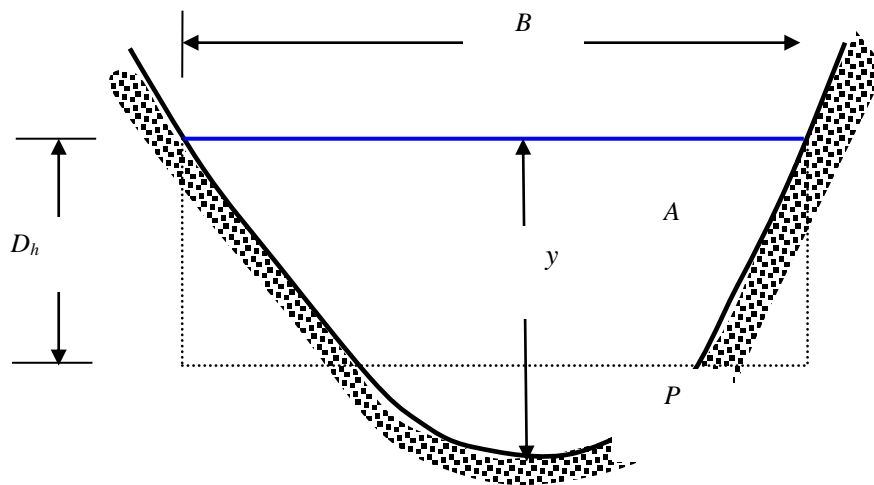
- **Ruisseau** : petit cours d'eau, de faible largeur et de longueur limitée, alimentée par des sources d'eau naturelles, souvent affluent d'un étang, d'un lac ou d'une rivière. Les ruisseaux se trouvent à la tête des bassins versants.
- **Rivière** : cours d'eau moyennement important, à écoulement continu ou intermittent, suivant un tracé défini et se jetant dans un autre cours d'eau, un lac, une mer. Abondant, et particulièrement celui qui se jette dans un fleuve.
- **Fleuve** : cours d'eau important, long et au débit élevé, comptant de nombreux affluents et se jetant dans la mer.
- **Oued** : terme d'origine arabe désignant un cours d'eau temporaire dans les régions arides ou semi-arides. Son écoulement dépend des précipitations et il peut rester à sec pendant de très longues périodes.
- **Torrent** : cours d'eau au débit rapide et régulier, situé sur une pente plus ou moins prononcée. Les torrents se retrouvent sur des terrains accidentés ou en montagne. Ce terme est utilisé principalement pour désigner les cours d'eau de montagne avec un lit rocheux et encaissé.

Les cours d'eau naturels présentent généralement des propriétés géométriques et hydrauliques très irrégulières. Lorsque l'on applique la théorie hydraulique à leur étude, les résultats obtenus ne sont que des approximations basées sur des hypothèses importantes (Petersen-øverleir, 2006).

### **I.2.4 Géométrie des canaux**

Les grandeurs géométriques les plus utilisées permettant de caractériser l'écoulement sont :

- Une section transversale d'un canal est une coupe plane perpendiculaire à la direction de l'écoulement (Carlier, 1986).
- La section ou surface mouillée,  $A$ , est la portion de la section transversale occupée par le liquide (Graf et Altinaka, 2000).



**Figure. I. 2.** Éléments géométriques de la section (Graf et Altinaka, 2000).

- Le périmètre mouillé,  $P$ , du canal, formé par la longueur de la ligne de contact entre la surface mouillée et le lit  $y$  compris les berges, mais ne comprenant pas la surface libre ;
- Le rayon hydraulique,  $R_h$ , donné par le quotient de la surface mouillée,  $A$ , et du périmètre mouillé,  $P$ , exprimé en (m) :

$$R_h = \frac{A}{P} \quad (\text{I.1})$$

- Le diamètre hydraulique  $D_h$  du canal défini par :

$$D_h = 4R_h \quad (\text{I.2})$$

- Largeur au miroir  $B$  : largeur de la section d'écoulement au niveau de la surface libre ;
- Profondeur d'eau est la hauteur d'eau mesurée à partir du point le plus bas de la section perpendiculaire à l'écoulement.

## **I.3. Cours d'eau**

### **I.3.1 Définition**

Bien qu'ils ne mobilisent qu'une part infime de l'eau terrestre, les cours d'eau constituent un maillon essentiel de la dynamique terrestre en assurant le transfert des eaux, sédiments et substances dissoutes des continents vers les océans. Fleuves et rivières collectent d'une part les eaux de nappes souterraines, des ruissellements de surface ; d'autre part, ils transportent les éléments issus de l'érosion mécanique et biochimique sur les versants. Ils participent donc activement au cycle de l'eau, mais aussi au cycle géologique de la planète. Les cours d'eau enfin constituent des éléments essentiels à la vie des sociétés : approvisionnement en cas, axe de transport, mais aussi égout. Ainsi les grandes plaines alluviales ont été depuis longtemps des pôles de développement urbain, industriel et énergétique. C'est pourquoi les cours d'eau ont été harnachés de levées et digues, équipés de barrages. Il est donc bien difficile de trouver aujourd'hui une rivière « naturelle » (Amat et al., 2008).

### **I.3.2 Dynamique du cours d'eau**

#### **I.3.2.1. Débit liquide**

Les précipitations tombent dans le bassin versant, qui est la zone de réception, et leur fréquence, intensité et durée dépendent du type de climat et des conditions météorologiques. Le ruissellement des précipitations varie en fonction des caractéristiques des roches, des sols et de la couverture végétale (comme la forêt, la prairie ou les cultures labourées). À l'exutoire de chaque bassin versant élémentaire (qui correspond à un point spécifique de la topographie et du réseau hydrographique), ainsi que du plus grand bassin, s'écoule un débit liquide ( $Q$ ), qui représente le volume d'eau qui sort de ce bassin par unité de temps (en mètres cubes par seconde, par jour ou par an) (Malavoi et Bravard, 2010).

Lorsque les précipitations tombent sur un substratum ou un sol imperméable (comme les roches métamorphiques, l'argile, les marnes ou les zones urbanisées), le ruissellement prédomine par rapport à l'infiltration. Ainsi, la même quantité de pluie entraîne un débit plus élevé à l'exutoire du bassin versant que si elle tombait sur un sol perméable (comme le sous-sol calcaire ou sableux, ou un sol humifère épais). De plus, un sol ayant les mêmes caractéristiques de perméabilité sera plus sujet au ruissellement s'il est utilisé pour des cultures plutôt que s'il est couvert de forêts ou de prairies. Si vous recherchez des informations plus détaillées sur le débit liquide et sa formation, nous vous

recommandons de consulter des manuels d'hydrologie tels que (Cosandey et Robinson, 2000 ; Musy et Higy, 2004 ; Malavoi et Bravard, 2010).

### **I.3.2.2. Débit solide**

Le débit solide d'un cours d'eau provient également du bassin versant, principalement par le biais des apports externes, tels que l'érosion des versants. Il est important de souligner que l'érosion latérale des berges alluviales, considérée comme un apport interne au système fluvial, contribue également de manière significative. Dans ce contexte, nous examinerons la formation et le fonctionnement du débit solide, en mettant l'accent sur la fraction grossière qui joue un rôle essentiel dans l'équilibre géodynamique et est communément appelée "charge de fond". Il est à noter que cette charge alluviale grossière est également à l'origine de nombreux habitats indispensables aux écosystèmes aquatiques et ripariens (Malavoi et Souchon, 1996 ; Malavoi et Bravard, 2010).

### **I.3.2.3. Érosion Hydrique**

L'érosion hydrique, qui englobe l'érosion de surface, est un processus complexe qui affecte les sols et les formations superficielles, ainsi que le substratum géologique. Ce phénomène peut être naturel ou accéléré par les activités humaines. Il est principalement causé par l'action des eaux courantes, telles que le ruissellement et les petites coulées boueuses. Contrairement aux grands mouvements de terrain, l'érosion de surface se produit sur une faible profondeur, généralement de quelques centimètres à quelques décimètres. Ce processus se déroule en deux phases : la première consiste en l'enlèvement du substrat, tandis que la deuxième implique le transport et le dépôt des sédiments érodés (Rey et al., 2004 ; Balasubramanian, 2017 ; Toy et al., 2002).

### **I.3.3 Morphologie**

La morphologie d'un cours d'eau correspond à ce qu'on peut observer de son tracé sur une carte topographique ou sur une photographie aérienne. Il peut être, par exemple, rectiligne (zones de montagne), méandreux (plaines) ou encore composé de chenaux d'écoulement multiples (tresses). La morphologie du cours d'eau dépend aussi des sédiments que celui-ci charrie et de la géologie de la région, impactant notamment les pentes des berges et la forme des zones inondables (Beaulieu et al., 2007).

### **I.3.3.1. Lit du cours d'eau**

Le niveau d'eau et la surface occupée du canal varie en fonction du débit. Un cours d'eau analysée à partir d'une section transversale présente une section principale occupée par les écoulements normaux (appelé lit mineur) et une plaine d'inondation occupée lorsque le cours d'eau est en crue (Verniers, 1995).

Les principaux termes sont :

- **Lit** : désigne tout l'espace occupé, en permanence ou temporairement, par un cours d'eau (Verniers, 1995).
- Le **lit mineur ou lit ordinaire** est calé en générale entre deux berges bien manquées dans la topographie ; il correspond au passage des hautes eaux. Il peut être formé d'un chenal unique, plus ou moins sinueux, ou être décomposé en plusieurs bras. A l'intérieur de ce lit mineur, les sédiments étant remaniés fréquemment par les forts débits, la croissance de la végétation est relativement limitée : on voit donc de grands bancs nus de sable ou de galets encadrant ou séparant les chenaux (Amat et al., 2008).
- Le **lit majeur** est appelé **aussi plaine d'inondation**. Il est occupé partiellement ou totalement par les eaux selon la puissance des crues. Un examen de la topographie de la plaine révèle qu'il ne s'agit pas d'un milieu plan et uniforme. On y trouve des dépressions, des bras morts au fond desquels affleure la nappe, des buttes à peine marquées qui correspondent à d'anciens bancs ou d'anciennes îles... Lorsque l'homme l'a épargnée, une forêt alluviale occupe la plaine. La végétation y montre une grande variété selon la fréquence de submersion et la nature des sédiments déposés. La largeur de la plaine dépend du contexte morpho structural et de la puissance des crues. D'un point de vue hydrologique, les plaines jouent un rôle majeur. En assurant la rétention des eaux de crue, elles permettent de diminuer la lame d'eau écoulée vers l'aval. A l'occasion des inondations, l'infiltration des eaux dans la plaine permet de recharger la nappe, qui, au moment des basses eaux, soutiendra à son tour l'écoulement. Les mécanismes de débordement, d'infiltration des eaux entre le lit mineur et la plaine assurent aussi l'épuration des eaux, les échanges de substances dissoutes et de matière organique... La proximité de l'eau et le fait que certaines plaines soient fertiles expliquent que les lits majeurs aient connu une mise en valeur très précoce par les premiers agriculteurs et qu'à l'heure actuelle les grandes plaines

consultent des sites de développement urbain et industriel de première importance (Amat et al., 2008).

- **Le chenal d'étiage**, qui est compris dans le lit mineur et qui est utilisé par l'écoulement de basses eaux. Souvent ce chenal serpente entre les berges du lit mineur. Il est généralement très irrégulier, avec de petits rapides à la traversés de bancs d'alluvions que le débit trop faible ne parvient pas à entailler, et des mares stagnantes, presque sans écoulement, où pullulent insectes et algues, dans les « mouilles » séparées par ces « seuils » (Tricart, 1960).
- **Plaine d'inondation** : Au cours du 20e siècle, de vastes secteurs de plaines d'inondation ont été drainés et isolés des cours d'eau d'alimentation par des travaux d'endiguement artificiel. L'eau qui se répandait lentement, pratiquement à la surface, sur toutes les plaines d'inondation se trouve aujourd'hui confinée dans des zones toujours plus restreintes. De plus ces espaces ont été aménagés notamment pour l'agriculture mais surtout pour permettre aux populations de s'y installer. En conséquence, les crues sont plus importantes et plus susceptibles de causer des dommages plus importants – parfois catastrophiques – lorsque les digues se rompent, (Agence européenne pour l'environnement, 2006)
- **Aval et amont** : L'amont (vers la montagne) est la partie la plus élevée du cours d'eau du point de vue de l'observateur et l'aval (vers la vallée) est la partie la plus basse. Il faut se tourner dans le sens de l'écoulement du cours d'eau (de l'amont vers l'aval) pour définir la rive droite et la rive gauche d'un cours d'eau.
- **La rive gauche et la rive droite** : la rive correspond à la limite du lit mineur. Pour le littoral, on parle de ligne littorale (Andréassian et al., 2012).

La rive gauche d'un cours d'eau se trouve à gauche d'un observateur qui est placé dans le sens de l'écoulement de l'eau, c'est-à-dire depuis l'amont vers l'aval.

La rive droite d'un cours d'eau se trouve à droite d'un observateur qui est placé dans le sens de l'écoulement de l'eau, c'est-à-dire depuis l'amont vers l'aval.

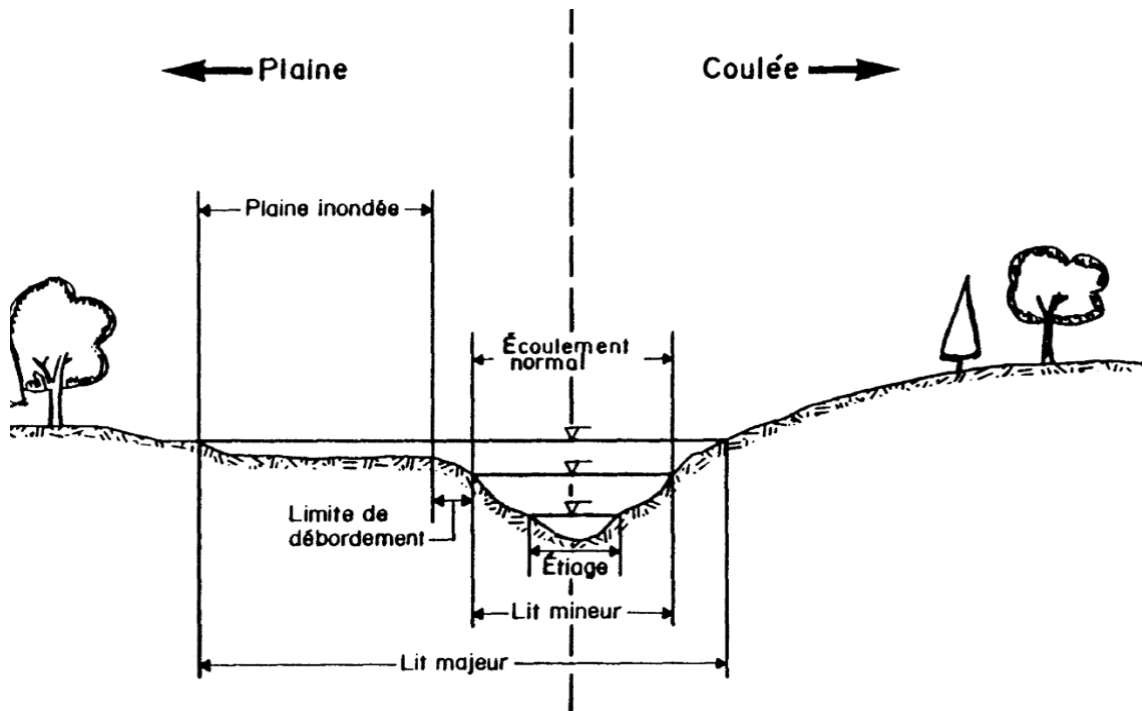


Figure. I. 3. Coupe transversale d'un cours d'eau (Verniers,1995).

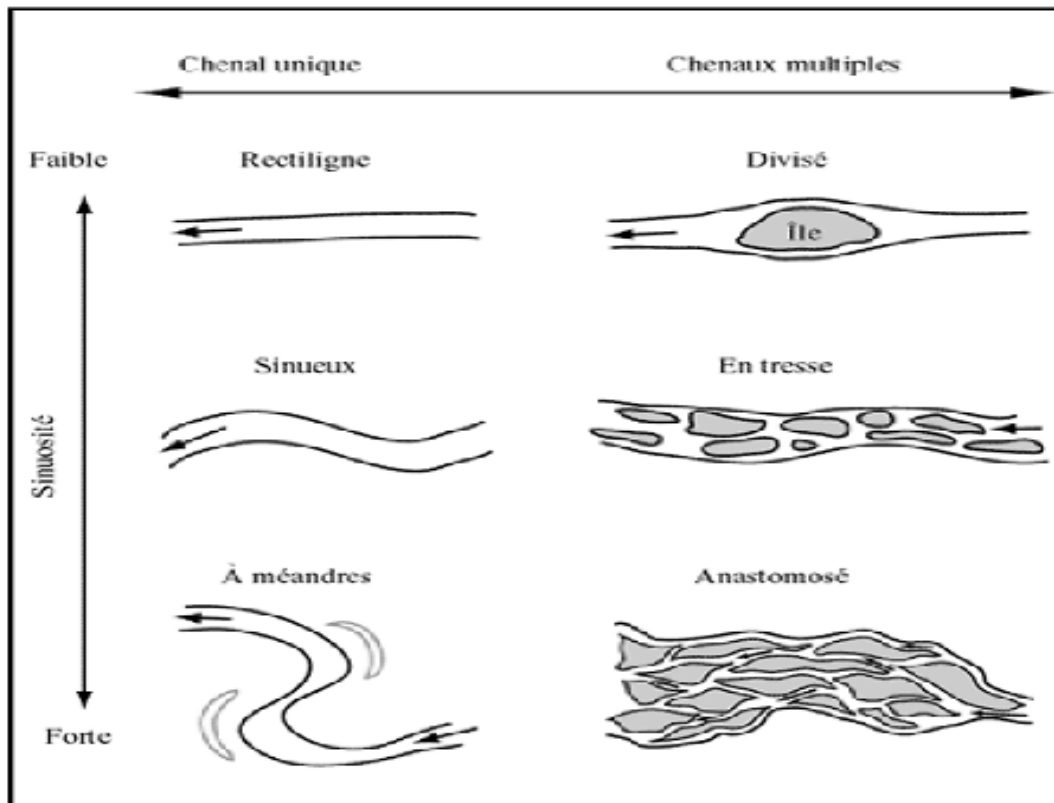
### I.3.4 Les principaux styles géomorphologiques des cours d'eau

On peut distinguer plusieurs styles géomorphologiques fluviaux caractérisés par des formes en plan différentes qui dépendent de niveaux d'énergie plus ou moins forts. Plus l'énergie est élevée, plus la vitesse de régénération des formes fluviales est grande (Amat et al., 2008) :

- **Les cours d'eau à chenal rectiligne** ou faiblement sinueux sont rares. Il s'agit en général de petits cours d'eau calmes, caractérisés par des débits assez faibles et peu contrariés. Ces petits appareils fluviaux transportent sur des pentes assez faibles une charge en suspension très nettement majoritaire. Les formes du chenal sont relativement stables. Cependant, il existe aussi des cours d'eau rectilignes particulièrement dynamiques : on les trouve en tête des cours d'eau de montagne, profondément encaissés entre des versants raides. Ils ont des pentes très raides et irrégulières (comme des marches) et transportent des blocs et gros galets.
- **Les chenaux à méandres** sont développés par des cours d'eau plus dynamiques, dont les débits peuvent être très importants et les crues puissantes (le Mississippi. Par exemple). L'indice de sinuosité, calculé en faisant le rapport entre la longueur de chenal et celle de son axe, est supérieur à 1,5. Il s'agit en général de cours d'eau de plaine aux pentes moyennes ou faible, caractérisés par une charge sédimentaire

mixte (transports en suspension et charge de fond). Les méandres évoluent par migration latérale : le chenal érode la berge concave et dépose les sédiments sur la berge convexe, construisant ainsi un banc de convexité. Le profil en travers montre une nette dissymétrie : du côté de la berge concave, le chenal est profond (la mouille) et le fond du chenal se relève en direction du banc de convexité. A l'occasion d'une forte crue, le méandre peut être recoupé par déversement, ou alors par tangence, Ce méandre est alors abandonné et forme un lac en ox-bow. Dans sa partie amont se forme un bouchon alluvial qui le déconnecte du nouveau bras, la partie médiane et aval restant souvent en eau grâce à la nappe. Les méandres peuvent connaître d'importants changements de tracé : la plaine est alors « tapissée » d'anciens méandres abandonnés, comme c'est le cas en Amazonie.

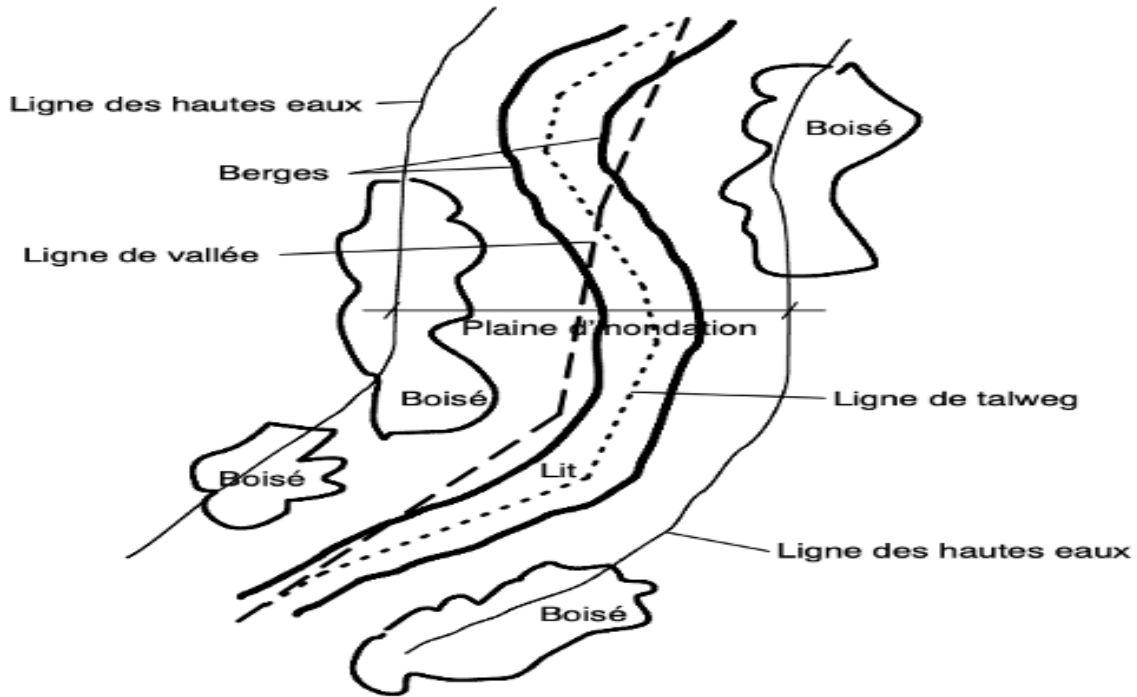
- **Les chenaux en tresse** sont constitués d'un lit large où de nombreux bras enserrant de grands bancs de galets ou de sable nus ; ils sont généralement mis en place par des cours d'eau à fort niveau d'énergie : pente forte, fortes variations de débit, crues violentes charriant sur le fond du chenal une charge solide abondante et de gros calibre. A l'occasion de crues aux écoulements rapides et très turbulents, les bancs migrent vers l'aval, et le chenal en érodant ses berges peut se déplacer latéralement.



**Figure. I. 4.** Classification des styles cheneaux (Church, 2006 ; Nanson et Knighton, 1996).

### **I.3.5 Vue longitudinale**

L'espace longitudinal et latéral qu'occupe un cours d'eau et ses composantes (chenal principal et plaine d'inondation) est appelé "corridor du cours d'eau" (figure I.5). La figure présente aussi les principaux termes utilisés.



**Figure. I. 5.** Concept de corridor d'un cours d'eau (Verniers, G. 1995).

La sinuosité est définie comme le rapport de la longueur de la ligne de talweg sur la longueur de la ligne de vallée pour un tronçon (Peiry, J. 1989).

$$\text{Sin} = \frac{L_{ta}}{L_{va}} \quad (\text{I.3})$$

Sin = sinuosité

$L_{ta}$  = Longueur de la ligne de talweg (L)

Lorsque cette valeur est supérieure à 1,3, le tronçon du cours d'eau est considéré comme sinueux.

## **I.4. Les écoulements dans les cours d'eau**

### **I.4.1 Classification d'écoulement**

Les écoulements à surface libre peuvent être classés suivant la variation de la profondeur d'écoulement et la vitesse en fonction du temps et de l'espace.

#### **✓ Variabilité dans le temps**

Si les vitesses moyennes  $V$  et la profondeur  $y$  restent constants en termes de magnitude et de direction au fil du temps, on dit que le mouvement est permanent (ou stationnaire). Par conséquent, le débit reste constant (Longuet et Michael, 1953).

La vitesse moyenne  $V$  de l'écoulement est liée au débit  $Q$  par la relation :

$$Q = VA \tag{I.4}$$

Le mouvement est non permanent dans le cas contraire. Au sens strict, l'écoulement dans les canaux est rarement permanent. Néanmoins les variations temporelles sont, dans certains cas, suffisamment lentes pour que l'écoulement puisse être considéré comme une succession de régime permanent. On peut alors définir ainsi le régime quasi-permanent.

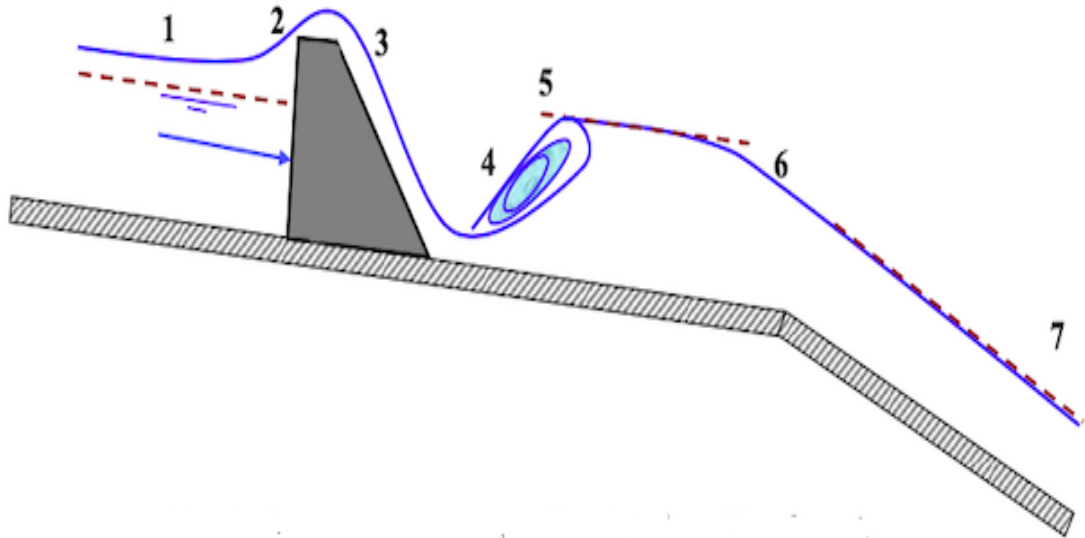
#### **✓ Variabilité dans l'espace**

Le mouvement est uniforme si les paramètres caractérisant l'écoulement restent invariables dans les diverses sections du canal. La ligne de la pente du fond est donc parallèle à la ligne de la surface libre (Chambonet al., 2014 ; Chaudhry, 2008).

Le mouvement est non-uniforme ou varié si les paramètres caractérisant l'écoulement changent d'une section à l'autre. La pente de la surface libre diffère de celle du fond (Whitham, 1958).

Un écoulement non-uniforme peut être accéléré ou décéléré suivant que la vitesse croît ou décroît dans le sens du mouvement (Kironoto et al., 1995)

- Lorsque le mouvement est graduellement varié, la profondeur ainsi que les autres paramètres varient lentement d'une section à l'autre.
- Lorsque le mouvement est rapidement varié, les paramètres caractérisant l'écoulement changent brusquement sur une courte distance.
- Lorsque le mouvement évolue graduellement, la profondeur, notée  $D_h(x) = D_h$ , ainsi que les autres paramètres, changent seulement de manière très lente d'une section à l'autre.



**Figure. I. 6.** Schéma représentant les différents régimes d'écoulement (Träsch, 2019).

- 1) uniforme fluvial.
- 2) fluvial graduellement décéléré.
- 3) rapidement accéléré (fluvial puis torrentiel).
- 4) ressaut.
- 5) uniforme fluvial.
- 6) rapidement accéléré (fluvial puis torrentiel).
- 7) uniforme torrentiel.

Lorsque le mouvement varie rapidement, la profondeur  $D_h(x)$  ainsi que les autres paramètres changent brusquement, parfois avec des discontinuités. Ces changements se manifestent généralement près d'une singularité telle qu'un déversoir, un rétrécissement, un ressaut hydraulique ou une chute abrupte (Boudjemia et Touati, 2016).

#### **I.4.2 Régime d'écoulement**

L'écoulement d'un fluide réel dans un canal à surface libre est soumis aux forces suivantes (Chaudhry, 2008) :

- Forces d'inertie ;
- Forces de gravité ;
- Forces de frottement (viscosité et rugosité).

Pour l'étude hydraulique des canaux, on définit habituellement les nombres adimensionnels suivants :

➤ **Nombre de Froude**

C'est le rapport entre les forces d'inertie et celles de gravité où (Donelan and Kram 2000) :

$$F_r = \frac{V}{\sqrt{gy}} \quad (\text{I.5})$$

Où V est la vitesse moyenne de l'écoulement, g est l'accélération de la pesanteur et y est la hauteur d'écoulement.

Le rôle du nombre de Froude est de permettre le classement des écoulements comme suit :

Écoulement fluvial :  $F_r < 1$  ;

Écoulement critique :  $F_r = 1$  ;

Écoulement torrentiel :  $F_r > 1$ .

➤ **Nombre de Reynolds**

Le passage d'un régime à l'autre dépend de la valeur d'un paramètre adimensionnel, le nombre de *Reynolds*. Celui-ci est par définition le rapport des forces d'inertie aux forces de viscosité, et s'écrit (Rott,1990) :

$$R = \frac{VD_h}{\nu} \quad (\text{I.6})$$

Où V est la vitesse moyenne de l'écoulement, Dh est le diamètre hydraulique et  $\nu$  est la viscosité cinématique du liquide en écoulement.

En tenant compte de la relation (I.4) et  $D_h = 4A/P$ , la relation précédente s'écrit :

$$R = \frac{4Q}{P\nu} \quad (\text{I.7})$$

## **I.5. Distribution de la vitesse dans les écoulements à surface libre**

### **I.5.1 Distribution des vitesses dans un écoulement uniforme**

La distribution de la vitesse dans un écoulement turbulent demeure quasi uniforme lorsque la couche limite est pleinement développée. La distribution des vitesses suit approximativement une loi logarithmique. La contrainte de cisaillement ou tangentielle en n'importe quel point de l'écoulement turbulent, se produisant au-dessus d'une paroi solide, est donnée par la relation de Prandtl (1926) :

$$\tau = \rho l^2 (dv / dy)^2 \quad (\text{I.8})$$

$\rho$  est la masse volumique du liquide en écoulement,  $l$  est une longueur caractéristique dite longueur de mélange,  $dv/dy$  est le gradient de vitesse à la hauteur  $y$  de la paroi solide et normale à celle-ci. La longueur de mélange peut être interprétée par la longueur au-delà de laquelle la particule liquide voit sa quantité de mouvement diminuée.

Dans la région proche de la paroi solide, Prandtl (1926) utilise deux approches :

1. La longueur de mélange est proportionnelle à  $y$ , soit  $l = Ky$  où  $K$  est le facteur de proportionnalité entre  $l$  et  $y$  et dont la valeur a été estimée à 0,40 environ.
2. La contrainte tangentielle est constante.

Puisque la contrainte tangentielle à la surface est égale à la force tractrice unitaire  $\tau_0$ , la constance de la contrainte tangentielle implique que  $\tau = \tau_0$ . La relation (I.8) devient :

$$dv = (1/k)(\tau_0/\rho)^{1/2}(ydy) \quad (\text{I.9})$$

L'intégration de la relation (I.9) mène à :

$$V = 2,5\sqrt{(\tau_0/\rho)}\ln(y/y_0) \quad (\text{I.10})$$

«Ln» désigne le logarithme népérien et  $y_0$  représente la constante d'intégration. La force tractrice  $\tau_0$  s'exprime par la relation (I.8).

On peut alors écrire que :

$$\sqrt{\tau_0/\rho} = V_f \quad (\text{I.11})$$

ou  $V_f$  a la dimension d'une vitesse et elle est connue sous le nom de vitesse de frottement.

Tenant compte de (I.10), la relation (I.8) permet d'écrire que :

$$V_f = \sqrt{gR_h i} \quad (\text{I.12})$$

La relation (I.9) devient alors :

$$V = \frac{5}{2}V_f \ln \frac{y}{y_0} \quad (\text{I.13})$$

La relation (I.12) indique que la vitesse dans un écoulement turbulent est une fonction logarithmique de la distance  $y$ . Elle est connue sous le nom de loi universelle de Prandtl Von – Karman de la distribution des vitesses. Cette loi a été vérifiée par plusieurs expériences et les résultats ont montré une remarquable similitude entre la distribution des vitesses observée expérimentalement et celle issue de la théorie. Lorsque la surface solide

est lisse, la constante  $y_0$  ne dépend que de la vitesse de frottement  $V_f$  et de la viscosité cinématique  $\nu$  du liquide :

$$y_0 = m_0 \frac{V}{V_f} \quad (\text{I.14})$$

$m_0$  Est une constante égale à 1/9 lorsque la surface solide est lisse. Cette constante a été déduite des essais de *Nikuradse* concernant les conduites lisses. La combinaison des relations (I.12) et (I.13) donne ainsi, pour les surfaces lisses, la répartition de la vitesse dans un écoulement turbulent :

$$V = \frac{5}{2} V_f L n \frac{9yV_f}{\nu} \quad (\text{I.15})$$

Lorsque la surface est rugueuse, la constante  $y_0$  dépend de la rugosité absolue  $\varepsilon$  :

$$y_0 \approx m_0 \varepsilon \quad (\text{I.16})$$

La constante  $m_0$  est approximativement égale à 1/30 et la relation (I.12) devient alors :

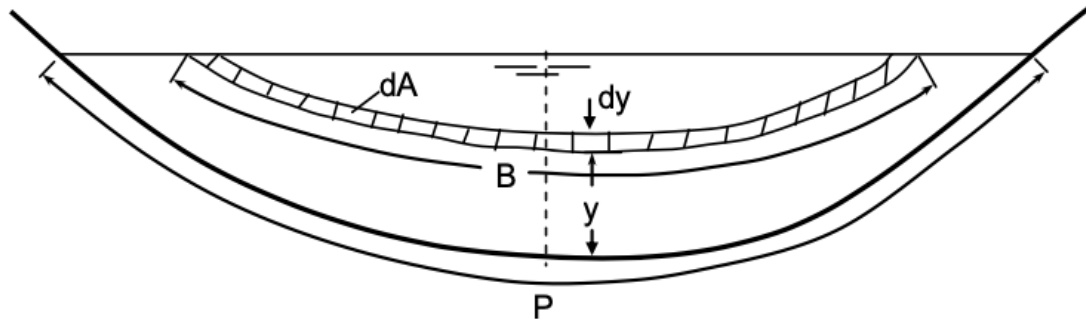
$$V = \frac{5}{2} V_f L n \frac{30y}{\varepsilon} \quad (\text{I.17})$$

### **I.5.2 Équation de Keulegan**

En utilisant la loi universelle de *Prandtl-Von-Karman* de la distribution des vitesses, *Keulegan* (1938) aboutit à des équations donnant la vitesse moyenne d'un écoulement turbulent dans les canaux ouverts par une approche théorique simple. En se basant sur l'équation de continuité, le débit volume  $Q$  passant par une section quelconque de l'écoulement peut être écrit :

$$Q = VA = \int_{\delta_0=0}^{y=h} VB dy \quad (\text{I.18})$$

$V$  est la vitesse moyenne de l'écoulement,  $h$  est la profondeur de l'écoulement,  $A$  est l'aire de la section mouillée,  $B$  est la longueur de la courbe d'égale vitesse (figure I.8) et  $y$  est la profondeur verticale comptée à partir de la surface jusqu'à la courbe d'égale vitesse. La sous couche limite laminaire d'épaisseur  $\delta_0$  est considérée comme étant très mince ( $\delta_0 = 0$ ).



**Figure. I. 7.** Schéma de définition ayant servi de base à l'établissement de l'équation de Keulegan. Le maximum de vitesse est supposé être à la surface et que la longueur  $B$  est proportionnelle à la distance verticale  $y$ . Ceci permet d'écrire :

$$B = P - \gamma y \quad (\text{I.19})$$

$P$  est le périmètre mouillé de la surface considérée et  $\gamma$  est une fonction dépendant de la forme de la section. Ainsi, l'aire de la section mouillée  $A$  est :

$$A = \int_0^h B dy = Ph - \frac{1}{2} \gamma h^2 \quad (\text{I.20})$$

La combinaison des relations (I.9), (I.17), (I.18) et (I.19) permet d'écrire, après intégration :

$$V = V_f \left[ 5,751 \log \left[ \frac{h}{m_0 R_h} \exp \left( -1 - \frac{\gamma h^2}{4A} \right) \right] + 5,75 \log \frac{m_0 R_h}{y_0} \right] \quad (\text{I.21})$$

Le premier terme du membre droit de l'équation (I.20) est une fonction de la forme de la section du canal considéré. Cependant, la variation que subit ce terme pour différentes formes de canaux est relativement faible et il a été remplacé par une constante désignée par  $A_0$ . La relation (I.20) prend alors une forme plus simplifiée et s'écrit :

$$V = V_f \left[ A_0 + 5,751 \log \left( \frac{m_0 R_h}{y_0} \right) \right] \quad (\text{I.22})$$

La relation (I.21) représente l'équation théorique générale de la vitesse moyenne de l'écoulement dans les canaux ouverts.

Pour les canaux à paroi lisse, l'étude de *Keulegan* (1938), basée sur les valeurs expérimentales de Nikuradse, montre que  $A_0 \approx 3,25$ . Ainsi, l'équation théorique générale de la vitesse moyenne de l'écoulement uniforme dans les canaux ouverts à paroi lisse est, en tenant compte de (I.13) :

$$V = V_f \left[ 3,25 + 5,751 \log \left( \frac{R_h V_f}{\nu} \right) \right] \quad (\text{I.23})$$

En ce qui concerne les canaux ouverts à paroi rugueuse, Keulegan (1938) montre, après avoir analysé les mesures de Bazin, que la constante  $A_0$  varie dans une large gamme en fonction de la forme de la section du canal ( $3,23 \leq A_0 \leq 16,92$ ), et la valeur  $A_0 = 6,25$  est alors adoptée. Ainsi, l'expression théorique générale de la vitesse moyenne de l'écoulement uniforme dans les canaux ouverts à parois rugueuses est, en tenant compte de (I.15) :

$$V = V_f \left[ 6,25 + 5,75 \log \left( \frac{R_h}{\varepsilon} \right) \right] \quad \text{Pour les canaux rugueux} \quad (\text{I.24})$$

En combinant les expressions  $V = C \sqrt{R_h i}$  de Chézy et  $\sqrt{g R_h i}$  de la vitesse de frottement, on peut écrire :

$$\frac{V}{V_f} = \frac{c}{\sqrt{g}} \quad (\text{I.25})$$

En tenant compte de la relation (I.24) et de la définition du nombre de *Reynolds* modifié tel que  $R_e = R_h V / \nu$ , les relations (I.22) et (I.23) permettent alors d'exprimer le coefficient  $C$ , respectivement pour un canal à paroi lisse et rugueuse :

$$c = 32,6 \log \frac{20,8 R_e}{c} \quad \text{Pour les canaux lisses} \quad (\text{I.26})$$

$$c = 32,6 \log \frac{12,2 R_e}{k} \quad \text{Pour les canaux rugueux} \quad (\text{I.27})$$

Une autre étude a été effectuée par Iwagaki sur des données expérimentales obtenues à partir de différentes sources. Les résultats de l'étude ont révélé que la résistance à l'écoulement turbulent dans les canaux ouverts devient évidemment plus grande que celle dans des conduites avec augmentation du nombre de Froude. Iwagaki conclut que la grande instabilité de la surface libre est due aux nombres de Froude élevés. Il est possible d'introduire l'effet de l'instabilité de surface libre dans les équations de Keulegan. L'équation (I.23) et (I.24) peut être exprimée comme suit :

$$V = V_f \left( A_s + 5,75 \log \frac{R V_f}{\nu} \right) \quad \text{Pour les canaux lisses} \quad (\text{I.28})$$

$$V = V_f \left( A_r + 5,75 \log \frac{R V_f}{\nu} \right) \quad \text{Pour les canaux rugueux} \quad (\text{I.29})$$

## **I.6. Mesure de l'écoulement dans les cours d'eau**

Pour mesurer le débit d'un écoulement naturel dans un cours d'eau ou un canal. Il existe quatre grandes catégories de méthodes (Musy et Higy, 2004).

- Le méthode volumétrique (ou jaugeage capacitif) permet de déterminer le débit directement à partir du temps nécessaire pour remplir d'eau un récipient d'un volume donné. Prise en compte des aspects pratiques spécifiques à la méthode de mesure (taille de conteneur requise, incertitude temporelle, emplacement spécifique possible). Cette procédure n'est généralement effectuée qu'à des débits très faibles, au plus quelques litres/seconde.
- Les méthodes hydrauliques tiennent compte des forces qui régissent l'écoulement (pesanteur. Inertie. Viscosité...). Ces méthodes obéissent aux lois de l'hydraulique
- La manière d'explorer le champ de vitesse consiste à déterminer la vitesse d'écoulement en différents points de la section, tout en mesurant la surface de la section mouillée. Ces techniques nécessitent du matériel spécifique (moulinets, perche, saumon, courantomètre... etc.) et du personnel formé à leur utilisation. Parmi les nombreuses méthodes d'étude des champs de vitesse, les jaugeages au moulinet et au flotteur, ainsi que les principes de fonctionnement des capteurs électromagnétiques.
- Les méthodes physico-chimiques tiennent compte de la variation. Certaines propriétés physiques d'un liquide (concentration de certains éléments dissous) lors de son écoulement. Ces méthodes consistent généralement à injecter des objets en solution dans les cours d'eau et à surveiller leur concentration dans le temps. Ce sont les méthodes dites chimiques ou de dilution.

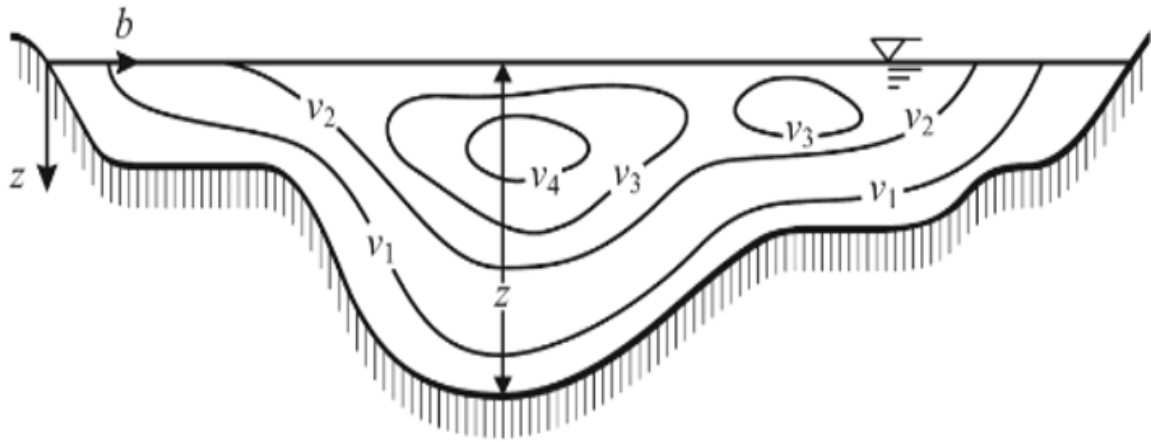
Toutes ces méthodes de mesures des débits nécessitent généralement un régime d'écoulement en régime fluvial, sauf les jaugeages chimiques, qui sont appropriés en cas d'écoulement torrentiel.

La mesure de l'écoulement dans un cours d'eau est une opération relativement aisée et routinière. On présente les pratiques de cette opération ci-dessous (Nicolas, 2012).

### **I.6.1 Mesure instantanée**

La mesure du débit d'un cours d'eau nécessite le déploiement d'une équipe de professionnels disposant de l'équipement nécessaire pour déterminer leurs positions dans l'espace et mesurer la vitesse et la profondeur de l'écoulement avec une plus grande précision possible. La principale difficulté opérationnelle, outre la sécurité du personnel, est la variance de vitesse des rivières, comme le montre la Figure I.9. Par exemple. Il est

donc nécessaire de déterminer la répartition des vitesses sur toute la section d'écoulement  
Les méthodes la plus courante (Nicolas, 2012) :



**Figure. I. 8.** Distribution typique des vitesses dans une section transversale (Nicolas, 2012).

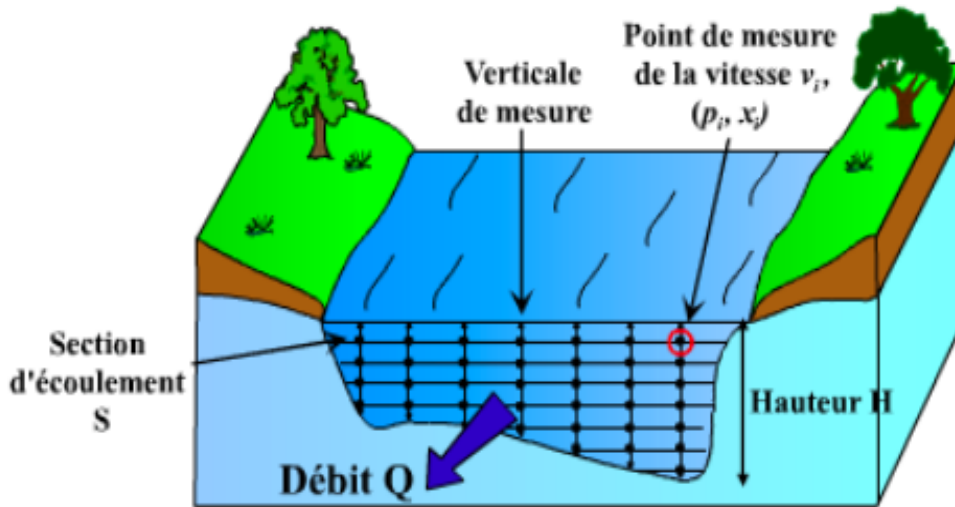
### **I.6.2 Jaugeage par exploration du champ de vitesse**

Rappelons que la vitesse d'écoulement n'est jamais uniforme dans la section transversale d'un cours d'eau. Le principe de cette méthode consiste donc à calculer. Le débit à partir du champ de vitesse déterminé dans une section transversale du cours d'eau (en un certain nombre de points. Situés le long de verticales judicieusement réparties sur la largeur du cours d'eau). Parallèlement à cette exploration du champ de vitesse, on relève le profil en travers du cours d'eau en mesurant sa largeur et en effectuant des mesures de profondeur. Le débit  $Q$  ( $m^3/s$ ) s'écoulant dans une section d'écoulement  $S$  ( $m^2$ ) d'une rivière (section mouillée) peut être défini à partir de la vitesse moyenne  $V$  [ $m/s$ ] perpendiculaire à cette section par la relation :  $Q = V \cdot S$ .

La section d'écoulement peut être évaluée en relevant la profondeur d'eau en diverses réparties régulièrement sur toute la largeur.

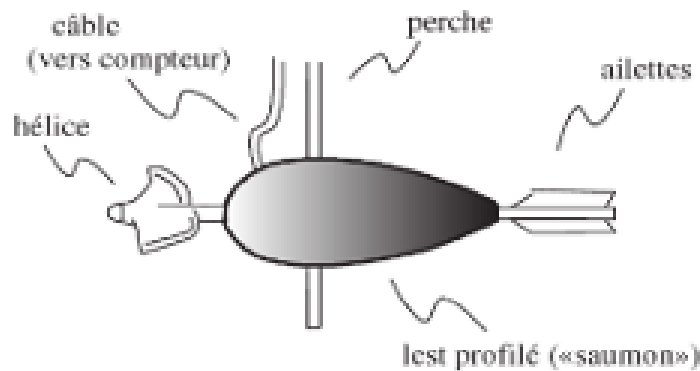
### **I.6.3 Le jaugeage au moulinet**

Le moulinet hydrométrique permet de mesurer la vitesse ponctuelle de l'écoulement. Le nombre de mesures sur une verticale est choisi de façon à obtenir une bonne description de la répartition des vitesses sur cette verticale. De manière générale, on fera entre 1, 3 ou 5 mesures suivant la profondeur du lit Figure I.9 (Diconne, 1972).



**Figure. I. 9.** Débits et champ de vitesse à travers une section (Ammari et Remini, 2014).

La vitesse d'écoulement est mesurée en chacun des points à partir de la vitesse de rotation de l'hélice située à l'avant du moulinet (nombre de tours  $n$  par unité de temps). La fonction  $v = f(n)$  est établie par une opération d'étalonnage (courbe de tarage du moulinet). Suivant le mode opératoire adopté pour le jaugeage. Le moulinet peut être monté sur une perche rigide ou sur un lest profilé appelé saumon Figure I.10.



**Figure. I. 10.** Jaugeage au moulinet (Subramanya, 2008).

#### **I.6.4 Mesure en continu**

L'implantation d'une station de jaugeage de l'écoulement, ou station hydrométrique, implique la mesure simultanée du débit et du niveau de l'eau à une même section transversale d'un cours d'eau, pour l'ensemble des débits possibles à ce site. Il faut donc

visiter le site à de nombreuses reprises et selon diverses conditions hydro climatiques avant d'établir une relation mathématique qui lie les niveaux et les débits. Cette relation s'appelle courbe de tarage. Il ne faut toutefois pas sous-estimer la difficulté d'obtenir des observations de débits à l'occasion d'événements extrêmes, puisqu'une équipe de terrain n'est pas nécessairement disponible à ce moment-là. Ces observations sont toutefois essentielles à la détermination d'une courbe de tarage de qualité, qui mènera à des erreurs d'estimation limitées à 5 ou 10 % en interpolation (Hirsch et Costa, 2004)

La forme généralement utilisée pour exprimer une courbe de tarage est une fonction de puissance, telle que :

$$Q = ah^b$$

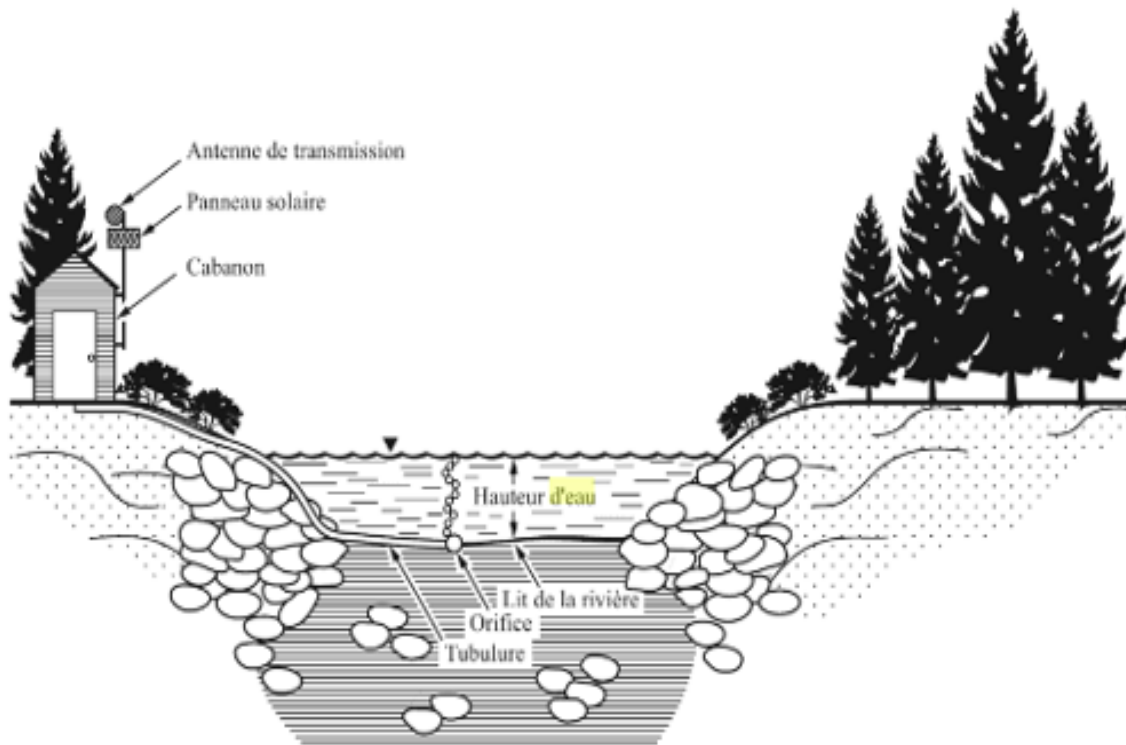
Où  $h$  est la hauteur de niveau de l'eau, et les paramètres  $a$  et  $b$  sont déterminés par régression linéaire. On peut employer aussi une forme polynomiale :

$$Q = a_0 + a_1h + a_2h^2 + a_3h^3 + \dots$$

Où encore la variante suivante de la fonction de puissance :

$$Q = a(h - h_{ref})^b$$

Où  $h_{ref}$  est une hauteur de référente.



**Figure. I. 11.** Station de jaugeage typique (Nicolas, 2012).

Le choix de la section de jaugeage reste une étape cruciale pour faciliter les opérations, mais surtout pour éviter de devoir réviser la courbe de tarage, à la suite de modifications qui touchent la section de contrôle. Il faut donc privilégier les sites (Aharony et al., 2013) :

- Facilement accessibles ;
- Immédiatement en amont d'une section de contrôle (section d'un canal découvert ou le débit est déterminé uniquement par le niveau d'eau) ;
- Situés le long d'un bief rectiligne (lignes de courant parallèles et normales à la section transversale) ;
- Dépourvus de plantes aquatiques envahissantes ;
- Dont le fond et les berges sont stables ;
- Dont la profondeur moyenne est supérieure à 0,3 m ;
- Dont la vitesse moyenne est supérieure à 0.10 m/s ;
- Sans débordement printanier dans une plaine d'inondation ;
- Sans accumulation de frasil en hiver.

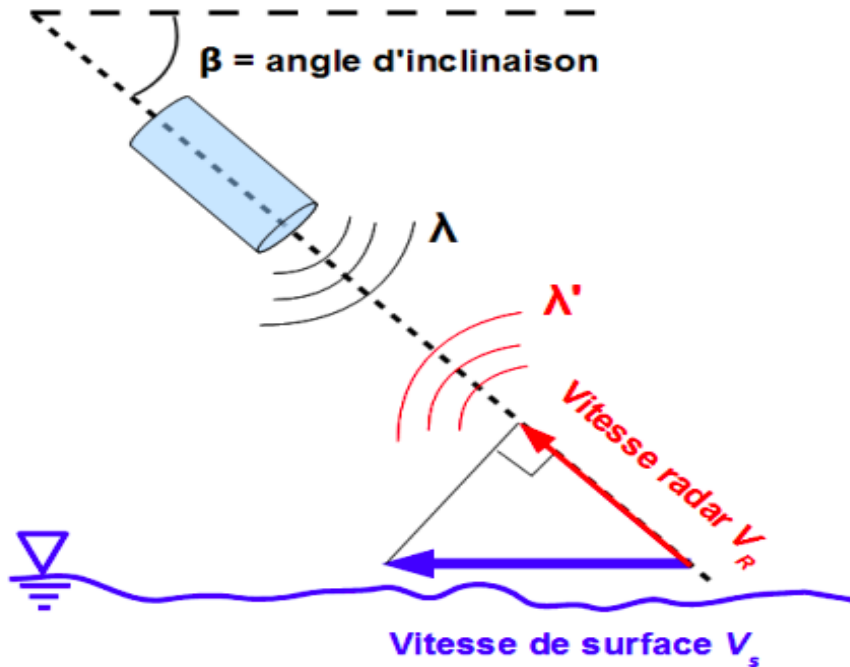
### **I.6.5 Jaugeage par radar mobile**

Le principe de mesure du capteur repose sur l'émission d'impulsions d'ondes électromagnétiques très courtes, suivies de l'interprétation du signal de retour. Ces capteurs émettent généralement dans les bandes de fréquence ISM (Industriel, Scientifique, et Médical). Le modèle de radar utilisé est le capteur SVR (Decatur Electronics) qui utilise la technologie FMCW (Frequency Modulated Continuous Waves) avec une fréquence de 24,1 GHz (Dramais et al., 2011)

Les ondes émises par le radar sont réfléchies par la rugosité de la surface de l'écoulement (comme illustré dans la figure 1), ce qui entraîne une modification de la fréquence rétrodiffusée proportionnelle à la vitesse de l'écoulement projetée le long de l'axe de visée, selon un effet Doppler. Il est donc essentiel de connaître l'angle d'inclinaison du radar par rapport à l'horizontale. Dans le cas du SVR Decatur, cet angle est fourni par un inclinomètre interne (Dramais et al., 2014).

Afin de mesurer correctement la vitesse de l'écoulement, le niveau de signal reçu doit être suffisant. Pour cela, la longueur d'onde des ondes réfléchies doit respecter la condition de Bragg (Costa et al., 2006). En pratique, cette technique fonctionne efficacement sur une surface d'écoulement suffisamment rugueuse (avec des vaguelettes en

surface d'une longueur d'onde de 0,9 mm pour le capteur SVR, en utilisant un angle incident de 45°). Augmenter l'angle de visée (compris entre 30° et 60° pour un SVR) permet d'augmenter le signal de retour radar, mais cela entraîne également une augmentation des erreurs de projection de la vitesse (Dramais et al., 2014).



**Figure. I. 12.** Mesure par radar mobile (Dramais et al., 2014).

$$V_s = \frac{V_R}{\cos \beta}$$

Avec  $V_s$  La vitesse de la surface,  $V_R$  La vitesse mesurée par le radar en m/s,  $\beta$  l'angle incident mesuré par le SVR en °.

Les mesures radar ne permettent que la mesure partielle du champ de vitesse, limitée à la surface de l'écoulement. Cependant, pour calculer un débit, il est nécessaire de déterminer une vitesse moyenne sur toute la section de l'écoulement. Pour ce faire, on utilise généralement un modèle de distribution verticale des vitesses basé sur une loi logarithmique ou une loi de puissance. Le rapport entre la vitesse moyenne et la vitesse de surface est appelé "coefficient de vitesse" et est représenté par le symbole " $\alpha$ " (Dramais et al., 2014).

$$V_m = V_s \times \alpha$$

Avec  $V_m$  la vitesse moyenne sur la verticale

$V_s$  : La vitesse de surface

$\alpha$  : Le coefficient de vitesse

Le coefficient de vitesse "  $\alpha$  " est évalué conformément à la norme ISO 748 (ISO, 2009). Selon cette norme, sa valeur varie généralement entre 0,80 pour les rivières peu profondes ayant un fond rugueux et 0,91 pour les chenaux profonds ayant un fond plus lisse. Pour des conditions moyennes, le coefficient de vitesse est généralement établi à 0,86. Ces valeurs permettent d'estimer la vitesse moyenne de l'écoulement à partir de la vitesse mesurée à la surface, facilitant ainsi le calcul du débit total de l'écoulement. (Dramais et al., 2014 ; Dramais et al., 2011).

## **I.7. Conclusion**

Dans ce chapitre, nous avons donné certaines notions et définitions générales, relevant les écoulements et les morphologies des cours d'eau, aussi que les méthodes de mesure des débits.

En comprenant les différents types d'écoulements, tels que les écoulements laminaires et turbulents, nous sommes mieux équipés pour analyser la dynamique complexe des cours d'eau. Nous avons également exploré les facteurs qui influencent la morphologie des cours d'eau, tels que le régime hydraulique, la sédimentation et l'érosion. Cette connaissance est essentielle pour appréhender comment les cours d'eau évoluent naturellement et comment les actions humaines peuvent affecter leur forme et leur fonctionnement.

Les méthodes de mesure des débits jouent un rôle critique dans la gestion des ressources en eau. En utilisant des techniques les plus utilisées, telles que les jauges à moulinet. Nous pouvons estimer avec précision le débit des cours d'eau à différents endroits et moments. Cela permet de surveiller l'écoulement de l'eau et de prendre des décisions éclairées en matière de gestion des ressources hydriques, que ce soit pour l'approvisionnement en eau potable, l'irrigation agricole ou la production d'énergie hydroélectrique.

# **CHAPITRE II**

## **ÉTUDE HYDROLOGIQUE DES BASSINS VERSANTS**

## **II.1. Introduction**

Ce chapitre est consacré à l'étude hydrologique et l'analyse des écoulements des cours d'eau naturels des bassins d'étude situés dans un climat méditerranéen au nord-centre de l'Algérie, Pour illustrer les caractéristiques de ces régimes hydrologiques.

L'étude porte sur l'estimation des débits liquides dans plusieurs stations hydrométriques du bassin côtier algérois. Ce bassin est drainé par plusieurs cours d'eau et présente une grande diversité géologique, géomorphologique, climatique et topographique.

Les cours d'eau examinés sont situés dans la partie centre-nord de l'Algérie, entourés par la mer Méditerranée au nord, le bassin côtier Constantinois à l'est, et les bassins des rivières Soummam, Isser et Cheliff au sud et à l'ouest. La frontière sud est principalement délimitée par les montagnes de l'Atlas Tellien (Ammari, 2012 ; Zeroual, 2009).

## **II.2. Cours d'eau étudiés**

L'Etude se focalise sur les cours d'eau naturel situés au Nord-Centre de l'Algérie. Les données dans ce travail ont été rassemblées par l'ANRH pendant une période moyenne 24 ans, pour cinq oued (Stations d'Amont des gorges oued « Chiffa », Station pont RN11 oued « Bellah », Station ARIB dans le bassin d'oued « Chillef », Station de Baghlia oued « Sebaou », ainsi que la station de Fer à Cheval oued « Mazafran »),

L'objet de cette étude se concentre sur les cours d'eau naturels situés au centre de l'Algérie. Les données utilisées dans ce travail ont été collectées sur une période de jaugeage pour cinq oueds spécifiques, à savoir les stations en amont des gorges de l'oued Chiffa, la station du pont RN11 de l'oued Bellah, la station ARIB dans le bassin de l'oued Chillef, la station de Baghlia de l'oued Sebaou, ainsi que la station de Fer à Cheval de l'oued Mazafran. Ces stations hydrométriques ont été sélectionnées en raison de la disponibilité des données de mesure des débits liquides, fournies par les services hydrologiques de l'Agence Nationale des Ressources Hydrauliques (ANRH) en Algérie.

L'ensemble des cours d'eau naturels en Algérie présente un régime d'écoulement irrégulier et très variable avec un volume de sédiments très important (Marouf et Remini, 2011 ; Marouf, 2012), surtout au début de la saison des pluies. De plus, le manque des ressources humaines et matérielles a aggravé la situation (Marouf, 2012 ; Marouf et Remini, 2011).

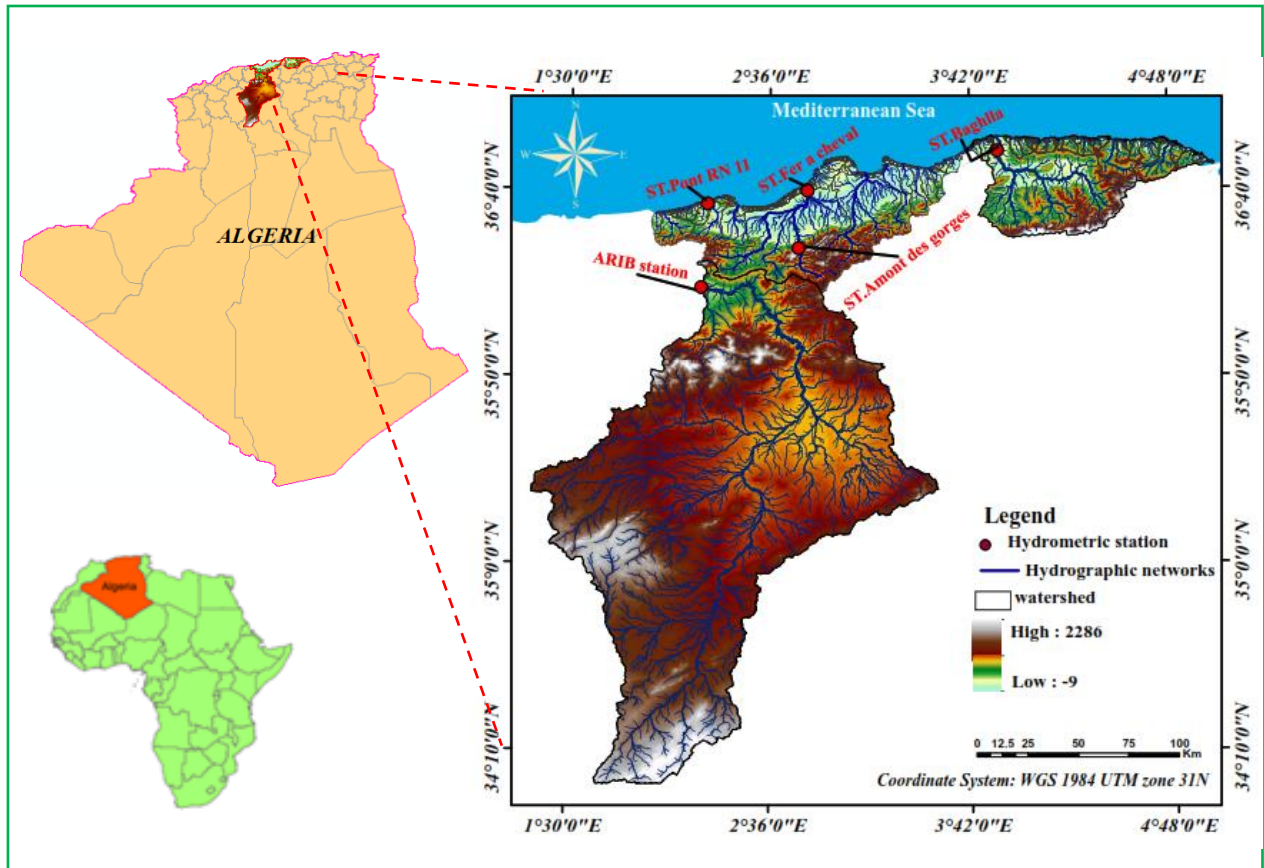


Figure. II. 1. Carte de la station hydrométrique des cinq cours d'eau étudiés dans le nord de l'Algérie.

### II.2.1 Bassin versant d'Oued Chiffa

Oued Chiffa situé au sud-Est de la capitale Alger. Il est limité au Nord par la commune d'oued-el Alleug, de l'Est par la ville de Blida, au Sud par wilaya de Médéa et à l'Ouest par la ville de Mouzaia. Le cours d'eau chiffa prend sa source près de Médéa, arrive dans les gorges, avec une superficie de 342 km<sup>2</sup> et d'un périmètre de 83 km. Ce cours d'eau est situé entre l'Atlas Bliden et le massif de Mouzaia, les hauteurs d'eau et les débits liquides sont mesurés au niveau de la station hydrométrique Amont des gorges (coordonnées Lambert (X = 506,15 m ; Y = 342,3 m).

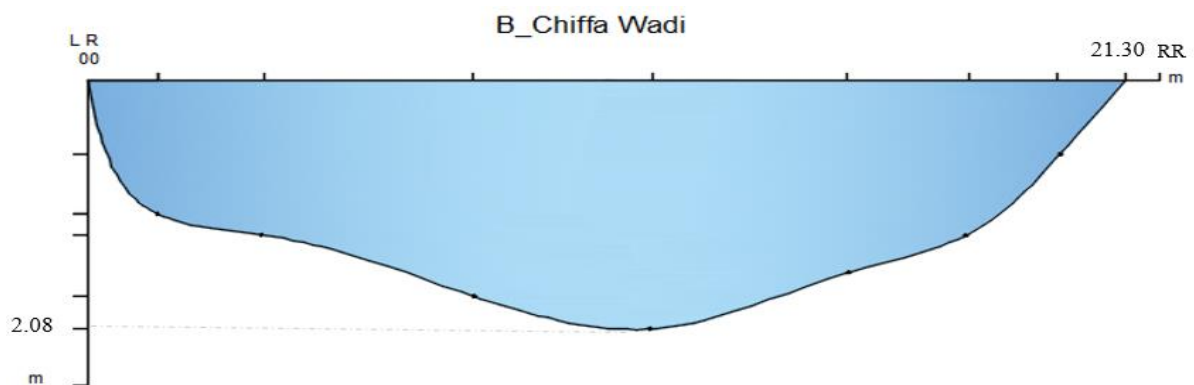


Figure. II. 2. Coupe transversale du cours d'eau naturel "Chiffa"(ANRH).

Tableau. II. 1. Caractéristiques de formes des sous bassins d’Oued chiffa (Station Amont des gorges).

Bassin versant d’Oued Chiffa	
Surface (Km <sup>2</sup> )	342
Périmètre du bassin versant (km)	83
L.thalweg (km)	35
L. rect. Equivalent (Km)	23,5
Altitude max. (m)	1629
Altitude moy. (m)	830
Altitude min. (m)	290
Indice de Pente Globale I <sub>PG</sub> (%)	5,67
Indice de pente de Roche I <sub>PR</sub>	0,2
Indice de Compacité I <sub>c</sub>	1,17
Coefficient de torrentialité C <sub>t</sub>	17,1
Densité de drainage D <sub>d</sub>	3,3

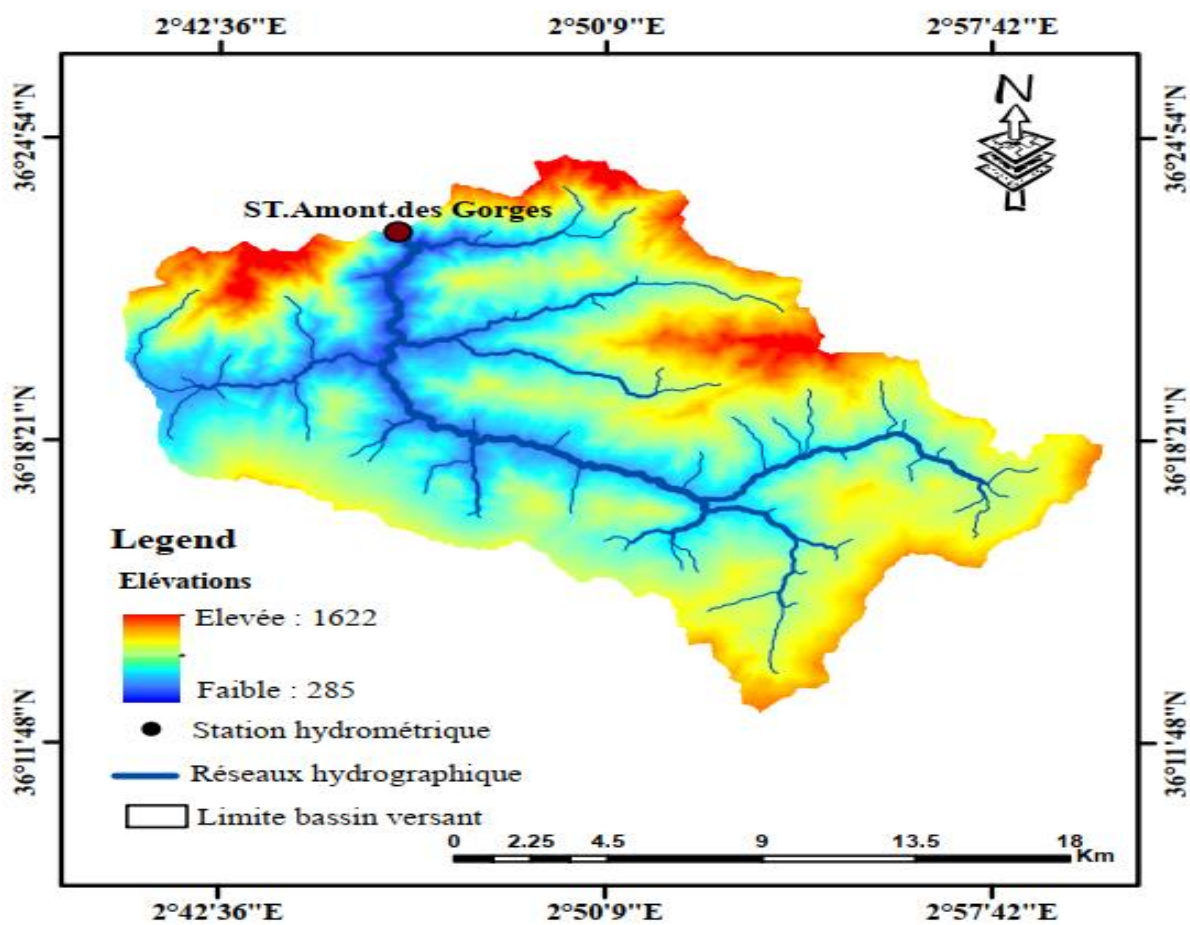
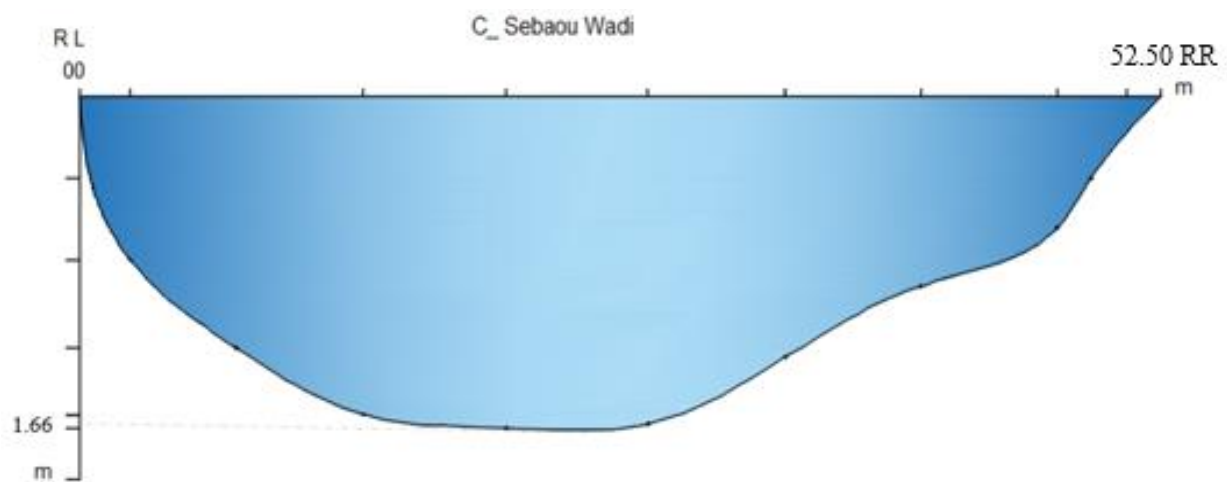


Figure. II. 3. Station Amont des gorges “Oued Chiffa“.

### II.2.2 Bassin versant d'Oued Sebaou

Le secteur d'étude s'étale sur le bassin versant de la Kabylie, celui de l'oued Sébaou (côtier algérois). Il est localisé à environ 80 Km, à l'est d'Alger, administrativement, il chevauche deux willaya, à savoir Tizi-ouzou (haut et moyen Sébaou) et Boumerdès (bas Sébaou). Le bassin versant de l'oued Sébaou est limité au nord par la chaîne littorale et la mer Méditerranée, à l'Est par le bassin versant de la Soummam, à l'Ouest par le bassin versant de l'Isser, au Sud par la chaîne du Djurdjura. La superficie totale du bassin versant de l'oued Sébaou est de 2500 Km<sup>2</sup>, occupant ainsi la majorité de la grete Kabylie et leur périmètre 200 km. L'oued Sébaou est le principal cours d'eau qui traverse la plaine alluviale, cette dernière porte d'ailleurs son nom. Les mesures des débits et la hauteur d'eau par la station Baghlia (coordonnées Lambert (X = 603,7 m ; Y = 390,05 m).



**Figure. II. 4.** Coupe transversale du cours d'eau naturel "Sebaou" (ANRH).

**Tableau. II. 2.** Caractéristiques de formes des sous bassins d'Oued Sebaou (Station Baghlia).

Bassin versant d'Oued Sebaou	
Surface (Km <sup>2</sup> )	2500
Périmètre du bassin versant (km)	200
L.thalweg (km)	92
L. rect. Equivalent (Km)	60,5
Altitude max. (m)	2305
Altitude moy. (m)	590
Altitude min. (m)	20
Indice de Pente Globale $I_{PG}(\%)$	3,78

Indice de pente de Roche $I_{PR}$	0,17
Indice de Compacité $I_c$	1,14
Coefficient de torrencialité $C_t$	19,6
Densité de drainage $D_d$	2,7

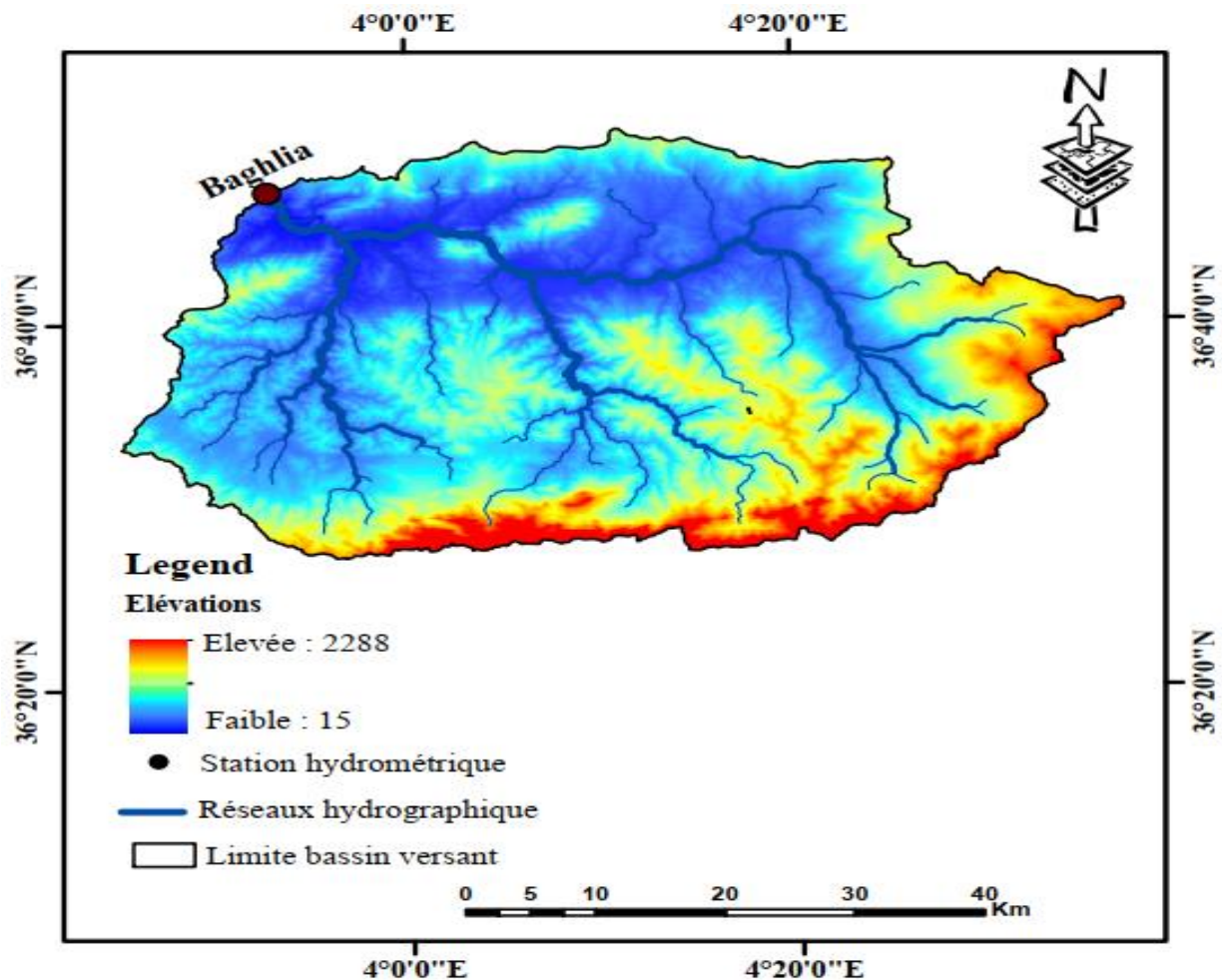


Figure. II. 5. Station Baghlia “Oued Sebaou “.

### II.2.3 Bassin versant d’Oued Chellif

Oued Chellif, est le plus important et le plus long oued d’Algérie, Situé au centre ouest de l’Algérie du Nord dans le bassin Chellif entre 2°25’ et 3°45’ de longitude Est et entre 35°45’ et 36°00’ de l’altitude Nord, il est limité au nord par la mer Méditerranée, à l’ouest par la région Oranie\_Chott\_Hodna, la région est bordée par deux chaines principales, l’Atlas Tellien au nord et l’Atlas Saharien au sud, d’une superficie de 43700 km<sup>2</sup> et d’un périmètre de 1383 km. Oued Chellif c’est un cours d’eau de 725 km de long parmi les grands oueds Nord-Africains, il prend sa source dans l’Atlas tellien plus précisément dans le Djebel Amour et se jette dans la Méditerranée. Le Chellif est caractérisé par une vallée très fertile. Il est le plus important fleuve

d'Algérie. Les mesures de hauteurs d'eau et les débits liquides sont mesurés au niveau de la station hydrométrique (coordonnées Lambert (X = 439,5 m ; Y = 332,58 m).

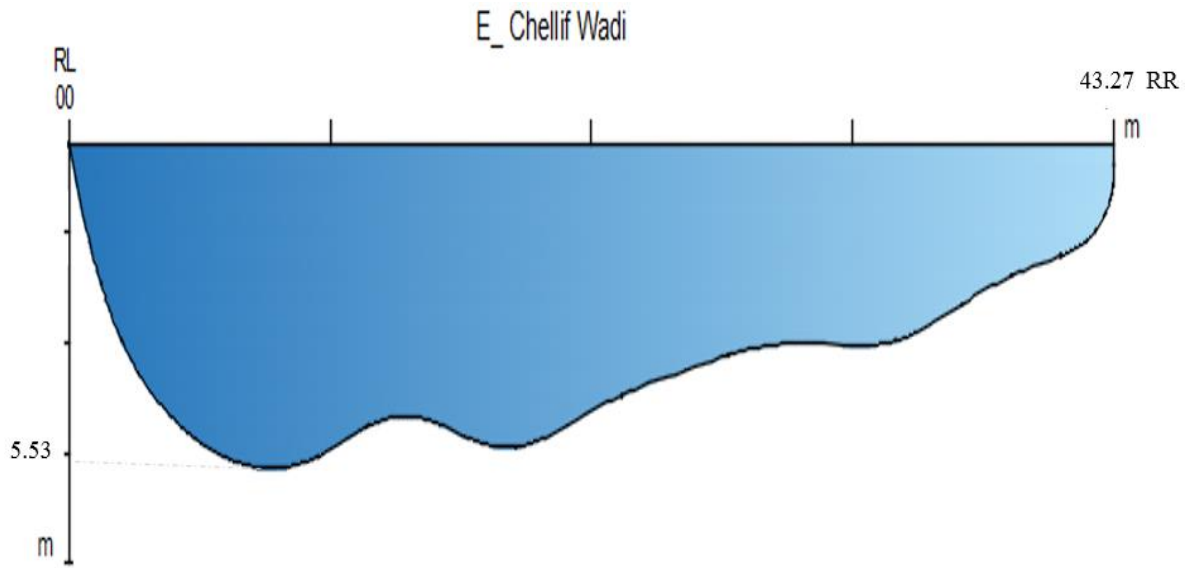


Figure. II. 6. Coupe transversale du cours d'eau naturel "Chellif"(ANRH).

Tableau. II. 3. Caractéristiques de formes de sous bassin d'Oued Chellif (Station Arib Chellif).

Bassin versant d'Oued Chellif	
Surface (Km <sup>2</sup> )	43700
Périmètre du bassin versant (km)	1383
L.thalweg (km)	476
L. rect. Equivalent (Km)	459
Altitude max. (m)	1983
Altitude moy. (m)	1021,5
Altitude min. (m)	230
Indice de Pente Globale I <sub>PG</sub> (%)	1,85
Indice de pente de Roche I <sub>PR</sub>	0,17
Indice de Compacité I <sub>c</sub>	1,85
Coefficient de torrentialité C <sub>t</sub>	19,6
Densité de drainage D <sub>d</sub>	2,7

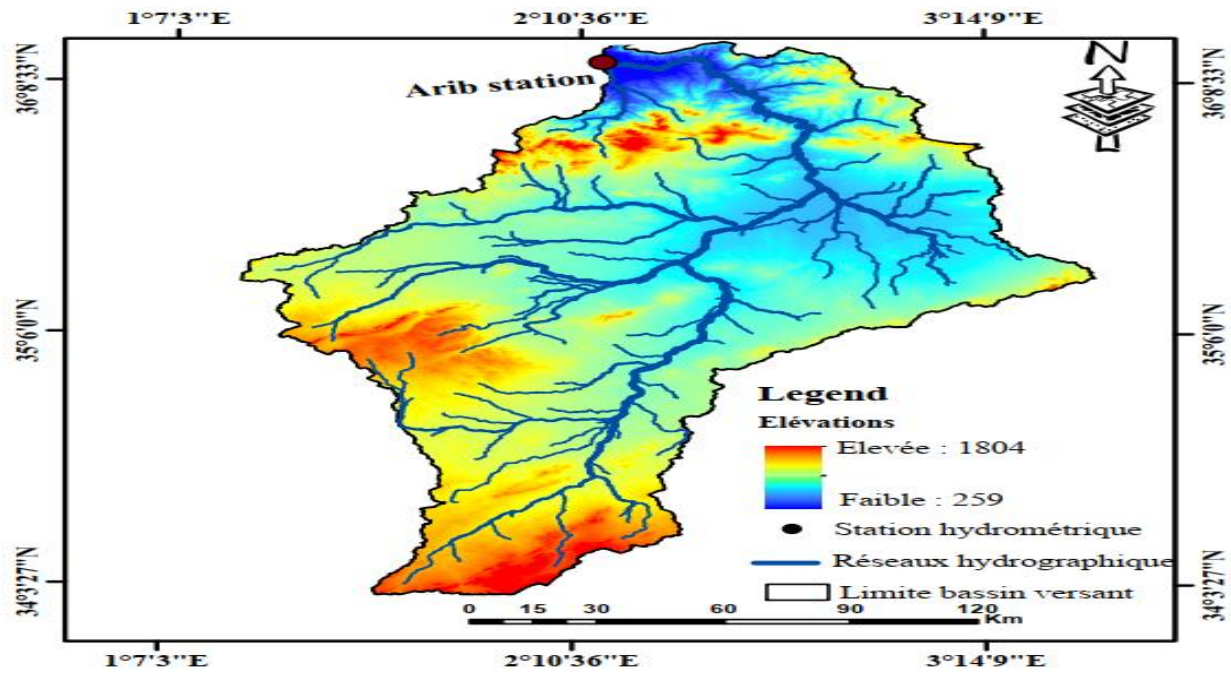


Figure. II. 7. Station Arib “Oued Chellif”.

### II.2.4 Bassin versant d’Oued Mazafran

Oued Mazafran est situé dans le bassin versant Mazafran au Côtier Méditerranéen à environ 30 km à l’ouest d’Alger séparant la wilaya de Tipaza et celle d’Alger de la plaine de Mitidja entre 36° 15’ et 36°45’ de latitude Nord et 2° 15’ et 3° de longitude Est. Le bassin versant de l’oued Mazafran est l’un des plus importants bassins du côté Algérois qui possède une superficie 1900 km<sup>2</sup> et de périmètre 200 km. Il est limité au nord par le Sahel de direction Est-Ouest, au sud par l’Atlas de Blida, à l’Ouest par l’oued el Harrach. Oued Mazafran prend sa source dans l’Atlas tellien, et traversent la plaine de la Mtidja suivant une direction ouest-est pour rejoindre à l’aval, dans la partie méridionale du Sahel, au Sud-Ouest de Koléa près du bois domanial de Soumia. On signale que les mesures de hauteurs d’eau et les débits liquides sont mesurés au niveau de la station Fer à Cheval qui situé à wilaya de Tipaza (Kolea). Coordonnées Lambert (X = 510,05 m ; Y = 373,2 m).

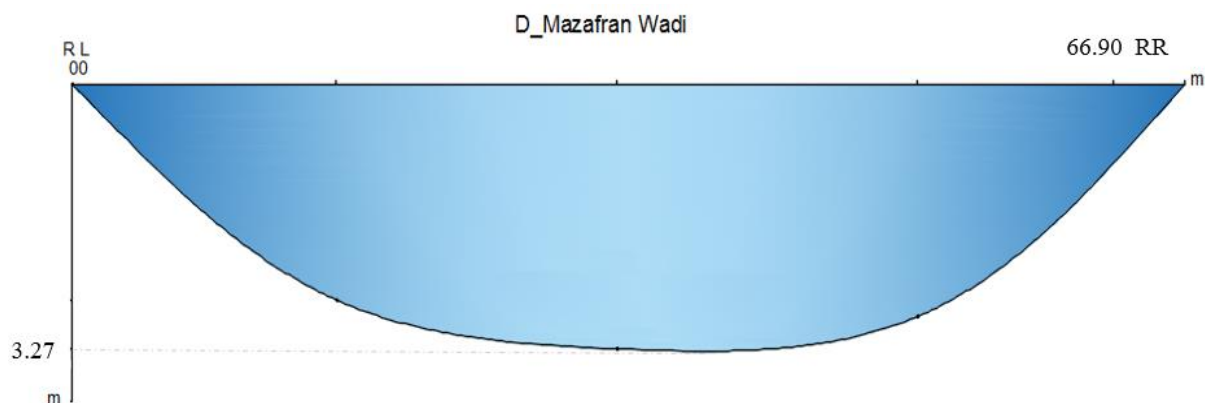


Figure. II. 8. Coupe transversale du cours d'eau naturel "Mazafran"(ANRH).

Tableau. II. 4. Caractéristiques de formes des sous bassins d’Oued Mazafran (Station Fer à Cheval).

Bassin versant d’Oued Mazafran	
Surface (Km <sup>2</sup> )	1900
Périmètre du bassin versant (km)	200
L.thalweg (km)	61,7
L. rect. Equivalent (Km)	96
Altitude max. (m)	1629
Altitude moy. (m)	446
Altitude min. (m)	10
Indice de Pente Globale I <sub>PG</sub> (%)	2,65
Indice de pente de Roche I <sub>PR</sub>	0,26
Indice de Compacité I <sub>c</sub>	1,19
Coefficient de torrentialité C <sub>t</sub>	21
Densité de drainage D <sub>d</sub>	2,8

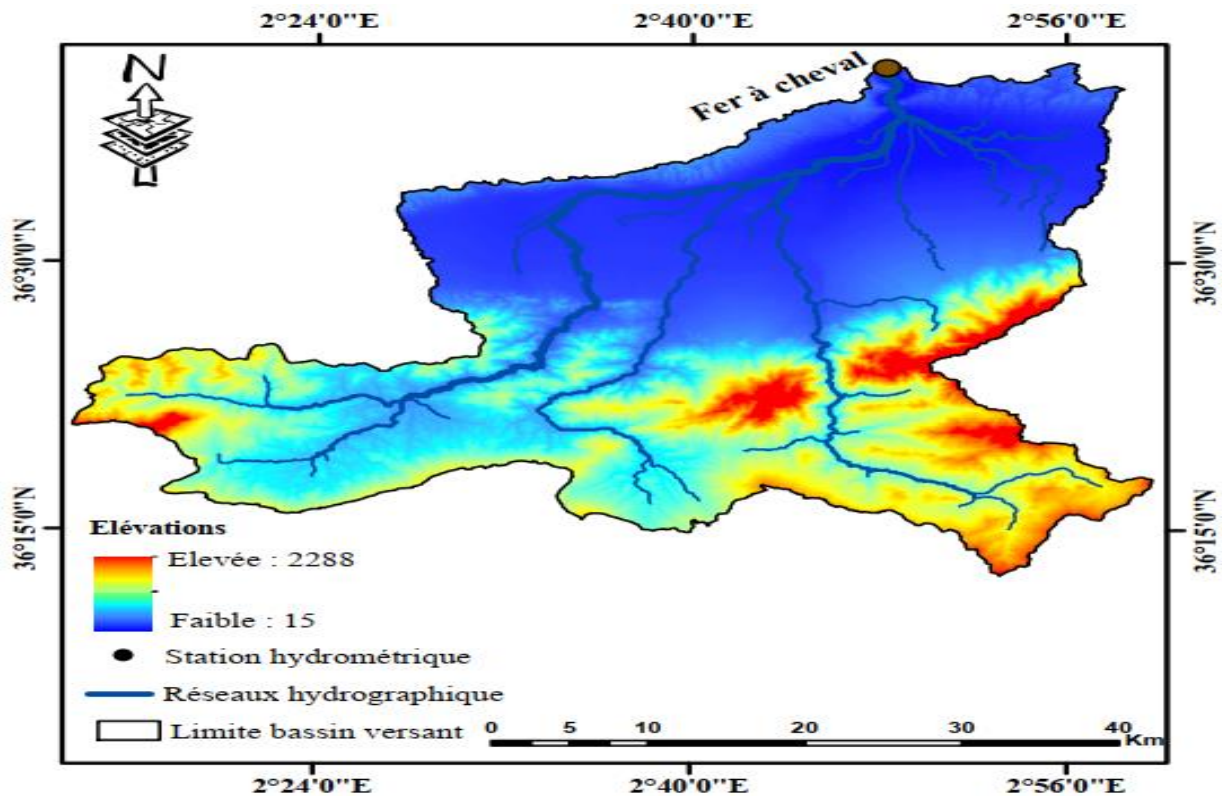
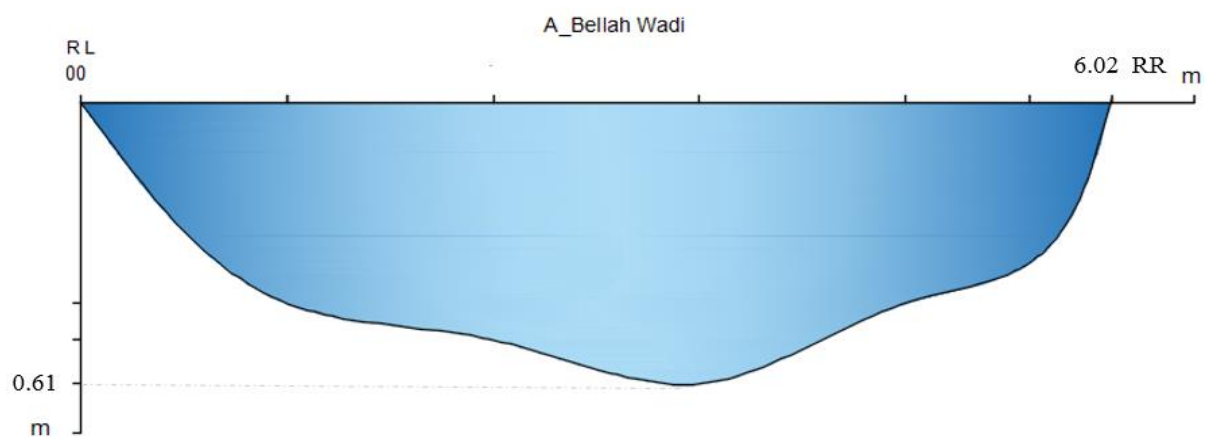


Figure. II. 9. Station Fer à cheval “Oued Mazafran”.

### II.2.5 Bassin versant d’Oued Bellah

Oued Bellah est situé dans le bassin versant du côtier algérois qui présente une superficie de 55 km<sup>2</sup> avec un périmètre de 38 km, dans climat semi-aride, humide et froid en hiver et sec

et chaud en été. Il est situé entre 2°9'33,84" et 2°16'1,2" de longitude Est et entre 36°30'7,56" et 36°36'41,76" de latitude Nord, à l'Ouest de la capitale Alger, qui limité au Nord par la mer méditerranée au Sud et Sud-Ouest par la wilaya de Ain Defla et à l'Ouest par la wilaya de Chlef, à l'Est par la wilaya de Tipaza, le cours d'eau Bellah faisant partie de la wilaya de Tipazza, ce dernier faisant parotide de wilaya de Tipaza et de type montagneux avec des pentes fortes supérieures à 12,5%. Le bassin est contrôlé par deux stations hydrométriques : la station de Sidi Ghiles et la station de Bellah Pont RN 11. On signale que les mesures de hauteurs d'eau et les débits liquides sont mesurés au niveau de la station Pont RN 11 qui est situé à l'exutoire (coordonnées Lambert (X = 458,65 m ; Y = 367,50 m).



**Figure. II. 10.** Coupe transversale du cours d'eau naturel "Bellah"(ANRH).

**Tableau. II. 5.** Caractéristiques de formes de sous bassin d'Oued Bellah (Station Pont RN 11).

<b>Bassin versant d'Oued Bellah</b>	
Surface (Km <sup>2</sup> )	55
Périmètre du bassin versant (km)	38
L.thalweg (km)	10,39
L. rect. Equivalent (Km)	96
Altitude max. (m)	736
Altitude moy. (m)	254
Altitude min. (m)	25
Indice de Pente Globale I <sub>PG</sub> (%)	6,84
Indice de pente de Roche I <sub>PR</sub>	0,22
Indice de Compacité I <sub>c</sub>	0,68
Coefficient de torrentialité C <sub>t</sub>	32,76
Densité de drainage D <sub>d</sub>	3,6

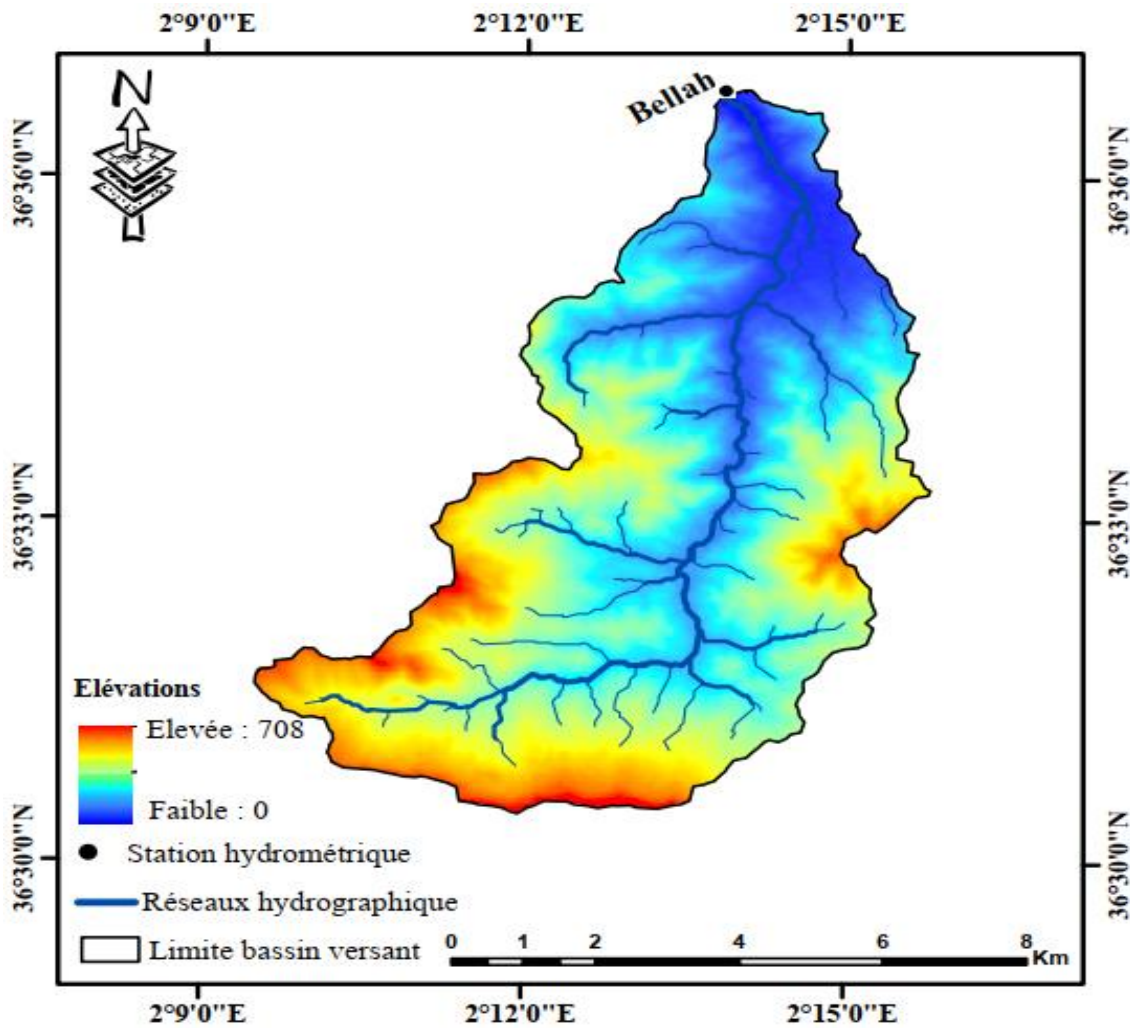


Figure. II. 11. Station Pont RN 11 “Oued Bellah”.

### II.3. Analyse de l'écoulement

Le débit a tendance à varier dans le temps en raison du caractère aléatoire de la variable d'entrée (précipitation). Comprendre l'ampleur de ces changements est important pour caractériser la dynamique fluviale sédimentaire du bassin dans les régions semi-arides.

La distribution du ruissellement annuel dépend des changements climatiques saisonniers. Dans la plupart des applications, une distribution mensuelle du débit est suffisante et peut être déduite des modifications des paramètres météorologiques. Pour certains problèmes, une distribution du débit à l'échelle journalière peut devoir être envisagée.

Les crues se produisent de manière aléatoire, il est donc nécessaire d'une part d'évaluer l'apport de surface dans les crues et d'autre part de quantifier la charge, des débits solide et liquide, causée par le ruissellement de surface.

Nous disposons des données des stations de mesure hydrométrique suivantes (Tableau II.6).

Tableau. II. 6. Identification des Stations hydrométriques des cours d'eau étudiées.

Station	Oued	Code	X (Km)	Y (Km)	Date
Arib Chellif	Chellif	01 17 02	0,1-6,70	0,225	1994-2013
Amont des gorges	Chiffa	021126	0,01-2,08	0,105	1997-2016
Baghlia	Sebaou	022001	0,01-1,77	0,377	1987-2001
Fer à cheval	Mazafran	021201	0,02-3.94	0,777	1985-2008
Pont RN 11	Bellah	02 03 23	0,008- 1,015	0,0022	1991-2015

L'analyse statistique des séries saisonnière des débits moyens mensuels s'avère indispensable pour mieux connaître les fluctuations des débits (Marouf, 2012).

Les caractéristiques de l'écoulement à savoir ; le débit minimum, maximum et moyens annuel, le module spécifique ( $M^\circ$ ), l'hydraulicité ( $K_i$ ), la lame d'eau ruisselée ( $L^\circ$ ), le coefficient d'écoulement, ainsi que les caractéristiques statistiques sont calculés pour les stations hydrométriques étudiées.

### II.3.1 Variation de L'écoulement

A partir des débits journaliers déterminés par l'Agence Nationale des Ressources Hydrique (ANRH), nous avons établis pour les cinq stations hydrométriques les variations des débits liquides ainsi que quelques caractéristiques principales de l'écoulement. (Tableaux II.7, II.8, II.9, II.10, II.11 et II.12).

Les paramètres statistiques les plus courant utilisés (écart-type, coefficient de variation) ont été présentés pour représenter la variabilité du débit.

Tableau. II. 7. Variation interannuelle de l'écoulement au droit de différentes stations.

Stations hydrométriques	Paramètres de variations	Sept	Oct	Nov	Déc	Janv	Fev	Mars	Avril	Mai	Juin	Juil	Août	Année
Arib Chellif	Qmoy (m <sup>3</sup> /s)	1,72	3,29	2,51	4,25	6,72	9,26	7,42	4,29	1,98	1,42	1,72	1,57	3,89
	Écart-type (m <sup>3</sup> /s)	1,914	6,142	2,195	8,717	10,43	13,44	9,45	7,288	2,524	1,709	2,104	2,108	3,98
	CV	1,1101	1,869	0,874	2,053	1,551	1,452	1,273	1,697	1,272	1,202	1,223	1,338	1,0250
Amont de groges	Qmoy (m <sup>3</sup> /s)	0,91	1,18	1,72	3,68	4,43	6,13	5,71	3,35	1,32	0,39	0,17	0,82	2,11
	Écart-type (m <sup>3</sup> /s)	4,88	4,73	5,16	6,32	6,95	9,95	9,10	5,97	1,42	0,35	0,17	4,85	2,14
	CV	5,35	4,02	3,00	1,72	1,57	1,62	1,60	1,78	1,07	0,91	0,99	5,88	1,01
Baghlia	Qmoy (m <sup>3</sup> /s)	0,08	0,51	4,75	32,47	41,35	44,86	48,03	28,29	14,34	11,25	0,94	1,80	19,01
	Ecart type (m <sup>3</sup> /s)	0,10	0,88	7,68	36,92	35,98	39,26	35,74	23,28	12,02	21,57	2,75	5,87	13,09
	CV	1,35	1,72	1,62	1,14	0,87	0,88	0,74	0,82	0,84	1,92	2,92	3,25	0,69
Fer à Cheval	Qmoy (m <sup>3</sup> /s)	0,27	1,55	2,05	4,82	6,73	8,90	9,59	4,16	2,38	0,54	0,19	0,13	3,62
	Écart-type (m <sup>3</sup> /s)	0,36	3,68	3,45	6,55	6,81	12,83	10,20	7,04	4,41	0,87	0,14	0,15	2,89
	CV	1,31	2,38	1,68	1,36	1,01	1,44	1,06	1,69	1,85	1,59	0,72	1,19	0,80
Pont RN 11	Qmoy (m <sup>3</sup> /s)	0,09	0,13	0,26	0,34	0,27	0,50	0,44	0,24	0,13	0,08	0,07	0,04	0,14
	Écart-type (m <sup>3</sup> /s)	1,82	1,63	1,32	1,25	0,82	1,22	1,13	1,22	1,06	1,61	2,25	2,37	0,75
	CV	0,09	0,13	0,26	0,34	0,27	0,50	0,44	0,24	0,13	0,08	0,07	0,04	0,14

Tableau. II. 8. Caractéristique de l'écoulement d'Oued Chellif (Station Arib Chellif).

Année	Q max (m <sup>3</sup> /s)	Q min (m <sup>3</sup> /s)	Q moy (m <sup>3</sup> /s)	Apport (Mm <sup>3</sup> )	M <sub>0</sub> (l/s.Km <sup>2</sup> )	Ki	Lr (mm)	Ce
1979 - 1980	149,80	1,18	5,48	171,40	3,96	1,83	123,87	0,28
1980 - 1981	59,25	0,65	4,10	128,40	2,96	1,37	92,80	0,21
1982 - 1983	125,50	0,98	4,30	136,20	3,11	1,44	98,43	0,22
1983 - 1984	65,60	0,09	1,87	58,79	1,35	0,63	42,49	0,10
1984 - 1985	131,00	0,11	3,11	98,72	2,25	1,04	71,35	0,16
1985 - 1986	132,00	0,21	4,54	142,10	3,28	1,52	102,70	0,23
1986 - 1987	268,80	2,11	8,60	262,30	6,21	2,88	189,57	0,43
1987 - 1988	18,86	0,64	1,72	54,55	1,24	0,58	39,42	0,09
1988 - 1989	130,00	0,01	1,51	47,88	1,09	0,51	34,60	0,08
1989 - 1990	18,90	0,23	0,86	27,03	0,62	0,29	19,53	0,04
1990 - 1991	56,88	0,64	1,58	49,64	1,14	0,53	35,88	0,08
1991 - 1992	304,00	0,37	7,37	232,80	5,33	2,47	168,25	0,38
1992 - 1993	10,30	0,08	0,60	18,80	0,43	0,20	13,59	0,03
1993 - 1994	29,22	0,00	0,23	7,07	0,17	0,08	5,11	0,01
1994 - 1995	9,40	0,17	5,33	18,00	3,85	1,78	13,01	0,03
1995 - 1996	108,90	0,32	2,60	80,15	1,88	0,87	57,93	0,13
1996 - 1997	32,60	0,29	0,90	28,57	0,65	0,30	20,65	0,05
2000 - 2001	95,80	0,00	2,02	63,70	1,46	0,68	46,04	0,11
2001 - 2002	46,30	0,12	0,66	20,75	0,48	0,22	14,99	0,03
2002 - 2003	110,55	0,00	2,38	73,91	1,72	0,80	53,41	0,12
Moyenne	134,65	0,50	5,23	164,24	0,97	1,00	62,18	0,14
Écart-type	173,66	0,41	3,69	117,32	0,68	0,78	52,20	0,12
Coeff d'assy	3,85	2,87	3,11	3,18	3,11	0,34	0,86	0,00
Max	942,83	2,24	21,15	674,31	3,92	2,88	189,57	0,43
Min	17,23	0,12	1,76	55,52	0,33	0,08	5,11	0,01
Coeff d'appl	17,53	11,65	12,28	12,75	12,28	0,31	0,80	0,00
Médiane	74,47	0,49	4,15	130,65	0,77	0,74	44,26	0,10
Variance	30156,86	0,17	13,63	13763,46	0,47	1,83	3770,91	8,60

Tableau. II. 9. Caractéristiques de l'écoulement d'Oued Mazafran (Station Fer à Cheval).

Année	Q max (m <sup>3</sup> /s)	Q min (m <sup>3</sup> /s)	Q moy (m <sup>3</sup> /s)	Apport (Mm <sup>3</sup> )	M <sub>0</sub> (l/s.Km <sup>2</sup> )	Ki	Lr (mm)	Ce
1981 - 1982	125,50	0,52	2,01	61,73	1,06	0,67	32,49	0,06
1982 - 1983	65,60	1,09	4,75	150,04	2,50	1,58	78,97	0,15
1983 - 1984	131,00	0,28	1,92	5,55	1,01	0,64	2,92	0,01
1984 - 1985	132,00	0,78	0,64	19,06	0,34	0,21	10,03	0,02
1985 - 1986	268,80	0,00	3,60	112,32	1,89	1,20	59,11	0,11
1986 - 1987	18,86	0,01	7,72	234,84	4,06	2,57	123,60	0,24
1987 - 1988	130,00	0,00	1,14	35,79	0,60	0,38	18,84	0,04
1988 - 1989	18,90	0,00	1,85	58,67	0,97	0,62	30,88	0,06
1989 - 1990	56,88	0,00	0,33	10,46	0,17	0,11	5,51	0,01
1990 - 1991	304,00	0,00	2,15	4,92	1,13	0,72	2,59	0,01

1991 - 1992	10,30	0,00	5,08	161,52	2,67	1,69	85,01	0,16
1992 - 1993	29,22	0,01	1,08	34,15	0,57	0,36	17,97	0,03
1993 - 1994	338,00	0,01	1,26	39,41	0,66	0,42	20,74	0,04
1994 - 1995	262,00	0,01	3,47	111,04	1,83	1,16	58,44	0,11
1995 - 1996	18,00	0,28	1,29	3,43	0,68	0,43	1,80	0,00
1996 - 1997	0,72	0,12	0,30	0,80	0,16	0,10	0,42	0,00
1998 - 1999	108,90	0,32	2,60	80,15	1,37	0,87	57,93	0,11
1999 - 2000	32,60	0,29	0,90	28,57	0,47	0,30	20,65	0,04
2000 - 2001	95,80	0,00	2,02	63,70	1,06	0,67	46,04	0,09
2001 - 2002	46,30	0,12	0,66	20,75	0,35	0,22	14,99	0,03
2002 - 2003	110,55	0,00	2,38	73,91	1,25	0,79	53,41	0,10
2003 - 2004	32,08	0,14	1,75	55,73	0,92	0,58	29,33	0,06
2004 - 2005	43,20	0,13	3,05	7,22	1,61	1,02	3,80	0,01
2005 - 2006	29,72	0,09	1,44	3,83	0,76	0,48	2,02	0,00
2006 - 2007	412,05	0,10	3,75	119,00	1,97	1,25	62,63	0,12
2007 - 2008	398,49	0,12	3,56	112,34	1,88	1,19	59,13	0,11
2008 - 2009	129,44	0,06	3,83	121,24	2,02	1,28	63,81	0,12
2009 - 2010	187,52	0,06	3,73	116,51	1,96	1,24	61,32	0,12
2010 - 2011	361,76	0,25	9,80	304,01	5,16	3,27	160,00	0,31
2011 - 2012	374,00	0,08	11,93	371,96	6,28	3,98	195,77	0,38
Moyenne	142,41	0,16	3,00	84,09	1,58	1,00	46,01	0,09
Écart-type	130,06	0,30	1,83	61,59	0,96	0,61	32,41	0,06
Coeff d'assy	1,08	2,02	1,60	1,38	1,60	0,53	1,12	0,00
Max	412,05	1,09	11,93	371,96	6,28	3,98	195,77	0,38
Min	0,72	0,00	0,30	0,80	0,16	0,10	0,42	0,00
Coeff d'appl	0,84	2,37	1,86	1,65	1,86	0,62	1,57	0,00
Médiane	109,73	0,08	2,09	60,20	1,10	0,70	31,68	0,06
Variance	17497,56	0,06	7,28	8088,18	2,02	2,43	2222,65	4,31

Tableau. II. 10. Caractéristiques de l'écoulement d'Oued Chiffa (Station Amont Des Gorges).

Année	Q max (m <sup>3</sup> /s)	Q min (m <sup>3</sup> /s)	Q moy (m <sup>3</sup> /s)	Apport (Mm <sup>3</sup> )	M <sub>0</sub> (l/s.Km)	Ki	Lr (mm)	Ce
1969 - 1970	130,00	0,00	4,13	131,35	0,00	0,87	415,68	0,67
1970 - 1971	140,00	0,00	1,96	61,57	0,00	0,41	194,85	0,31
1971 - 1972	366,00	0,03	5,93	186,77	0,08	1,26	591,03	0,95
1972 - 1973	351,50	0,11	7,61	235,21	0,36	1,61	744,35	1,20
1973 - 1974	731,50	0,18	7,71	243,57	0,57	1,63	770,79	1,24
1974 - 1975	138,00	0,12	2,02	63,34	0,39	0,43	200,44	0,32
1975 - 1976	313,00	0,08	4,02	124,31	0,25	0,85	393,39	0,63
1976 - 1977	138,80	0,00	1,63	51,30	0,00	0,35	162,35	0,26
1977 - 1978	98,75	0,05	1,09	34,29	0,16	0,23	108,51	0,17
1978 - 1979	172,12	0,06	0,07	3,86	0,19	0,01	12,22	0,02
1980 - 1981	60,92	0,09	3,25	7,86	0,28	0,69	24,87	0,04
1981 - 1982	47,96	0,02	47,96	30,53	0,06	10,16	96,61	0,16
1982 - 1983	60,92	0,02	9,08	24,33	0,06	1,92	76,99	0,12
1983 - 1984	3,77	0,04	33,30	86,31	0,13	7,06	273,13	0,44
1984 - 1985	19,45	0,01	0,10	53,36	0,03	0,02	168,86	0,27

1985 - 1986	377,00	0,01	1,78	145,50	0,03	0,38	460,44	0,74
1986 - 1987	199,27	0,05	4,86	55,54	0,16	1,03	175,76	0,28
1987 - 1988	2,55	0,01	0,17	147,23	0,03	0,04	465,93	0,75
1988 - 1989	377,00	0,01	1,34	5,42	0,02	0,28	17,14	0,03
1989 - 1990	142,88	0,01	0,26	42,83	0,02	0,06	135,55	0,22
1990 - 1991	116,60	0,02	1,31	8,22	0,06	0,28	26,03	0,04
1991 - 1992	401,90	0,04	12,10	40,55	0,13	2,56	128,31	0,21
1992 - 1993	52,56	0,02	0,85	31,36	0,06	0,18	99,24	0,16
1993 - 1994	270,00	0,01	0,82	55,54	0,04	0,17	175,76	0,28
1994 - 1995	228,00	0,00	2,70	147,23	0,00	0,57	465,93	0,75
1995 - 1996	127,28	0,01	2,17	5,42	0,03	0,46	17,14	0,03
1996 - 1997	24,68	0,00	0,07	42,83	0,01	0,01	135,55	0,22
2003 - 2004	31,00	0,15	1,04	8,22	0,47	0,22	26,03	0,04
2004 - 2005	22,07	0,13	1,65	40,55	0,41	0,35	128,31	0,21
2005 - 2006	17,47	0,07	5,82	26,87	0,22	1,23	85,02	0,14
2006 - 2007	140,30	0,05	0,91	26,18	0,16	0,19	82,86	0,13
2007 - 2008	51,02	0,05	0,79	86,14	0,16	0,17	272,60	0,44
2008 - 2009	114,04	0,03	0,60	68,05	0,09	0,13	215,34	0,35
2009 - 2010	/	0,37	3,05	2,06	1,17	0,65	6,52	0,01
2010 - 2011	101,00	0,21	1,58	18,60	0,67	0,33	58,86	0,09
2011 - 2012	42,90	0,07	0,48	24,90	0,21	0,10	78,80	0,13
2013 - 2014	159,95	0,01	0,50	18,79	0,03	0,11	59,46	0,10
Moyenne	160,00	0,06	4,72	64,49	0,18	1,00	204,07	0,33
Écart-type	156,04	0,08	9,30	63,57	0,24	1,97	201,18	0,32
Coeff d'assy	1,75	2,46	3,77	1,46	7,79	0,80	4,61	0,01
Max	731,50	0,37	47,96	243,57	1,17	10,16	770,79	1,24
Min	2,55	0,00	0,07	2,06	0,00	0,01	6,52	0,01
Coeff d'appl	1,68	2,36	3,62	1,40	7,47	0,77	4,42	0,01
Médiane	127,28	0,03	1,65	42,83	0,09	0,35	135,55	0,22
Variance	24341,65	0,01	86,57	4041,47	0,02	18,34	12789,47	20,56

Tableau. II. 11. Caractéristiques de l'écoulement d'Oued Sebaou (Station Baghlia).

Année	Q max (m <sup>3</sup> /s)	Q min (m <sup>3</sup> /s)	Q moy (m <sup>3</sup> /s)	Apport (Mm <sup>3</sup> )	M <sub>0</sub> (l/s.Km <sup>2</sup> )	Ki	Lr (mm)	Ce
1976 - 1977	710,00	0,09	13,73	433,05	7,08	0,77	290,64	0,40
1977 - 1978	640,00	0,07	13,45	424,05	6,93	0,75	284,60	0,39
1978 - 1979	710,00	0,02	14,98	473,69	7,72	0,84	317,91	0,43
1979 - 1980	1610,00	0,02	25,39	800,73	13,09	1,42	537,40	0,73
1980 - 1981	1483,20	0,05	20,20	636,98	10,41	1,13	427,50	0,58
1981 - 1982	1020,00	0,00	17,58	554,47	9,06	0,98	372,13	0,51
1982 - 1983	725,00	0,00	27,57	871,98	14,21	1,54	585,22	0,80
1983 - 1984	921,20	0,00	29,05	916,00	14,97	1,62	614,77	0,84
1984 - 1985	2500,00	0,00	17,44	546,8	8,99	0,97	366,98	0,50
1985 - 1986	1254,00	0,00	34,22	1048	17,64	1,91	703,36	0,96
1986 - 1987	1282,00	0,65	36,75	1129,36	18,94	2,05	757,96	1,03
1987 - 1988	81,14	0,20	4,18	132,40	2,16	0,23	88,86	0,12
1988 - 1989	230,84	0,14	8,41	266,08	4,34	0,47	178,58	0,24

<b>1989 - 1990</b>	285,30	0,77	2,44	76,91	1,26	0,14	51,62	0,07
<b>1990 - 1991</b>	564,00	4,50	26,75	838,38	13,79	1,49	562,67	0,77
<b>1991 - 1992</b>	585,00	0,19	17,04	537,82	8,79	0,95	360,95	0,49
<b>1992 - 1993</b>	406,52	0,11	11,77	374,02	6,07	0,66	251,02	0,34
<b>1993 - 1994</b>	488,00	0,03	4,87	150,01	2,51	0,27	100,68	0,14
<b>1994 - 1995</b>	500,00	0,02	17,24	550,09	8,89	0,96	369,19	0,50
<b>1995 - 1996</b>	1060,00	1,10	17,17	537,71	8,85	0,96	360,88	0,49
<b>1996 - 1997</b>	86,44	0,04	2,34	73,63	1,21	0,13	49,42	0,07
<b>1997 - 1998</b>	1275,00	0,10	27,49	870,39	14,17	1,53	584,15	0,80
<b>1998 - 1999</b>	362,72	0,00	21,91	674,66	11,29	1,22	452,79	0,62
<b>Moyenne</b>	816,54	0,35	17,91	539,16	9,23	1,00	376,92	0,51
<b>Écart-type</b>	568,04	0,95	9,77	297,91	5,04	0,55	203,48	0,28
<b>Coeff d'assy</b>	1,21	4,17	0,10	0,06	0,05	0,01	0,03	0,00
<b>Max</b>	2500,00	4,50	36,75	1129,36	18,94	2,05	757,96	1,03
<b>Min</b>	81,14	0,00	2,34	73,63	1,21	0,13	49,42	0,07
<b>Coeff d'appl</b>	1,13	3,90	0,09	0,05	0,05	0,01	0,02	0,00
<b>Médiane</b>	710,00	0,05	17,24	537,82	8,89	0,96	366,98	0,50
<b>Variance</b>	322674,88	0,90	95,53	104809,76	49,24	5,33	61694,84	84,1

Tableau. II. 12. Caractéristiques de l'écoulement d'Oued Bellah (Station Pont RN 11).

Année	Q max (m <sup>3</sup> /s)	Q min (m <sup>3</sup> /s)	Q moy (m <sup>3</sup> /s)	Apport (Mm <sup>3</sup> )	M <sub>0</sub> (l/s.Km <sup>2</sup> )	Ki	Lr (mm)	Ce
<b>1974 - 1975</b>	36,00	0,04	0,22	6,93	4,00	1,10	125,95	0,24
<b>1975 - 1976</b>	85,08	0,09	0,51	15,83	9,28	2,55	287,88	0,55
<b>1976 - 1977</b>	25,64	0,04	0,16	5,11	2,94	0,81	92,89	0,18
<b>1977 - 1978</b>	38,22	0,04	0,20	6,27	3,64	1,00	113,96	0,22
<b>1978 - 1979</b>	111,80	0,04	0,42	13,19	7,67	2,11	239,81	0,46
<b>1979 - 1980</b>	50,75	0,05	0,27	8,73	5,00	1,37	158,78	0,31
<b>1980 - 1981</b>	54,84	0,19	0,32	9,86	5,73	1,58	179,22	0,35
<b>1981 - 1982</b>	0,45	0,06	0,10	0,27	1,87	0,52	4,89	0,01
<b>1982 - 1983</b>	39,33	0,05	0,22	6,91	4,03	1,11	125,62	0,24
<b>1983 - 1984</b>	15,74	0,00	0,12	3,91	2,26	0,62	71,01	0,14
<b>1984 - 1985</b>	26,10	0,00	0,29	9,16	5,24	1,44	166,60	0,32
<b>1985 - 1986</b>	20,28	0,00	0,18	5,69	3,31	0,91	103,42	0,20
<b>1986 - 1987</b>	37,44	0,00	0,59	18,37	10,71	2,95	334,02	0,64
<b>1987 - 1988</b>	29,28	0,00	0,31	9,61	5,56	1,53	174,82	0,34
<b>1988 - 1989</b>	19,93	0,00	0,34	10,76	6,19	1,70	195,64	0,38
<b>1989 - 1990</b>	18,91	0,00	0,21	6,62	3,82	1,05	120,43	0,23
<b>1990 - 1991</b>	17,62	0,00	0,29	8,89	5,19	1,43	161,61	0,31
<b>1991 - 1992</b>	61,70	0,00	0,16	5,15	2,95	0,81	93,72	0,18
<b>1992 - 1993</b>	12,58	0,00	0,09	2,80	1,61	0,44	51,00	0,10
<b>1993 - 1994</b>	10,26	0,00	0,08	2,64	1,54	0,42	48,05	0,09
<b>1994 - 1995</b>	16,44	0,00	0,11	3,64	2,08	0,57	66,17	0,13
<b>1995 - 1996</b>	54,60	0,01	0,16	5,07	2,95	0,81	92,26	0,18
<b>1996 - 1997</b>	30,68	0,00	0,03	1,05	0,62	0,17	19,16	0,04

<b>1997 - 1998</b>	10,41	0,00	0,03	0,86	0,50	0,14	15,73	0,03
<b>1998 - 1999</b>	41,40	0,00	0,20	6,06	3,57	0,98	110,23	0,21
<b>1999 - 2000</b>	76,82	0,00	0,17	5,45	3,10	0,85	99,12	0,19
<b>2000 - 2001</b>	7,20	0,00	0,02	0,69	0,40	0,11	12,46	0,02
<b>2001 - 2002</b>	187,44	0,00	0,07	2,33	1,36	0,37	42,43	0,08
<b>2002 - 2003</b>	23,39	0,00	0,12	3,81	2,24	0,62	69,27	0,13
<b>2003 - 2004</b>	13,40	0,01	0,06	1,96	1,12	0,31	35,55	0,07
<b>2004 - 2005</b>	21,04	0,04	0,16	4,93	2,89	0,80	89,58	0,17
<b>2005 - 2006</b>	48,70	0,02	0,18	5,67	3,28	0,90	103,06	0,20
<b>2006 - 2007</b>	9,75	0,04	0,12	3,93	2,27	0,62	71,39	0,14
<b>Moyenne</b>	39,15	0,02	0,20	6,13	3,60	0,99	111,39	0,21
<b>Écart-type</b>	36,24	0,04	0,13	4,20	2,39	0,66	76,45	0,15
<b>Coeff d'assy</b>	2,60	2,94	1,23	1,10	22,29	6,13	20,06	0,04
<b>Max</b>	187,44	0,19	0,59	18,37	10,71	2,95	334,02	0,64
<b>Min</b>	7,20	0,00	0,02	0,27	0,40	0,11	4,89	0,01
<b>Coeff d'appl</b>	2,48	2,81	1,17	1,05	21,27	5,85	19,14	0,04
<b>Médiane</b>	27,69	0,00	0,17	5,45	3,10	0,85	99,12	0,19
<b>Variance</b>	1313,23	0,00	0,02	17,68	0,31	0,09	321,46	0,62

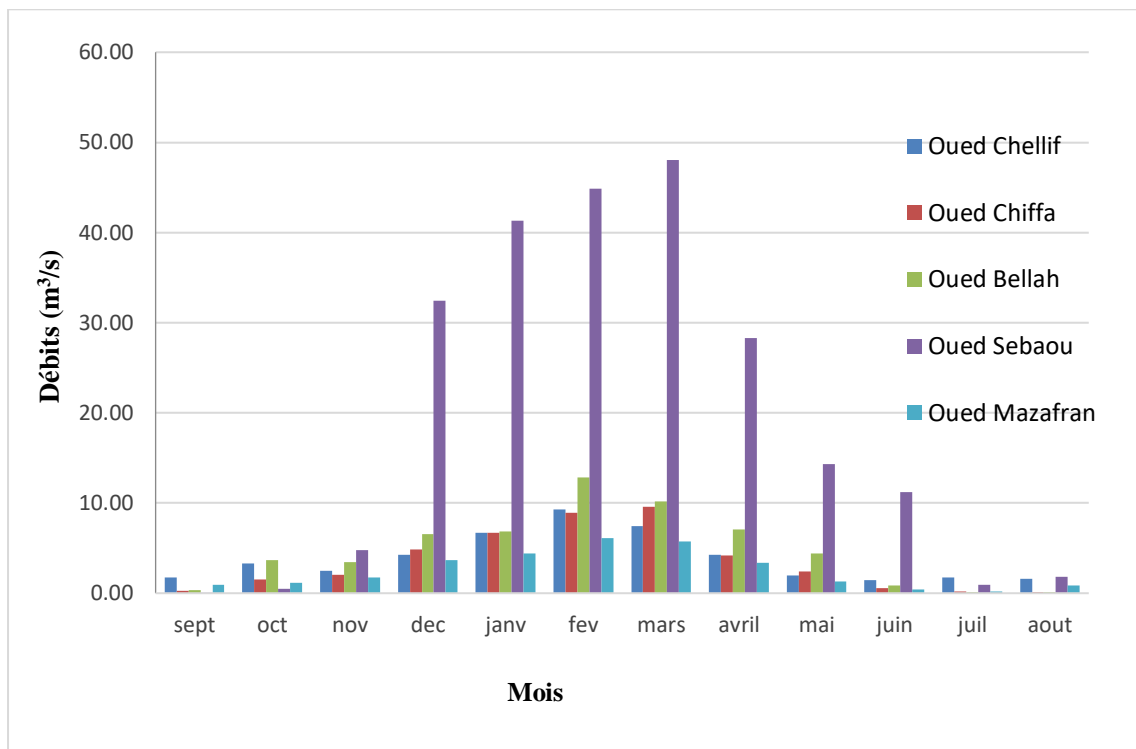


Figure. II. 12. Histogramme de débits moyens mensuels.

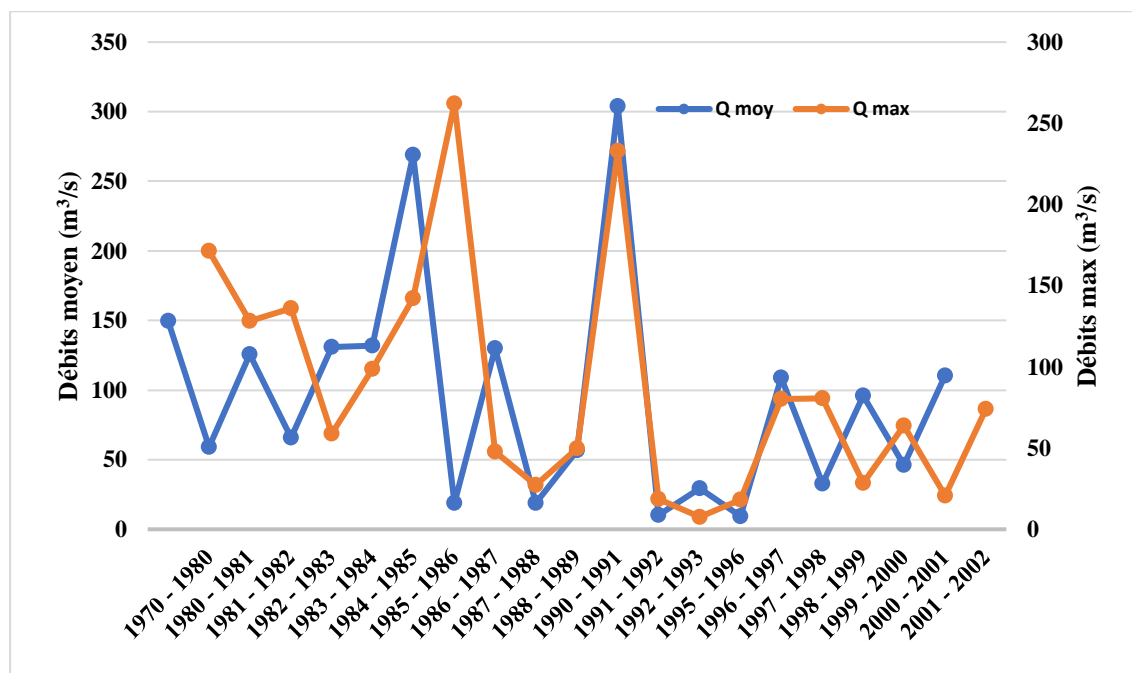
**II.3.2 Débit moyen et max annuel**

Les variations annuelles des débits dans les cours d'eau étudiés ont été présentées, les stations hydrométriques représentent les séries statistiques appropriées allant de (20 à 38 ans).

Les débits des cours d'eau étudiés et plus irréguliers, sont très variables d'une année à l'autre (figures II.13, II.14, II.15, II.16 et II.17).

L'écoulement exceptionnel de l'année 1981/1982 pour la station Oued Chiffa (station Amont Des Gorges) présente un coefficient d'hydraulicité de 10 et qui peut être attribué aux fortes crues de cette année.

Les paramètres statistiques les plus courants utilisés pour représenter la variabilité du débit sont présentés ci-dessus. Les valeurs des coefficients hydraulicité analysées nous ont permis de distinguer les saisons sèches et humides dans différentes stations étudiées (voir tableaux II.7, II.8, II.9, II.10 et II.12).



**Figure. II. 13.** Variation annuelle des débits moyen et max (m³/s) d'écoulement d'Oued Chellif (Station Arib Chellif).

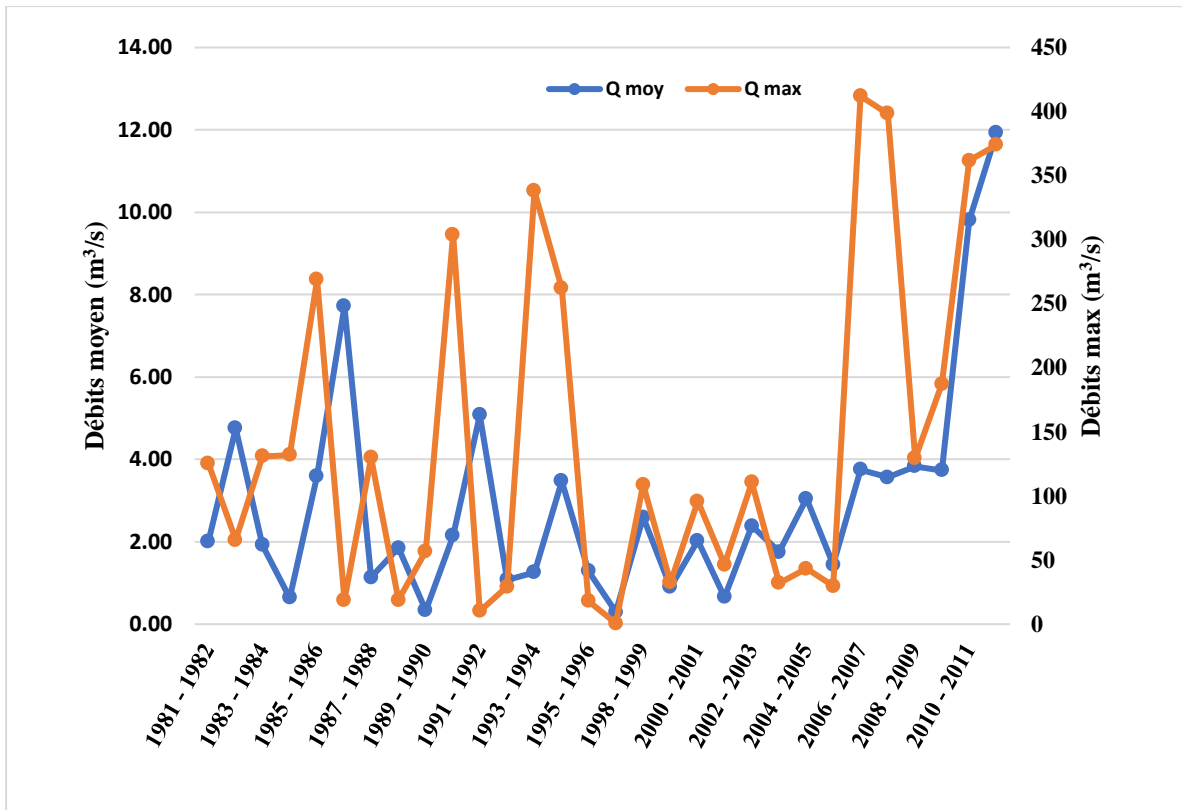


Figure. II. 14. Variation annuelle des débits moyen et max (m<sup>3</sup>/s) d'écoulement d'Oued Mazafran (Station Fer à Cheval).

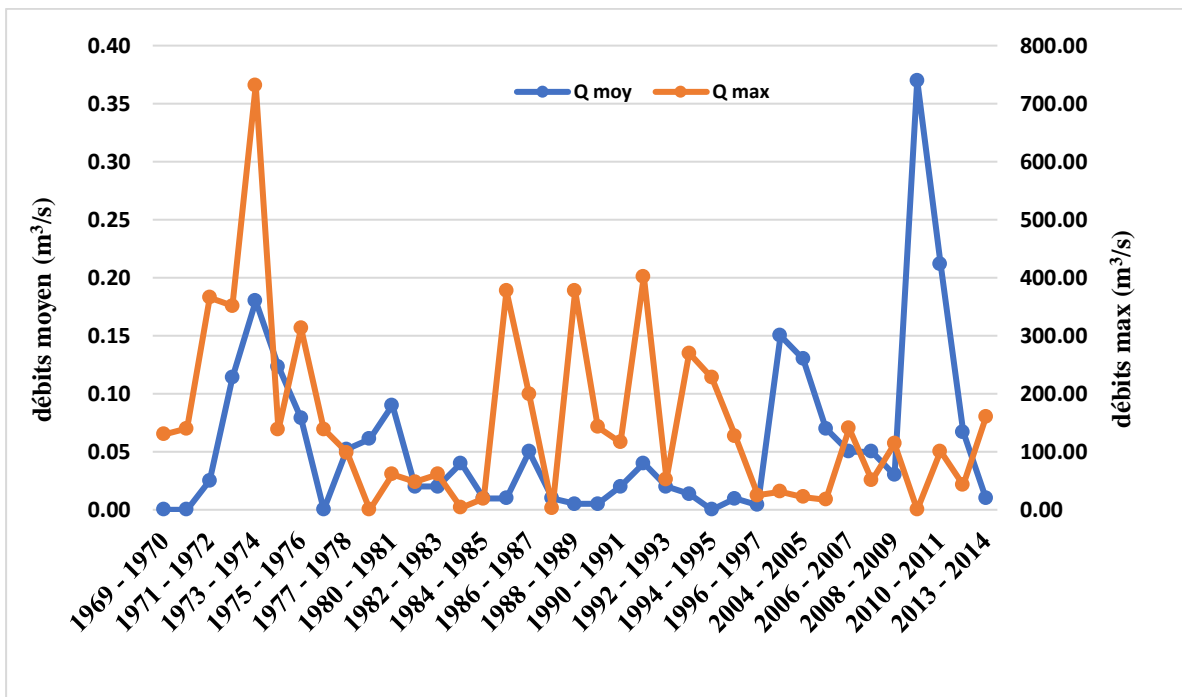


Figure. II. 15. Variation annuelle des débits moyen et max (m<sup>3</sup>/s) d'écoulement d'Oued Chiffa (Station Amont Des Gorges).

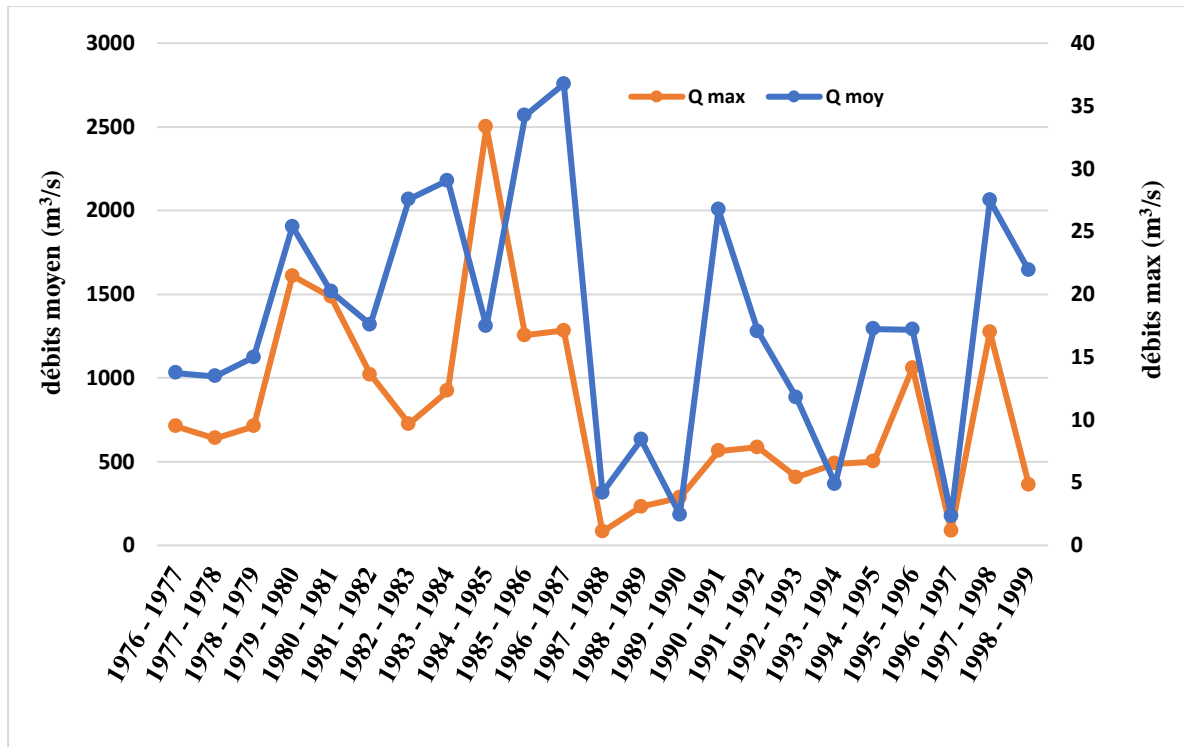


Figure. II. 16. Variation annuelle des débits moyen et max (m³/s) d'écoulement d'Oued Sebaou (Station Baghlia).

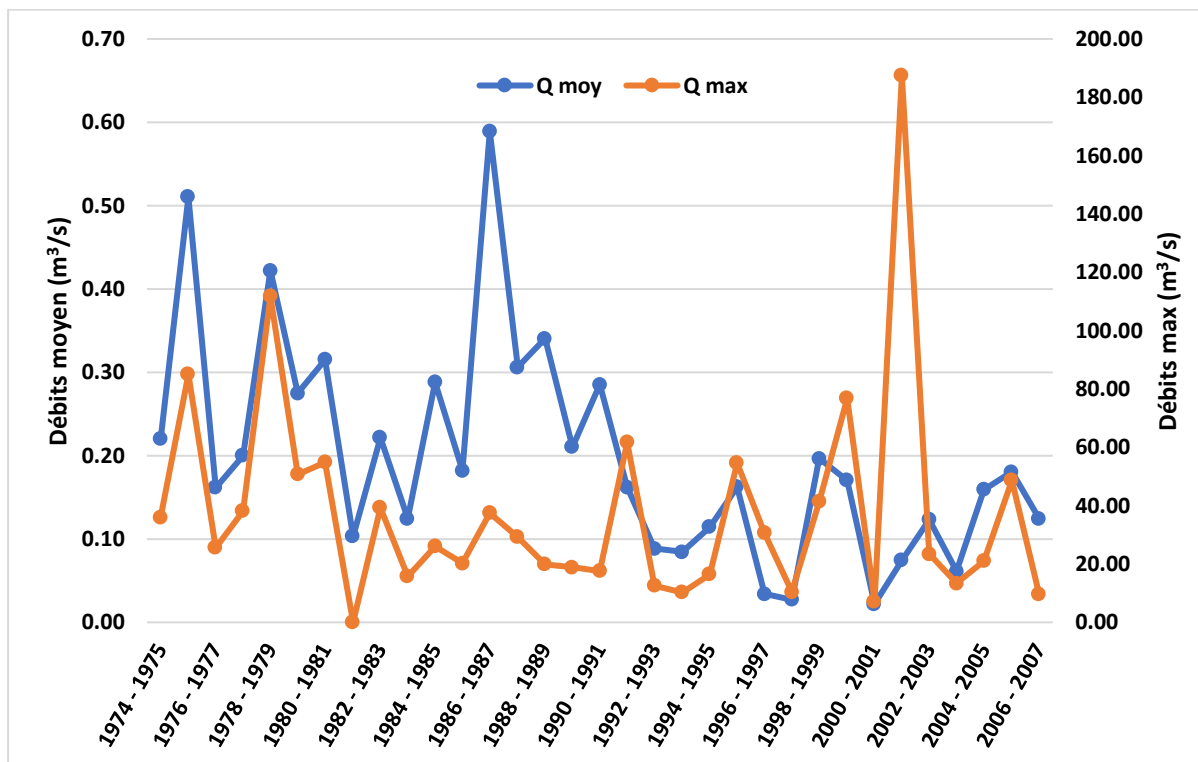
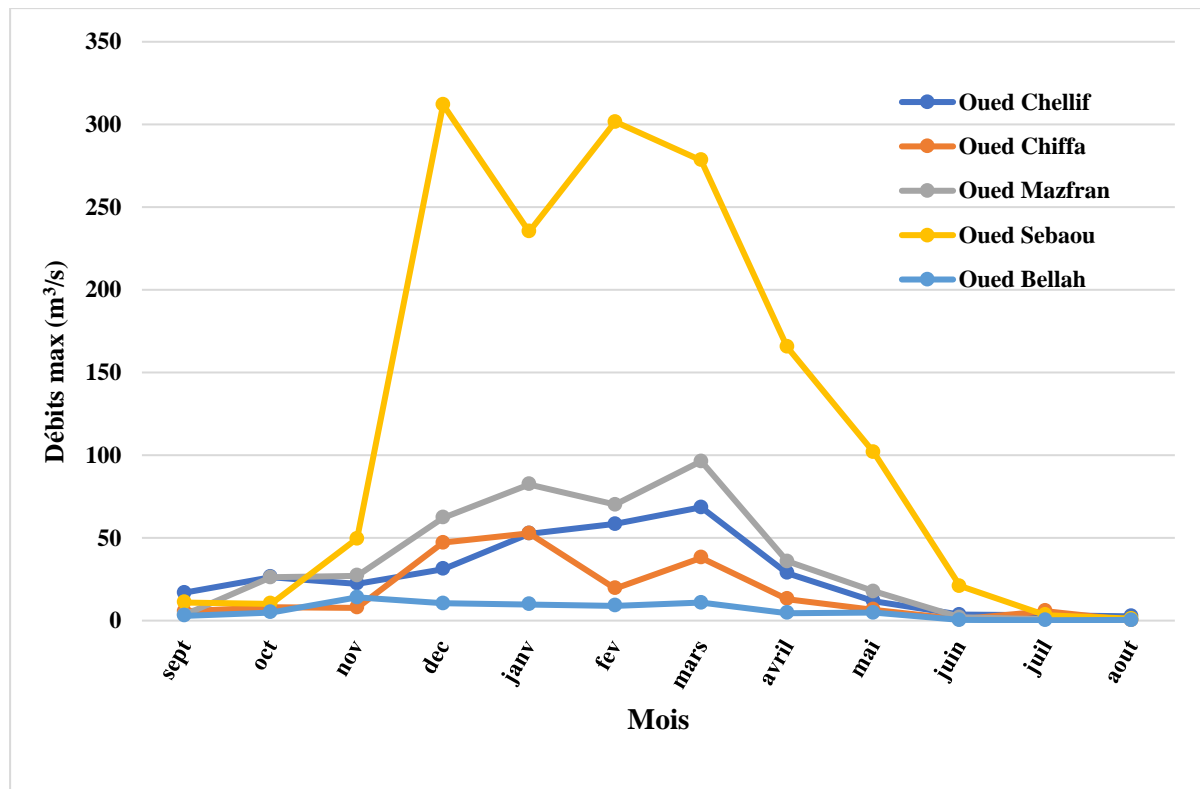


Figure. II. 17. Variation annuelle des débits moyen et max (m³/s) d'écoulement d'Oued Bellah (Station Pont RN 11).



**Figure. II. 18.** Variation saisonnière des débits maxima mensuels.

L'étude des écoulements dans les cours d'eau étudiés montre que les débits importants sont observés à la Station Oued Sebaou et qui sont enregistrés généralement entre les mois de Décembre et Mai. Le débit maximal atteint les 311,92 m<sup>3</sup>/s durant le mois de Décembre.

Les crues se produisent lorsque le débit d'un cours d'eau augmente de manière soudaine, entraînant une élévation significative du niveau de l'eau. En général, les crues entraînent des inondations lorsque les eaux débordent du lit du cours d'eau et recouvrent les zones environnantes.

Contrairement aux crues, les étiages correspondent aux plus bas niveaux d'eau atteints par un cours d'eau, Lorsque tous les courants de surface ont cessé et que seule l'eau souterraine est alimentée.

## II.4. Conclusion

Ce chapitre montre l'analyse des débits et les écoulements des cours d'eau naturels dans plusieurs stations hydrométriques du bassin des côtières d'Alger a permis de mieux comprendre les caractéristiques hydrologiques spécifiques de la région étudiée et d'apporter des connaissances précieuses pour la gestion efficace des ressources en eau superficielles. Ces résultats fournissent une meilleure connaissance de régime hydrologique des cours d'eau et une

base solide pour des études futures et des mesures concrètes visant à optimiser l'utilisation de l'eau et à assurer une gestion durable de nos ressources hydriques dans la région d'étude semi-aride et méditerranéenne.

# **CHAPITRE III**

## **TRAITEMENT STATISTIQUE ET ANALYSE DES DONNÉES**

### **III.1. Introduction**

Les données des séries de jaugeage dans les différentes stations étudiées proviennent des services de l'Agence Nationale des Ressources Hydrauliques (A.N.R.H). Cependant, la disparité entre ces sources pose souvent un problème en termes de qualité des données. Ainsi, il est nécessaire d'effectuer une critique et un contrôle de la qualité des données de jaugeage afin de garantir l'homogénéité et la fiabilité des séries utilisées pour une étude hydrologique spécifique. Ces traitements peuvent prendre différentes formes, allant d'un simple traitement préliminaire à des ajustements, tels que le comblement des données manquantes, l'extension de séries de données courtes ou encore l'homogénéisation des séries étudiées grâce à des méthodes statistiques.

### **III.2. Collecte et contrôle des données**

Nous possédons Cinq stations réparties dans l'espace étudié et sont relativement bien distribuées.

L'Agence Nationale des Ressources Hydraulique (A.N.R.H).

- Choix des stations remplissant les conditions, les plus longues, les moins lacuneuses et représentatives du sous bassin correspondant ;
- Comblement de lacune dans les séries chronologiques des stations choisies ;
- Détections des ruptures.

Avant de pouvoir les exploiter et bien qu'elles soient dans un format adéquat, il importe de contrôler leur fiabilité et leur précision. Le contrôle de la validité des données d'observation est un travail préalable indispensable à toute analyse correcte malgré les nombreux contrôles manuels intervenant dans la chaîne de traitement des données de jaugeage.

Le contrôle manuel s'effectue à plusieurs niveaux : d'abord au niveau de la station, puis au niveau du service central, où il est fréquemment possible de détecter des erreurs de natures très diverses.

Le contrôle manuel se fait à plusieurs niveaux, d'abord au niveau de la station, puis au niveau du service central où souvent encore des erreurs de nature fort différente sont susceptibles d'être relevées.

Il existe deux types d'erreurs (Capaccioli et Vaucouleurs,1983) :

**A. Les erreurs aléatoires (accidentelles)**

Elles affectent la précision des données, et sont dues à des raisons nombreuses et variées, généralement inconnues, affectant différemment chaque mesure individuelle. Ces erreurs étant inévitables, il faut en estimer l'importance afin de pouvoir en tenir compte lors de l'évaluation de l'incertitude finale.

**B. Les erreurs systématiques**

L'origine des erreurs systématiques est le plus souvent liée à la calibration de l'appareil de mesure qui n'est pas parfaite ou à un phénomène extérieur qui perturbe la mesure (erreur d'appareillage, changement d'observateur...).

Comme il est connu, la base des études climatologiques ou hydrologiques nécessite des séries de données régulières continues et de longue durée ; l'application des méthodes statistiques sur ces séries impose un certain nombre de conditions entre autres, la continuité des observations dans le temps et la pérennité des méthodes d'observations.

L'historique du réseau climatique de l'ANRH montre que la longueur des séries est inversement proportionnelle à leur nombre. Nous ne disposons que de successions de séries de courte durée, et ce pour plusieurs raisons dont les plus importantes sont :

- La fermeture de certaines stations.
- Les déplacements de certaines autres.
- L'implantation récente de nouvelles stations.

**III.3. Caractéristiques des stations**

Pour un bon traitement des données recueillies, il est nécessaire et indispensable que la période de collecte soit suffisamment longue, autrement dit l'échantillon doit être suffisamment grand (Ardilly, 2006). Cependant, la réalité est différente à cause des lacunes que nous trouvons souvent dans les séries d'observations. (Achite et Meddi, 2004). Pour être utilisables, les observations doivent être aussi exactes que possible et comparables les unes aux autres, donc synchrones et faites dans les mêmes conditions.

**III.4. Choix des stations**

Deux critères nous ont permis de sélectionner les stations choisies : la longueur des séries et leur continuité dans le temps, puis la localisation géographique afin d'avoir une bonne représentativité spatiale. Nous avons ainsi retenu seulement les stations possédant les séries de données les plus complètes possibles (Whitfield et al., 2012).

Tableau. III. 1. Caractéristiques des données des stations hydrométriques étudiées.

<b>Oued Chellif - Station Arib Chellif (1994 – 2009)</b>									
	Nombre des données	Min	Max	Moyenne	Écart-type	Médiane	Coefficient de variation (Cv)	Coefficient d'asymétrie (Cs)	Coefficient d'aplatissement (Ck)
<b>Vitesse (m/s)</b>	230	0,119	1,57	0,849	0,278	0,849	0,328	0,0589	3,00
<b>Rh (m)</b>	193	0,04	0,591	0,253	0,1000	0,240	0,395	0,790	3,94
<b>Débit (m<sup>3</sup>/s)</b>	189	0,038	67,5	2,57	6,81	0,882	2,65	6,23	49,7
<b>Hauteur (cm)</b>	225	0,140	49,9	11,1	10,4	7,90	0,937	2,00	7,06
<b>Largeur (m)</b>	194	0,931	300	137	42,6	131	0,311	0,662	6,00
<b>Section (m<sup>2</sup>)</b>	191	0,121	25,8	3,42	5,26	1,68	1,54	3,01	11,0

<b>Oued Mazafran - Station Fer à Cheval (1985 – 2015)</b>									
	Nombre des données	Min	Max	Moyenne	Écart-type	Médiane	Coefficient de variation (Cv)	Coefficient d'asymétrie (Cs)	Coefficient d'aplatissement (Ck)
<b>Vitesse (m/s)</b>	305	0,149	3,90	0,934	0,489	0,874	0,523	1,84	8,62
<b>Rh (m)</b>	259	0,00	1,50	0,230	0,195	0,170	0,849	3,02	15,5
<b>Débit (m<sup>3</sup>/s)</b>	347	0,006	71,8	3,72	11,4	0,565	3,06	4,35	21,4

<b>Hauteur (cm)</b>	350	2,00	767	88,5	108	47,5	1,22	3,49	19,2
<b>Largeur (m)</b>	324	0,800	96	8,93	8,92	6,65	0,998	4,66	36,00
<b>Section (m<sup>2</sup>)</b>	353	0,047	147	5,65	16,1	1,15	2,85	5,18	34,7

<b>Oued Chiffa - Station Amont Des Gorges (1997 – 2016)</b>									
	<b>Nombre des données</b>	<b>Min</b>	<b>Max</b>	<b>Moyenne</b>	<b>Écart-type</b>	<b>Médiane</b>	<b>Coefficient de variation (Cv)</b>	<b>Coefficient d'asymétrie (Cs)</b>	<b>Coefficient d'aplatissement (Ck)</b>
<b>Vitesse (m/s)</b>	261	0,211	3,97	1,00	0,680	0,815	0,679	1,79	5,95
<b>Rh (m)</b>	224	0,0700	0,600	0,210	0,0984	0,200	0,468	1,27	4,69
<b>Débit (m<sup>3</sup>/s)</b>	240	0,0180	9,82	0,751	1,32	0,372	1,76	3,96	20,3
<b>Hauteur (cm)</b>	241	1,0	95	40,6	17,6	45,0	0,434	-0,111	3,03
<b>Largeur (m)</b>	268	0,700	19,3	5,45	4,39	4,00	0,806	1,60	4,58
<b>Section (m<sup>2</sup>)</b>	234	0,0930	8,17	1,13	1,43	0,726	1,27	3,30	14,4

<b>Oued Sebaou - Station Baghlia (1987 – 2013)</b>									
	<b>Nombre des données</b>	<b>Min</b>	<b>Max</b>	<b>Moyenne</b>	<b>Écart-type</b>	<b>Médiane</b>	<b>Coefficient de variation (Cv)</b>	<b>Coefficient d'asymétrie (Cs)</b>	<b>Coefficient d'aplatissement (Ck)</b>
<b>Vitesse (m/s)</b>	310	0,155	2,65	1,10	0,539	1,06	0,489	0,371	2,35
<b>Rh (m)</b>	151	0,00500	4,05	0,491	0,719	0,130	1,47	2,42	9,09
<b>Débit (m<sup>3</sup>/s)</b>	262	0,0200	29,6	6,33	7,26	2,92	1,15	1,36	4,09
<b>Hauteur (cm)</b>	339	-177	187	3,72	59	5,00	15,9	0,142	3,73
<b>Largeur (m)</b>	284	0,600	95,6	28,0	21,0	26,6	0,750	0,579	2,63
<b>Section (m<sup>2</sup>)</b>	318	0,0480	98,8	16,9	20,5	10,3	1,21	1,85	6,03

<b>Oued Bellah - Station Pont RN 11 (1991 – 2015)</b>									
	<b>Nombre des données</b>	<b>Min</b>	<b>Max</b>	<b>Moyenne</b>	<b>Écart-type</b>	<b>Médiane</b>	<b>Coefficient de variation (Cv)</b>	<b>Coefficient d'asymétrie (Cs)</b>	<b>Coefficient d'aplatissement (Ck)</b>
<b>Vitesse (m/s)</b>	85	0,282	1,45	0,681	0,243	0,630	0,356	0,850	3,51
<b>Rh (m)</b>	151	0,00500	4,05	0,491	0,719	0,130	1,47	2,42	9,09
<b>Débit (m<sup>3</sup>/s)</b>	151	0,0010	1,49	0,196	0,247	0,112	1,26	2,55	10,00
<b>Hauteur (cm)</b>	150	1,00	102	56,5	25,6	62,0	0,453	0,499	2,46
<b>Largeur (m)</b>	150	0,400	9,80	2,69	1,81	2,10	0,675	1,38	4,89
<b>Section (m<sup>2</sup>)</b>	150	0,0150	2,17	0,395	0,368	0,278	0,932	1,96	7,51

### **III.5. Test d'homogénéité**

Un échantillon de données aléatoires est dit homogène" lorsque les données qui le constituent proviennent de la même distribution (prises dans les mêmes conditions). La vérification de cette hypothèse est fondée sur la comparaison des moyennes des débits obtenus avant et après la date de changement dans la procédure d'acquisition des données, lorsque cette dernière date de changement est connue.

Toutes les études sont basées sur des séries de données d'observations, qui ne sont pas dépourvues d'erreurs et de lacunes. L'absence de certaines informations (absence de mesures, défaut de transmission, erreur de stockage, manipulation erronée...) peut occasionner au niveau des fichiers disponibles, certaines données erronées ou manquantes, aussi, chaque étude hydrologique nécessite une critique des données afin d'éliminer les données les plus suspectes.

Pour travailler sur une base fiable et loin des anomalies, l'étude de l'homogénéité est une étape indispensable pour s'assurer la fiabilité de ces données (Ondo et al., 1997 ; Brunet-Moret, 1977, 1979). Pour cela, nous avons testé l'homogénéité d'une série temporelle par la méthode du simple cumul (méthode absolue) et appliqué la méthode du double cumul qui a été proposée par Kohler (1949) et du cumul des résidus (Bois, 1971,1986).

On vérifie l'homogénéisation des totaux annuels qui se situent en abscisses de la station de référence, et en ordonnées les totaux annuels en station à contrôler. Lorsque la distribution des points est linéaire, la série est homogène et lorsqu'elle connaît des ruptures, la série est hétérogène.

### **III.6. Choix du modèle d'ajustement**

L'ajustement d'un modèle fréquentiel à un échantillon donné a pour but, le regroupement en une fonction et ses paramètres, les informations se trouvant au sein de cet échantillon même. Le choix du modèle en analyse fréquentielle est sans doute l'étape la plus critique introduisant les plus grands incertitudes (Meylan et al., 2012). L'adoption d'un modèle fréquentiel pour étudier et décrire les phénomènes hydrologiques est donc une décision, un choix. Dubreuil (1974) donne des éléments d'aide au choix de la loi théorique à choisir :

1. Dans une région les données, une variable hydrologique déterminée suit généralement la même loi en tout site d'observation, d'où l'intérêt des études systématiques et celui de la connaissance de toute étude antérieure.
2. En l'absence d'information régionale, on fait un essai de report graphique des points observés sur papier à échelle gaussienne des abscisses, ce qui permet d'envisager une loi normale en une loi dissymétrique. En effet, l'anamorphose d'échelle permet l'alignement des points suivant une loi normale, tandis que ceux suivant une loi dissymétrique se présenteraient selon une courbe à concavité tournée vers le haut (loi hypergaussienne) ou vers le bas (loi hypogaussienne).

Selon (Musy et Meylan, 1999), le choix du modèle fréquentiel est déterminant quant à la validité des résultats d'une analyse fréquentielle, seulement il n'existe pas de formule universelle pour choisir un modèle.

La raison pour laquelle il serait recommandé de considérer la distribution de Gumbel, est qu'elle n'implique que deux paramètres (la distribution de GEV à trois paramètres) et qui est traduite par une expression mathématique plus simple. Cependant, des études récentes ont montré que sur le plan de la pratique hydrologique, la distribution de Gumbel, bien que plus simple, ne fournit pas nécessairement les meilleurs résultats (Malamud et Turcotte, 2006 ; Coles et al., 2003 ; Koutsoyannis, 2004 a & b ; Sisson et al., 2006 ; Goula et al., 2007 ; Goula et al., 2008 ).

Pour l'analyse des dépassements au-delà d'un seuil assez élevé, Belkma et Haans (1974) et Pickands (1975), ont introduit la loi de Preto généralisée, qui est une loi paramétrique considéré comme étant une distribution limite normalisée adaptée aux excès au-delà d'un seuil et est devenue d'une grande importance dans l'étude des valeurs extrêmes, La modélisation fondée sur la loi GPD comme composante d'un modèle de renouvellement dépassement est préférable à celle n'utilisant que la loi exponentielle: cette dernière n'est en effet qu'un cas particulier de la GPD (Bacro et Chaouche, 2006).

Nous retenons deux familles de loi pour l'étude des valeurs extrêmes :

- Lois GEV (Loi Généralisée des valeurs extrêmes et loi de Gumbel) pour la méthode des maxima annuels.
- Lois de Pareto Généralisée pour la méthode des excès au-delà d'un seuil.

La représentation graphique de la loi de distribution choisie est procédée comme suit :

Les débits moyens observés dans les différentes stations étudiées s'ajustent parfaitement aux lois de Gumbel et Lognormal (figures III.1, III.2, III.3, III.4 et III.5).

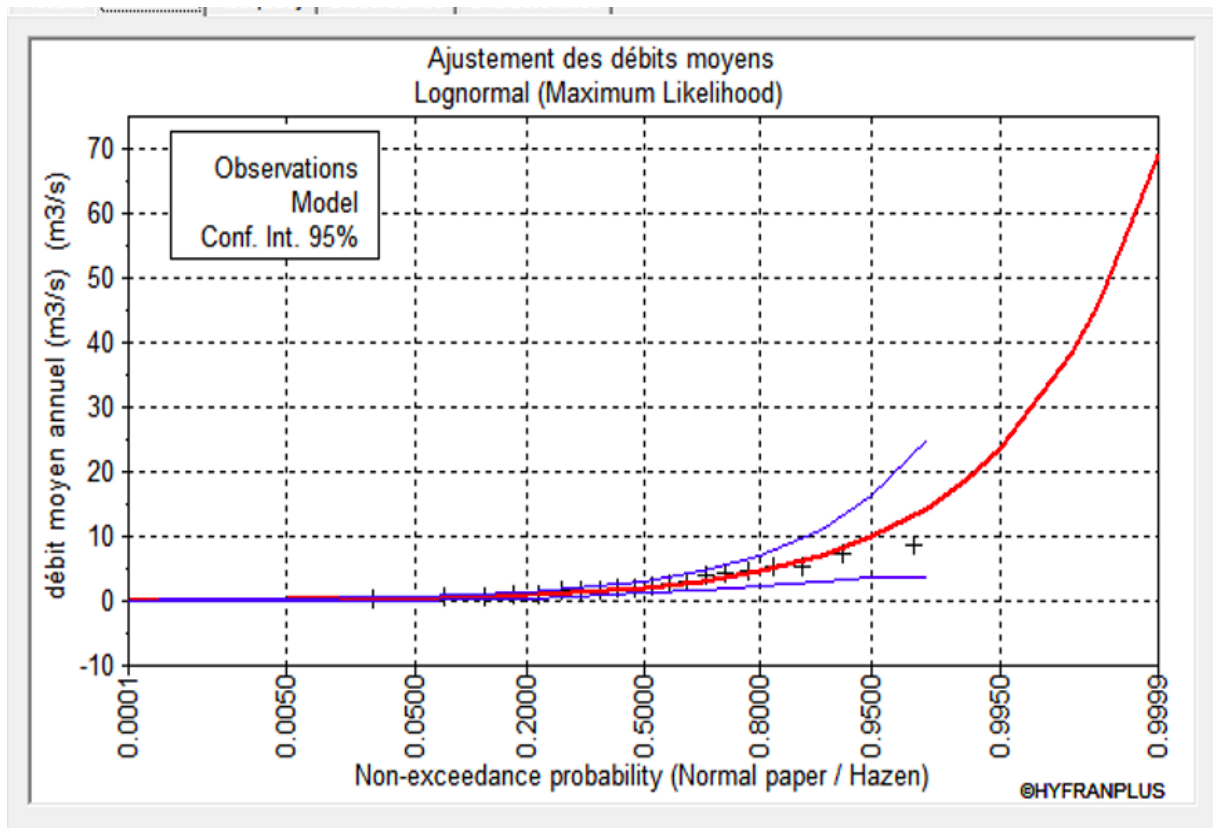


Figure. III. 1. Ajustement des débits moyens à la loi log normal d'Oued Chellif (Station Arib Chellif).

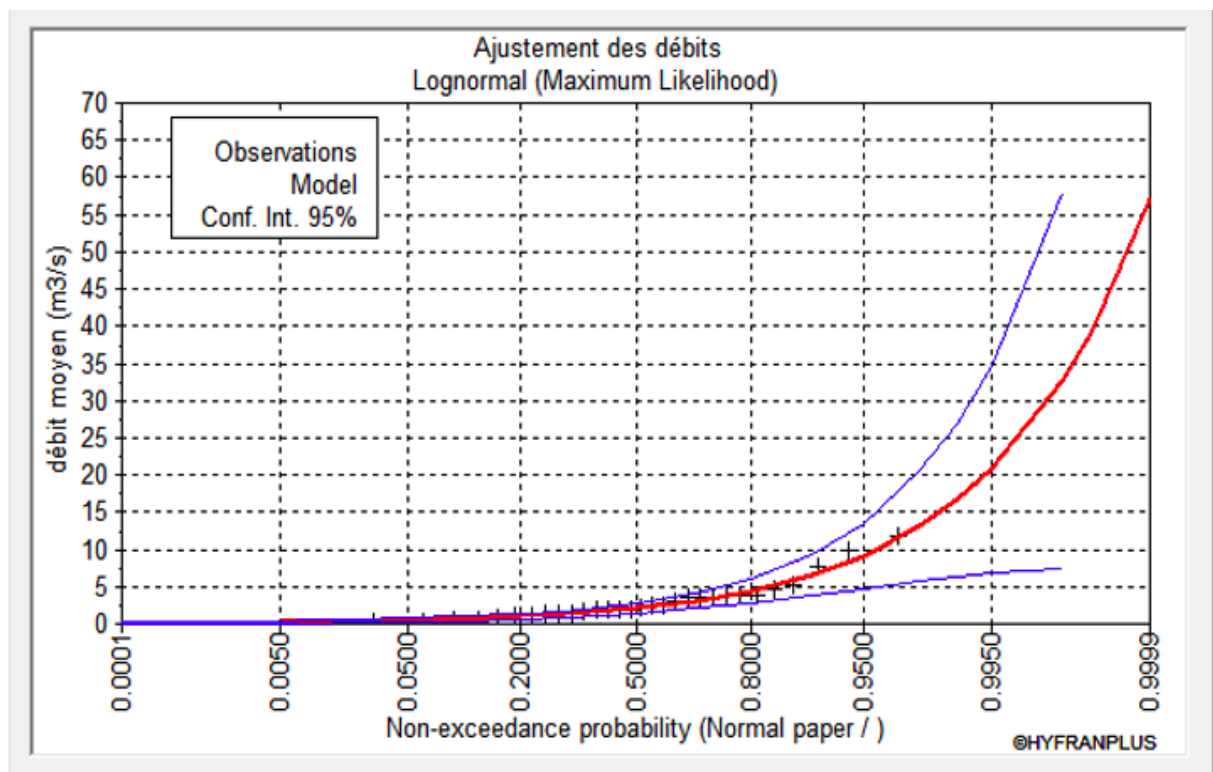


Figure. III.2. Ajustement des débits moyens à la loi log normal d'Oued Mazafran (Station Fer à Cheval).

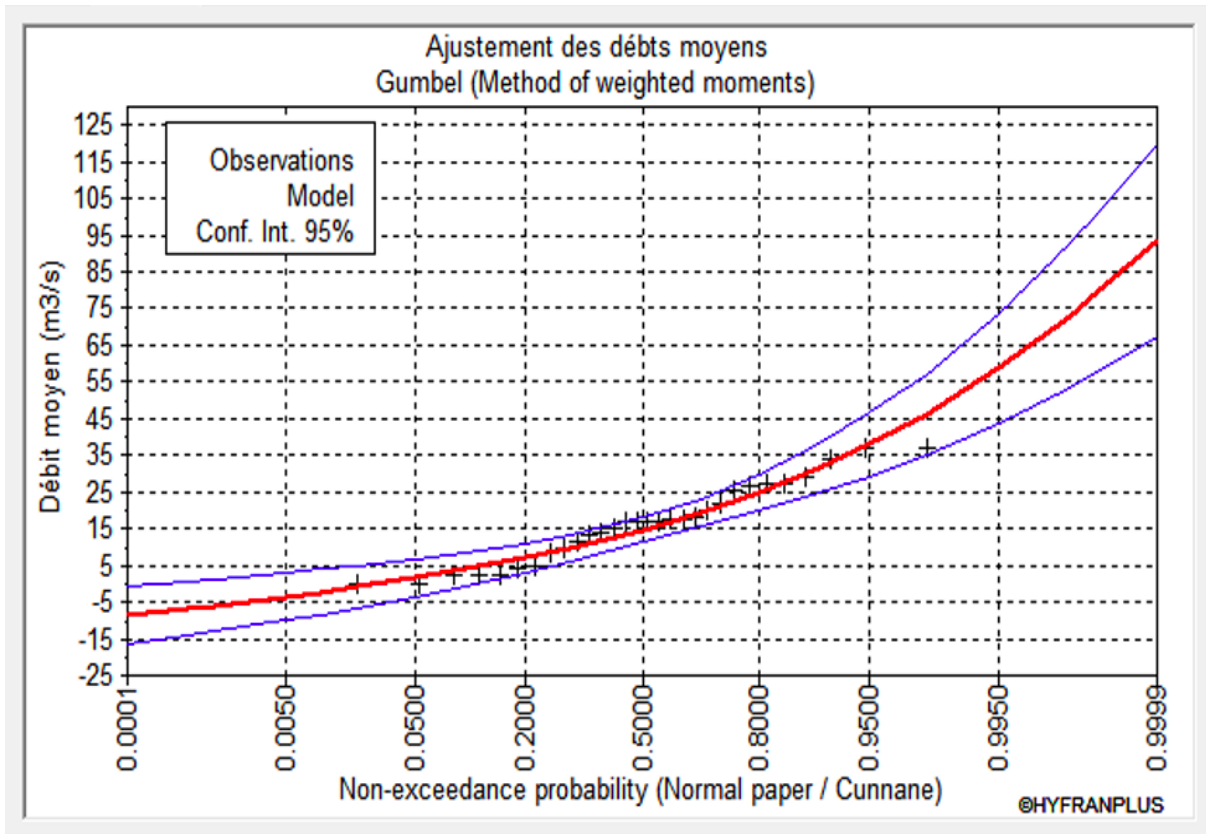


Figure. III.3. Ajustement des débits moyens à Gumbel d'Oued Chiffa (Station Amont Des Gorges).

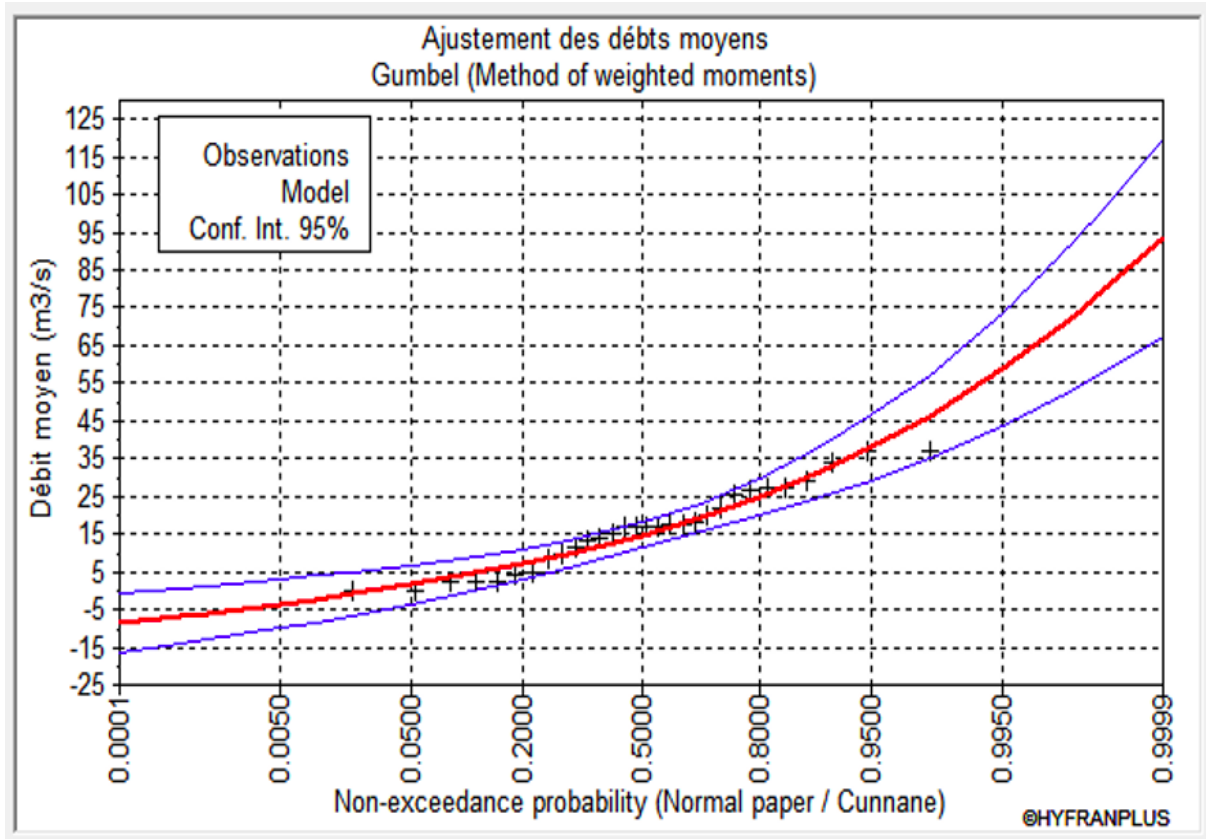


Figure. III.4. Ajustement des débits moyens à Gumbel d'Oued Sebaou (Station Baghlia).

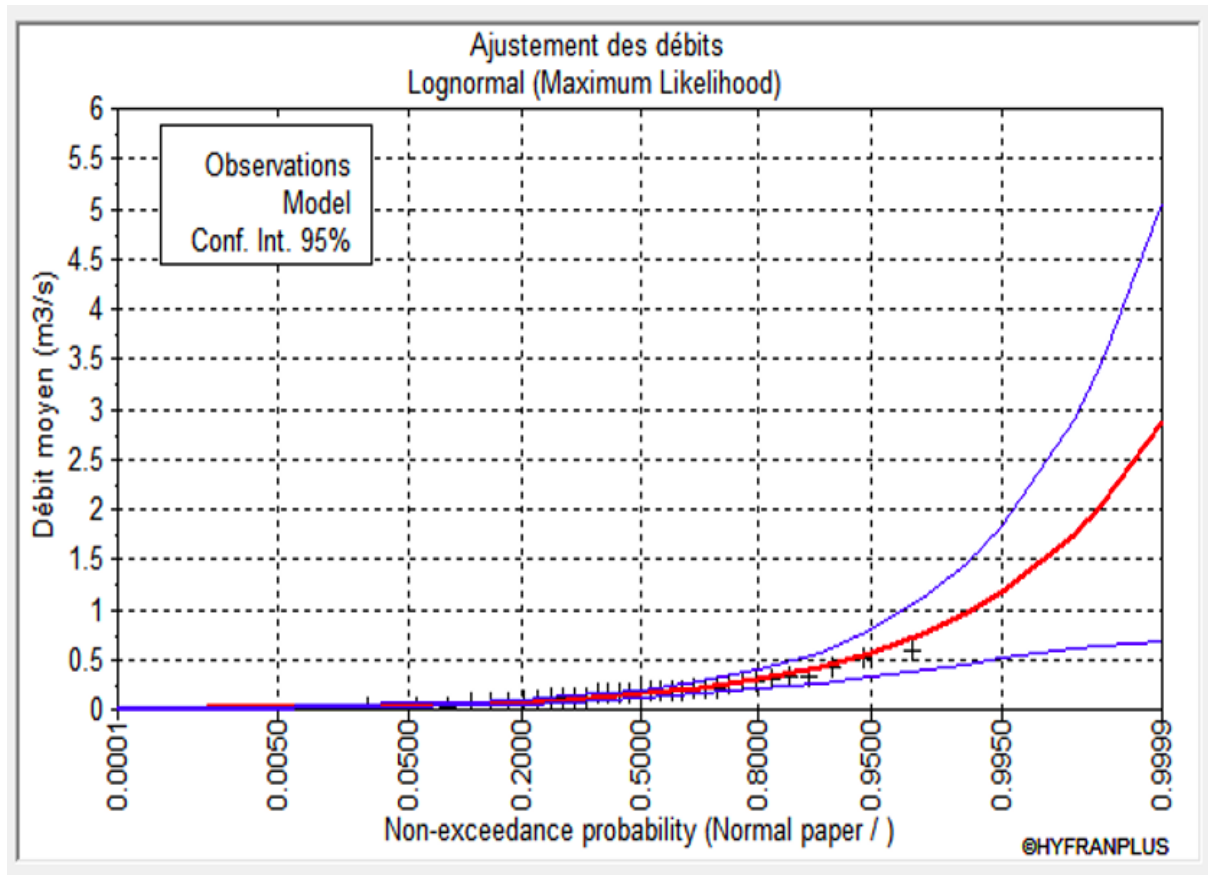


Figure. III.5. Ajustement des débits moyens à la loi log normal d'Oued Bellah (Station Pont RN 11).

### III.7. Test de Wilcoxon

Le test de Wilcoxon est un des tests le plus utilisé pour vérifier l'homogénéité d'un échantillon de données. Il permet d'effectuer des comparaisons entre deux sous échantillons et de vérifier si leurs moyennes sont significativement différentes. Considérons une série de données notées respectivement  $x_1, x_2, x_3, \dots, x_n$  et  $x_{n+1}, x_{n+2}, \dots, x_m$ , dont les moyennes sont respectivement  $M_1$  et  $M_2$ . Le test permet de comparer les hypothèses suivantes :

$$H_0 : M_1 = M_2 ;$$

$$H_1 : M_1 \neq M_2$$

La statistique du test utilisée ( $V$ ) est distribuée asymptotiquement selon une loi normale centrée réduite. Cette statistique s'exprime comme suit :

$$V = [W - 0.5(n + 1) + 0.5] / \sqrt{Var(W)} \tag{III.1}$$

Où

$$W = \sum_{i=1}^m R_i S(R_i) \tag{III.2}$$

$R_j$  : est le rang du classement des variables aléatoires des deux échantillons combinés et triés en ordre croissant.

$(R_i)$  : facteur multiplicateur égal à (0) si la donnée est effectuée après le changement,  $s(R_i)$  est le rang de la donnée si elle est effectuée avant le changement.

La variance de la statistique  $s$  s'exprime par la formule suivante :

$$Var(V) = \frac{n(m-n)}{12} - \frac{n(m-n) \sum_{i=1}^k d_k^3 - d_k}{12m(m-1)} \quad (III.4)$$

Où :  $k$  : étant le nombre de groupes distinct contenant des observations égales.  $d_k$  : est le nombre d'observations égales dans chacun des groupes.

Les règles de décision à un seuil  $\alpha$  donné pour effectuer ce test sont les mêmes que pour une loi normale : si  $Abs(V) > Z_{\alpha/2}$ , on rejette  $H_0$ , sinon, nous acceptons  $H_0$ .

$Abs(V)$  : Valeur absolue de la statistique  $V$ .

$Z_{\alpha/2}$  est le quantile de probabilité au dépassement  $\alpha/2$  de la loi normale centrée réduite.

Les résultats de l'homogénéité et l'ajustement des séries hydrométriques étudiées ont été présentées dans le tableau ci-après.

Tableau. III. 2. Homogénéité et ajustement des données des stations hydrométriques étudiées.

Oued Chellif - Station Arib Chellif (1994 – 2009)									
	Nombre des données	Moyenne	Écart-type	Médiane	Coefficient de variation (Cv)	Coefficient d'asymétrie (Cs)	Coefficient d'aplatissement (Ck)	Intervalle de confiance	Loi d'ajustement
Vitesse (m/s)	230	0,849	0,278	0,849	0,328	0	3,00	Accepté 98%	Normal
Rh (m)	193	0,253	0,1000	0,240	0,395	0,790	3,94	Accepté 98%	Gamma
Débit (m <sup>3</sup> /s)	189	2,17	4,95	0,869	2,29	18,8	2100	Accepté 98%	Log normal
Hauteur (cm)	225	11,6	16,9	7,81	1,46	1	-1	Accepté 98%	GEV
Largeur (m)	194	137	40,4	130	0,295	1,14	2,4	Accepté 98%	Gumbel
Section (m <sup>2</sup> )	191	3,25	4,94	1,78	1,52	8,09	224	Accepté 98%	Log normal

Oued Mazafran - Station Fer à Cheval (1985 – 2015)									
	Nombre des données	Moyenne	Écart-type	Médiane	Coefficient de variation (Cv)	Coefficient d'asymétrie (Cs)	Coefficient d'aplatissement (Ck)	Intervalle de confiance	Loi d'ajustement
Vitesse (m/s)	305	0,938	0,500	0,828	0,533	1,75	8,89	Accepté 98%	Log normal
Rh (m)	259	0,230	0,221	0,175	0,963	23,4	40,6	Accepté 98%	GEV
Débit (m <sup>3</sup> /s)	347	1,39	4,26	0,539	N/D	N/D	86,7	Accepté 98%	GEV
Hauteur (cm)	350	90,6	129	52,0	1,43	7,17	165	Accepté 98%	Log normal
Largeur (m)	324	8,80	7,50	6,70	0,852	3,18	25,1	Accepté 98%	Log normal
Section (m <sup>2</sup> )	353	8,23	29,2	1,11	3,55	6,30	46,6	Accepté 98%	GEV

<b>Oued Chiffa - Station Amont Des Gorges (1997 – 2016)</b>									
	Nombre des données	Moyenne	Écart-type	Médiane	Coefficient de variation (Cv)	Coefficient d'asymétrie (Cs)	Coefficient d'aplatissement (Ck)	Intervalle de confiance	Loi d'ajustement
<b>Vitesse (m/s)</b>	261	0,994	0,643	0,835	0,647	2,21	12,8	Accepté 98%	Log normal
<b>Rh (m)</b>	224	0,210	0,0980	0,191	0,466	1,50	7,25	Accepté 98%	Log normal
<b>Débit (m<sup>3</sup>/s)</b>	240	0,737	1,46	0,331	1,99	13,8	91,2	Accepté 98%	Log normal
<b>Hauteur (cm)</b>	241	40,6	17,6	40,6	0,434	0,00	3,00	Accepté 98%	Normal
<b>Largeur (m)</b>	268	5,40	4,42	4,18	0,818	3,00	22,4	Accepté 98%	Gumbel
<b>Section (m<sup>2</sup>)</b>	234	1,09	1,32	0,699	1,20	5,30	80,6	Accepté 98%	Log normal

<b>Oued Sebaou - Station Baghlia (1987 – 2013)</b>									
	Nombre des données	Moyenne	Écart-type	Médiane	Coefficient de variation (Cv)	Coefficient d'asymétrie (Cs)	Coefficient d'aplatissement (Ck)	Intervalle de confiance	Loi d'ajustement
<b>Vitesse (m/s)</b>	310	1,10	0,539	1,07	0,489	0,371	2,73	Accepté 98%	Generalized Gamma
<b>Rh (m)</b>	325	0,482	0,271	0,424	0,562	2,02	12,1	Accepté 98%	Log normal
<b>Débit (m<sup>3</sup>/s)</b>	262	6,33	8,84	4,12	1,40	8,91	1	Accepté 98%	GEV
<b>Hauteur (cm)</b>	339	3,72	59	3,72	15,9	0,00	3,00	Accepté 98%	Normal
<b>Largeur (m)</b>	284	28,0	21,0	26,6	0,750	0,579	2,63	Accepté 98%	Pareto
<b>Section (m<sup>2</sup>)</b>	320	17,5	29,4	10,6	1,68	1	1	Accepté 98%	GEV

<b>Oued Bellah - Station Pont RN 11 1991 – 2015</b>									
	<b>Nombre des données</b>	<b>Moyenne</b>	<b>Écart-type</b>	<b>Médiane</b>	<b>Coefficient de variation (Cv)</b>	<b>Coefficient d'asymétrie (Cs)</b>	<b>Coefficient d'aplatissement (Ck)</b>	<b>Intervalle de confiance</b>	<b>Loi d'ajustement</b>
<b>Vitesse (m/s)</b>	85	0,684	0,249	0,643	0,364	1,14	5,39	Accepté 98%	Log normal
<b>Rh (m)</b>	151	0,497	1,11	0,203	2,23	17,8	1820	Accepté 98%	Log normal
<b>Débit (m<sup>3</sup>/s)</b>	151	0,244	0,254	0,106	1,29	2,58	10,2	Accepté 98%	GEV
<b>Hauteur (cm)</b>	150	56,5	25,9	59,7	0,458	-0,658	3,30	Accepté 98%	GEV
<b>Largeur (m)</b>	150	2,71	2,00	2,18	0,738	2,61	17,2	Accepté 98%	Log normal
<b>Section (m<sup>2</sup>)</b>	144	4,26	0,501	0,280	1,19	5,25	76,2	Accepté 98%	Log normal

D'après l'analyse statistique et le traitement des données élaborés, les résultats des tests confirment que les données collectées sont satisfaisantes aux critères de qualité requis, ce qui atteste de leur pertinence et de leur fiabilité pour une bonne exploitation.

### **III.8. Conclusion**

Dans ce chapitre, nous avons réalisé une analyse rigoureuse de la qualité des données de jaugeage des cours d'eau naturels étudiés dans le cadre de notre étude hydrologique. Les séries de données ont été soumises à un processus d'homogénéisation et de normalisation afin d'assurer leur cohérence et leur fiabilité. Grâce à ces étapes, nous avons éliminé les disparités entre les différentes sources de données, comblé les lacunes et ajusté certains points pour les séries de données courtes. Ce travail de vérification et de traitement a permis d'obtenir des séries de jaugeage homogènes et normalisées, prêtes à être utilisées dans notre étude.

Ces résultats garantissent la crédibilité et la précision de nos analyses hydrologiques, nous permettant ainsi de prendre des décisions éclairées basées sur des données de haute qualité. Il est essentiel de souligner que ce processus de vérification et de normalisation des séries de données constitue une étape cruciale pour garantir la validité de nos résultats et renforcer la pertinence de notre étude hydrologique.

# **CHAPITRE IV**

MÉTHODES MMR ET CHIU

## IV.1. Introduction

Ce chapitre a pour objectif de proposer une synthèse des méthodes d'estimation des débits liquides pour des cours d'eau naturels situés dans un climat semi-aride et méditerranéen.

La première partie de ce chapitre est consacrée à la présentation de la méthode du modèle rugueux de référence (MMR). Le modèle rugueux de référence est défini et les caractéristiques hydrauliques de l'écoulement qui s'y produit sont présentées et discutées. L'intérêt du modèle rugueux de référence choisi est mis en évidence et servira de base au calcul du débit liquide faisant l'objet de ce travail.

La deuxième partie de ce chapitre est réservée à la méthode de Chiu. L'application de la Méthode Chiu de l'information et du concept de l'entropie dans le domaine de l'hydrologie et l'hydraulique.

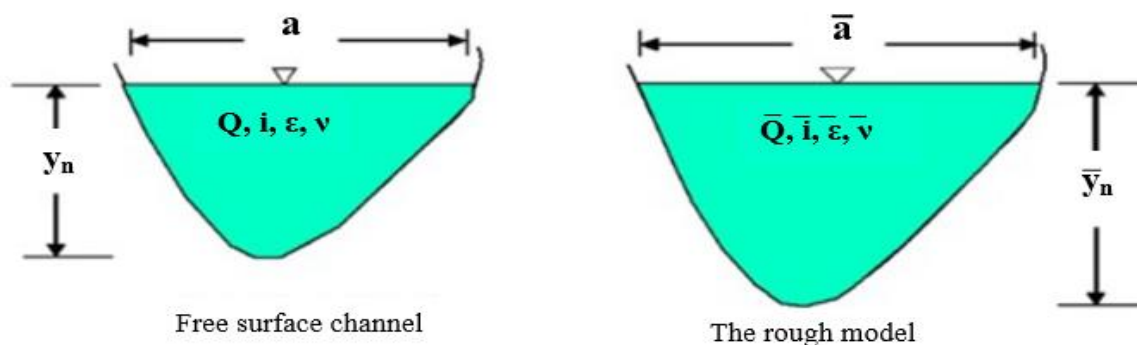
## IV.2. Méthode du modèle rugueux de référence (MMR)

### IV.2.1 Historique

Établie et développée par Achour Bachir entre les années 2002 et 2006 au laboratoire de recherche en hydraulique souterraine et de surface (LARHYSS) de l'université de Biskra, la méthode du modèle rugueux de référence (MMR) est destinée au calcul des conduites et canaux en charge ou à surface libre (Achour. B, 2007).

### IV.2.2 Modèle rugueux de référence

Considérons un canal à surface libre de forme quelconque, comme le montre la figure IV.1. Le canal écoule le débit volume  $Q$ , d'un liquide de viscosité cinématique  $\nu$ , sous une pente longitudinale  $i$ . Il est caractérisé par la dimension linéaire  $a$ , la hauteur d'eau  $y_n$  et. L'état de la paroi interne du canal considéré est caractérisé par la rugosité absolue  $\varepsilon$  (ou  $k$  selon Strickler).



**Figure. IV. 1.** Schéma de la définition d'un canal de forme quelconque et de son modèle rugueux.

D'autre part, on associe au canal considéré un canal de référence (ou modèle de référence) de même forme mais dont les caractéristiques géométriques sont différentes. Ses dimensions linéaires  $\bar{a}$  et  $\bar{y}_n$  (Figure IV.1) sont telles que :  $\bar{a} > a$  et  $\bar{y}_n > y_n$ .

En outre, le canal de référence écoule le débit volume  $\bar{Q}$ , d'un liquide de viscosité cinématique  $\bar{\nu} = \nu$ , sous la pente  $\bar{i}$  qui correspond à la pente géométrique de l'ouvrage. L'état de la paroi interne du canal considéré est caractérisé par la rugosité absolue  $\bar{\varepsilon}$ .

Les caractéristiques du canal de référence sont  $\bar{A}$ ,  $\bar{P}$ ,  $\bar{D}_h$  et  $\bar{R}$ , qui désignent respectivement l'aire de la section mouillée, le périmètre mouillé, le diamètre hydraulique et le nombre de Reynolds.

On affecte au canal de référence une forte rugosité relative  $\bar{\varepsilon}/\bar{D}_h$  tel que : (Achour, 2007)

$$\frac{\bar{\varepsilon}}{\bar{D}_h} = 3,7 \cdot 10^{-2} \quad (\text{IV.1})$$

Le régime de l'écoulement dans le canal de référence est ou peut être considéré comme étant turbulent rugueux de sorte que le coefficient de frottement  $\bar{f}$  soit donné par la relation de Colebrook-White pour  $\bar{R} \rightarrow \infty$ . Il vient que :

$$\frac{1}{\sqrt{\bar{f}}} = -2 \log \left( \frac{\bar{\varepsilon}/\bar{D}_h}{3,7} \right) = -2 \log \left( \frac{3,7 \cdot 10^{-2}}{3,7} \right) = 4 \quad (\text{IV.2})$$

Ou bien :

$$\bar{f} = \frac{1}{16} \quad (\text{IV.3})$$

Ainsi, le coefficient de frottement  $\bar{f}$  de l'écoulement dans le modèle rugueux de référence est constant et vaut 1/16.

Le tableau IV.1 résume les caractéristiques géométriques du modèle de référence ainsi que ses caractéristiques hydrauliques.

Tableau. IV. 1. Caractéristiques géométriques et hydrauliques du MMR.

Paramètre	Symbole
Aire de la section mouillée	$\bar{A}$
Périmètre mouillé	$\bar{P}$
Rayon hydraulique	$\bar{R}_h = \bar{A}/\bar{P}$
Diamètre hydraulique	$\bar{D}_h = 4\bar{R}_h$
Rugosité absolue	$\bar{\varepsilon}$
La pente géométrique	$\bar{i}$
Débit volume	$\bar{Q}$
Coefficient de frottement	$\bar{f}$
Nombre de Reynolds	$\bar{R}$

### IV.2.3 Relation régissant l'écoulement dans le modèle rugueux

L'écoulement dans le modèle rugueux de référence étant à surface libre, il convient de lui appliquer la relation de *Darcy-Weisbach*. La relation de *Darcy-Weisbach* s'applique également aux canaux à ciel ouvert, dans laquelle le gradient  $J$  de la perte de charge linéaire correspond à la pente géométrique  $i$  du canal étudié, soit :

$$i = f \frac{1}{D_h} \frac{Q^2}{2gA^2} \quad (\text{IV.4})$$

Où  $f$  est le coefficient de frottement.

En tenant compte du fait que  $D_h = 4A/P$ , la relation (IV.4) peut alors s'écrire plus simplement

$$i = \frac{f}{8g} \frac{P}{A^3} Q^2 \quad (\text{IV.5})$$

Appliquées au modèle rugueux de référence, la relation précédente mène à écrire que la pente longitudinale  $\bar{i}$  du modèle rugueux est :

$$i = \frac{\bar{f}}{8g} \frac{\bar{P}}{\bar{A}^3} \bar{Q}^2 \quad (\text{IV.6})$$

Où  $\bar{A}$  et  $\bar{P}$  désignent respectivement l'aire de la section mouillée et le périmètre mouillé de l'écoulement dans le modèle rugueux. La relation (IV.6) s'écrit alors, avec  $\bar{f} = 1/16$ :

$$i = \frac{1}{128g} \frac{\bar{P}}{\bar{A}^3} \bar{Q}^2 \quad (\text{IV.7})$$

En ayant recours à la relation de Darcy-Weisbach, la relation (IV.7) constitue alors la relation de base régissant l'écoulement dans le modèle rugueux de référence, applicable à tout profil géométrique.

D'autre part, le nombre de Reynolds  $\bar{R}$  caractérisant l'écoulement dans le modèle rugueux de référence s'écrit par définition :

$$\bar{R} = \frac{\bar{V} \bar{D}_h}{\nu} \quad (\text{IV.8})$$

Où  $\bar{V}$  est la vitesse moyenne de l'écoulement dans le modèle rugueux, de viscosité cinématique  $\nu$ .

Or la vitesse moyenne  $\bar{V}$  est liée au débit volume  $\bar{Q}$  par une relation identique à la relation soit :

$$\bar{Q} = \bar{V} \bar{A} \quad (\text{IV.9})$$

La relation (IV.8) s'écrit alors :

$$\bar{R} = \frac{4\bar{Q}}{\bar{P}\bar{V}} \quad (\text{IV.10})$$

#### IV.2.4 Calcul de l'écoulement turbulent par le modèle rugueux de référence

Dans le présent paragraphe, l'écoulement turbulent en canal à surface libre est calculé en ayant recours aux caractéristiques de l'écoulement dans le canal rugueux de référence.

Il s'agit d'établir principalement les relations permettant de déterminer le débit volume  $Q$  écoulé par un canal à surface libre. Cependant, des relations intéressantes seront également proposées pour le calcul du débit volume  $Q$  telles que le nombre de Reynolds  $R$  et le coefficient de frottement  $f$ .

##### IV.2.4.1. Expression du nombre de Reynolds

Il s'agit d'établir l'expression du nombre de Reynolds  $R$  caractérisant l'écoulement turbulent dans un canal à surface libre écoulant un débit volume  $Q$  inconnu. Les paramètres connus du problème sont les dimensions linéaires du canal (c'est-à-dire  $A$  et  $P$ ), la pente  $i$  du canal, la rugosité absolue  $\varepsilon$  et la viscosité cinématique  $\nu$  du liquide en écoulement.

Assumons les égalités suivantes :  $A = \bar{A}$ ;  $P = \bar{P}$ ;  $R_h = \bar{R}_h$  et  $i = \bar{i}$ .

Sous ces conditions, il est évident que les débits volume et les nombres de Reynolds sont différents, soient  $Q \neq \bar{Q}$ ,  $R \neq \bar{R}$ . La relation (IV.7) devient alors :

$$i = \frac{1}{128g} \frac{P}{A^3} \bar{Q}^2 \quad (\text{IV.11})$$

Le débit  $\bar{Q}$  s'exprime alors, en vertu de la relation (IV.11), par

$$\bar{Q} = 8\sqrt{2} \frac{A^{3/2}}{P^{1/2}} \sqrt{ig} \quad (\text{IV.12})$$

D'autre part, la relation (IV.10) devient, pour  $P = \bar{P}$ :

$$\bar{R} = \frac{4\bar{Q}}{Pv} \quad (\text{IV.13})$$

En insérant les relations (IV.12) dans la relation (IV.13), celle-ci devient :

$$\bar{R} = 32\sqrt{2} \left(\frac{A}{P}\right)^{3/2} \frac{\sqrt{ig}}{v} \quad (\text{IV.14})$$

En tenant compte du fait que  $R_h = A/P$ , la relation (IV.14) peut alors s'écrire plus simplement :

$$\bar{R} = 32\sqrt{2} \frac{\sqrt{igR_h^3}}{v} \quad (\text{IV.15})$$

En comparant les relations (I.6) et (IV.13), nous pouvons conclure que le nombre de Reynolds caractérisant l'écoulement turbulent dans le modèle rugueux de référence est différent de celui de l'écoulement dans la conduite considérée  $R \neq \bar{R}$ .

En égalant les relations (IV.5) et (IV.11), nous pouvons écrire :

$$\frac{f}{8g} \frac{P}{A^3} Q^2 = \frac{1}{128g} \frac{P}{A^3} \bar{Q}^2 \quad (\text{IV.16})$$

Soit :

$$Q = \frac{1}{4\sqrt{f}} \bar{Q} \quad (\text{IV.17})$$

Désignons par  $\psi_Q$  la quantité  $[1/(4\sqrt{f})]$  figurant dans la relation (IV.17), soit :

$$\psi_Q = \frac{1}{4\sqrt{f}} \quad (\text{IV.18})$$

La relation (IV.17) s'écrit donc :

$$Q = \psi_Q \bar{Q} \quad (\text{IV.19})$$

La relation (IV.19) indique que le débit volume  $Q$  écoulé par le canal considéré est égal au débit  $\bar{Q}$  écoulé par le modèle rugueux de référence corrigé par les effets d'un facteur  $\psi_Q$ . Celui-ci peut donc être considéré comme étant le facteur de correction des débits volume.

D'autre part, le rapport entre les relations (I.6) et (IV.13) mène à écrire que :

$$\frac{R}{\bar{R}} = \frac{Q}{\bar{Q}} \quad (\text{IV.20})$$

Or, selon la relation (IV.19), le rapport  $Q/\bar{Q}$  n'est autre que le facteur de correction des débits volume  $\psi_Q$ . Ainsi :

$$R = \psi_Q \bar{R} \quad (\text{IV.21})$$

La relation (IV.21) traduit le fait que le nombre de Reynolds  $R$  caractérisant l'écoulement dans un canal semi circulaire à surface libre est égal au nombre de Reynolds  $\bar{R}$  caractérisant l'écoulement dans le modèle rugueux de référence, corrigé par les effets du facteur de correction des débits volumes  $\psi_Q$ .

#### IV.2.4.2. Formule modifiée de Colebrook-White

La formule universellement connue de *Colebrook-White*, exprimant de manière implicite le coefficient de frottement  $f$  s'écrit :

$$f = \left[ -2 \log \left( \frac{\varepsilon/D_h}{3.7} + \frac{2.51}{R\sqrt{f}} \right) \right] \quad (\text{IV.22})$$

En substituant dans cette dernière relation les formules (IV.18, IV.21 et  $D_h = 4R_h$ ), nous pouvons écrire que :

$$\frac{1}{16\psi_Q^2} = \left[ -2 \log \left( \frac{\varepsilon/D_h}{14.8} + \frac{2.51 \times 4 \psi_Q}{\psi_Q \bar{R}} \right) \right]^{-2} \quad (\text{IV.23})$$

Après simplifications et réarrangements, il vient que :

$$\psi_Q = -\frac{1}{2} \log \left( \frac{\varepsilon/R_h}{14.8} + \frac{10.04}{\bar{R}} \right) \quad (\text{IV.24})$$

Nous pouvons conclure que le facteur de correction des débits  $\psi_Q$  est fonction de la rugosité relative  $\varepsilon/R_h$  de la conduite et du nombre de Reynolds  $\bar{R}$  caractérisant l'écoulement dans le modèle rugueux de référence.

La relation (IV.24) est applicable dans tous les domaines de l'écoulement turbulent (lisse, de transition et rugueux).

Pour le domaine de l'écoulement lisse correspondant à  $\varepsilon/R_h = 0$  ou pour celui de l'écoulement pratiquement lisse correspondant à  $\varepsilon/R_h \rightarrow 0$ , la relation (IV.24) s'écrit :

$$\psi_Q = -\frac{1}{2} \log \left( \frac{10.04}{\bar{R}} \right) \quad (\text{IV.25})$$

Par contre, pour le domaine turbulent rugueux correspondant à  $\nu \rightarrow 0$  ou à  $\bar{R} \rightarrow \infty$ , la relation (IV.24) devient :

$$\psi_Q = -\frac{1}{2} \log \left( \frac{\varepsilon/R_h}{14.8} \right) \quad (\text{IV.26})$$

#### IV.2.4.3. Expression du coefficient de frottement

Même si la valeur du débit volume  $Q$  n'est pas une donnée du problème, la valeur exacte du coefficient de frottement  $f$  peut être calculée, à condition que les paramètres  $i, R_h, \varepsilon$  et  $\nu$ , soient connus.

Pour établir l'expression du coefficient de frottement  $f$ , il suffit d'éliminer le facteur de correction des débits volume  $\psi_Q$  entre les relations (IV.18) et (IV.24), soit :

$$\frac{1}{4\sqrt{f}} = -\frac{1}{2} \log \left( \frac{\varepsilon/R_h}{14.8} + \frac{10.04}{\bar{R}} \right) \quad (\text{IV.27})$$

Ou bien :

$$f = \left[ -2 \log \left( \frac{\varepsilon/D_h}{14.8} + \frac{10.04}{\bar{R}} \right) \right]^{-2} \quad (\text{IV.28})$$

La relation (IV.28) est applicable dans tout le domaine de l'écoulement turbulent. Le nombre de Reynolds  $\bar{R}$  est calculé selon la relation la relation (IV.15) pour les valeurs connues des paramètres  $i, R_h$  et  $\nu$ .

#### IV.2.4.4. Expression du débit volume $Q$

L'expression du débit volume  $Q$  écoulé par un canal à surface libre peut être déterminée après avoir éliminé le facteur de correction des débits volume  $\psi_Q$  des relations (IV.19) et (IV.24), soit :

$$Q = -\frac{\bar{Q}}{2} \log \left( \frac{\varepsilon/R_h}{14.8} + \frac{10.04}{\bar{R}} \right) \quad (\text{IV.29})$$

Dans la relation (IV.29), le débit volume  $\bar{Q}$  est donné par la relation (IV.12), soit :

$$\bar{Q} = 8\sqrt{2}A\sqrt{igR_h} \quad (\text{IV.30})$$

Les relations (IV.28) et (IV.29) permettent d'écrire que :

$$Q = -4\sqrt{2}A\sqrt{igR_h} \log \left( \frac{\varepsilon/R_h}{14.8} + \frac{10.04}{\bar{R}} \right) \quad (\text{IV.31})$$

La relation (IV.31) est la relation générale qui permet le calcul du débit volume  $Q$  dans une conduite demi circulaire à surface libre. Elle est applicable à l'ensemble du domaine de l'écoulement turbulent correspondant à  $R > 2300$  et  $0 \leq \varepsilon/R_h \leq 0.05$  (Achour.B, 2007).

### IV.3. Méthode CHIU

La méthode Chiu permet facilement de déterminer les débits des cours d'eau naturels. L'application de cette méthode permet de définir un modèle linéaire entre la vitesse moyenne et la vitesse maximale d'une section d'écoulement. De plus, il détermine les débits avec une précision très satisfaisante, en utilisant les valeurs des paramètres d'entropie  $M$  et expressions pour l'estimation de l'aire de la section transversale domaine des cours d'eau naturels. Cette méthode a montré des bonnes performances dans l'estimation du débit liquide des cours d'eau naturels aux stations mal jaugées en climat semi-aride, comme dans le cas d'un certains bassins algériens (Ammari et Remini, 2010 ; Moramarco et al., 2008 ; Ammari et al., 2017).

Basé sur le concept d'entropie, Chiu (1988, 1989) a développé une expression simple, qui décrit le profil de vitesse illustré à la figure IV.2. Cette méthode a été utilisée aux USA (Chiu et Chen, 1999) et à Taïwan (Chen et al., 2002), ainsi qu'en Italie (Moramarco et al., 2004), donnant de très bonnes performances.

#### IV.3.1 Historique de la théorie de l'information (l'entropie)

La théorie de l'Information trouve son origine dans les travaux de Ronald Aylmer Fisher (1948), un statisticien qui a formulé une définition précise de l'information. Selon Fisher, l'information est équivalente à la valeur moyenne du carré de la dérivée du logarithme de la loi de probabilité étudiée. A partir de l'inégalité de Cramer, la valeur d'une telle information est proportionnelle à la faible variabilité des conclusions résultantes. En termes simples, moins une observation est probable, plus son observation apporte d'information.

D'autres modèles mathématiques ont complété et étendu de façon formelle la définition de l'information.

Claude Shannon et Warren Weaver (1950) renforcent le paradigme. Ils sont des ingénieurs en télécommunication et se préoccupaient de mesurer l'information pour en déduire les fondamentaux de la Communication (et non une théorie de l'information). Dans Théorie Mathématique de la Communication en 1948, ils modélisent l'information pour étudier les lois correspondantes : bruit, entropie et chaos, par analogie générale aux lois d'énergétique et de thermodynamique. Leurs travaux complétant ceux d'Alan Turing, de Norbert Wiener et de John Von Neumann (pour ne citer que les principaux) constituent le socle initial de la théorie du signal et des « Sciences de l'Information ».

La théorie de l'information est due à Shannon (1948), avec bien sûr l'influence des grands théoriciens de l'informatique (Turing, Von Neumann, Wiener). À noter des convergences avec les travaux de Fisher.

Le terme entropie a été forgé en 1865 par le physicien allemand Clausius à partir de la racine grecque *trope* qui évoque l'idée de transformation ou de retour en arrière. Il introduisit cette grandeur afin de caractériser mathématiquement l'irréversibilité de processus physiques, tels qu'une transformation de travail en chaleur. Nous verrons en particulier que la notion de quantité d'information, utile en théorie de la communication ou en informatique, est étroitement apparentée à celle d'entropie.

### IV.3.2 Définition de la théorie de l'information

L'entropie est une mesure de l'incertitude de l'état d'un système physique (Tung, 1990). L'entropie est une grandeur thermodynamique. C'est une quantité physique, mesurable, associée au degré de désordre d'un système macroscopique, ou au manque d'informations sur son état microscopique. Son unité est le joule par kelvin (J/K) en physique.

Deux définitions de l'entropie peuvent être données :

L'une est celle de la thermodynamique classique. On la doit à Rudolf Clausius dont l'œuvre s'appuyait sur celles de nombreux devanciers, dont Sadi Carnot et Rudolf (1899).

L'entropie est l'une des grandeurs les plus fondamentales de la thermodynamique. Elle intervient dans le "second" principe, associé à l'intuition de la flèche du temps. De nombreux phénomènes se produisent spontanément toujours dans un sens et sans jamais revenir à leur point de départ. Ils sont dits irréversibles.

L'entropie de Shannon, due à Claude Shannon (1951), est une fonction mathématique qui correspond à la quantité d'information contenue ou délivrée par une source d'information. Une langue, un signal électrique, un fichier informatique quelconque, constituent des sources d'informations. La définition de l'entropie d'une source selon Shannon est telle que plus la source est redondante, moins elle contient d'information au sens de Shannon. L'entropie est ainsi maximale pour une source dont tous les symboles sont équiprobables.

Cette définition est utilisée en électronique numérique pour numériser une source en utilisant le minimum possible de bits sans perte d'information.

Enfin, elle sert à connaître sur combien de bits au minimum on peut coder un fichier, ce qui est très utile pour savoir quelle limite on peut espérer atteindre les algorithmes de compression qui ne perdent pas d'information. Il existe de tels algorithmes dits optimaux, c'est-à-dire qui compressent le fichier en un fichier d'entropie minimale.

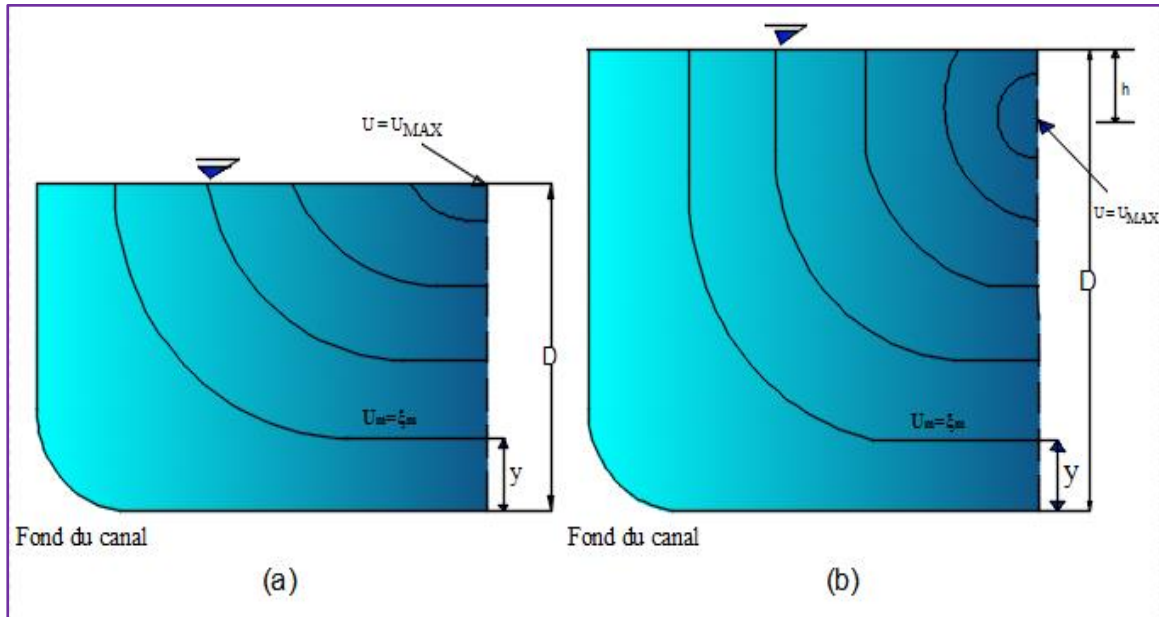
L'entropie permet donc de mesurer la quantité d'information moyenne d'un ensemble d'évènements (en particulier de messages) et de mesurer son incertitude. On la note  $H$  :

$$H(I) = -\sum_{i \in I} P_i \log_2 P_i \quad (\text{IV.32})$$

Qui est la somme des produits des probabilités des différents états du système par les logarithmes de ces probabilités prise avec le signe inverse.

### IV.3.3 Dérivations des équations analytiques pour une distribution de vitesse bidimensionnelle dans un canal ouvert

Dans un canal ouvert, qui n'est pas "large", la vitesse change dans le vertical ( $y$ ) et transversales ( $z$ ) des directions. Les isovels (surfaces d'iso vitesse) sont courbées vers le haut en direction de la surface de l'eau. La vitesse maximale peut se produire au-dessous de la surface de l'eau, de sorte que la vitesse puisse avoir une valeur égale en deux valeurs différentes de  $y$  (la distance verticale du lit de canal). En modélisant la distribution de vitesse il est, donc, logique d'abord de transformer le  $y$  cartésien dans un autre système, dit,  $\xi$  - $\eta$ , dans lequel le  $\xi$  a une relation unique, linéaire avec une valeur de vitesse. Pour clarifier cette idée, des modèles de la distribution bidimensionnelle de vitesse sont décrits dans la figures IV.2 (a - b) où les isovels sont représentés par un système de coordonnées de  $\xi$  courbe ;  $\eta$  est la trajectoire orthogonale. Une fois que l'équation du  $\xi$  en fonction de  $y$  et de  $z$  est déterminée, l'équation de  $\eta$  peut être dérivée d'elle. L'idéal est d'employer les coordonnées de  $\xi$  et de  $\eta$  ressemblent à celles utilisées dans le cadre des coordonnées cylindriques lors de l'analyse de l'écoulement dans une conduite circulaire. (Chiu et Chao, 1988).



**Figure. IV. 2.** Modèle de distribution de vitesse et système de coordonnées curvilignes (Chiu et Chen 1999).

Où  $u$  ( $0 \leq u \leq u_{max}$ ) être la vitesse moyenne ponctuelle dans le temps et, en conséquence, temps est invariable sur une isovelle, qui est assigné d'une valeur de  $\xi$  ( $\xi_0 \leq \xi \leq \xi_{max}$ ). le terme  $u$  est presque zéro le long d'une isovelle qui a la valeur de  $\xi$  égale à  $\xi_0$ , qui a une petite valeur et représente le lit du canal (le fond du canal). En outre,  $u$  égale  $u_{max}$  qui la valeur maximale de  $u$ , au  $\xi$  serait égale à  $\xi_{max}$ , qui lequel peut se produire sur ou au-dessous de la surface libre de l'eau, la vitesse  $u$  augmente avec la coordonnée spatiale  $\xi$  de  $\xi_0$  jusqu'au  $\xi_{max}$ , Une fonction déterministe  $G$  devrait donc exister qui est liée  $u$  à  $\xi$  :

$$u = G(\xi) \tag{IV.33}$$

Dans la formulation probabiliste, on considère  $\xi$  et  $u$  comme des variables aléatoires avec les fonctions  $q$  et  $p$  de densité de probabilité, respectivement. Supposant que  $\xi$  est une variable spatiale entre  $\xi_0$  et le  $\xi_{max}$ , qui sont également probable, de sorte que la fonction de densité de probabilité de  $\xi$  soit uniforme sur l'intervalle  $(\xi_0, \xi_{max})$  :

$$q = (\xi) = \frac{1}{\xi_{max} - \xi_0} \tag{IV.34}$$

Avec  $\xi_{max} \neq \xi_0$ , Puis, la probabilité que la vitesse soit inférieure ou égale à  $u$ , est une ou la fonction de distribution (cumulative)  $P$  de  $u$ , qui peut être dérivée (d'après Chiu, 1989) comme

$$p(u) = \int_0^u p(u) du = \int_0^u p[G(\xi)] du = \int_{\xi_0}^{G^{-1}(u)} q(\xi) d\xi = \frac{\xi - \xi_0}{\xi_{max} - \xi_0} \tag{IV.35}$$

Ceci signifie que, si le  $\xi$  est aléatoirement prélevé entre  $\xi_0$  et le  $\xi_{max}$  et  $u=G(\xi)$  est obtenu, la probabilité de la vitesse variante entre  $u$  et  $u+$  du est  $p(u)du$ . On peut montrer empiriquement, en utilisant une méthode développée par Chiu et Chiou (1986), que  $(\xi-\xi_0) / (\xi_{max} - \xi_0)$  de l'Équation (IV.35) est (approximativement) égale au rapport de la surface entre l'isovel  $\xi$  et le lit de canal, et la surface libre du canal. Selon l'Équation (IV.35) et la définition de la fonction de densité de probabilité

$$p(u) = \frac{dP(u)}{du} = \frac{dP(u)}{d\xi} \frac{d\xi}{du} = \left[ (\xi_{max} - \xi_0) \frac{d\xi}{du} \right] \quad (IV.36)$$

L'Équations (IV.35) et (IV.36) indiquent que, une fois que des équations de  $p(u)$  et de  $\xi$  sont déterminées, une équation de distribution de vitesse peut être dérivée pour décrire la structure de la variation de  $u$  avec le  $\xi$  et, par conséquent, les coordonnées  $y$  et  $z$ .

Une approche pour déterminer la fonction de densité de probabilité  $p(u)$  est le principe de maximisation de l'entropie, La méthode dans ce cas actuel est de choisir  $p(u)$  qui maximise l'entropie.

$$H(u) = - \int_0^{u_{max}} p(u) \ln p(u) du \quad (IV.37)$$

Avec

$$\int_0^{u_{max}} p(u) du = 1 \quad (IV.38)$$

$$\text{Et } \int_0^{u_{max}} up(u) du = \bar{u} \quad (IV.39)$$

Dans les Équations (IV.36) et (IV.38),  $u_{max}$  est la vitesse maximale et  $\bar{u}$  est

La vitesse moyenne égale à  $Q/A$ , où  $Q$  est le débit et  $A$  est la section. L'équation (IV.39) est satisfaite la condition que  $u$  doit être réparti sur la section de sorte que  $\bar{u} A=Q$ . la méthode est basée sur le postulat qu'un système dans une condition régulière d'équilibre tend à maximiser l'entropie. Dans la théorie de l'information, l'entropie définie par l'Équation (IV.38) est interprété comme une teneur moyenne en information par message (données d'échantillon) concernant de  $u$  et employée comme mesure d'incertitude, ou la maximisation d'entropie est équivalente à rendre la distribution de probabilité aussi uniforme possible tandis que satisfaire les interprétations additionnelles régnautes de contraintes de l'entropie peut être trouvé dans la littérature appropriée (Goldman1953 ; Leopold et Langbein1962 ; Chiu 1987). L'Application d'un tel concept d'entropie a augmenté récemment en hydrologie (Rao et Hsieh 1987). La méthode de calcul de variations qui peuvent être utiliser pour déterminer le  $p(u)$  dans des équations précédentes est de résoudre l'équation suivante :

$$\frac{\partial}{\partial p} [-p \ln p + \lambda_1 p + \lambda_2 (up)] = 0 \quad (\text{IV.40})$$

$\lambda_1$  et  $\lambda_2$  = multiplicateur de Lagrange. L'équation (IV.40) donne :

$$p(u) = e^{\lambda_1 - 1} e^{\lambda_2 u} \quad (\text{IV.41})$$

Et

$$V = V_f \left\{ 5.75 \left[ \frac{h}{mR} \exp \left( -1 - \frac{\gamma h^2}{4A} \right) \right] + 5.75 \log \frac{mR}{y_0} \right\} \quad (\text{IV.42})$$

En substituant de l'Équation (4.41) dans les équations (IV.37) et (IV.38) on obtient :

$$e^{\lambda_1 - 1} = \lambda_2 (e^{\lambda_2 u_{max}} - 1)^{-1} \quad (\text{IV.43})$$

$$\text{Et } \bar{u} = u_{max} e^{\lambda_2 u_{max}} (e^{\lambda_2 u_{max}} - 1)^{-1} - \frac{1}{\lambda_2} \quad (\text{IV.44})$$

La solution de l'équation différentielle de l'Équation (IV.35) et en utilisant les conditions aux limites ( $u=0$  à  $\xi=\xi_0$ ).

$$u = \frac{1}{\lambda_2} \ln \left( 1 + \frac{\lambda_2}{e^{\lambda_1 - 1}} \frac{\xi - \xi_0}{\xi_{max} - \xi_0} \right) \quad (\text{IV.45})$$

Si  $\bar{u}$  et  $u_{max}$  sont connus (mesurées), les deux paramètres  $\lambda_1$  et  $\lambda_2$  peuvent être déterminés par les Équations (IV.43) et (IV.37) ; si non, les paramètres peuvent être déterminés par la méthode des moindres carrés utilisant les données de distribution de vitesse.

Le terme  $e^{\lambda_1 - 1}$ , dans l'Équation (IV.41), est la valeur de la fonction de densité de probabilité  $p(u)$  à  $u=0$ , et lié à  $\lambda_2$  et  $u_{max}$  par l'Équation (IV.43). Donc on peut écrire l'Équation (4.44). Dans la forme suivante (Chiu et Chen, 1999) :

$$\frac{u}{u_{max}} = \frac{1}{M} \ln \left[ 1 + (e^M - 1) \frac{\xi - \xi_0}{\xi_{max} - \xi_0} \right] \quad (\text{IV.45})$$

En général ( $\xi_0=0$ ) donc on peut écrire l'Équation (IV.45) comme (Chiu et Chen, 1999) :

$$\frac{u}{u_{max}} = \frac{1}{M} \ln \left[ 1 + (e^M - 1) \frac{\xi}{\xi_{max}} \right] \quad (\text{IV.46})$$

Avec  $M=\lambda_2 u_{max}$ , paramètre sans dimension qui, selon Équations (IV.37) et (IV.40) est

$$M = \ln \frac{p(u_{max})}{p(0)} = \ln \frac{\left( \frac{du}{d\xi} \right)_{\xi=\xi_0}}{\left( \frac{du}{d\xi} \right)_{\xi=\xi_0}} \quad (\text{IV.47})$$

$M$  est une mesure de l'uniformité de la probabilité de distribution de vitesse. L'Équation (IV.46) est une équation bidimensionnelle de distribution de vitesse dans les canaux ouverts, dans lequel le point de la vitesse maximale peut produire au-dessous de la surface de l'eau. On a deux paramètres,  $\lambda_2$  et  $u_{max}$ . Dans les canaux large  $(\xi - \xi_0) / (\xi_{max} - \xi_0)$  peut être remplacée par  $y/D$  où  $y$  est la profondeur de l'eau. Du point considéré l'Équation (IV.48)

avec  $(\xi - \xi_0) / (\xi_{max} - \xi_0)$  remplacée par  $y/D$  a été trouvée plus avantageuse que l'équation universelle de la distribution de vitesse de Pradth-Von Karman (1935).

#### IV.3.4 Mesures de l'homogénéité de distribution de vitesse

Chiu et Chen 1999, Dans un isovel (iso vitesse) où  $\xi = \bar{\xi}$  et  $u = \bar{u}$ , l'Équation IV.45 devient

$$\frac{\bar{u}}{u_{max}} = \frac{1}{M} \ln \left[ 1 + (e^M - 1) \frac{\bar{\xi} - \xi_0}{\xi_{max} - \xi_0} \right] \quad (IV.49)$$

D'autre part, en divisant les deux côtés de l'Équation (IV.44) par  $u_{max}$  on trouve

$$\phi = \frac{\bar{u}}{u_{max}} = e^M (e^M - 1)^{-1} - \frac{1}{M} \quad (IV.50)$$

Des Équations (IV.41) et (IV.43), on trouve :

$$\frac{\bar{\xi} - \xi_0}{\xi_{max} - \xi_0} = \frac{\exp[M e^M (e^M - 1)^{-1} - 1] - 1}{e^M - 1} \quad (IV.51)$$

En utilisant les Équations (IV.41), (IV.43), et (IV.48), l'intégration de l'Équation (IV.36) donne l'entropie

$$H(u) = H\left(\frac{u}{u_{max}}\right) + \ln u_{max} \quad (IV.52)$$

Dans le quel  $H(u/u_{max})$  est une fonction d'entropie

$$H(u) = \phi(M) - \ln M \quad (IV.53)$$

Où

$$\phi(M) = 1 + \ln(e^M - 1) - M e^M (e^M - 1)^{-1} \quad (IV.54)$$

La probabilité fonction de densité  $p(u/u_{max})$  peut seulement être exprimée en fonction d'un seul paramètre de l'entropie  $M$

$$p\left(\frac{u}{u_{max}}\right) = u_{max} p(u) = M (e^M - 1)^{-1} \exp\left(\frac{Mu}{u_{max}}\right) \quad (IV.55)$$

En se basant sur le principe l'entropie maximum, les quatre fonctions donnent  $\bar{u}/u_{max}$ ,  $(\bar{\xi} - \xi_0)/(\xi_{max} - \xi_0)$ ,  $H(u/u_{max})$ , et  $p(u/u_{max})$  ces derniers ont un paramètre commun,  $M$ , ce terme décrit donc la forme de la distribution des vitesses dans une section d'écoulement.

$(\bar{\xi} - \xi_0)/(\xi_{max} - \xi_0)$ , et  $H(u/u_{max})$ , sont des mesures de l'homogénéité (ou uniformité) de la distribution spatiale de la vitesse ; et  $p(u/u_{max})$  donne la fonction de densité de probabilité de  $(u/u_{max})$ .

Le long de l'axe de  $y$  où  $z=0$ , (Chiu, 1989).

$$\xi = \frac{y}{D-h} \exp\left(1 - \frac{y}{D-h}\right) \quad (\text{IV.56})$$

Dans le quel  $D$  = profondeur de l'eau à l'axe de  $y$  ;  $y$  = distance verticale du lit de canal ; et  $h$  = paramètre. Si  $u_{\max}$  se produit au-dessous de la surface de l'eau,  $h$  est la profondeur réelle de  $u_{\max}$  au-dessous de la surface.

### IV.3.5 Relation entre la vitesse moyenne et maximale

Selon l'étude de Chiu et Said réalisée en 1995, il existe une relation entre la vitesse moyenne ( $U_m$ ) et la vitesse maximale ( $U_{\max}$ ) dans une section de mesure :

$$U_m = \Phi(M) U_{\max} \quad (\text{IV.57})$$

Où  $\Phi(M)$  est une fonction du paramètre d'entropie  $M$ .

Sur cette base, la méthode de Chiu,  $\Phi(M)$  est définie en maximisant l'entropie, soit  $M$  fonction de la probabilité  $f(U / U_{\max})$  (Chiu et Hsu, 2006) :

$$\phi = \frac{U_m}{U_{\max}} = \frac{e^M}{e^M - 1} - \frac{1}{M} \quad (\text{IV.58})$$

Si  $\phi(M)$  est constant dans la section de mesure, alors  $M$  peut décrire le profil de vitesse (Chiu 1988 ; 1989).

Le paramètre d'entropie  $M$  peut être considéré comme une fonction des paramètres morphologiques des cours d'eau, lesquels peuvent être résumés par la classification de Rosgen (1994).

Le paramètre d'entropie  $M$  peut être considéré comme une fonction des paramètres morphologiques des cours d'eau, résumés par la classification de Rosgen (1994).

Moramarco et al. (2008) ont démontré que lorsque les cours d'eau partagent la même classification de Rosgen (1994), ils peuvent avoir des paramètres d'entropie similaires.

### IV.3.6 Détermination de la surface de la coupe transversale

La surface de la section transversale peut être estimée par la hauteur de jauge, la profondeur sur l'axe  $y$  et le produit du canal largeur multipliée par la profondeur sur l'axe  $y$ . Si le lit du cours d'eau est exempt d'affouillement et de dépôts est stable, il est plus pratique d'utiliser la hauteur de la jauge pour estimer la surface de la section transversale.

L'évaluation de la vitesse moyenne  $U_m$  est la première étape de l'estimation du débit, puis la surface de la section d'écoulement est définie. Afin de déterminer l'aire de la section transversale ( $S$ ), Chiu a proposé différents modèles d'estimation (Chiu et al., 1999) :

$$S = a_1 (G - b_1)^{c_1} \quad (\text{IV.59})$$

$$S = a_2 (H - b_2)^{c_2} \quad (\text{IV.60})$$

$$S = a_3 (HL - b_3)^{c_3} \quad (\text{IV.61})$$

Où  $G$  est le niveau d'eau aux échelles,  $H$  est la profondeur de l'eau,  $L$  est la largeur du chenal à la surface de l'eau,  $a_1$ ,  $a_2$ ,  $a_3$ ,  $b_1$ ,  $b_2$ ,  $b_3$ ,  $c_1$ ,  $c_2$  et  $c_3$  sont des coefficients qui peuvent être déterminés par analyser les données.

### IV.3.7 Expression du débit Chiu

Étant donné que  $U_m$  peut être calculée à partir de  $U_{\max}$  et  $S$  peut être estimée à partir de  $H$  et  $L$ , le débit peut être facilement évalué :

$$Q_{\text{Chiu}} = U_m S = \Phi(M) U_{\max} S \quad (\text{IV.62})$$

La vitesse moyenne ( $U_m$ ) a été déterminée par le rapport entre le débit mesuré et la section transversale du débit.  $U_{\max}$  a été obtenu directement à partir des données de jaugeage ; elle a été fixée égale à la valeur maximale sur toute la surface de la section bien que la vitesse maximale réelle soit inconnue (Chiu 1988).

## IV.4. Conclusion

L'objectif principal de ce chapitre est de présenter la théorie du modèle rugueux de référence et la méthode de Chiu.

La méthode du modèle rugueux de référence ressort que tout profil géométrique de canal possède un modèle rugueux de même forme et dont les caractéristiques sont bien définies et regroupées dans le tableau IV.1. La relation de base du calcul de l'écoulement par la théorie du modèle rugueux de référence est celle de Darcy-Weisbach. Lors de l'application de la formule de Darcy-Weisbach, il en est ressorti que le coefficient de frottement était directement lié au débit volume  $Q$  écoulé par le canal considéré et le débit  $\bar{Q}$  écoulé par le modèle rugueux de référence. Le paramètre fondamental liant ces débits est le paramètre  $\psi_Q$  et dont la signification physique a été mise en évidence. Ce paramètre est régi par une relation implicite [relation (IV.24)] et se présente en fonction des caractéristiques de l'écoulement

dans le modèle rugueux de référence. Lorsque ce paramètre est connu, le débit recherché est immédiatement déduit.

Dans la méthode de Chiu ses équations ont été développées en se basant sur le concept d'entropie pour décrire la répartition bidimensionnelle des vitesses à travers la section d'un canal ouvert. L'équation (IV.49 ou Eq. IV.50) permet de rendre compte des variations de vitesse dans les directions verticale et transversale, en prenant en compte la vitesse maximale qui peut se produire sur ou sous la surface de l'eau.

La définition et l'utilité de  $M$  en tant que nouveau paramètre hydraulique met en évidence l'importance de la vitesse maximale ( $U_{\max}$ ) dans un profil de canal. Cela souligne le besoin d'études expérimentales et analytiques approfondies sur  $U_{\max}$  pour mieux comprendre son influence sur le comportement hydraulique du canal et améliorer la gestion des ressources en eau superficielle. Les modèles de Chiu ont été ainsi utilisés pour estimer rapidement le débit d'écoulement en se basant sur la section d'écoulement et le rapport entre la vitesse moyenne et maximale, contribuant à une bonne maîtrise des écoulements dans les cours d'eau et les canaux.

# **CHAPITRE V**

**MODÈLES D'INTELLIGENCE ARTIFICIELLE**

## V.1. Introduction

L'intelligence artificielle constitue un domaine scientifique qui vise à trouver des méthodes pour résoudre des problèmes complexes sur le plan logique ou algorithmique (Jean, 2020 ; Hutter, 2004).

L'apprentissage automatique est une branche de l'intelligence artificielle qui étudie comment les machines peuvent apprendre par elles-mêmes. L'apprentissage profond, également connu sous le nom de deep learning, deep structured learning ou apprentissage hiérarchique, regroupe un ensemble de techniques d'apprentissage automatique qui cherche à modéliser des données avec un haut niveau d'abstraction en utilisant des architectures composées de transformations non linéaires (Najafabadi et al., 2015 ; Saritha et al., 2019).

Les progrès rapides réalisés dans le domaine de l'apprentissage automatique (ML) sont sans doute le développement actuel le plus pertinent pour le domaine de l'hydrologie. L'utilisation de méthodes de ML est essentielle pour exploiter la quantité croissante de données et pour faire face aux défis du changement climatique et à un impact humain toujours croissant sur l'environnement. Ces dernières années, les chercheurs ont commencé à utiliser des modèles d'intelligence artificielle pour résoudre des problèmes de prédiction dans de nombreux domaines (Jaafari et al., 2019), y compris la recherche hydrologique. Ces techniques d'apprentissage automatique ont été utilisées ainsi pour modéliser le transport sédimentaire dans les cours d'eau naturels (Hafied, et Marouf, 2019 ; Tamrabet et Marouf, 2019, 2022 ; Samantaray et Sahoo ; 2021, Samantaray et al., 2020 ; Afan et al., 2017 ; Rahman et Chakrabarty 2020), la modélisation du ruissellement, (Chang et al., 2014 ; Samantaray et Sahoo., 2020 ; Van et al., 2020), et la prévision des écoulements, (Aoulmi et al., 2012, 2022 et 2023., Sahoo et al., 2019 ; Samantaray et Sahoo A2021 ; Dalkiliç et Hashimi., 2020).

Les méthodes et l'application de l'intelligence artificielle ont récemment montré une grande contribution à la modélisation et à la prédiction des processus hydrologiques, du changement climatique et des systèmes terrestres (Ardabili et al., 2002 ; Huntingford et al., 2019 ; Nourani et al., 2014). Parmi eux, les méthodes d'apprentissage automatique se sont principalement révélées essentielles pour obtenir une précision, une robustesse, une efficacité, un coût de calcul et des performances globales du modèle plus élevés. Dans des nombreux cas, l'apprentissage automatique fournit des meilleures performances de prédiction à court et moyen terme par rapports aux modèles traditionnels (Ahmadlou et al., 2019).

Plusieurs études scientifiques sur les apports liquides et solides dans divers bassins versants au nord de l'Algérie ont approfondi l'analyse du bilan régissant le processus de transport pluie-débit-sédiment, en utilisant différents modèles (FARTAS et al., 2017 ; Marouf et Remini, 2011, 2019 ; Hafied et al. 2019 ; TAMRABET et al., 2019 ; Zeghmar et al., 2022 ; Batout et al. 2022 ; Zeyneb et al., 2022). Dans ce contexte, les travaux d'Aoulmi et al. (2020, 2021 et 2023) sont bien ainsi remarquables, leurs recherches se concentrent sur la modélisation pluie-débit, en intégrant des approches modernes d'optimisation heuristique et tirant parti de l'apprentissage en profondeur via un réseau neuronal convolutif (CNN).

Dans ce chapitre, nous présentons une étude comparative de huit modèles différents basé sur l'intelligence artificielle pour la prédiction des débits des écoulements des cours d'eau naturels situés dans un climat semi-aride et méditerranéen, le cas du nord de l'Algérie.

Les modèles étudiés comprennent les réseaux de neurones artificiels (ANN), système d'inférence neuro floue adaptative (ANFIS), les forêts aléatoires (Random Forest), arbre d'élagage à erreur (Rep Tree), arbre M5p, la régression par processus gaussien (GPR), les machines à vecteurs de support (SVM) et la machine à vecteurs de support libérale (LibSVM). Les performances de ces modèles ont été évaluées à l'aide de différents indicateurs statistiques, tels que le coefficient de corrélation (R), le coefficient d'efficacité (NSE), l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE) et le diagramme de Taylor pour les cinq stations de jaugeage étudiées.

## V.2. Historique d'intelligence artificielle

L'intelligence artificielle (IA) représente la quatrième révolution dans l'histoire de l'humanité (World Economic Forum, 2016). Le terme "intelligence artificielle" a été introduit pour la première fois par John McCarthy en 1956 lors de sa première conférence universitaire. Dans son travail séminaire intitulé "As We May Think" (Bush, 1945), Vannevar Bush a proposé un système visant à amplifier la connaissance et la compréhension individuelles. Cinq ans plus tard, en 1950, Alan Turing écrivit un article sur la possibilité des machines de simuler des êtres humains et d'accomplir des tâches intelligentes (Turing, 1950).

Cependant, la quête visant à déterminer si les machines peuvent véritablement penser a débuté bien avant cette période. L'intelligence artificielle (IA), se concentre sur la conception de programmes informatiques et de machines capables d'effectuer des tâches pour lesquelles les humains sont naturellement bons, notamment la compréhension du langage naturel, la compréhension de la parole et la reconnaissance d'images. Au milieu du XXe

siècle, l'apprentissage automatique est apparu comme un sous-ensemble de l'IA, offrant une nouvelle direction à la conception de l'IA en s'inspirant d'une compréhension conceptuelle du fonctionnement du cerveau humain (Rosenblatt, 1958 ; McCulloch et al., 1943). Aujourd'hui, l'apprentissage automatique reste étroitement lié à la recherche sur l'IA. Cependant, l'apprentissage automatique (ML) est souvent considéré plus largement comme un domaine scientifique qui se concentre sur la conception de modèles informatiques et d'algorithmes capables d'effectuer des tâches spécifiques, impliquant souvent la reconnaissance de formes, sans qu'il soit nécessaire de les programmer explicitement.

### V.2.1 Apprentissage automatique (Machine Learning)

ML est une branche de l'IA qui développe des algorithmes d'auto-apprentissage et d'auto-améliorations. Le terme "Machine Learning" a été utilisé pour la première fois par Arthur Samuel en 1959 pour décrire des algorithmes offrant "aux ordinateurs la capacité d'apprendre sans être explicitement programmés" (Turing, 1950 ; Rosenblatt, 1958).

L'apprentissage automatique traite de la question de savoir comment créer des programmes informatiques qui améliorent leur performance à certaines tâches grâce à l'expérience (Mitchell, 1997).

Les algorithmes d'apprentissage automatique ont été utilisés dans (Osisanwo et al., 2017 ; Mitchell, 1997 ; Latif et al., 2019) :

- 1- Des problèmes d'exploration de données où de grandes bases de données peuvent contenir des régularités implicites qui peuvent être découvertes automatiquement ;
- 2- Des domaines mal compris où les humains pourraient ne pas avoir les connaissances nécessaires pour développer des algorithmes efficaces ;
- 3- Des domaines où les programmes doivent s'adapter dynamiquement aux conditions changeantes.

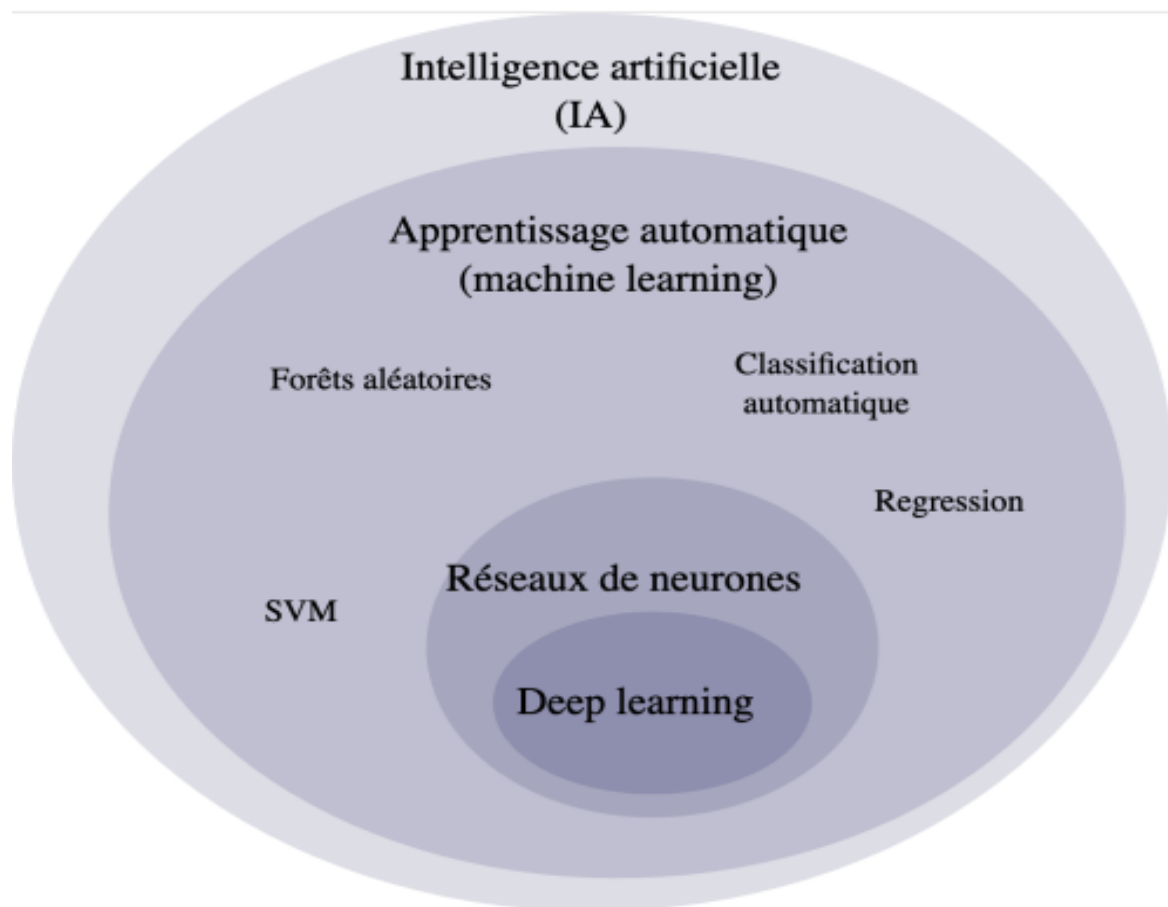
L'apprentissage d'une fonction cible à partir de données d'entraînement pose de nombreux problèmes (fonction représentation, comment et quand générer la fonction, avec quelle entrée donnée, comment évaluer les performances de la fonction générée, etc.).

Les principaux types d'apprentissage comprennent : l'apprentissage de concepts (CL), les arbres de décision (DT), réseaux de neurones (ANN), réseaux de croyances bayésiennes (BBN), apprentissage par renforcement (RL), algorithmes génétiques (GA) et programmation génétique (GP), apprentissage par instance (IBL), La programmation logique inductive (ILP)

et l'apprentissage analytique (AL), Support Vector Machine SVM (Zhang, 2000 ; Blum et al., 1997 ; Brazdil et al., 2003).

### V.2.2 Apprentissage profond (Deep Learning)

(DL) est une classe de techniques d'apprentissage automatique de pointe qui a suscité un énorme intérêt mondial au cours des dernières années (LeCun et al., 2015). DL utilise des méthodes d'apprentissage par représentation avec plusieurs niveaux d'abstraction pour traiter les données d'entrée sans avoir besoin d'une intervention manuelle. L'ingénierie des caractéristiques, reconnaissant automatiquement les structures complexes des données de grande dimension par projection sur une variété de dimension inférieure. Par rapport aux techniques conventionnelles, il a été démontré que DL atteint des précisions nettement plus élevées dans de nombreux domaines, En tant que composant fondamental de l'intelligence artificielle (IA) moderne, DL a fait des progrès substantiels ces dernières années et a aidé à résoudre des problèmes qui ont résisté à l'IA pendant des décennies (LeCun et al., 2015).



**Figure.V.1.** Relation entre l'intelligence artificielle, l'apprentissage automatique, les réseaux de neurones et l'apprentissage profond (Bunod et al., 2022).

### V.3. Contexte de la théorie des modèles

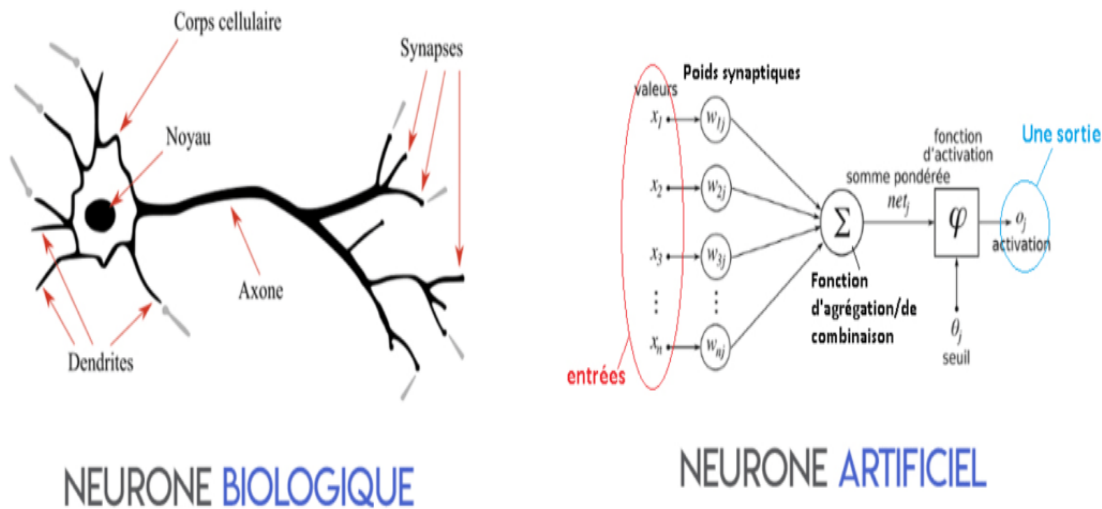
#### V.3.1 Réseaux de neurones artificiels (Artificial neural networks)

Les réseaux de neurones artificiels (ANN) sont devenus un modèle populaire et utile pour la classification, le regroupement, la reconnaissance de formes et la prédiction dans de nombreuses disciplines. Les RNA sont un type de modèle pour l'apprentissage automatique (ML) et sont devenus relativement compétitifs par rapport aux modèles de régression et statistiques conventionnels en ce qui concerne l'utilité (Dave et Dutta, 2014 ; Basu et al., 2010).

Les réseaux de neurones artificiels (RNA) sont des modèles mathématiques qui cherchent à imiter la structure et les fonctions des réseaux de neurones biologiques. Leur fondement central réside dans le neurone artificiel, qui est en réalité un modèle mathématique simple sous forme de fonction. Ce formulaire contient trois ensembles simples de règles : multiplication, addition et activation. A l'entrée d'un neurone artificiel, les entrées sont pondérées ce qui signifie que chaque valeur d'entrée est multipliée par le poids individuel. Dans la section médiane du neurone artificiel, il y a une fonction de sommation qui collecte toutes les entrées pondérées et biaisées. A la sortie du neurone artificiel, la somme des entrées préalablement pondérées et biaisées passe par la fonction d'activation du creux aussi appelée fonction de transfert (Krenker et al., 2011 ; Agatonovic et Beresford, 2000).

##### V.3.1.1. Neurone artificiel

Le neurone artificiel est un élément de base de tout réseau de neurones artificiels. Sa conception et ses fonctionnalités sont dérivées de l'observation d'un neurone biologique qui est la pierre angulaire des réseaux (systèmes) de neurones biologiques qui comprennent le cerveau, la moelle épinière et les ganglions périphériques. Des similitudes dans la conception et les fonctionnalités peuvent être observées dans la Figure V.2. Où le côté gauche d'une figure représente un neurone biologique avec son soma, ses dendrites et son axone et où le côté droit d'une figure représente un neurone artificiel avec ses entrées, poids, fonction de transfert, biais et sorties (Abraham, 2005 ; Messalti, 2015).



**Figure.V. 2.** Conception de neurones biologiques et artificiels (Akgün et Demir, 2018).

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,  $f$  est la fonction d'activation.

Dans le cas où des informations sur les neurones biologiques pénètrent dans le neurone via la dendrite, le soma traite les informations et les transmet via l'axone. Dans le cas d'un neurone artificiel, l'information entre dans le corps d'un neurone artificiel via des entrées pondérées (chaque entrée peut être multipliée individuellement par un poids). Le corps d'un neurone artificiel additionne alors les entrées pondérées, biaise et « traite » la somme avec une fonction de transfert. À la fin, un neurone artificiel transmet les informations traitées via une ou plusieurs sorties. L'avantage de la simplicité du modèle de neurone artificiel peut être vu dans sa description mathématique ci-dessous (Krenker et al., 2011) :

$$y(k) = f(\sum_{i=0}^m w_i(k) \cdot x_i(k) + b) \tag{V.1}$$

Où :

$x_i(k)$  Est la valeur d'entrée en temps discret  $k$  où  $i$  va de 0 à  $m$ ,

$w_i(k)$  Est la valeur de poids en temps discret  $k$  où  $i$  va de 0 à  $m$ ,

$b$  est biaisé,

$f$  est une fonction de transfert,

$y(k)$  Est la valeur de sortie en temps discret  $k$ .

Comme le montre le modèle d'un neurone artificiel et son équation (V.1), la principale variable inconnue de notre modèle est sa fonction de transfert. La fonction de transfert définit les propriétés du neurone artificiel et peut être n'importe quelle fonction mathématique. Nous la choisissons en fonction du problème que le neurone artificiel (réseau neuronal artificiel)

doit résoudre et, dans la plupart des cas, nous la choisissons parmi les fonctions suivantes : Fonction en escalier, fonction linéaire et fonction non linéaire (sigmoïde).

La fonction d'étape est une fonction binaire qui n'a que deux valeurs de sortie possibles (par exemple, zéro et un). Cela signifie que si la valeur d'entrée atteint un seuil spécifique, la valeur de sortie est égale à une valeur et que si le seuil spécifique n'est pas atteint, la valeur de sortie est différente. La situation peut être décrite par l'équation (V.2).

$$y = \begin{cases} 1 & \text{if } w_i x_i \geq \text{threshold} \\ 0 & \text{if } w_i x_i < \text{threshold} \end{cases} \quad (\text{V.2})$$

Lorsque ce type de fonction de transfert est utilisé dans un neurone artificiel, on parle de perceptron. Le perceptron est utilisé pour résoudre les problèmes de classification et, à ce titre, on le trouve le plus souvent dans la dernière couche des réseaux neuronaux artificiels. Dans le cas d'une fonction de transfert linéaire, le neurone artificiel effectue une simple transformation linéaire sur la somme des entrées pondérées et du biais. Un tel neurone artificiel contraste avec le perceptron, le plus souvent utilisé dans la couche d'entrée des réseaux neuronaux artificiels. Lorsque nous utilisons une fonction non linéaire, la fonction sigmoïde est la plus couramment utilisée. La fonction sigmoïde a une dérivée facile à calculer, ce qui peut être important lors du calcul des mises à jour des poids dans le réseau neuronal artificiel (Krenker et al., 2011).

Le perceptron multicouche (MLP) est le type de réseau neuronal le plus couramment utilisé. Il est composé de plusieurs couches interconnectées : une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. L'apprentissage du MLP est supervisé et se fait en minimisant une fonction de coût qui calcule la somme des erreurs au carré entre les valeurs observées et simulées du réseau. Pour cela, l'algorithme de rétropropagation est utilisé pour mesurer la différence entre les valeurs observées et simulées du réseau et propager cette différence à travers le réseau (Kim et al., 2010 ; Rezaeianzadeh et al., 2014).

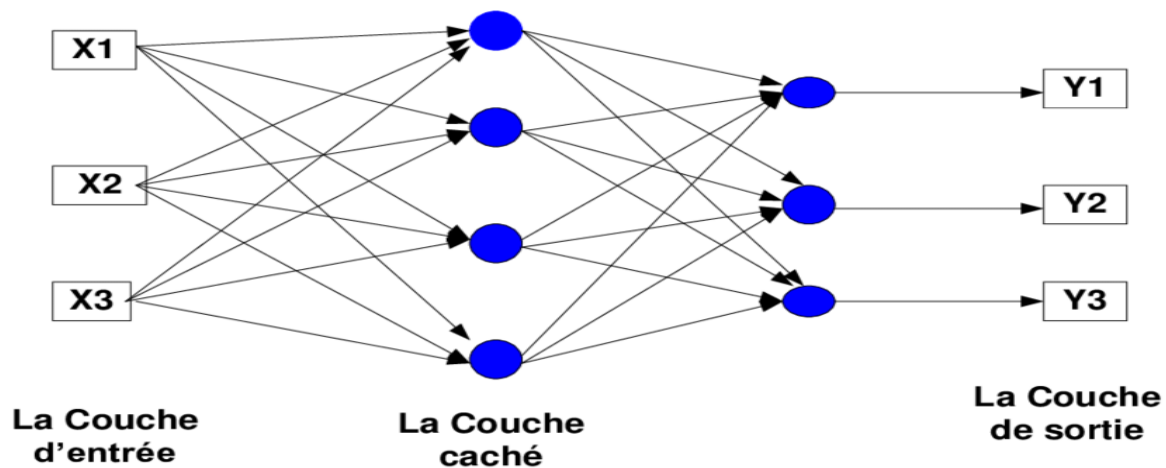


Figure.V. 3. Structure d'un MLP (Abroudi et al., 2013).

### V.3.2 Système d'inférence neuro floue adaptative (Adaptive neural-based fuzzy inference system)

La technique Adaptive Neuro-Fuzzy Inference System (ANFIS) a été initialement présentée par Jang en (1993). ANFIS est une technique d'apprentissage de données simple qui utilise la logique floue pour transformer des entrées données en une sortie souhaitée via des éléments de traitement de réseau de neurones hautement interconnectés et des connexions d'informations, qui sont pondérées pour mapper les entrées numériques en une sortie (Rajasekaran et Pai, 2003 ; Walia et al., 2015)

ANFIS combine les avantages des deux techniques d'apprentissage automatique (Fuzzy Logic et Neural Network) en une seule technique (Jang, 1993). Un ANFIS fonctionne en appliquant des méthodes d'apprentissage de réseau de neurones pour régler les paramètres d'un système d'inférence floue (FIS). Plusieurs fonctionnalités permettent à ANFIS d'obtenir un grand succès (Jang et Sun, 1995 ; Jang et al., 1997).

Il affine les règles floues IF-THEN pour décrire le comportement d'un système complexe ;

- Il ne nécessite pas d'expertise humaine préalable ;
- Il est facile à mettre en œuvre ;
- Il permet un apprentissage rapide et précis ;
- Il offre un ensemble de données souhaité ; un plus grand choix de fonctions d'appartenance à utiliser ; de fortes capacités de généralisation ; D'excellentes possibilités d'explication grâce aux règles floues ;
- Il est facile d'incorporer des connaissances à la fois linguistiques et numériques linguistiques et numériques pour la résolution de problèmes.

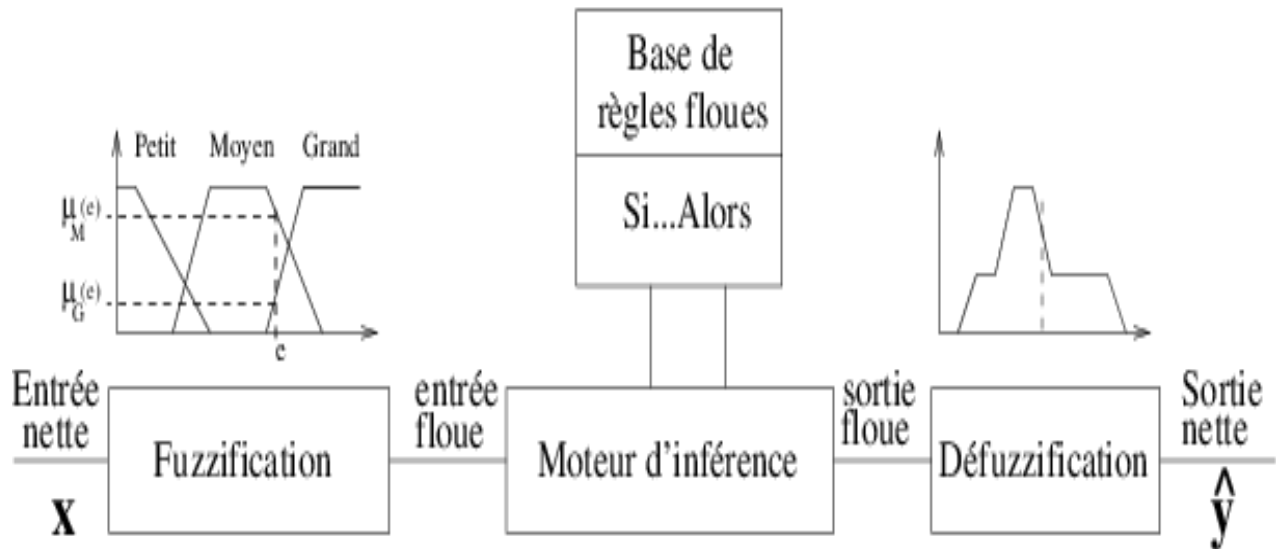


Figure. V. 4. Structure d'un système d'inférence floue (FIS) (Voiron et al., 2013).

### V.3.2.1. Architecture de l'ANFIS

Différentes règles ne peuvent pas partager la même fonction d'appartenance de sortie. Le nombre de fonctions d'appartenance doit être égal au nombre de règles. Pour présenter l'architecture ANFIS, deux règles floues IF-THEN basées sur un modèle de Sugeno du premier ordre sont considérées (Hmouz et al., 2011) :

Rule (1): **IF** x est  $A_1$  **AND** y est  $B_1$ , **THEN**

$$f_1 = p_1x + q_1y + r_1. \tag{V.3}$$

Rule (2): **IF** x est  $A_2$  **AND** y est  $B_2$ , **THEN**

$$f_2 = p_2x + q_2y + r_2. \tag{V.4}$$

Où :

- $x$  et  $y$  sont les entrées,
- $A_i$  et  $B_i$  sont les ensembles flous,
- $f_i$  sont les sorties à l'intérieur de la région floue spécifiée par la règle floue,
- $p_i$ ,  $q_i$  et  $r_i$  sont les paramètres de conception qui sont déterminés au cours du processus de formation.

La figure V.5. Illustre le mécanisme de raisonnement de ce modèle de Sugeno, qui est à la base du modèle ANFIS.

L'architecture ANFIS utilisée pour mettre en œuvre ces deux règles est représentée sur la figure V.5. Sur cette figure, un cercle indique un nœud fixe, tandis qu'un carré indique un nœud adaptatif. ANFIS a une architecture à cinq couches. Chaque couche est expliquée en détail ci-dessous (Hmouz et al., 2011 ; Vieira et al., 2004 ; Rules, 1993).

Dans couche (1), tous les nœuds sont des nœuds adaptatifs.

Les sorties de couche (2) sont le degré d'appartenance floue des entrées, qui sont donnés par les équations suivantes :

$$O_{1,i} = \mu_{A_i}(x), i = 1,2, \tag{V.5}$$

$$O_{1,i} = \mu_{B_i} - 2(y), i = 3,4, \tag{V.6}$$

Où  $x$  et  $y$  sont les entrées du nœud  $i$ , et  $A_i$  et  $B_i$  sont les étiquettes linguistiques (haut, bas, etc.) associées à cette fonction de nœud.  $\mu_{A_i}(x)$  et  $\mu_{B_i} - 2(y)$  peuvent adopter n'importe quelle fonction d'appartenance floue.

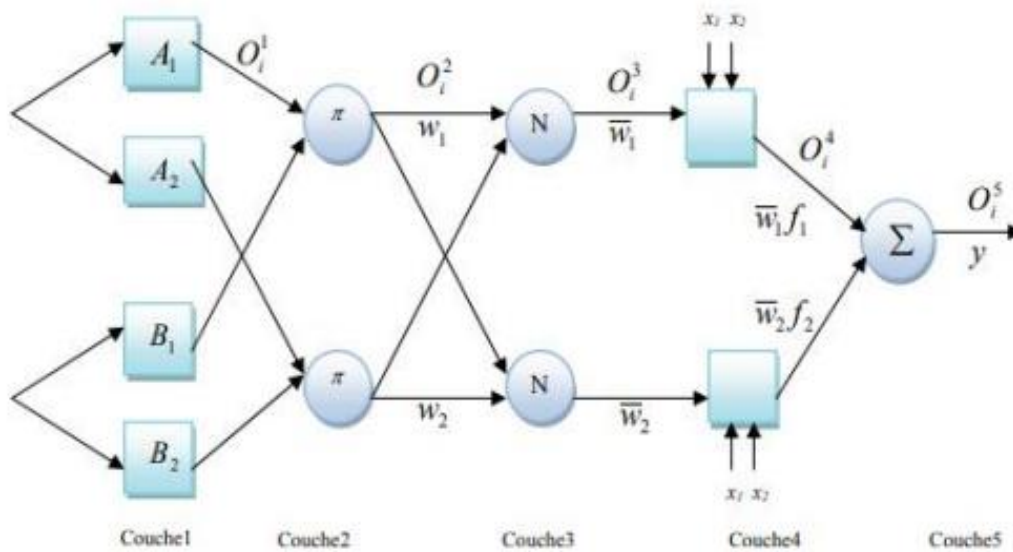


Figure.V. 5. Architecture équivalente d'ANFIS pour deux règles (Belhachat et al., 2007).

### V.3.3 Machine à vecteurs de support (Support vector machine)

La machine à vecteurs de support (SVM) a été entendue pour la première fois en 1992, présentée par Boser, Guyon et Vapnik lors de la conférence COLT-92. Les machines à vecteurs de support (SVM) sont un ensemble de méthodes d'apprentissage supervisé utilisées pour la classification et la régression (Boser et al., 1992). Ils appartiennent à une famille de classificateurs linéaires généralisés. En d'autres termes, la machine à vecteur de support (SVM) est un outil de classification et de prédiction de la régression qui utilise la théorie de l'apprentissage automatique pour maximiser la précision de la prédiction tout en évitant automatiquement un surajustement aux données. Les machines à vecteurs de support peuvent être définies comme des systèmes qui utilisent l'espace d'hypothèse d'une fonction linéaire dans un espace de caractéristiques à haute dimension, formés avec un algorithme d'apprentissage issu de la théorie de l'optimisation et qui met en œuvre un biais d'apprentissage dérivé de la théorie de l'apprentissage statistique. Le SVM est devenu célèbre lorsque, en utilisant des cartes de pixels en entrée, il a donné une précision comparable à celle

de réseaux neuronaux sophistiqués avec des caractéristiques élaborées dans une tâche de reconnaissance de l'écriture manuscrite (Jakkula, 2006). Les fondements des machines à vecteurs de support (SVM) ont été développés par Vapnik (Vapnik, 1995), et ont gagné en popularité grâce à de nombreuses caractéristiques prometteuses, telles que de meilleures performances empiriques. La formulation utilise le principe de minimisation structurelle du risque (SRM), qui s'est avéré supérieur (Burges, 1998). Au principe traditionnel de minimisation empirique du risque (ERM), utilisé par les réseaux neuronaux conventionnels. La MRS minimise une limite supérieure du risque attendu, alors que la MRE minimise l'erreur sur les données d'apprentissage. C'est cette différence qui confère aux SVM une plus grande capacité de généralisation, ce qui est l'objectif de l'apprentissage statistique. Les SVM ont été développés pour résoudre les problèmes de classification, mais ils ont récemment été étendus pour résoudre les problèmes de régression (Vapnik et al., 1996).

### V.3.3.1. SVM (principe de fonctionnement général)

SVM est une méthode puissante pour construire un classifieur. Il vise à créer une frontière de décision entre deux classes qui permet la prédiction d'étiquettes à partir d'un ou plusieurs vecteurs de caractéristiques (Noble, 2006). Cette frontière de décision, appelée hyperplan, est orientée de telle sorte qu'elle soit aussi éloignée que possible des points de données les plus proches de chacune des classes. Ces points les plus proches sont appelés vecteurs supports. Étant donné un ensemble de données d'entraînement étiqueté (Bennett et al., 2000 ; Huang et al., 2018) :

$$(x_1, y_1), \dots, (x_n, y_n), x_i \in R^d \text{ and } y_i \in (-1, +1)$$

Où  $x_i$  est une représentation vectorielle de caractéristiques et  $y_i$  l'étiquette de classe (négative ou positive) d'un composé d'apprentissage  $i$ .

L'hyperplan optimal peut alors être défini comme :

$$wx^T + b = 0 \tag{V.7}$$

Où  $w$  est le vecteur de poids,  $x$  est le vecteur de caractéristiques d'entrée et  $b$  est le biais.

Le  $w$  et  $b$  satisferaient les inégalités suivantes pour tous les éléments de l'ensemble d'apprentissage :

$$wx_i^T + b \geq +1 \text{ if } y_i = 1 \tag{V.8}$$

$$wx_i^T + b \geq -1 \text{ if } y_i = -1 \tag{V.9}$$

L'objectif de la formation d'un modèle SVM est de trouver le  $w$  et  $b$  de sorte que l'hyperplan sépare les données et maximise la marge  $1 / \|w\|_2$ .

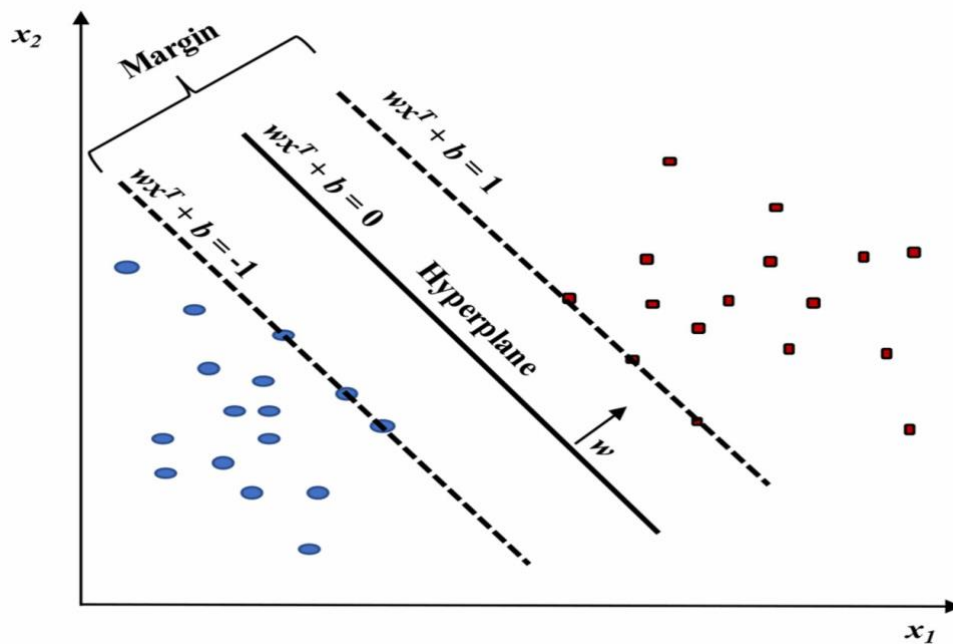
Vecteurs  $x_i$  pour lesquels  $|y_j e| (w x_i T + b) = 1$  sera appelé vecteur support (figure V.6).

L'algorithme SVM a été initialement proposé pour construire un classificateur linéaire en 1963 par Vapnik (Vapnik, 1963). Une autre utilisation de SVM est la méthode du noyau, qui nous permet de modéliser des modèles non linéaires de plus grande dimension (Aizerman., 1964). Dans un problème non linéaire, une fonction noyau pourrait être utilisée pour ajouter des dimensions supplémentaires aux données brutes et ainsi en faire un problème linéaire dans l'espace dimensionnel supérieur résultant (figure V.6). En bref, une fonction du noyau pourrait aider à effectuer certains calculs plus rapidement, ce qui nécessiterait autrement des calculs dans un espace de grande dimension.

Il est défini comme :

$$K(x, y) = \langle f(x), f(y) \rangle \quad (\text{V.10})$$

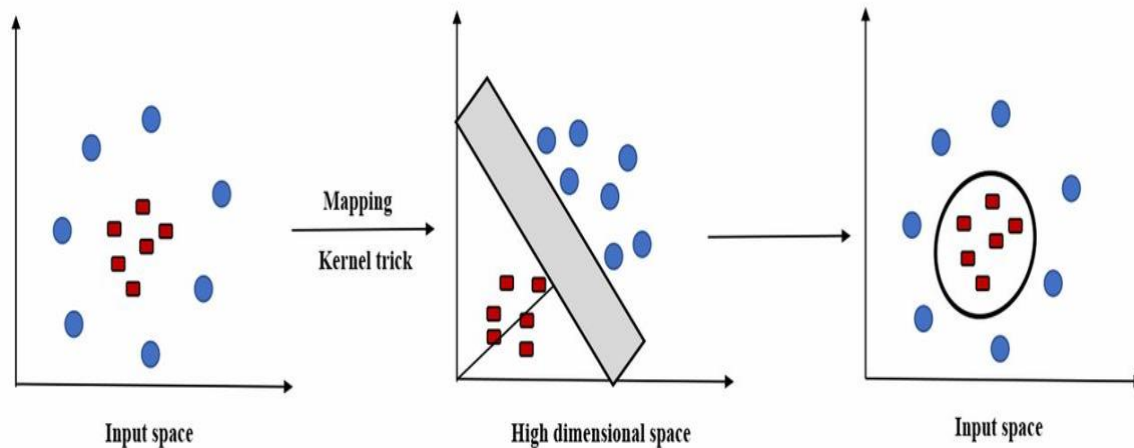
Ici  $K$  est la fonction noyau,  $x, y$  sont des entrées à  $n$  dimensions.  $f$  est utilisé pour mapper l'entrée de l'espace  $n$  dimensionnel à l'espace  $m$  dimensionnel.  $\langle x, y \rangle$  désigne le produit scalaire. Avec les fonctions du noyau, nous pourrions calculer le produit scalaire entre deux points de données dans un espace de dimension supérieure sans calculer explicitement le mappage de l'espace d'entrée à l'espace de dimension supérieure. Dans de nombreux cas, le calcul du noyau est facile alors qu'il est difficile d'aller dans l'espace de grande dimension pour calculer le produit interne de deux vecteurs caractéristiques. Le vecteur de caractéristiques (comprenant la taille, le nombre de cycles, la direction, etc.) peut rapidement augmenter en dimension, et cette augmentation est encore plus marquée pour des noyaux tels que le noyau de la fonction de base radiale (RBF) ( $K_{\text{RBF}}(x, y) = \exp(-\gamma \|x - y\|^2)$ ). Le vecteur caractéristique correspondant à une dimension infinie. Pourtant, le calcul du noyau reste presque trivial.



**Figure.V. 6.** Représentation géométrique de la marge du SVM (Huang et al., 2018).

Il existe deux types de cas pour les modèles SVM : les cas linéairement séparables et les cas non linéairement séparables. Les cas linéairement séparables sont les plus simples pour les SVM, car ils permettent de trouver facilement un classificateur linéaire. Cependant, dans la plupart des problèmes réels, il n'est pas possible de séparer linéairement les données. Par conséquent, l'utilisation du classificateur à marge maximale est limitée, car il fonctionne uniquement lorsque les classes d'apprentissage sont linéairement séparables (Golub et al., 1999).

**Noyau :** Si les données sont linéaires, un hyperplan séparateur peut être utilisé pour diviser les données. Cependant, il arrive souvent que les données soient loin d'être linéaires et que les ensembles de données soient inséparables. Dans ce cas, des noyaux sont utilisés pour cartographier de manière non linéaire les données d'entrée dans un espace à haute dimension. La nouvelle cartographie est alors linéairement séparable (Jakkula, 2006). Une illustration très simple est présentée ci-dessous dans la figure V.7 (Jakkula, 2006 ; Mitchell, 1997 ; Vapnik, 2006).



**Figure.V. 7.** Fonction noyau. Les données qui ne peuvent pas être séparées par SVM linéaire peuvent être transformées et séparées par une fonction noyau (Jakkula, 2006).

### V.3.4 Bibliothèque pour les machines à vecteurs de support (Library for Support Vector Machines)

La bibliothèque SVM (LibSVM) est un algorithme d'apprentissage automatique destiné à la classification de vecteurs de support et à la régression (Hsu et al., 2003). Elle a été développée en tant que progiciel intégré pour les tâches d'apprentissage automatique, telles que la classification et la régression SVM, depuis l'an 2000. Le but de LibSVM est de permettre aux utilisateurs d'explorer et d'expérimenter avec SVM pour leurs besoins spécifiques (Chang & Lin, 2011). Elle est efficace pour résoudre des problèmes de classification multi-output / multi-classes. Elle offre différentes formulations SVM, divers types d'approches de validation croisée pour la sélection du modèle, des estimations de probabilité et une interface utilisateur graphique simple pour démontrer la classification et la régression SVM. LibSVM prend en charge différents types de SVM, tels que SVM à une classe, SVM de régression et SVM NU.

- LibSVM prend en charge trois fonctions, à savoir

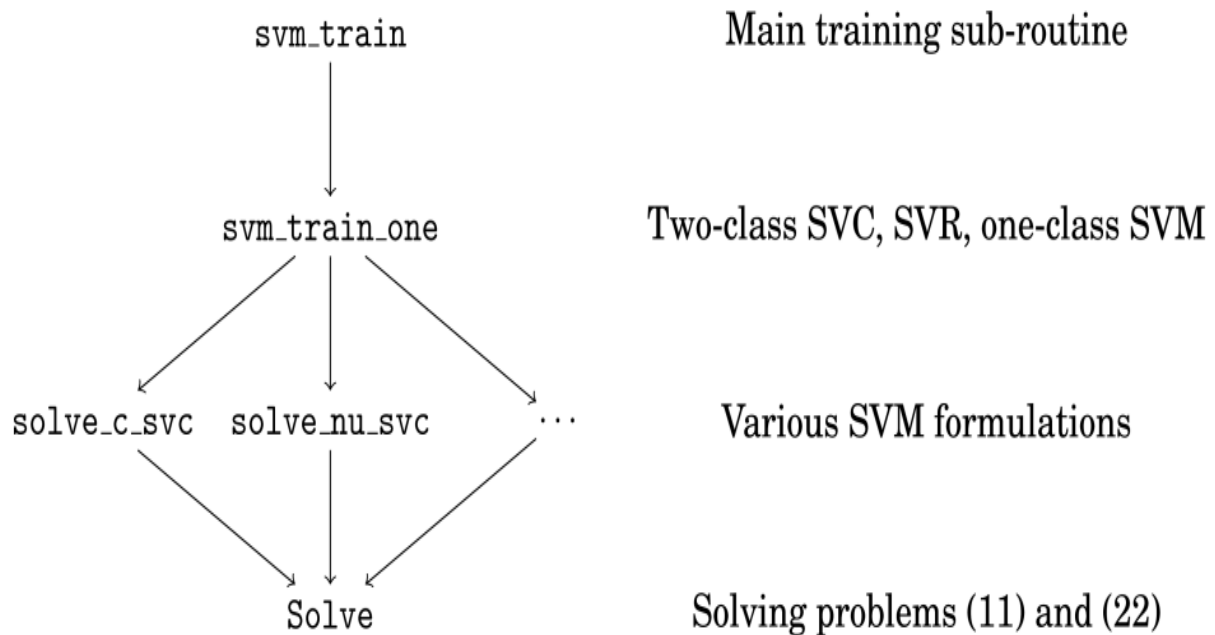
1- SVC (Support Vector Classification - binary-class et multi-class), qui peut être utilisée pour les tâches de classification. Multi-classes), qui peut être utilisée pour les tâches de classification ;

2- SVR (Support Vector Regression), utilisée pour les tâches de régression ;

3- SVM à une classe, utilisée pour l'estimation de la distribution.

Dans la figure V.8, nous pouvons voir l'organisation de la bibliothèque dans LibSVM pour le processus d'apprentissage. SVM\_train est une routine principale utilisée pour effectuer

l'apprentissage SVM. `SVM_train_one` qui se trouve en dessous est une routine pour sélectionner l'une des trois fonctions LibSVM (SVC, SVR et une classe SVM). Sous `SVM_train_one`, il existe différents types de fonctions SVM qui peuvent être utilisées en fonction du choix de `SVM_train_one`. Ces options produisent un modèle de résolution pour les données qui ont été formées précédemment (Abdiansah et Wardoyo, 2015 ; Horvath et al., 2014).



**Figure.V. 8.** Organisation de la bibliothèque dans LibSVM (Abdiansah et Wardoyo, 2015).

### V.3.5 Régression de processus gaussien (gaussian process regression)

GPR présente une approche d'apprentissage supervisé probabiliste et non paramétrique pour généraliser le mappage de fonctions non linéaires et complexes cachées dans des ensembles de données. Cette approche a récemment fait l'objet d'une grande attention de la part des chercheurs de diverses disciplines d'étude (Omran et al., 2016, Cheng et al., 2013). GPR est très efficace pour gérer les données non linéaires grâce à l'utilisation des fonctions du noyau. De plus, un avantage significatif du GPR est que le modèle peut fournir une réponse de fiabilité à une donnée d'entrée (Pal et al., 2010 ; Nguyen et al., 2009).

Étant donné un ensemble d'apprentissage, les données d'entrée sont appelées la matrice de conception et sont le vecteur de la sortie souhaitée. L'hypothèse principale de GPR est que la sortie est calculée comme (Rasmussen et Williams, 2006 ; Ebden, 2015).

$$D = \{(x_i, y_i) \mid i = 1, \dots, n\} X \in R^{D \times n} y \in R^n \quad (\text{V.11})$$

$$y = f(x) + \varepsilon, \tag{V.12}$$

Où représente un bruit homoscedastique pour tout échantillon.  $\varepsilon \sim N(0, \sigma^2) \in R x_i$

Dans la méthodologie GPR, les observations dans l'ensemble de données d'intérêt sont considérées comme un seul point échantillonné à partir d'une distribution gaussienne multivariée. De plus, on peut supposer que cette distribution gaussienne a la moyenne des zéros. La fonction de covariance dicte la relation d'une observation à une autre observation. La fonction de covariance exponentielle au carré est souvent sélectionnée dans GPR pour la tâche d'approximation de la fonction (Cheng et al., 2013, Ebden, 2015) :

$$y = \{y_1, \dots, y_n\} \quad K(x, x') \tag{V.13}$$

$$K(x, x') = \sigma_f^2 \times \exp\left(\frac{(x-x')^2}{2l^2}\right) + \sigma_n^2 \delta(x, x') \tag{V.14}$$

Où la covariance maximale admissible est définie comme. On note que n'atteint cette covariance maximale admissible que lorsqu'il est très proche de et donc presque parfaitement corrélé avec. Pendant ce temps, désigne le paramètre de longueur de la fonction noyau. De plus, représente une fonction delta de Kronecker ; si et si.

$$\sigma_f^2 k(x, x') x x' f(x) f(x') l \delta(x, x') \delta_{ij} = l i = j \delta_{ij} = 0 i \neq j \tag{V.15}$$

Compte tenu de l'ensemble de données d'apprentissage, le but ultime du processus d'apprentissage est de prédire la valeur de sortie d'un nouveau modèle d'entrée interrogé. Pour atteindre cet objectif, il est nécessaire d'établir trois matrices de covariance comme suit :

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & K(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & K(x_2, x_n) \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & K(x_n, x_n) \end{bmatrix} \tag{V.16}$$

$$k_* = [k(x_*, x_1) \quad k(x_*, x_2) \quad \dots \quad K(x_n, x_n)] \tag{V.17}$$

$$K_{**} = k(x_*, x_*) \tag{V.18}$$

En raison de l'hypothèse que les données sont échantillonnées à partir d'une distribution gaussienne multivariée, nous avons l'expression suivante :

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix}\right) \tag{V.19}$$

Puisqu'il  $y_*|y$  est prouvé qu'il est généré à partir d'une distribution gaussienne multivariée avec la moyenne de  $k_*k^{-1}y$  et la variance de  $k_{**} - k_*k^{-1}k_*^T$ , la moyenne et la variance estimées de la sortie prédite  $y_*$  sont données comme suit :

$$E(y_*) = k_*k^{-1}y, \quad (\text{V.20})$$

$$\text{var}(y_*) = k_{**} - k_*^{-1}k_*^T \quad (\text{V.21})$$

Lorsque les hyperparamètres de la fonction noyau sont spécifiés, les paramètres du modèle, y compris  $k$  et, peuvent être déterminés par inférence bayésienne. Cette tâche revient à maximiser une fonction log-postérieure (Ebden, 2015). Une fois la phase d'apprentissage terminée, le modèle GPR peut être utilisé pour prédire un échantillon de données d'entrée inconnu (Lázaro et al., 2013 ; Hoang et al., 2016)

### V.3.6 Forêt aléatoire (Random Forest)

La forêt aléatoire (RF) est une méthode d'apprentissage automatique développée par Breiman (Breiman, 2001); Elle a été largement appliquée dans différents domaines en raison de sa bonne stabilité et de sa généralisation (Were et al., 2015 ). La RF est une combinaison d'arbres de décision simples, dans lesquels chaque arbre est généré en sélectionnant des échantillons aléatoires et des caractéristiques de tous les prédicteurs (Oliveira et al., 2012, Naghibi et al. 2017). La majorité des résultats collectés à partir des arbres de décision sont considérés comme le résultat final du RF (Al-Abadi et Shahid, 2016). Cet algorithme est une extension du bagging, et un concurrent du boosting (Polikar, 2012). Les arbres de régression n'impliquent aucune hypothèse de distribution des données (Francke et al. 2008). Pour chaque arbre, les données de réponse ont été regroupées en deux nœuds descendants pour maximiser l'homogénéité, et la meilleure division binaire a été sélectionnée. Chaque nœud descendant de la scission sélectionnée a été traité de la même manière que le nœud d'origine, et le processus s'est poursuivi de manière récursive jusqu'à ce qu'un critère d'arrêt soit satisfait. Tous les arbres ont atteint leur taille maximale et les prédictions finales ont été obtenues à partir des résultats moyens (Breiman 2001).

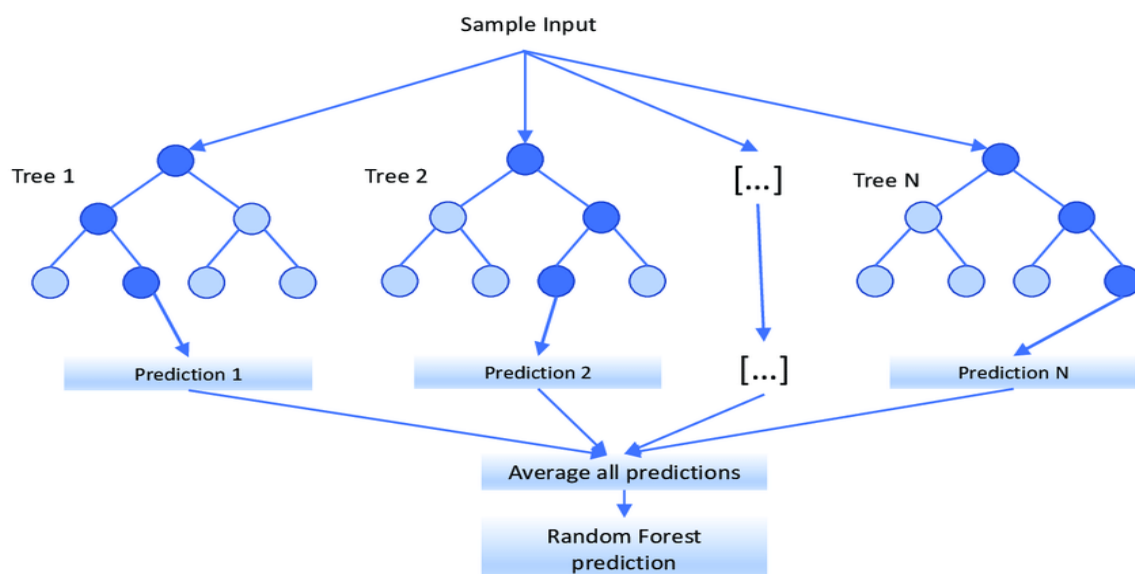
En règle générale, une partie distincte des échantillons doit être omise de l'ensemble de données d'origine, et ils sont appelés les échantillons hors sac (OOB). L'erreur quadratique moyenne (MSE) des échantillons OOB (c'est-à-dire la somme des résidus au carré des échantillons OOB divisée par la taille de l'échantillon) est souvent utilisée pour évaluer les performances de la méthode RF. Le sous-ensemble de formation et le sous-ensemble de test

(c'est-à-dire les échantillons OOB) sont cohérents pour différents modèles, comme mentionné précédemment, afin de comparer les performances des ELM et des RF. Deux paramètres importants doivent être déterminés au cours du processus de calibrage du modèle RF : le nombre d'arbres inclus dans la forêt ( $n$  arbre) et le nombre de prédicteurs testés à chaque nœud ( $m_{try}$ ). Pour obtenir l'erreur de généralisation minimale et la corrélation entre les arbres de décision, la valeur de  $m_{try}$  a été calculée selon des formules empiriques présentées dans des études antérieures (Huang et al., 2016 ; Al-Abadi et Shahid, 2016), comme suit :

$$\begin{cases} m_{try} = \log_2(M + 1) \\ m_{try} = \sqrt{M} \\ m_{try} = \frac{M}{3} \end{cases} \quad (V.22)$$

Où  $M$  est le nombre de variables d'entrée définies dans l'ensemble de données d'origine. La valeur de  $n$  arbre affecterait de manière significative les résultats des prévisions ; par conséquent, une expérience a été conçue pour choisir l'arbre  $n$  le plus raisonnable. Partant d'une valeur initiale de  $n$  arbre = 1, le modèle RF est entraîné et testé avec un nombre croissant d'arbres jusqu'à ce que  $n$  arbre = 500 soit atteint (avec une taille de pas de 1).

De plus, l'importance de chaque prédicteur est mesurée par l'augmentation des erreurs quadratiques moyennes (MSE) car les prédicteurs ont été exclus un par un des modèles RF. L'importance relative de chaque prédicteur est déterminée à partir de 100 exécutions des modèles RF et normalisée à 100 % pour fournir une base simple de comparaison dans différentes stations.



**Figure.V. 9.** Démonstration de la méthodologie RF ( Malekipirbazari & Aksakalli 2015 ).

L'apprentissage par arbre de décision soulève deux questions essentielles : (1) comment choisir l'attribut de division à chaque nœud interne et (2) combien de niveaux il faut avoir à chaque branche de l'arbre, c'est-à-dire quand arrêter la division. Dans le contexte des forêts aléatoires, qui sont des collections d'arbres de décision, la division est effectuée en fonction de l'indice de Gini, qui est décrit ci-dessous, et le nombre de niveaux dans chaque branche de l'arbre est contrôlé par un paramètre d'algorithme (Breiman, 2001).

L'indice de Gini au niveau d'un nœud d'arbre interne est calculé comme suit : Pour un attribut fractionné candidat (nominal)  $X_i$ , notez les niveaux possibles comme  $L_1, \dots, L_J$ . L'indice de Gini pour cet attribut est calculé comme suit :

$$G(x_i) := \sum_{j=1}^J \Pr(x_i = L_j)(1 - \Pr(x_i = L_j)) = 1 - \sum_{j=1}^J \Pr(x_i = L_j)^2 \quad (\text{V.23})$$

Une fois les indices de Gini calculés pour chaque attribut candidat à la scission, la scission est effectuée sur l'attribut qui a l'indice de Gini le plus élevé (Malekipirbazari et Aksakalli, 2015).

Les classificateurs à base d'arbres de décision présentent plusieurs propriétés intéressantes : ils sont faciles à interpréter, ils peuvent traiter des données numériques et nominales et ils sont faciles à construire. Néanmoins, les arbres de décision ne sont pas toujours compétitifs par rapport à d'autres techniques de classification. Ainsi, pour améliorer la précision des arbres, il est parfois nécessaire d'utiliser des méthodes d'ensemble telles que le boosting (apprentissage itératif à partir d'instances mal classées) et le bagging (construction de plusieurs arbres et combinaison des résultats). Les forêts aléatoires (FA), qui peuvent être considérées comme une technique de regroupement améliorée, sont une méthode puissante pour construire une forêt d'arbres de décision aléatoires. Les forêts aléatoires dé-corrèlent les arbres de décision de la forêt par le biais de la randomisation des attributs divisés, ce qui entraîne une amélioration par rapport aux arbres ensachés traditionnels et réduit la variance lorsque la moyenne est calculée sur l'ensemble des arbres (Breiman, 2001 ; Biau 2016 ; Svetnik et al., 2004).

### V.3.7 Arbre d'élagage à erreur (Reduced Error Pruning Tree)

Fondamentalement, Reduced Error Pruning Tree (REP Tree) est un arbre décisionnel rapide qui construit un arbre décisionnel basé sur le gain d'informations ou la réduction de la variance (Devasena, 2014 ; Quinlan, 1987). L'arbre REP est un arbre de décision rapide qui construit un arbre de décision/régression en utilisant le gain d'information comme critère de division, et l'élagage en utilisant l'élagage à erreur réduite. Il ne trie qu'une seule fois les

valeurs des attributs numériques. Les données incomplètes sont gérées grâce à la méthode C4.5, qui utilise des instances fractionnaires pour prendre en compte les valeurs manquantes.

Dans l'arbre de régression RT [E ; Y], E est la feuille de l'arbre où se termine l'arbre et Y est la variable réponse. Il convient de trouver une question binaire qui donne le maximum d'informations sur Y et de répéter le processus pour tous les niveaux de l'arbre (Kalmegh, 2015). Ici, Y est considéré comme la branche "spam" de l'arbre. La feuille E doit donner le maximum d'informations sur cette branche qui discrimine mieux les sites de spam et les sites authentiques. Dans chaque nœud enfant, le processus doit être répété de manière avide. Enfin, on obtient un arbre contenant le maximum d'informations sur les sites de spam. Comme l'algorithme est récursif, il nécessite un critère d'arrêt (Srinivasan et Mekala, 2014). Il s'agit ici d'un seuil. La somme des erreurs quadratiques pour un arbre RT est définie comme suit :

$$S = \sum_{E \in \text{Elevés}(RT)} \sum_{i \in E} (y_i - N_T)^2 \quad (\text{V.24})$$

Où,  $N_T$  est défini comme,

$$N_T = \frac{1}{P_C} \sum_{i \in T} Y_i \quad (\text{V.25})$$

L'équation ci-dessus est pour  $N_T$ . La formule devient alors,

$$S = \sum_{E \in \text{Elevés}(RT)} P_C V_C \quad (\text{V.26})$$

Est  $V_C$  considéré comme la variance interne et  $P_C$  est la prédiction de la classe.

### V.3.8 Arbre M5P (M5p tree)

Le M5P, une variante de l'algorithme M5 de Quinlan (1992), a été développée par Wang et Witten (1997) pour induire des arbres de modèles de régression. Dans cette méthode, un arbre de décision conventionnel est combiné avec des fonctions de régression linéaire aux nœuds feuilles (Braga et al., 2007). Tout comme dans les arbres de décision classiques, le modèle M5P développe un modèle en cassant l'espace d'entrée de l'ensemble de données d'apprentissage à l'aide de nœuds de décision, puis en attribuant un modèle linéaire à chaque sous-domaine. Le modèle final est un ensemble de sous-modèles linéaires qui capturent la non-linéarité nécessaire pour modéliser le problème de régression. Les valeurs prédites par le modèle et le chemin de décision de la racine de l'arbre au nœud feuille sont ensuite fournis. Pour développer le modèle M5P, trois étapes doivent être suivies : (i) la construction de l'arbre, (ii) l'élagage de l'arbre et (iii) le lissage de l'arbre (Sepahvet et al., 2019). Lors de la construction de l'arbre M5P, la maximisation d'un indice appelé réduction de l'écart type (SDR) est nécessaire (Behnood et al., 2017). Le SDR est calculé de la manière suivante :

$$SDR = sd(Q) - \sum_{i=1}^N \left| \frac{Q_i}{Q} \right| sd(Q) \quad (V.27)$$

Où Q est l'ensemble d'échantillons de données,  $Q_i$  est i sous ensemble de T, et sd : est l'écart type et N est le nombre d'observations. Pour éviter le problème de surajustement, les sous-arbres inutiles doivent être éliminés lors de l'arborescence taille. De cette manière, les variables d'entrée sont éliminées individuellement pour minimiser l'erreur du modèle. Enfin, les fortes discontinuités liées aux petits échantillons d'apprentissage sont compensées lors du lissage des arbres.

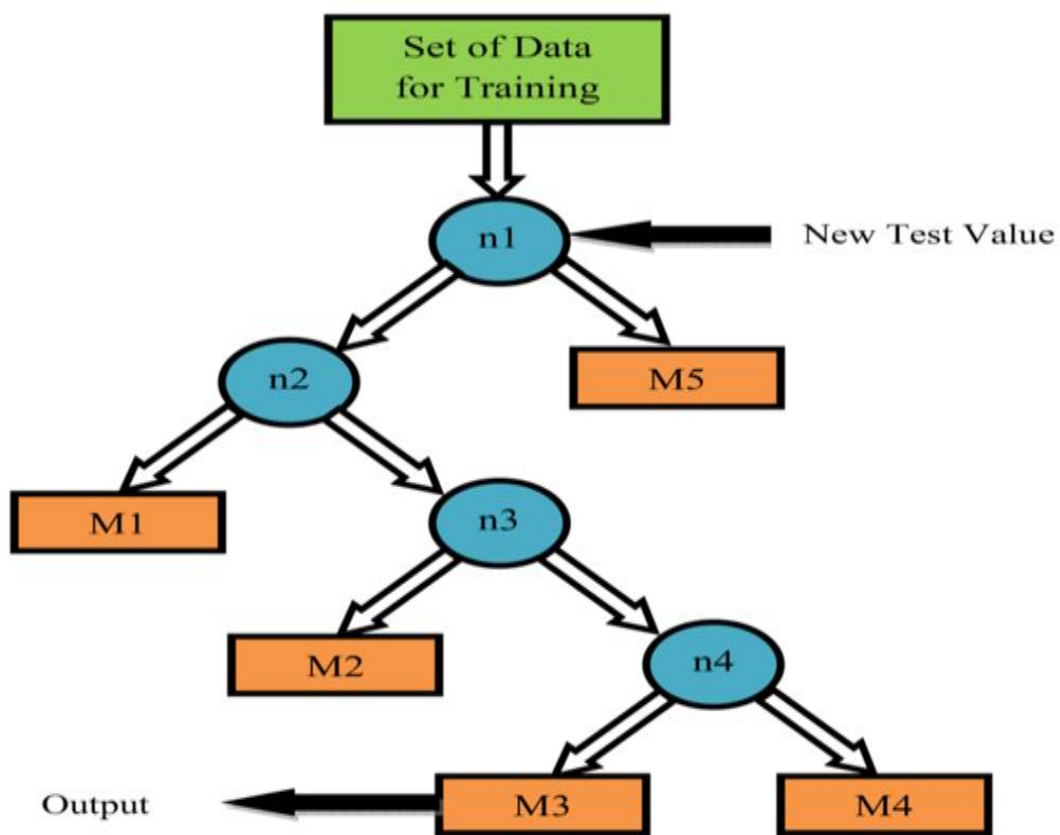


Figure.V. 10. Organigramme du fonctionnement M5p tree.

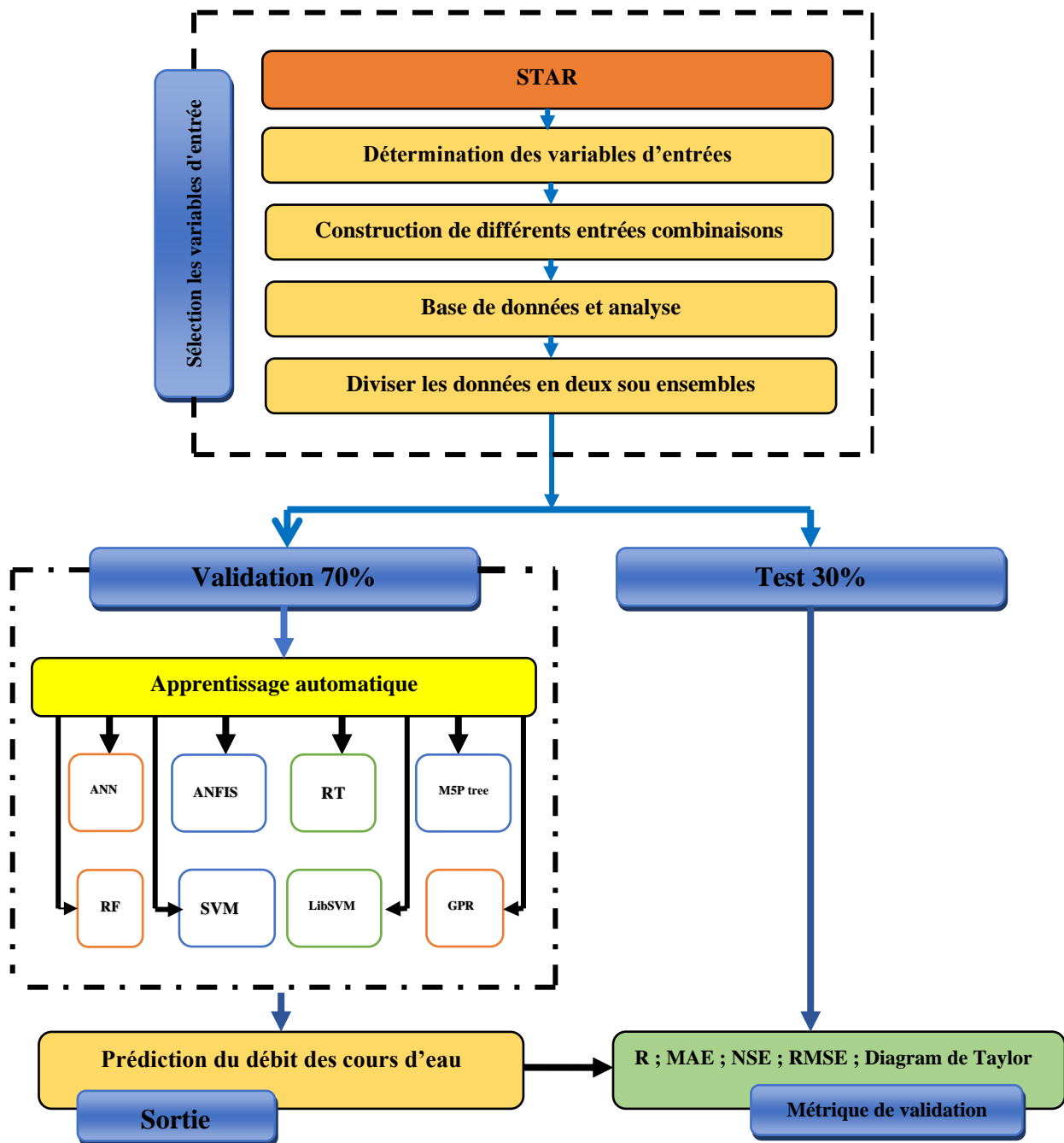


Figure.V. 11. Diagramme de méthodologie appliquée pour les modèles d'apprentissages utilisés.

#### V.4. Conclusion

À cet égard, la principale contribution dans cette partie de cette recherche est l'utilisation des modèles basés sur l'apprentissage automatique pour améliorer la précision de l'estimation des débits dans les cours d'eau mal jaugés situés dans un climat semi-aride et méditerranéen. Une description détaillée des modèles appliqués a été mentionnée dans ce chapitre et les résultats obtenus sont prometteurs dans cette étude. Ces modèles ont été validés en termes de critères statistiques et d'interprétations graphiques.

# **CHAPITRE VI**

## **MÉTHODES HYBRIDES DU MODÈLE SVM**

## VI.1. Introduction

La simulation d'écoulement surtout dans les bassins versants non jaugés est actuellement considérée comme l'une des tâches les plus difficiles en hydrologie des eaux de surface (Manguerra et Engel, 1998). Ces dernières années, différentes procédures ont été proposées pour estimer le débit des rivières et les cours d'eau naturels, parmi lequel les modèles hybrides d'intelligence artificielle ont suscité une attention notable (Madvar et al, 2020 ; Olyaie et al, 2015 ; Buyukyildiz et al, 2017 ; Madvar et al, 2020)

Malgré la pertinence des modèles ML dans les études hydrologiques, ils présentent encore plusieurs inconvénients, tels qu'un temps de calcul long et une capacité d'extraction de caractéristiques insuffisante (Nourani et al., 2014 ; Yang et al, 2002 ; Tao et al, 2020). Pour améliorer encore la précision des prévisions, les modèles hybrides deviennent des alternatives populaires, car ils peuvent surpasser les modèles simples (Ni et al., 2020) ; Afin de répondre aux demandes des utilisateurs en matière de précision et de simulation des modèles hydrologiques, plusieurs modèles ont été développés avec de nouveaux logiciels et couplages (Pham et al., 2020 ; Aswad et al., 2021 ; Samantaray et al., 2023 ; Agnihotri et al., 2022 ; Ehteram et al., 2020)

Dans ce chapitre, quatre modèles d'ensemble ont été développés intégrant une machine à vecteurs de support (SVM) avec plusieurs techniques d'apprentissage, à savoir, Bagging, Dagging, Random subspace et Rotation Forest, pour prédire le débit dans les cours d'eau naturelles dans un climat méditerranéen en Algérie. Les données de jaugeage de la station hydrométrique « Amont des gorges » ont été utilisées pour les modèles hybridés du SVM. Les performances des modèles développés ont été comparées à l'aide de mesures statistiques standard et graphiques afin de sélectionner le meilleur modèle pour une prédiction précise et de fournir une méthode fiable pour résoudre un des problèmes hydrologiques complexes.

## VI.2. Apprentissage hybride (Hybrid Learning)

L'apprentissage hybride fait référence aux architectures qui utilisent des composants génératifs (non supervisés) ainsi que discriminatifs (supervisés). La combinaison de différentes architectures peut être utilisée pour concevoir un réseau de neurones profond hybride (Deng et Yu, 2014 ; Mathew et al., 2021 ; Sarker, 2021).

Bien que l'apprentissage d'ensemble soit apparu comme un soulagement pour les chercheurs confrontés aux problèmes courants de complexité de calcul, de surajustement et de

respect des minima locaux dans les algorithmes de classification, les chercheurs ont trouvé des problèmes avec l'apprentissage d'ensemble (Zhang et Ma, 2012 ; Dietterich, 1997). L'ensemble compliqué de plusieurs classificateurs rend difficile la mise en œuvre et l'analyse des résultats. Au lieu d'améliorer la précision du modèle, les ensembles peuvent avoir tendance à augmenter l'erreur au niveau de l'apprenant de base individuel. Les ensembles peuvent entraîner une mauvaise précision en raison de la sélection de mauvais classificateurs en combinaison. Une approche récente pour traiter de tels problèmes est l'hybridation, c'est-à-dire la création d'un ensemble de modèles hétérogènes. En cela, plus d'une méthode est combinée, par exemple, en combinant le clustering et l'arbre de décision ou le clustering et l'exploration d'associations, etc. (Alzubi et al., 2018 ; Khashei et Bijari, 2011 ; Woźniak et al., 2014)

De tous les paradigmes d'apprentissage mentionnés ci-dessus, l'apprentissage supervisé est de loin le plus populaire auprès des chercheurs et des praticiens (Alzubi et al., 2018)

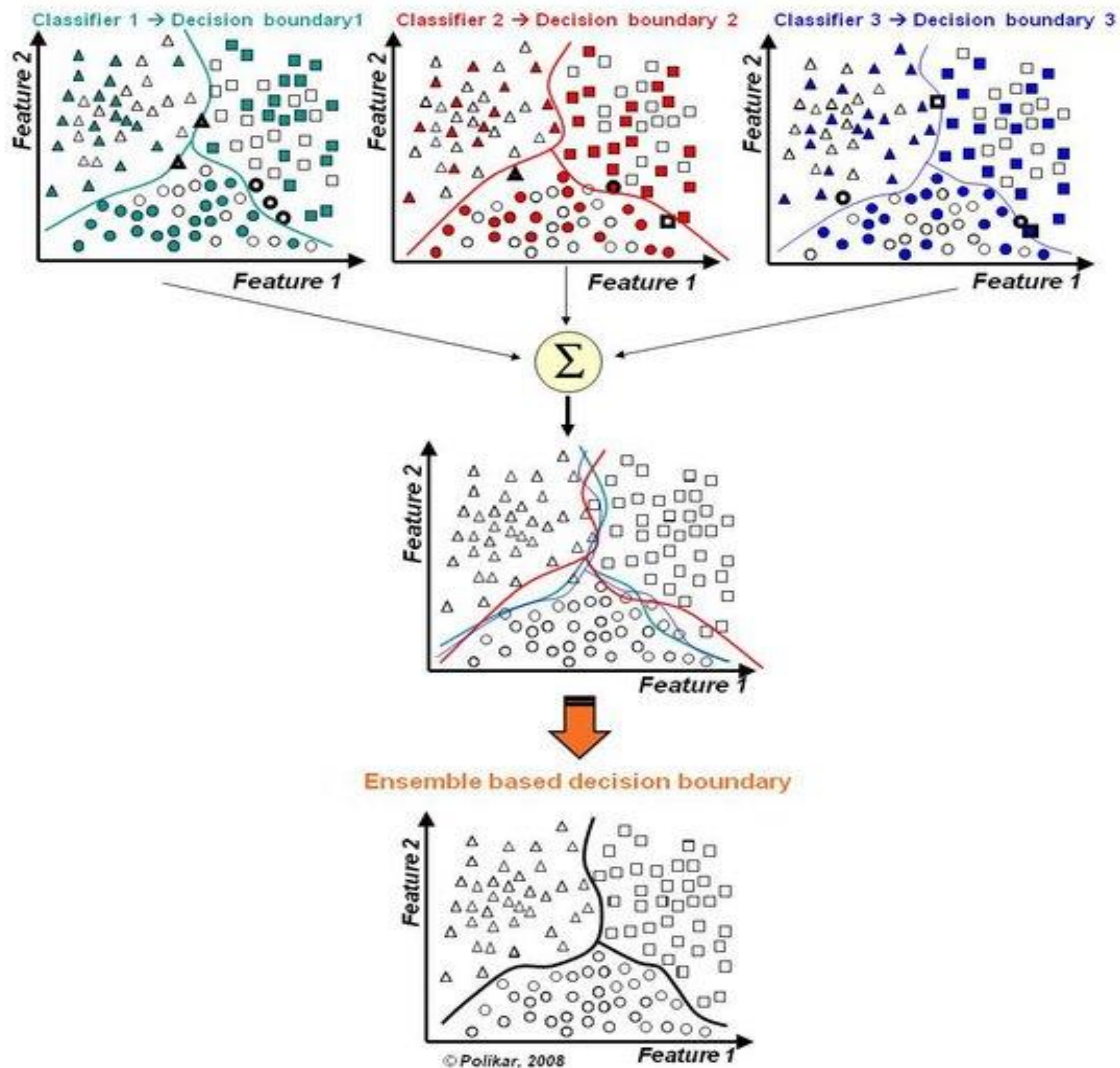
Dans la littérature, l'hybridation améliore les performances des classificateurs uniques. Les systèmes hybrides peuvent traiter des tâches relativement plus complexes car ils combinent différentes techniques (Lenard et al., 1998 ; Henderson et al., 2008). Généralement, les modèles hybrides sont basés sur la combinaison de deux ou plusieurs techniques d'apprentissage automatique (par exemple, les techniques de clustering et de classification) (Chou et al., 2013 ; Dang et al., 2021)

Selon Tsai et Chen (2010), deux méthodes peuvent être appliquées pour construire des modèles hybrides de classification la combinaison séquentielle de techniques de regroupement et de classification et la combinaison séquentielle de différentes techniques de classification.

### **VI.3. Apprentissage d'ensemble (Ensemble Learning)**

L'apprentissage en ensemble représente une approche puissante et polyvalente en apprentissage automatique. Cette méthode consiste à combiner de manière stratégique plusieurs modèles de base, pouvant être des arbres de décision, des réseaux de neurones ou d'autres traditionnellement utilisés dans l'apprentissage supervisé (Oza et Russell, 2001). L'objectif principal de l'apprentissage en ensemble est d'apporter une amélioration significative aux performances des modèles, que ce soit dans le contexte de la classification, de la prédiction, de l'approximation de fonctions ou d'autres tâches similaires, tout en minimisant le risque de choisir un modèle de qualité médiocre. De plus, cette approche trouve des applications variées, allant de la fusion de données à l'attribution de confiance aux décisions des modèles, en passant par la sélection optimale de caractéristiques, l'apprentissage continu et adaptatif, ainsi que la

correction d'erreurs. L'apprentissage en ensemble constitue ainsi une réponse efficace à de nombreux défis en apprentissage automatique, permettant aux modèles de s'adapter à une grande variété de scénarios et de problèmes, tout en renforçant la robustesse et la fiabilité des prédictions (Polikar, 2012 ; Turner, 1999).



**Figure. VI. 1.** Combinaison d'un ensemble de classificateurs pour réduire les erreurs de classification et/ou la sélection de modèles (Polikar, 2012).

#### VI.4. MACHINE A VECTEURS DE SUPPORT (SUPPORT VECTOR MACHINE)

SVM est un algorithme d'apprentissage automatique supervisé utilisé pour la classification et l'analyse de régression (Osisanwo et al., 2017). C'est un algorithme puissant qui peut gérer des ensembles de données linéaires et non linéaires. L'idée de base derrière SVM est de trouver la meilleure frontière possible qui peut séparer les points de données en différentes classes. Cette frontière s'appelle l'hyperplan. L'algorithme SVM essaie de maximiser la marge entre l'hyperplan et les points de données les plus proches de chaque classe. La marge

est définie comme la distance entre deux hyperplans parallèles les plus proches l'un de l'autre mais ne contenant aucun point de données. Pour trouver cet hyperplan, nous devons résoudre un problème d'optimisation qui consiste à minimiser une fonction de coût soumise à certaines contraintes. La fonction de coût pénalise les erreurs de classification des points de données tout en maximisant leur distance à l'hyperplan (Vincent et Bengio, 2001 ; Bhavsar et Panchal., 2012).

Une fois que nous avons trouvé cet hyperplan optimal, nous pouvons l'utiliser pour classer de nouveaux points de données en vérifiant simplement de quel côté ils tombent. Dans l'ensemble, SVM est un algorithme puissant qui peut gérer des ensembles de données complexes avec une grande précision et une capacité de généralisation (Cervantes et al., 2020).

### VI.5. Dagging

Le Dagging est une technologie d'intégration par rééchantillonnage proposée pour la première fois par Ting et Witten et destinée à améliorer la précision de prédiction des classificateurs de base individuels (Kotsianti et Kanellopoulos, 2007 ; Ting et Witten, 1997).

Le dagging (Directed Acyclic Graphs) est une méthode d'ensemble utilisée pour les problèmes de classification (Ting et Witten, 1997). Cela implique la création de plusieurs arbres de décision, chacun formé sur un sous-ensemble différent des données de formation. Les sous-ensembles sont créés en sélectionnant de manière aléatoire un sous-ensemble des fonctionnalités et un sous-ensemble des instances d'apprentissage. Les arbres de décision sont ensuite combinés dans un graphe acyclique dirigé, où chaque nœud représente un arbre de décision et les arêtes représentent les relations entre eux. Le graphe est construit de manière à ce que chaque nœud ne dépende que de ses nœuds parents, garantissant qu'il n'y a pas de cycles dans le graphe. Pour faire des prédictions, les données d'entrée sont passées à travers chaque arbre de décision dans le graphique, et la prédiction finale est faite en combinant les prédictions de tous les arbres à l'aide d'un schéma de vote pondéré. Il a été démontré que le dagging améliore la précision de la classification par rapport à l'utilisation d'un arbre de décision unique ou d'autres méthodes d'ensemble telles que le bagging ou le boosting (Kotsianti et Kanellopoulos, 2007). Il est particulièrement efficace lorsqu'il s'agit de données bruyantes ou incomplètes, car il peut aider à réduire le surajustement et à améliorer les performances de généralisation (Pham et al., 2020).

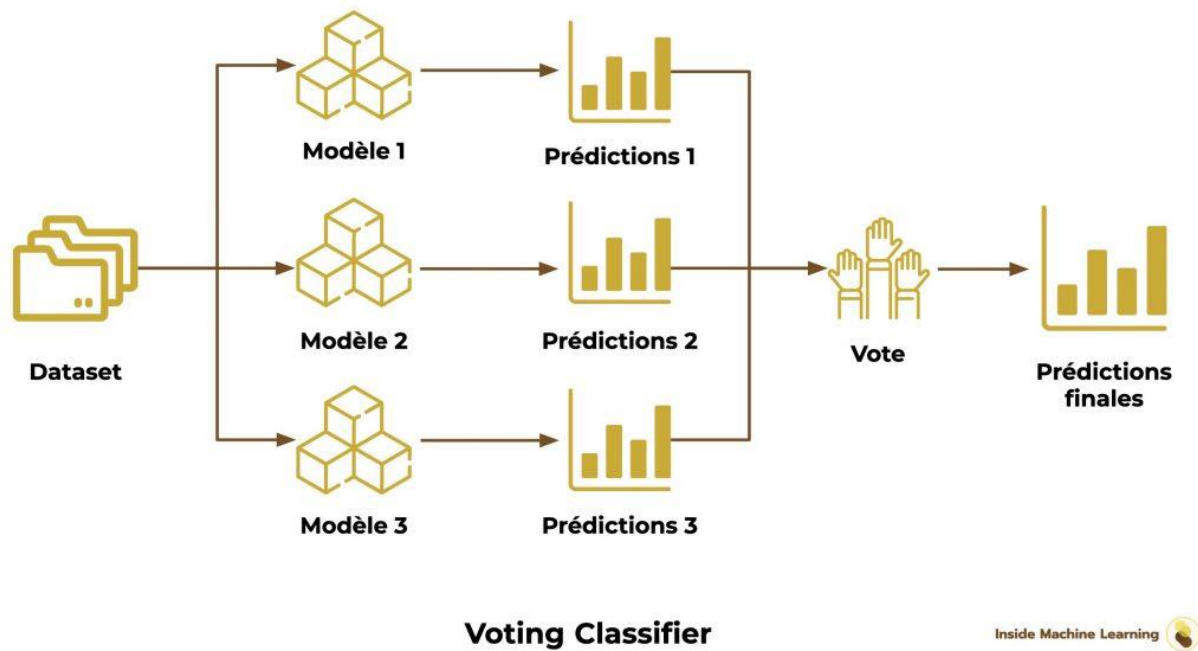
## VI.6. Bagging

Le bagging (Bootstrap Aggregating) est une méthode d'ensemble qui combine plusieurs modèles pour améliorer la précision et la stabilité des prédictions (Breiman, 1996). Cela fonctionne en créant plusieurs sous-ensembles de données de formation, en formant un modèle sur chaque sous-ensemble, puis en combinant les prédictions de tous les modèles pour faire une prédiction finale. Pour faire une prédiction à l'aide du modèle d'ensemble  $F(x)$ , nous appliquons simplement chaque modèle individuel  $f_m$  à l'entrée  $x$  et combinons leurs prédictions à l'aide d'une fonction d'agrégation (par exemple, vote majoritaire pour la classification ou moyenne pour la régression). Le bagging peut être utilisé avec n'importe quel algorithme d'apprentissage de base  $A$ , tel que les arbres de décision, les réseaux de neurones ou les machines à vecteurs de support. Il est particulièrement efficace lorsque l'algorithme de base est instable ou sujet au surajustement, car il contribue à réduire la variance et à améliorer les performances de généralisation. Dans l'ensemble, le bagging est une technique puissante pour améliorer la précision et la robustesse des modèles d'apprentissage automatique en combinant plusieurs modèles plus faibles en un ensemble plus fort. (Bauer et Kohavi, 1999 ; Bühlmann, 2002)

Le bagging adapte chaque classificateur de base à des sous-ensembles aléatoires de l'ensemble de données initial, puis agrège leurs prédictions individuelles pour former une prédiction finale (soit par vote, soit par calcul de la moyenne).

Tout d'abord, les échantillons bootstrap  $S(x_i, y_i)$  sont rééchantillonnés de manière aléatoire à partir d'un ensemble d'entraînement  $(x_i, y_i)$ , formant ainsi un ensemble de sous-ensembles d'apprentissage, où  $x_i \in R, y_i \in Y$  (estimation du débit) (Breiman, 1996). Ensuite, plusieurs modèles basés sur un classificateur sont construits en fonction de chaque sous-ensemble,  $C_i(x)$  est un classificateur construit à partir de chaque échantillon directeur. Tous les modèles basés sur le classificateur  $(L_i)$  sont agrégés pour générer le modèle final  $(L_0)$ , où  $L_1, L_2, \dots, L_n$  génère un classificateur combiné  $(L_0)$ .  $L_0$  prédit l'étiquette de classe d'une instance  $x$  donnée en calculant les votes à l'aide de l'équation suivante (Bauer et Kohavi, 1999).

$$L'(x) = \arg \max_{y \in Y} \sum_{i=1}^t L(C_i(x) = y) \quad (\text{VI.1})$$



**Figure. VI. 2.** Organigramme du fonctionnement Bagging (Salcedo-Sanz et al.,2022).

### VI.7. Sous-espace aléatoire (Random Subspace)

RSS est une technique utilisée dans l'apprentissage automatique pour améliorer la précision et la robustesse des modèles de classification. Cela implique la sélection aléatoire d'un sous-ensemble de caractéristiques à partir de l'ensemble de données d'origine et la formation de plusieurs modèles sur ces sous-ensembles (Bertoni et al., 2005). Les prédictions de ces modèles sont ensuite combinées pour faire une prédiction finale. Dans cette méthode, le nombre de modèles et la taille des sous-ensembles d'entités peuvent être ajustés en fonction de la complexité de l'ensemble de données et du niveau de précision souhaité. La méthode d'ensemble Random Subspace est particulièrement utile lorsqu'il s'agit d'ensembles de données de grande dimension où le surajustement est un problème, car elle aide à réduire la variance dans le modèle en introduisant un caractère aléatoire dans le processus de sélection des caractéristiques (Pham et al., 2020 ; Saha et al., 2023).

Cette technique a été appliquée à un nombre important de problèmes non linéaires (García et Ortiz, 2008), dans diverses disciplines (science médicale, informatique et bancaire). L'application de la technique RSS dans l'ingénierie des transports est encore limitée (Li et al., 2011). Une présentation détaillée de la technique RSS est la suivante :

Étant donné un ensemble d'échantillons de formation  $G$  de taille  $n$ , l'ensemble  $S = (S_1, S_2, \dots, S_n)$  avec chaque objet de formation  $S_i$  ( $i = 1, 2, \dots, n$ ) est un vecteur à  $q$  dimensions  $S_i = (S_{i1}, S_{i2}, \dots, S_{iq})$ , décrit par  $q$  caractéristiques. Si l'on sélectionne  $r < q$  caractéristiques, on

obtient un sous-espace aléatoire  $r$ -dimensionnel de l'espace de caractéristiques original  $q$  dimensionnel. Par conséquent, chaque objet d'apprentissage modifié  $S_i = (S_{i1}, S_{i2}, \dots, S_{in})$  ( $i = 1, \dots, n$ ) est une unité de l'ensemble d'échantillons d'apprentissage  $S = (S_1, S_2, \dots, S_n)$ . La combinaison de l'ensemble d'échantillons de formation  $S$  et des classificateurs pour créer des modèles hybrides qui sont votés dans la règle de décision finale. L'algorithme du RSS peut être interprété comme suit (Skurichina et Duin, 2002)

$$\gamma(s) = \operatorname{argmax} \sum_d \delta_{sn} (C^d(s)), y \quad (\text{VI.2})$$

Où  $\delta_{i,j}$  est le symbole de Kronecker, et  $y = (-1, 1)$  est une décision ou une étiquette de classe du classificateur et  $C^d(s)$  sont les classificateurs ( $d = 1, 2, \dots, D$ ).

## VI.8. Forêt de rotation (Rotation Forest)

Rotation Forest est une méthode d'ensemble qui combine la puissance des arbres de décision et de l'analyse en composantes principales (ACP) pour améliorer la précision des modèles de classification. L'idée de base derrière Rotation Forest est de créer plusieurs sous-ensembles de l'ensemble de données d'origine, d'appliquer l'ACP à chaque sous-ensemble, puis de construire un arbre de décision sur chaque sous-ensemble transformé (Kuncheva et Rodríguez, 2007). La prédiction finale est faite en prenant un vote majoritaire de tous les arbres de décision. Dans cette méthode, le nombre de modèles et la taille des sous-ensembles d'entités peuvent être ajustés en fonction de la complexité de l'ensemble de données et du niveau de précision souhaité. La méthode d'ensemble Random Subspace est particulièrement utile lorsqu'il s'agit d'ensembles de données de grande dimension où le surajustement est un problème, car elle aide à réduire la variance dans le modèle en introduisant un caractère aléatoire dans le processus de sélection des caractéristiques (Saha et al., 2023 ; Rodriguez ET al., 2006).

La structure de l'algorithme RF\* est donnée comme suit (Ozcift et Gulten, 2011).

Soit le jeu de données d'entraînement  $X$  un  $N \times n$  matrice, et laissez les étiquettes correspondantes avoir  $F$  nombre de fonctionnalités. De plus, supposons que les étiquettes de classe sont  $\{\omega_1, \dots, \omega_c\}$ . Supposons que l'ensemble de fonctionnalités de l'ensemble de données est partitionné de manière aléatoire en  $K$  sous-ensembles et  $L$  DTs de l'algorithme RF\* désigné par  $\{D_1, \dots, D_L\}$ . Les trois étapes suivantes sont nécessaires pour construire l'ensemble de données d'apprentissage d'un classifieur.

$F$  est divisé au hasard en  $K$  ensembles de fonctionnalités où chaque sous-ensemble à  $M = n/k$  nombre de fonctionnalités.

Laissez  $F_{ij}$  Soit  $j$ th sous-ensemble de fonctionnalités pour former le classifieur  $D_i$ , et laissez  $X_{ij}$  soit l'ensemble de données pour  $X$ . Ensuite, un sous-ensemble aléatoire non vide est conçu pour  $X_{ij}$  pour former un nouvel ensemble d'apprentissage via un rééchantillonnage Bootstrap avec une taille d'échantillon généralement inférieure à celle de  $X_{ij}$ . Ensuite, une transformation linéaire est appliquée à  $X_{ij}$  produire les coefficients de la matrice  $C_{ij}$ , La taille de chaque matrice est  $T \times 1$  avec les coefficients de  $a_{ij}^1, \dots, a_{ij}^{Mj}$ .

La matrice de rotation creuse  $R_i$  avec les coefficients obtenus de la matrice  $C_{ij}$  est donnée comme suit :

$$R_i = \begin{bmatrix} a_{i1}^{(1)}, L, a_{i1}^{(M1)} & 0 & L & 0 \\ 0 & a_{i2}^{(1)}, L, a_{i2}^{(M2)} & L & 0 \\ M & M & 0 & M \\ 0 & 0 & L, a_{ik}^{(1)}, L, a_{ik}^{(Mk)} \end{bmatrix} \quad (VI.3)$$

Alors, la matrice de rotation creuse  $R_i$  est arrangée avec les coefficients obtenus qui sont calculés pour chaque classe en utilisant la méthode de combinaison moyenne dans l'échantillon de test donné comme suit :

$$m_j(x) = \frac{1}{L} \sum_{i=1}^L d_{ij}(xR_i^a), j = 1, \dots, c \quad (VI.4)$$

Où  $d_{ij}(xR_i^a)$  est la probabilité générée par le classifieur  $C_i$  pour l'hypothèse que  $x$  appartient à la classe  $j$ . Enfin,  $x$  sera assigné à la classe avec la plus grande confiance (Pham et al., 2016).

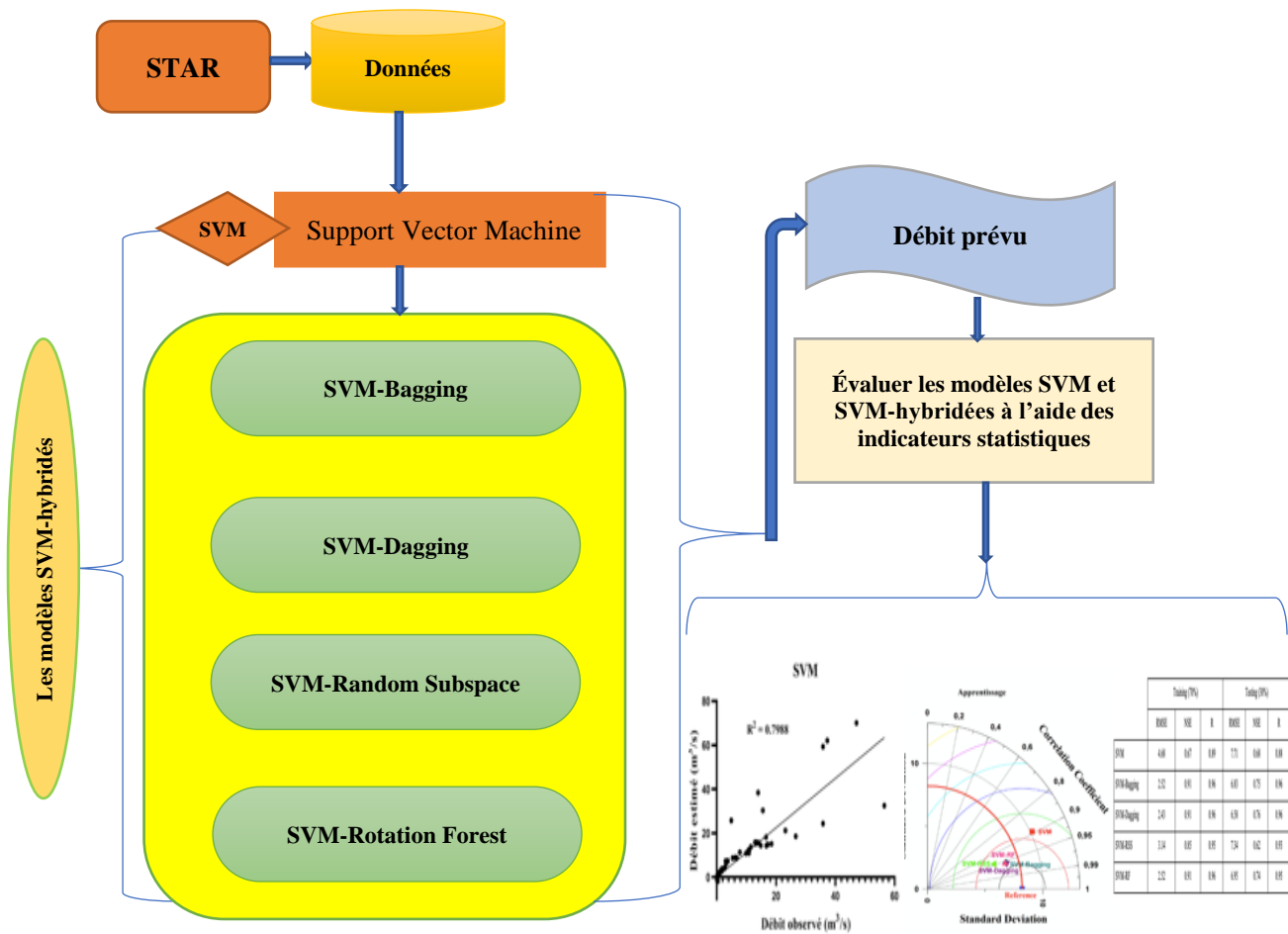


Figure. VI. 3. Diagramme de méthodologie des méthodes SVM-Hybridés.

### VI.9. Évaluation et comparaison des modèles

Le choix des indicateurs de performance appropriés est crucial car chaque indicateur a ses propres propriétés. De plus, connaître les points forts de chaque mesure statistique peut fournir un meilleur résultat.

Par conséquent, pour évaluer l'efficacité des différents modèles prédictifs est évaluée par une interprétation graphique (ligne, nuage de points et diagramme de Taylor...) et par des indicateurs de performance.

Dans cette étude, la performance prédictive du modèle a été évaluée en utilisant de nombreux indicateurs statistiques bien connus. Ces indicateurs sont définis ci-dessous.

Le coefficient d'efficacité du modèle de Nash-Sutcliffe est une mesure statistique utilisée pour évaluer la précision d'un modèle prédictif. Le NSE est calculé en comparant les valeurs observées et prédites d'une variable sur une période donnée. Une adéquation parfaite entre les valeurs observées et prédites se traduirait par une valeur NSE de 1, tandis qu'une valeur

NSE de 0 indique que les prédictions du modèle ne sont pas meilleures que l'utilisation de la valeur moyenne observée.

L'erreur quadratique moyenne (RMSE) est une autre mesure statistique utilisée pour évaluer la précision d'un modèle hydrologique. Elle mesure le degré de concordance entre les valeurs prédites et les valeurs réelles en calculant la racine carrée de la différence quadratique moyenne entre les deux. Une RMSE plus faible indique une meilleure concordance entre les valeurs observées et les valeurs prédites.

Le coefficient de corrélation (R) mesure le degré de corrélation entre deux variables. En hydrologie, il peut être utilisé pour évaluer la corrélation entre une variable hydrologique, telle que le débit et la vitesse d'écoulement. Le coefficient de corrélation est compris entre -1 et +1, les valeurs proches de +1 indiquant une forte corrélation positive et les valeurs proches de -1 indiquant une forte corrélation négative. Une valeur de 0 indique qu'il n'y a pas de corrélation. Un coefficient de corrélation élevé indique que les changements d'une variable sont fortement associés aux changements de l'autre variable. Toutefois, une corrélation élevée n'implique pas nécessairement un lien de causalité.

La Mean Absolute Error (MAE) est une métrique essentielle pour évaluer la précision des modèles prédictifs, notamment en hydrologie. En calculant la moyenne des écarts absolus entre les valeurs prédites et observées, la MAE offre une mesure directe de la précision moyenne du modèle, sans accentuer les erreurs extrêmes. Son utilisation dans cette thèse vise à obtenir une évaluation fine des performances des modèles hydrologiques, en tenant compte des erreurs spécifiques à chaque observation. La MAE est particulièrement pertinente dans le contexte de la préservation de la qualité de l'eau et de la gestion des débits, contribuant ainsi à éclairer les décisions basées sur des modèles prédictifs et à améliorer continuellement les approches de modélisation hydrologique.

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (Q_{o,i} - Q_{p,i})^2}{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2} \quad (\text{VI.5})$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{o,i} - Q_{p,i})^2} \quad (\text{VI.6})$$

$$\text{R} = \frac{\frac{1}{n} \sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)(Q_{p,i} - \bar{Q}_p)}{\sqrt{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2 \sum_{i=1}^n (Q_{p,i} - \bar{Q}_p)^2}} \quad (\text{VI.7})$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Q_{o,i} - Q_{p,i}| \quad (\text{VI.8})$$

Où :  $n$  est le nombre de données dans la période d'essai,  $Q_{o,i}$  est le débit observé ;  $Q_{p,i}$  est le débit prédit ;  $Q_o$  est la valeur moyenne de la valeur observée du débit ;  $\overline{Q_p}$  est la valeur moyenne du débit prédiction.

Le diagramme de Taylor est une représentation graphique de  $R$ , RMSE et de l'écart-type d'un ensemble de prédictions de modèles par rapport aux observations (Taylor, 2001). Le diagramme consiste en un système de coordonnées polaires avec l'écart-type observé en abscisse et le coefficient de corrélation en ordonnée. Le RMSE est représenté par des contours qui sont des cercles concentriques autour de l'origine. Chaque prédiction du modèle est représentée par un point sur le diagramme, sa distance par rapport à l'origine indiquant son RMSE et son angle indiquant sa corrélation avec les observations. Plus un point est proche du point d'observation (situé à  $x=1$ ,  $y=0$ ), plus il correspond aux observations en termes de corrélation et de RMSE. Le diagramme de Taylor permet de comparer plusieurs modèles ou simulations aux observations sur un seul graphique, ce qui permet d'identifier facilement les modèles les plus performants dans l'ensemble ou pour des variables spécifiques.

## VI.10. Conclusion

Ce chapitre consiste à développer une approche pour la simulation d'écoulement dans les cours d'eau naturels du bassin des côtières d'Alger, le nord centre de l'Algérie, où les modèles de prédiction hybridés précis ont été étudiés dans le but de bien prédire les débits et maîtriser leurs comportements dans un cours d'eau naturel.

L'approche de modélisation proposée intègre une machine à vecteurs de support (SVM) avec quatre techniques d'apprentissage d'ensemble différentes, notamment le Bagging, le Dagging, le sous-espace aléatoire et la forêt de rotation.

Les modèles ont été développés et évalués à partir des données de jaugeage de la station hydrométrique « Amont des gorges » en tenant compte les paramètres quantitatifs du cours d'eau, tels que le débit, la vitesse, la profondeur, la largeur et le rayon hydraulique.

# **CHAPITRE VII**

## **RÉSULTATS ET DISCUSSION**

## **VII.1. Introduction**

La fourniture de prévisions précises et fiables sur les débits futurs est cruciale pour des applications pratiques en ingénierie, telles que la conception d'infrastructures de contrôle des crues et la gestion des bassins versants (Blöschl et al., 2013 ; Lou et al., 2020).

La simulation d'écoulement, en particulier dans les bassins versants non jaugés, est une tâche difficile en hydrologie des eaux de surface. Ces dernières années, différentes approches ont été proposées pour estimer les débits naturels des rivières et des cours d'eau.

Le présent document de recherche tente d'appliquer une nouvelle méthode de calcul des débits qui est la méthode du modèle rugueux (RMM). Cette dernière est appliquée pour la première fois aux cours d'eau naturels de la zone méditerranéenne et semi-aride. La logique de la recherche découle de la nécessité de réduire les coûts des stations hydrométriques et d'augmenter la fréquence des mesures de débit, en particulier en cas de fortes crues. Cela nécessite le développement des méthodes efficaces pour l'estimation des débits des cours d'eau naturels. Sur cette base, cinq stations hydrométriques sont utilisées aux fins de collecte de données. Afin de réaliser une analyse comparative des performances, une méthode bien connue de la littérature, la méthode de Chiu (Chiu et Said, 1995), a également été appliquée pour estimer les débits aux cinq stations hydrométriques étudiées.

Ensuite, nous présentons une étude comparative de huit modèles différents basés sur l'intelligence artificielle pour la prédiction des débits des écoulements dans les cours d'eau naturels étudiés. Les modèles étudiés comprennent les réseaux de neurones artificiels (ANN), système d'inférence neuro floue adaptative (ANFIS), les forêts aléatoires (Random Forest), arbre d'élagage à erreur (Rep Tree), arbre M5p, la régression par processus gaussien (GPR), les machines à vecteurs de support (SVM) et la machine à vecteurs de support libérale (LibSVM). Les performances de ces modèles ont été évaluées à l'aide de différents indicateurs statistiques, tels que le coefficient de corrélation (R), le coefficient d'efficacité (NSE), l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE) et le diagramme de Taylor pour l'ensemble des stations de jaugeage.

A la fin, une approche de modélisation hybridé a été proposé pour une prédiction plus précise des débits dans les cours d'eau naturels étudiés et qui intègre une machine à vecteurs de support (SVM) avec plusieurs techniques d'ensemble, à savoir, Bagging, Dagging, Random subspace et Rotation Forest. Les données de jaugeage de la station hydrométrique « Amont des gorges » ont été pris en considération pour l'application des modèles SVM-hybridés. Les

paramètres quantitatifs utilisés sont (le débit, la vitesse, la profondeur, la largeur et le rayon hydraulique).

Les résultats obtenus indiquent que les modèles proposés ont surpassé les autres modèles appliqués en se basant sur l'efficacité de Nash-Sutcliffe (NSE), l'erreur quadratique moyenne (RMSE) et le coefficient de corrélation (R). Ces résultats indiquent vraiment que le modèle SVM autonome et ses modèles hybridés montrent un niveau de précision adéquat pour la prédiction des apports liquides.

## VII.2. Méthode MMR et CHIU

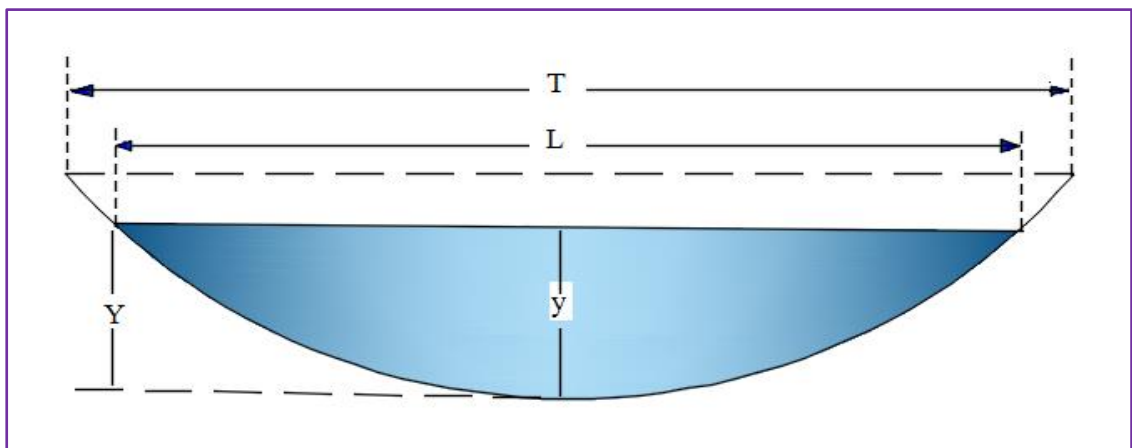
### VII.2.1. MMR

Cette section rapporte et discute les résultats obtenus par la méthode MMR pour chaque station hydrométrique. Afin d'évaluer la validité de la MMR, les valeurs des paramètres hydrauliques estimées par cette méthode ont été comparées à celles mesurées pour chacune des cinq stations incluses dans cette étude. Les résultats obtenus ont été présentés dans les figures (VII.2, VII.3 et VII.4).

Le canal de forme semi circulaire est l'un des ouvrages le plus connu et le plus rencontré dans les aménagements hydrauliques. La section transversale de chaque station hydrométrique a été déterminée en tenant compte des caractéristiques des cours d'eau naturels de la zone d'étude, à savoir de grandes largeurs et de faibles profondeurs (Marouf, 2012). En référence à la coupe transversale des cours d'eau schématisée à la figure VII.1, l'équation suivante a été utilisée (Verniers, 1995).

$$A = \frac{2}{3}y.L$$

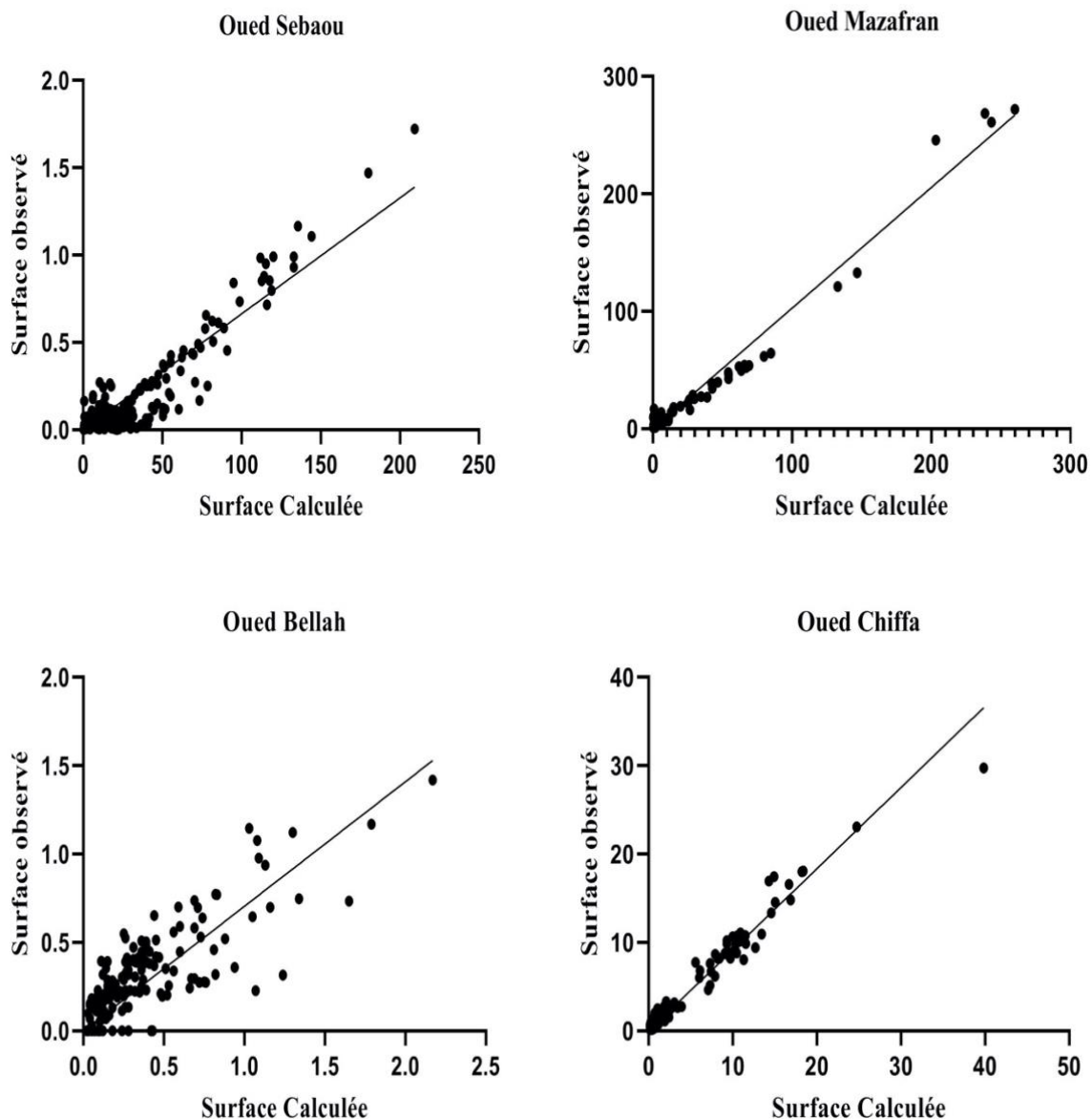
VII.1



**Figure. VII. 1.** Coupe transversale d'un cours d'eau générique avec une profondeur d'écoulement  $y$ , une largeur de canal  $T$  et une largeur de surface  $L$ .

Plus précisément, la figure VII.2 rapporte la comparaison en termes de section transversale pour l'ensemble des stations, montrant le bon ajustement obtenu pour chaque station étudiée. En effet, le coefficient de corrélation  $R$  varie entre 0.835 et 0.979. Ces résultats sont également résumés dans le tableau VII.1

La figure VII.3 montre la relation entre l'écart relatif du débit mesuré et calculé par la méthode MMR [équation (IV.31)] et le rayon hydraulique, ce qui montre la géométrie adimensionnelle des cours d'eau à étudier. Les graphiques montrent d'excellentes tendances logarithmiques obtenus dans la majorité des stations étudiées (Oued Ballah ; Oued Chiffa ; Oued Mazafran ; Oued Chellif), les coefficients de corrélation varient entre (0,510 et 0,921). De plus, l'erreur relative maximale constatée entre le débit mesuré et le débit calculé ( $\Delta Q/Q$ ) est inférieure à 20 % (voir Tableau VII.1).



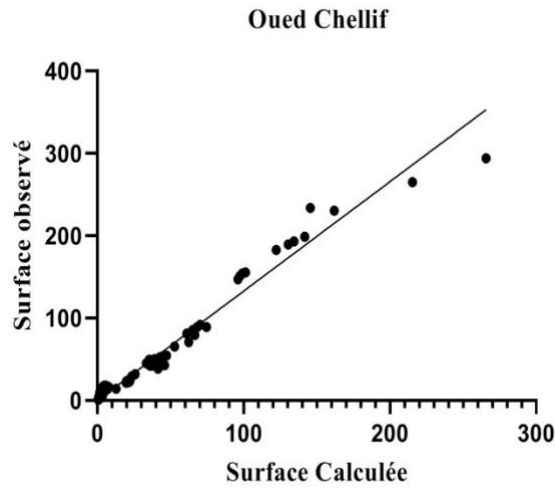
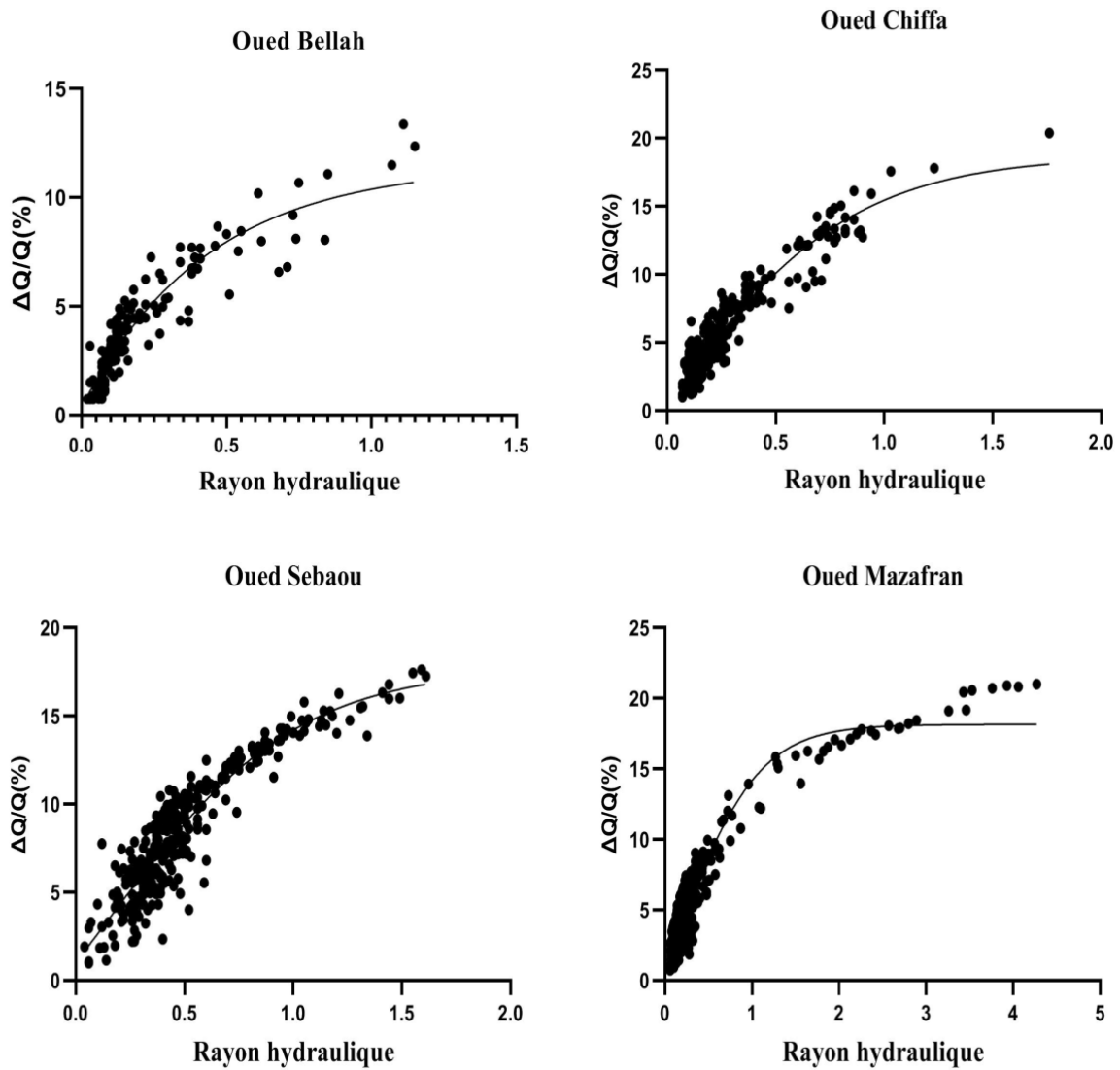


Figure. VII. 2. Relation surface mesurée ( $A_{mes}$ ) et surface calculée ( $A_{cal}$ ) de la section transversale pour la méthode MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).



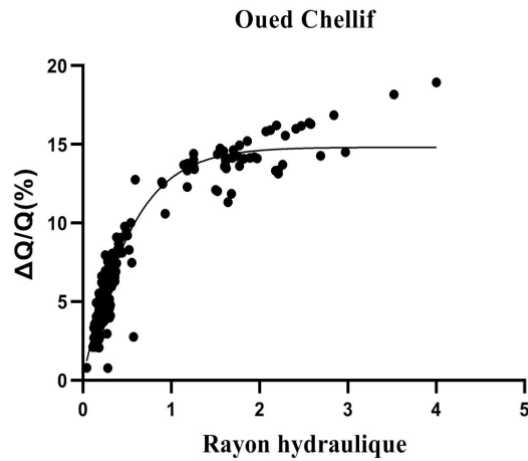
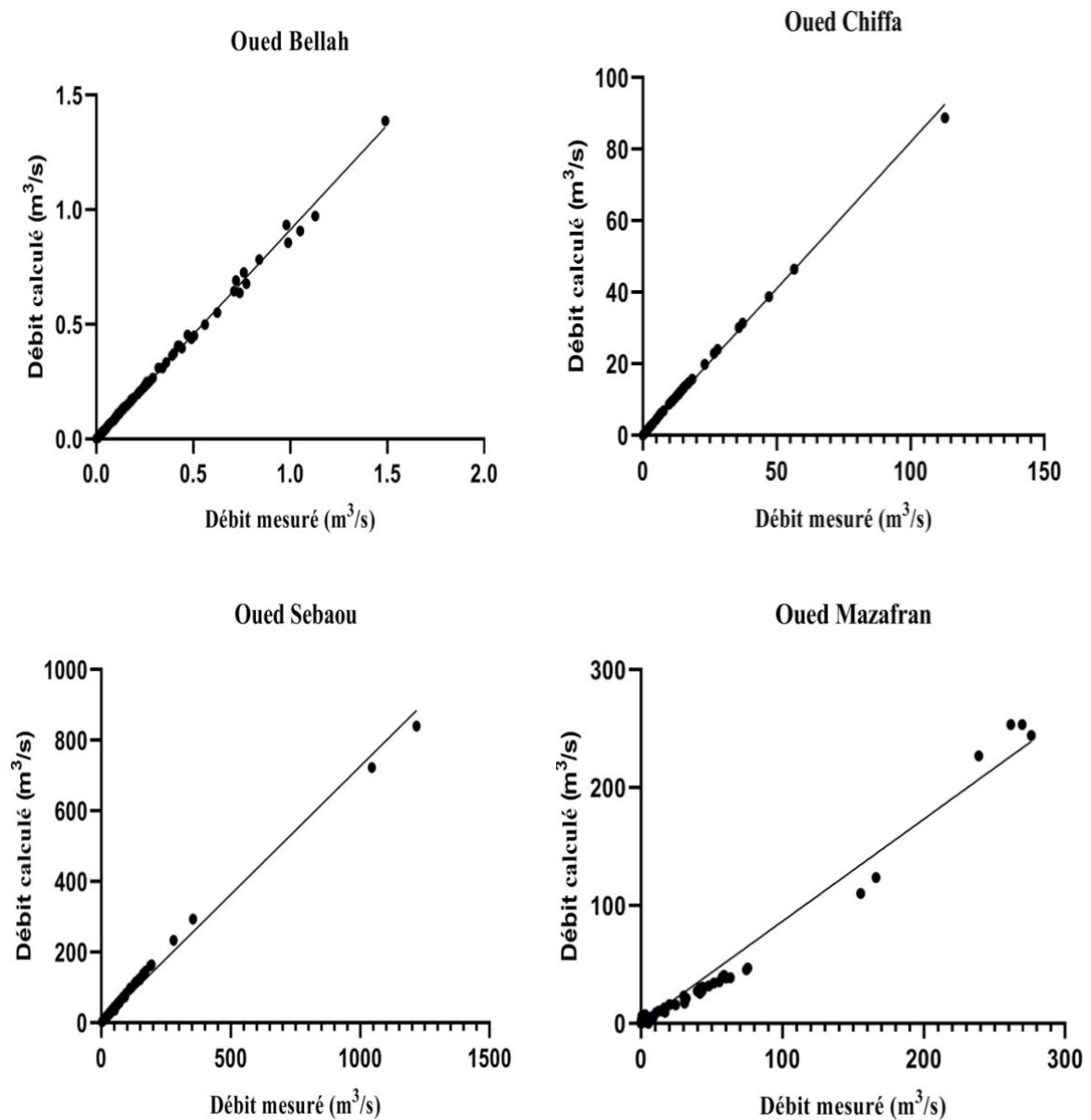
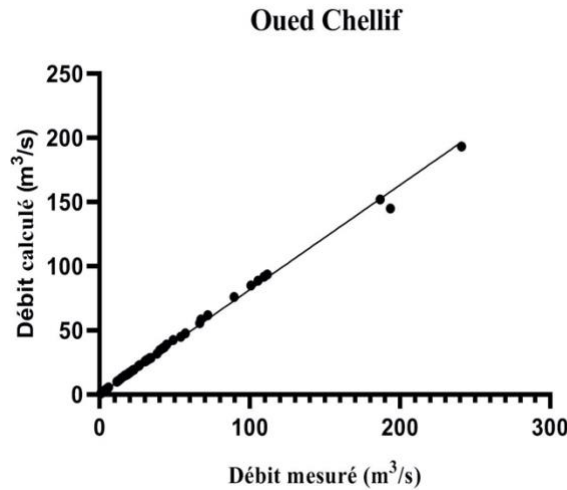


Figure. VII. 3. Relation entre l'écart relative du débit ( $\Delta Q/Q$ ) et le rayon hydraulique ( $R_h$ ) obtenus par la méthode MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).





**Figure. VII. 4.** Relation débits mesurés et débits calculés de la méthode MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).

Enfin, la figure VII.4 montre la comparaison entre les valeurs de débit mesurés et calculés pour chaque station hydrométrique. Les graphiques montrent que la MMR a donné de bonnes estimations des valeurs de débit pour toutes les stations étudiées. En effet, le coefficient de corrélation varie entre 0,925 et 0,972, les résultats de l'équation (IV.31) sont en bon accord avec les données réelles. Sur la base de la méthode proposée, le débit calculé montre une amélioration significative par rapport au débit mesuré pour les cinq stations hydrométriques, voir le tableau VII.1.

**Tableau. VII. 1.** Les valeurs de R des figures (VII.2, VII.3, et VII.4) pour les différentes stations étudiées

Oued	R		
	$A_{cal} - A_{mes}$	$R_h - (\Delta Q/Q)$	$Q_{cal} - Q_{mes}$
<b>Chellif</b>	0,979	0,9218	0,929
<b>Chiffa</b>	0,974	0,8321	0,925
<b>Sebaou</b>	0,897	0,5107	0,936
<b>Mazafran</b>	0,976	0,888	0,972
<b>Bellah</b>	0,835	0,7178	0,954

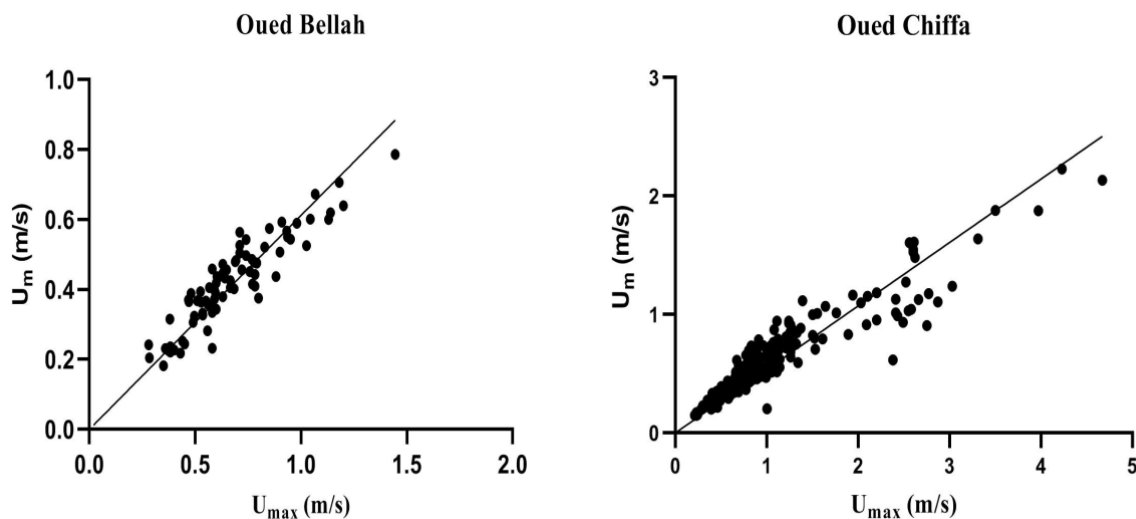
Globalement, les résultats obtenus ont montré de bonnes performances du MMR dans l'estimation des paramètres hydrauliques des cours d'eau naturels étudiés. Par conséquent, cette méthode se révèle être une solution efficace pour estimer les débits avec une précision très adaptée, notamment durant les périodes de crue, où les méthodes de jaugeage traditionnelles peuvent présenter des difficultés d'application.

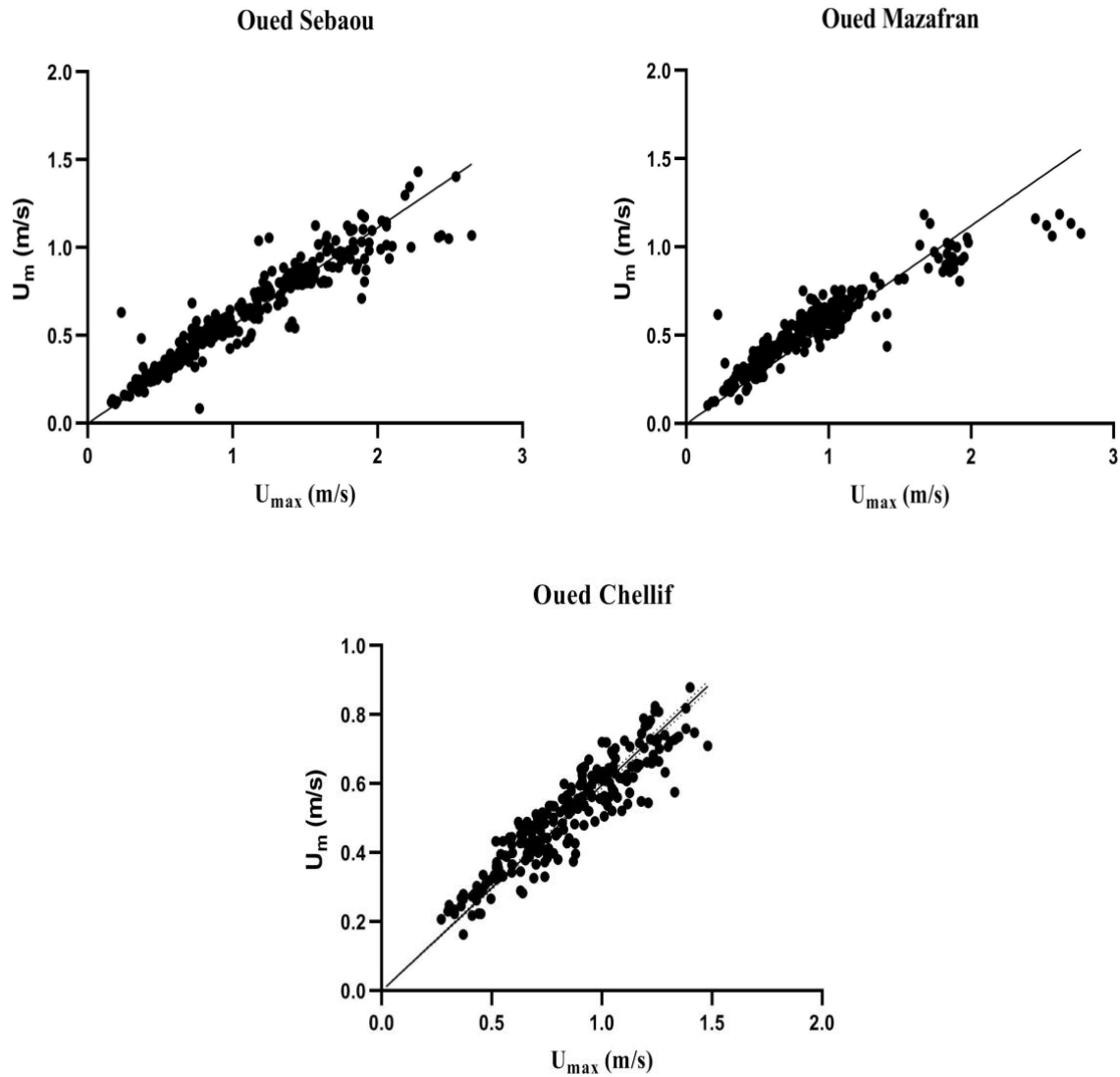
### VII.2.2. Méthode de Chiu

Dans cette section, les résultats obtenus pour chaque station hydrométrique par la méthode de Chiu sont mentionnés ci-dessous, la figure VII.5 représente la relation entre la vitesse moyenne et maximale pour chaque station étudiée, tandis que le tableau VII.2 rapporte les résultats obtenus en termes de ( $\phi$ , M et R). Les graphiques de la figure VII.5 montrent une meilleure tendance et de corrélation entre  $U_m$  et  $U_{max}$  pour toutes les stations étudiées. En effet, le coefficient de corrélation varie entre 0,84 et 0,89 (tableau VII.2). L'excellente corrélation entre  $U_m$  et  $U_{max}$  indique que la valeur de  $\phi$  est constante à la section du canal sur une large gamme de débits, de hauteur d'eau et de concentration en sédiments, que le débit soit permanent ou non, et que le cours d'eau soit confiné au lit majeur ou non. Ceci confirme les résultats obtenus par Chiu et Said (1995).

Par conséquent, la valeur de M pourrait être supposée constante à n'importe quel endroit des cours d'eau étudiés. Ainsi, il peut être utilisé pour décrire facilement le profil de vitesse.

La figure VII.6 représente l'aire de la section d'écoulement en fonction de la hauteur d'eau H et de la largeur du canal L. Les graphiques montrent une bonne corrélation entre l'aire de la section d'écoulement et le produit HL. Par conséquent, l'équation (IV.61) a été utilisée pour déterminer l'aire de la section transversale, ce qui s'est avéré être la méthode la plus adaptée pour les cas étudiés. Les expressions de la section transversale d'écoulement sont reportées dans le tableau VII.3. Avec les valeurs de R. Les oueds Chellif, Mazafran et Chiffa affichent les meilleures performances. Les caractéristiques de la section transversale à ces stations étaient plus homogènes, permettant de meilleures mesures. Quoiqu'il en soit, les expressions de Chiu ont permis aux chercheurs de calculer l'aire de la section transversale bien que les sections transversales aient été de forme irrégulière.

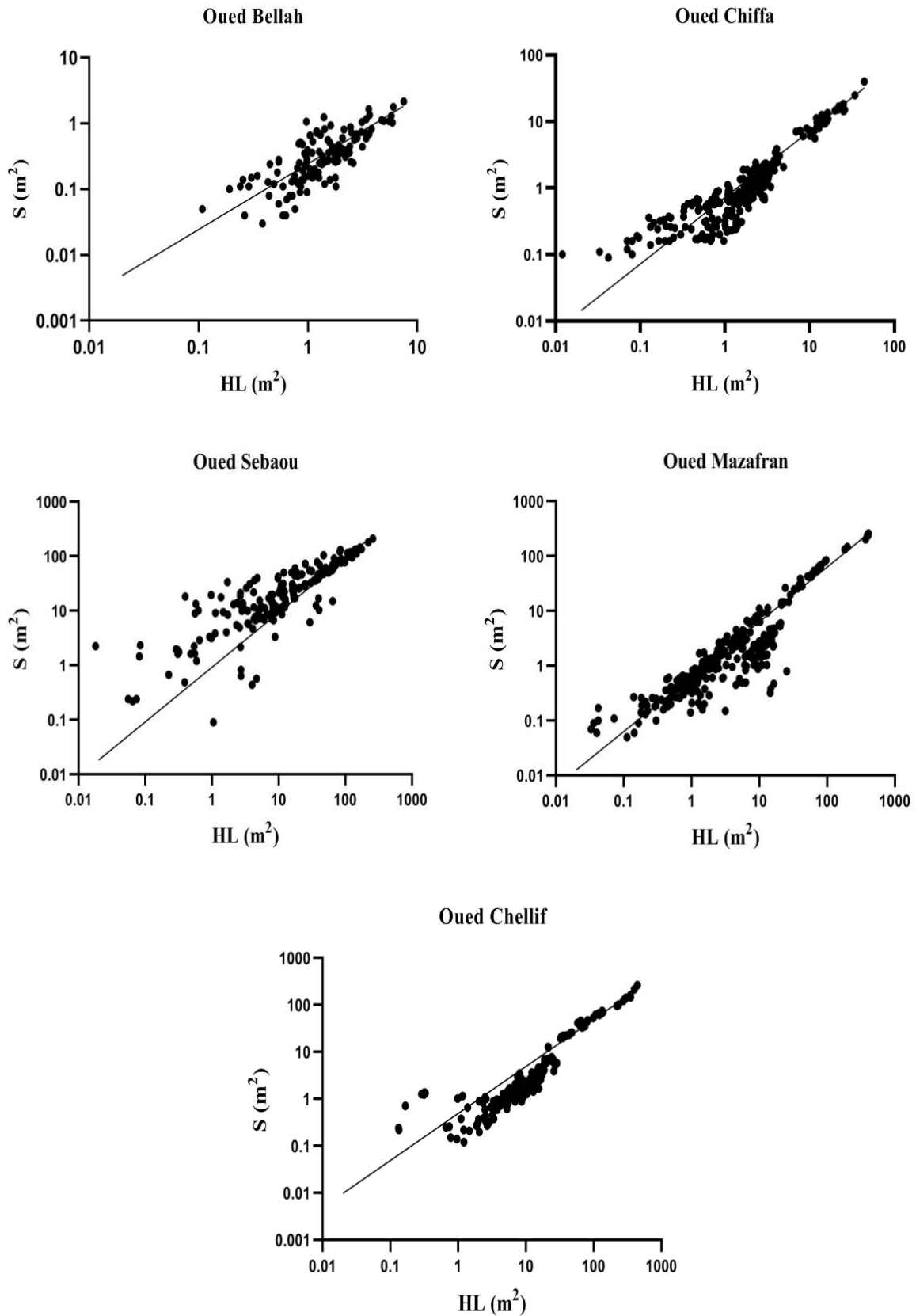




**Figure. VII. 5.** Relation entre la vitesse moyenne ( $U_m$ ) et la vitesse maximale ( $U_{max}$ ) obtenue par la méthode de Chiu (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).

**Tableau. VII. 2.** Valeurs de  $\Phi$ , M et R pour chaque station étudiée.

Oued	$\phi$	M	R
Chellif	0,6053	1,3	0,849
Chiffa	0,5579	0,712	0,882
Mazafran	0,5774	0,94	0,860
Sebaou	0,5486	0,59	0,893
Ballah	0,5890	1,1	0,845



**Figure. VII. 6.** Relation entre la surface de la section transversale et le produit de la hauteur d'eau (H) et la largeur du canal (L) par la méthode de Chiu (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).

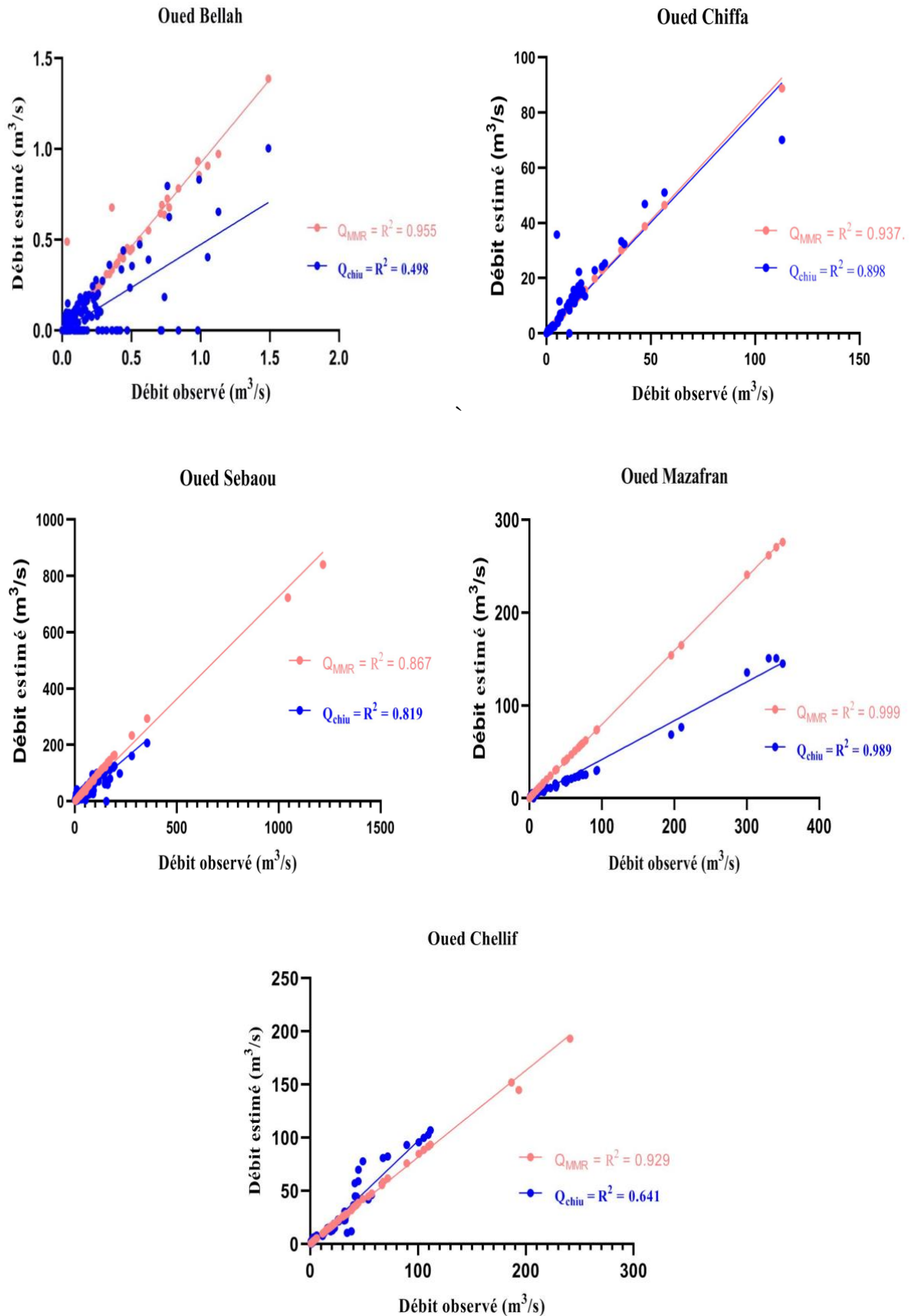
**Tableau. VII. 3.** Expressions des sections estimées pour l'ensemble des stations étudiées selon la méthode de Chiu.

Oued	Expression	R
Chellif	$S = 0,2507(HL)^{1,04}$	0,841
Chiffa	$S = 0,6338(HL)^{0,873}$	0,828
Mazafran	$S = 0,5353(HL)^{0,867}$	0,811
Sebaou	$S = 3,8494(HL)^{0,658}$	0,706
Ballah	$S = 0,2333(HL)^{0,8955}$	0,576

### VII.2.3. Comparaison entre les méthodes MMR et CHIU

Les représentations des débits calculés ( $Q_{calc}$ ) par rapport aux débits mesurés ( $Q_{mes}$ ) pour les méthodes de Chiu et MMR sont illustrées à la Figure VII.7. Les deux méthodes ont démontré une grande précision pour les oueds Mazafran et Chiffa, avec un coefficient de corrélation R excédant 0,89 pour la méthode Chiu et atteignant 0,93 pour la méthode MMR. Bien que, la MMR a obtenu de meilleurs résultats par rapport à la méthode Chiu pour chaque oued étudié. Les résultats obtenus à partir de l'application de la MMR pour l'estimation des débits dans les cours d'eau naturels sont très encourageants et elle sera très indispensable en particulier pour les bassins non jaugés situés en climat semi-aride.

À l'avenir, les débits optimaux calculés devraient être étudiés plus avant lorsque davantage de données seront disponibles. Une étude plus approfondie des débits mesurés et calculés devrait être menée.



**Figure. VII. 7.** Relation débits observés et débits estimés pour les deux méthodes CHIU et MMR (a) St. Ballah, b) St. Chiffa, c) St. Sebaou, d) St. Mazafran et f) St. Chellif).

#### VII.2.4. Le critère de Nash et Sutcliffe

La précision des résultats obtenus par la MMR et la méthode de Chiu a été évaluée en comparant les valeurs de débit calculées ( $Q_{\text{calc},i}$ ) et mesurées ( $Q_{\text{mes},i}$ ) à travers un critère de qualité. Plus précisément, le critère de Nash et Sutcliffe (1970) en se basant sur la somme des erreurs quadratiques.

Le tableau VII.4 montre les performances des méthodes en termes de NSE. Les performances étaient excellentes pour la méthode MMR, résultant en  $\text{NSE} > 95$  pour l'ensemble des stations. Par rapport à la méthode MMR, les performances de la méthode Chiu étaient inférieures pour toutes les stations étudiées, soit  $\text{NSE} < 70\%$  pour les stations Oued Chellif et Oued Mazafran.

**Tableau. VII. 4.** Résultats du critère de NSE pour les modèles obtenus par les méthodes (Chiu et MMR).

Cours d'eau	Méthode	NSE	
Oued Sebou	Chiu	91,45%	Excellent
	MMR	97,19%	Excellent
Oued Chellif	Chiu	67,11%	Satisfaisant
	MMR	96,56%	Excellent
Oued Chiffa	Chiu	93,52%	Excellent
	MMR	96,12%	Excellent
Oued Mazafran	Chiu	64,03%	Satisfaisant
	MMR	95,56%	Excellent
Oued Ballah	Chiu	81,19%	Satisfaisant
	MMR	95,19%	Excellent

### VII.3. Exemple d'application (Méthode MMR et CHIU)

Dans le contexte de l'estimation du débit des cours d'eau naturels, explorons de manière concrète, de la station amont des gorges de l'Oued Chiffa comme un exemple d'application pour une meilleure compréhension des méthodes utilisées et appliquées dans cette thèse. Les données de jaugeage, soigneusement collectées par l'Agence Nationale des Ressources Hydrauliques (ANRH), constituent la base pour l'application des diverses méthodes et modèles, tels que la méthode du modèle rugueux de référence (MMR), Chiu, les modèles d'intelligence artificielle, et le modèle SVM hybridé.

À titre d'illustration, la méthode MMR et le modèle Chiu sont des modèles empiriques. Les modèles d'intelligence artificielle s'appuient sur des algorithmes avancés. Huit modèles distincts (SVM, LIBSVM, ANN, REP TREE, ANFIS, M5p, GPR, Random Forest) sont employés pour analyser les données de jaugeage et déterminer le débit correspondant. En parallèle, le modèle SVM hybridé fusionne les avantages des machines à vecteurs de support (SVM) avec d'autres techniques ensemblistes (Dagging, Bagging, Rotation Forest, Random subspace) et cela dans le but de bien fournir une estimation robuste des débits liquides.

Dans cet exemple concret, on met en lumière l'application pratique de la diversité des approches utilisées dans la détermination de ces débits des cours d'eau naturels étudiés, contribuant ainsi à une compréhension approfondie de la dynamique hydraulique des Oueds.

**- Méthode MMR**

Dans le cadre de la méthode MMR, les paramètres tels que la section, le rayon hydraulique et la pente sont déduits à partir de diverses mesures de jaugeage en vue d'estimer le débit. De plus, ce dernier est évalué à l'aide des expressions suivantes :

1. **Nombre de Reynolds** 
$$\bar{R} = 32\sqrt{2} \frac{\sqrt{igR_h^3}}{\nu}$$
2. **Coefficient de Manning Strickler** 
$$k_s = \frac{\sqrt{igR_h^3}}{\nu}$$
3. **Rugosité absolue** 
$$\varepsilon = \left(\frac{8,2\sqrt{g}}{K_s}\right)^6$$
4. **Expression du débit** 
$$Q = -4\sqrt{2}A\sqrt{igR_h} \log\left(\frac{\varepsilon/R_h}{14,8} + \frac{10,04}{\bar{R}}\right)$$

Ou  $R_h$  est le rayon hydraulique,  $i$  est la pente du cours d'eau,  $\nu$  et la viscosité de l'eau  $\nu = 10^6$ ,  $\bar{R}$  est le Nombre de Reynolds du modèle rugueux,  $\varepsilon$  est la rugosité.

**Tableau. VII. 5.** Représentation des données observées par l'ANRH et les résultats obtenus par l'application de la méthode MMR "Oued Chiffa".

Données					Résultats			
Date	Section (m <sup>2</sup> )	Rayon hydraulique (m)	Q <sub>mes</sub> (m <sup>3</sup> /s)	Pente %	Ks	ε	$\bar{R}$	Q <sub>cal</sub> (m <sup>3</sup> /s)
07/04/1997	0,27	0,29	0,06	0,0038	76,669	0,00141	134688,254	0,053
22/03/1999	2,72	0,48	2,02	0,0189	88,037	0,00061	648610,932	1,833

09/06/1999	0,52	0,24	0,36	0,0394	89,798	0,00054	335139,228	0,338
06/07/1999	0,40	0,12	0,15	0,0310	85,382	0,00074	103877,665	0,141
01/08/1999	0,33	0,18	0,14	0,0249	85,704	0,00072	171149,346	0,137
15/08/1999	0,23	0,14	0,11	0,0004	88,072	0,00061	154499,516	0,107
03/10/1999	0,16	0,13	0,07	0,0003	87,127	0,00065	132572,494	0,070
18/10/1999	0,50	0,19	0,26	0,0003	87,313	0,00064	212756,465	0,244
27/12/1999	1,46	0,42	1,32	0,0003	90,716	0,00051	684667,782	1,194
19/01/2000	1,12	0,24	0,59	0,0002	86,754	0,00067	260718,709	0,552

**- Méthode de Chiu**

Dans le cadre de la méthode Chiu, la vitesse, le débit, la hauteur et la largeur sont des paramètres essentiels utilisés pour différentes mesures de jaugeage, contribuant ainsi à l'estimation du débit.

**1. Relation entre la vitesse moyenne et maximale**

La relation entre la vitesse moyenne ( $U_m$ ) et la vitesse maximale ( $U_{max}$ ) dans une section de mesure peut se mettre sous la forme :

$$U_m = \Phi(M)U_{max}$$

Où  $\Phi(M)$  est une fonction du paramètre d'entropie M.

**Tableau. VII. 6.** Valeurs de  $\Phi$  à "Oued Chiffa".

Oued	$\phi$
Chellif	0,6053

**2. Détermination de la surface de la coupe transversale**

$$S = a_1 (HL - b_1)^{C1}$$

H est la profondeur de l'eau, L est la largeur du chenal à la surface de l'eau,  $a_1, a_2, a_3, b_1$ , sont des coefficients qui peuvent être déterminés par analyser les données.

**Tableau. VII. 7.** Expression de section estimée par la méthode de Chiu à "Oued Chiffa".

Oued	Expression	R
Chiffa	$S = 0,6338(HL)^{0,873}$	0,828

### 3. Expression du débit Chiu

Étant donné que,  $U_m$  peut être calculée à partir de  $U_{max}$  et  $S$  peut être estimée à partir de  $H$  et  $L$ , Donc le débit peut être facilement évalué :

$$Q_{Chiu} = U_m S = \Phi(M) U_{max} S$$

**Tableau. VII. 8.** Représentation des données observées par l'ANRH et les résultats obtenus par l'application de la méthode Chiu à "Oued Chiffa".

Données						Résultats			
Date	Section (m <sup>2</sup> )	Q <sub>mes</sub> (m <sup>3</sup> /s)	Largeur (m)	Hauteur (cm)	Vitesse Max	Débit / sections	Hauteur×Largeur	S <sub>est</sub>	Q <sub>cal</sub>
07/04/1997	0,27	0,06	2,10	19	0,43	0,219	0,399	0,2680	0,0645
22/03/1999	2,72	2,02	8,00	51	0,85	0,472	4,08	2,7407	1,3052
09/06/1999	0,52	0,36	2,80	28	0,71	0,494	0,784	0,5266	0,2095
06/07/1999	0,40	0,15	2,60	23	0,57	0,438	0,598	0,4017	0,1282
01/08/1999	0,33	0,18	2,50	20	0,43	0,282	0,5	0,3358	0,0809
15/08/1999	0,23	0,11	1,70	20	0,59	0,474	0,34	0,2283	0,0755
03/10/1999	0,16	0,07	1,30	19	0,50	0,427	0,247	0,1659	0,0464
18/10/1999	0,50	0,26	3,00	25	0,62	0,488	0,75	0,5038	0,1750
27/12/1999	1,46	1,32	4,30	51	0,92	0,817	2,193	1,4731	0,7593
19/01/2000	1,12	0,59	4,30	39	0,82	0,494	1,677	1,1265	0,5175

## VII.4. Modèle d'intelligence artificielle

Dans cette étude, nous avons évalué plusieurs algorithmes d'apprentissage automatique pour la prédiction du débit des cours d'eau naturels, notamment ANN, ANFIS, SVM, LibSVM, Rep Tree, GPR, RF et M5P. Ces modèles ont été vérifiés par leurs performances pour les phases d'entraînement et de test.

En phase d'entraînement (Tableau VII.9), on peut observer que l'ensemble des modèles ont donné des résultats satisfaisants pour toutes les stations d'étude, les modèles RF, ANN et ANFIS offrent approximativement des débits simulés plus performants que les modèles, M5P, Rep Tree et SVM, LibSVM et GPR; à titre d'illustration, pour la station 'oued Sebaou', le

modèle ANFIS a donné un coefficient de corrélation R très élevé (0,999), un bon coefficient d'efficacité NSE (0,999) et une faible erreur quadratique moyenne RMSE (0,004) et même pour l'erreur absolue moyenne MAE (0,003). Donc s'est avéré être le meilleur modèle pour cette station. Alors que l'ANN est devenu l'excellent dans la station 'Oued Chiffa', où R est très fort (0,999), RMSE très bon (0,004) et MAE faible (0,001) et NSE idéal (0,999). Le modèle RF a également donné des meilleures valeurs de performance (R, NSE, RMSE, MAE) par rapport aux autres modèles pour les stations 'Oued Chellif', 'Oued Bellah' et 'Mazafran', sachant que les autres modèles tels que M5P, Rep Tree, SVM, et LibSVM ont également affiché des performances remarquables, tandis que le modèle GPR a présenté des résultats moins satisfaisants pour l'ensemble des stations étudiées.

Une tendance identique est également constatée pour la phase de test, où les modèles RF, ANN et ANFIS ont obtenu de bien meilleurs résultats que les modèles, Rep Tree, M5P, SVM, LibSVM et GPR sur tous les indicateurs d'efficacité effectués. En gardant le même exemple de stations qu'en phase d'apprentissage, l'application du modèle ANFIS à la station 'oued Sebaou' a donné le meilleur résultat avec un très fort coefficient de corrélation R (0,968), un MAE (0,003), un NSE (0,905) et un RMSE de 0,005, alors que l'utilisation du modèle ANN dans la station 'Oued Chiffa' a obtenu d'excellent résultat avec R=0,991, NSE à 0,978 et RMSE diminuait (0,006). Cependant, RF a fourni les valeurs de performance la plus précise de (R, MAE, NSE, RMSE), ces valeurs confirment l'excellence de ce modèle Pour les Stations 'Oued Chellif', 'Oued Bellah' et 'Oued Mazafaran' successivement.

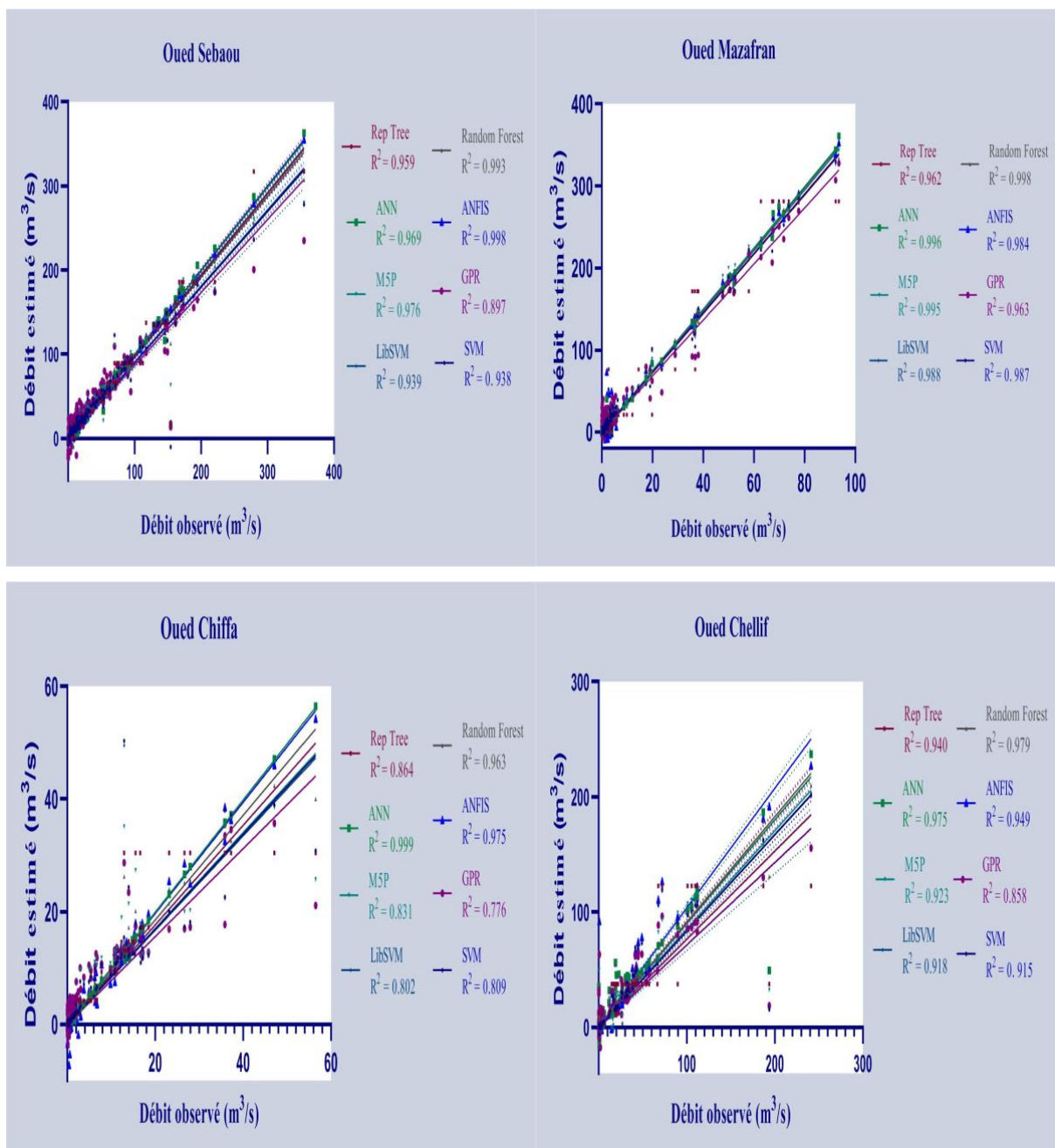
**Tableau. VII. 9.** Indicateurs de performance des modèles (ANN, ANFIS, SVM, LibSVM, Random Forest, GPR, Rep Tree, M5P) optimaux sur les périodes d'apprentissage et de test pour les cinq stations étudiées.

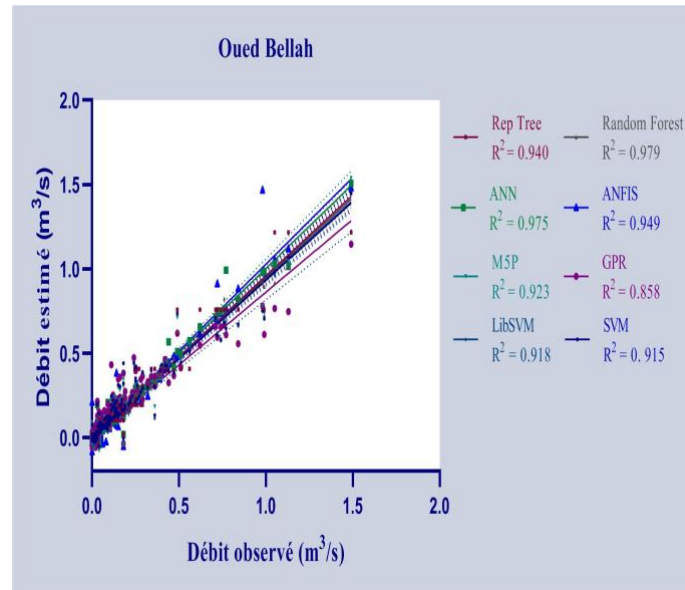
Station	Model	Apprentissage 70%				Test 30%			
		R	MAE	NSE	RMSE	R	MAE	NSE	RMSE
Oued Sebaou	ANN	0,983	0,008	0,966	0,023	0,960	0,003	0,927	0,0046
	ANFIS	0,999	0,003	0,999	0,004	0,968	0,003	0,905	0,005
	SVM	0,967	0,012	0,930	0,039	0,933	0,006	0,708	0,009
	Lib SVM	0,967	0,012	0,932	0,038	0,934	2,280	0,709	3,299
	Random Forest	0,996	0,0051	0,991	0,0144	0,993	0,0014	0,985	0,0021
	GPR	0,950	0,029	0,885	0,0503	0,774	0,0022	-1,527	0,0274
	Rep Tree	0,978	0,0115	0,957	0,0308	0,936	0,0034	0,861	0,0064
	M5P	0,987	0,0099	0,974	0,024	0,975	0,0026	0,944	0,004

<b>Oued Chiffa</b>	<b>ANN</b>	0,999	0,001	0,999	0,004	0,991	0,0004	0,978	0,0006
	<b>ANFIS</b>	0,992	0,010	0,985	0,018	0,126	0,018	-35,49	0,027
	<b>SVM</b>	0,889	0,019	0,790	0,068	0,862	0,002	0,330	0,003
	<b>Lib SVM</b>	0,890	0,019	0,791	0,068	0,852	0,002	0,308	0,003
	<b>Random Forest</b>	0,981	0,008	0,959	0,030	0,665	0,002	0,319	0,003
	<b>GPR</b>	0,888	0,044	0,753	0,074	0,796	0,012	-11,10	0,015
	<b>Rep Tree</b>	0,925	0,017	0,856	0,056	0,831	0,002	0,667	0,002
	<b>M5P</b>	0,906	0,023	0,821	0,063	0,822	0,002	0,341	0,003
<b>Oued Chellif</b>	<b>ANN</b>	0,929	0,017	0,860	0,050	0,979	0,008	0,954	0,012
	<b>ANFIS</b>	0,997	0,006	0,993	0,010	0,904	0,019	0,011	0,058
	<b>SVM</b>	0,913	0,010	0,825	0,056	0,911	0,017	0,068	0,056
	<b>Lib SVM</b>	0,914	0,010	0,825	0,056	0,912	0,0172	0,110	0,055
	<b>Random Forest</b>	0,984	0,007	0,962	0,025	0,996	0,003	0,989	0,005
	<b>GPR</b>	0,888	0,031	0,740	0,068	0,920	0,029	0,425	0,044
	<b>Rep Tree</b>	0,867	0,019	0,744	0,067	0,931	0,013	0,528	0,040
	<b>M5P</b>	0,923	2,774	0,850	12,537	0,917	0,008	0,755	0,028
<b>Oued Bellah</b>	<b>ANN</b>	0,989	0,022	0,977	0,038	0,983	0,023	0,966	0,040
	<b>ANFIS</b>	0,997	0,019	0,994	0,019	0,952	0,108	0,751	0,108
	<b>SVM</b>	0,958	0,045	0,915	0,074	0,965	0,037	0,910	0,065
	<b>Lib SVM</b>	0,960	0,045	0,920	0,071	0,962	0,043	0,908	0,066
	<b>Random Forest</b>	0,990	0,018	0,976	0,012	0,988	0,015	0,975	0,034
	<b>GPR</b>	0,933	0,069	0,837	0,102	0,930	0,083	0,854	0,083
	<b>Rep Tree</b>	0,970	0,003	0,941	0,062	0,974	0,029	0,940	0,053
	<b>M5P</b>	0,965	0,046	0,929	0,068	0,963	0,944	0,905	0,067
<b>Oued Mazafren</b>	<b>ANN</b>	0,998	0,002	0,997	0,0081	0,686	0,00383	0,0400	0,011
	<b>ANFIS</b>	0,998	0,0033	0,998	0,006	0,3020	0,0175	-7,954	0,033
	<b>SVM</b>	0,996	0,006	0,992	0,015	0,500	0,0096	-2,0539	0,1096
	<b>Lib SVM</b>	0,996	0,007	0,992	0,015	0,525	0,009	-1,882	0,019
	<b>Random Forest</b>	0,999	0,003	0,998	0,007	0,878	0,0033	0,662	0,0065
	<b>GPR</b>	0,985	0,023	0,963	0,034	0,239	0,024	-6,678	0,319
	<b>Rep Tree</b>	0,983	0,012	0,9672	0,0323	0,500	0,0074	-2,4667	0,0209
	<b>M5P</b>	0,998	0,006	0,996	0,0105	0,648	0,0061	0,298	0,0094

La figure VII.8. Présente une comparaison détaillée des débits observés et prédites pour les cinq stations d'étude fournies par les huit modèles d'intelligence artificielle appliqués.

Le degré de dispersion des points de données est un indicateur de la qualité de la prédiction : plus les points de données ce sont concentrés sur la courbe, la performance de prédiction est meilleure. Les modèles ANFIS, ANN, RF, SVM, LibSVM et M5P montrent un excellent degré d'ajustement, tandis que les modèles GPR et Rep Tree présentent un degré de dispersion inférieur par rapport aux autres modèles. Le modèle Random Forest (RF) a les meilleures performances de prédiction pour les stations (Mazafran, Chellif et Bellah), L'ANFIS pour la station Oued Sebau et L'ANN pour la station Oued Chiffa surpassent les autres modèles.





**Figure. VII. 8.** Relation débits observés et débits simulés (prévues) pour les cinq cours d'eau naturels étudiés (a- St. Sebaou, b- St. Mazafran, c- St. Chiffa, d- St. Chellif, e- St. Bellah).

**Le diagramme de Taylor** (Taylor 2001) est utilisé pour représenter visuellement les mesures d'efficacité des résultats de la modélisation en traçant un ensemble des points sur un tracé polaire. Ce diagramme a bien mis en évidence la variation spatiale du débit attendu par les modèles par rapport à la valeur observée pour les stations étudiées.

Le diagramme de Taylor définit l'écart type (SD) entre les valeurs observées et prédites le long des intervalles radiaux avec des racines, tandis que les valeurs R sont déterminées par les angles de direction. Le diagramme suppose que les valeurs observées sont affichées indépendamment et que les performances du modèle sont meilleures lorsque les indicateurs de performance de prédiction sont plus proches des valeurs observées.

La figure VII.9 montre que, sur les cinq stations d'étude (Oued Bellah, Mazafran, Sebaou, Chiffa et Chellif), le modèle Random Forest donne des résultats plus étroitement alignés sur les données observées par rapport aux modèles dans les stations (Chellif, Bellah et Mazafaran) tandis que le modèle ANN et ANFIS ont également montrés leurs excellences performances sur le Chiffa' et le Sebaou. Bien que, le modèle GPR a présenté les mauvais résultats obtenus en termes de précision de prédiction.

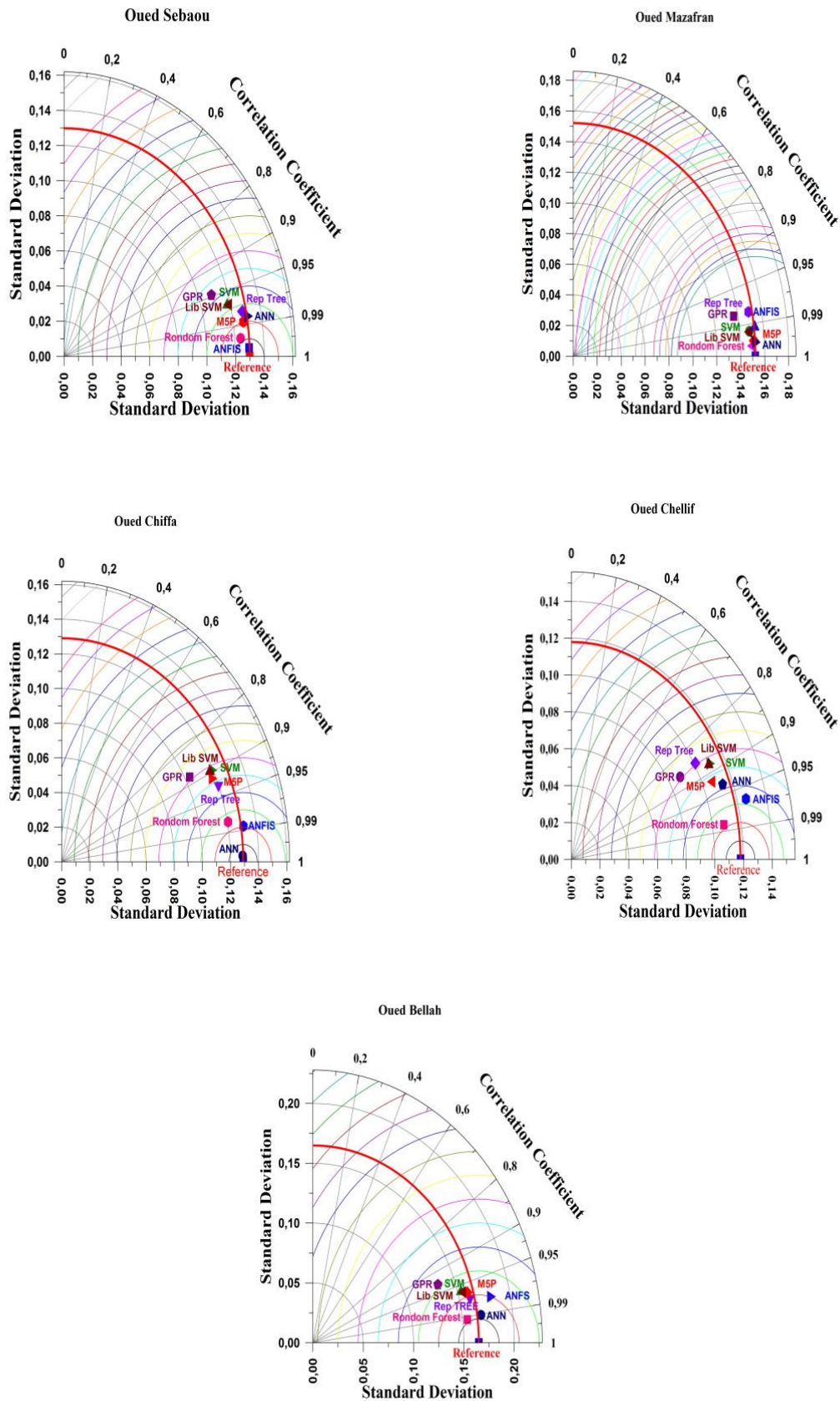


Figure. VII. 9. Diagrammes de Taylor affichant une comparaison statistique des huit modèles pour les cours d'eau naturels des oueds (a) Sebaou, b) Mazafran, c) Chiffa, d) Chellif, e) Bellah).

Les modèles ANN, RF et ANFIS sont des choix appropriés pour traiter de données complexes avec des interactions non linéaires entre les caractéristiques. Les réseaux de neurones artificiels (ANN) sont flexibles dans leur conception d'architecture, ce qui permet de capturer des relations complexes dans les données. Les forêts aléatoires (RF) sont efficaces pour gérer de données bruyantes et des interactions non linéaires, ce qui est important pour la prédiction du débit. L'ANFIS fusionne les réseaux de neurones artificiels et les systèmes de logique floue afin de réaliser des prédictions précises des débits, en capturant de manière efficace les relations complexes entre les caractéristiques. Cependant, le modèle GPR peut être moins performant avec les données de jaugeage des cours d'eau naturels du cas d'étude en raison de son hypothèse de distribution normale. D'autres modèles, tels que M5P, Rep Tree, SVM et LibSVM peuvent offrir des performances satisfaisantes, mais peuvent être moins adaptés aux données complexes, comme celles des cours d'eau naturels. Le choix du modèle dépend des caractéristiques des données, des interactions entre les variables, des hypothèses du modèle, des performances requises et des contraintes de temps et des ressources. Les modèles ANN et RF sont appropriés pour les données de jaugeage des cours d'eau naturels du cas étudié en raison de leur capacité à traiter de données complexes avec des interactions non linéaires, mais il est important de comprendre leurs avantages et limites avant de faire ce choix.

#### **VII.4.1. Modélisation des débits par l'intégration des données hydrométriques et l'algorithmes d'apprentissage automatique**

Dans cette partie nous avons combiné, les données des débits de cinq stations hydrométriques situées dans le nord de l'Algérie pour développer un modèle d'apprentissage automatique global capable de prédire les débits des cours d'eau en Algérie. Nous avons commencé par collecter et vérifier les données de chaque station pour éliminer les erreurs et les valeurs aberrantes. Ensuite, nous avons utilisé une méthode de fusion de données pour combiner les séries temporelles de chaque station en une seule série composite. Nous avons utilisé différents modèles d'apprentissage automatique, tels que (ANN, ANFIS, SVM, LibSVM, M5P, Rep Tree, GPR et Random Forest), pour effectuer la prédiction. Enfin, nous avons validé notre modèle en utilisant la validation croisée et en mesurant les performances avec des indicateurs, tels que R, RMSE, NSE, MAE et diagramme de Taylor.

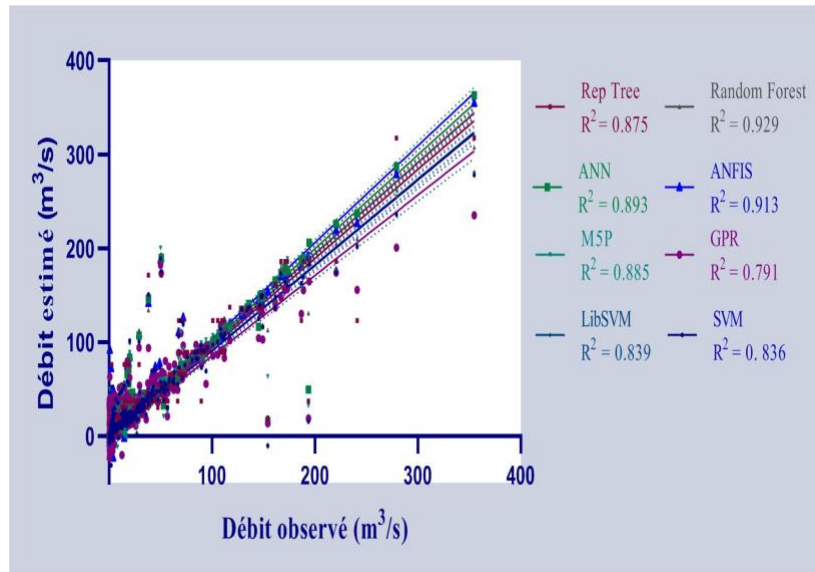
Dans la phase d'entraînement (Tableau VII.10), tous les modèles d'ensemble ont principalement surpassé les modèles GPR, qui ont montré les valeurs NSE et R les plus basses (0,741 et 0,438) et la plus grande erreur de modélisation (RMSE = 4,68 et MAE=5,461). Le modèle RF a montré les meilleures performances et R = 0,885, MAE = 2,012, NSE = 0,695 et RMSE =9,088, suivi du ANFIS, et les autres modèles.

En termes de performances de validation, une fois de plus, le modèle GPR autonome a montré les performances les plus faibles avec les valeurs R et NSE les plus faibles (0,948 et 0,885) et la plus grande erreur de modélisation (NSE = 9,708 et RMSE = 15,502). Le modèle RF a montré la prédiction la plus précise avec R = 0,997, MAE = 1,679, NSE= 0,99 et RMSE = 4,348, suivi les autres modèles étudiés.

**Tableau. VII. 10.** Indicateurs de performance des modèles : ANN, ANFIS, SVM, LibSVM, Random Forest, GPR, Rep Tree, M5P, optimaux sur les périodes d'Apprentissage et de Test pour l'ensemble des cours d'eau étudiés.

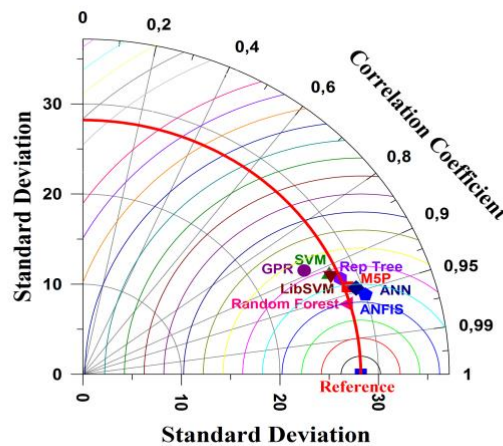
Station	Modèle	Apprentissage 70%				Test 30%			
		R	MAE	NSE	RMSE	R	MAE	NSE	RMSE
Toutes les Stations	ANN	0,844	2,517	0,605	10,337	0,985	2,552	0,971	7,787
	ANFIS	0,875	2,395	0,639	9,881	0,991	2,418	0,980	6,323
	SVM	0,791	2,786	0,491	11,733	0,996	3,766	0,9141	11,047
	Lib SVM	0,793	2,791	0,499	11,645	0,972	3,815	0,943	10,932
	Random Forest	0,885	2,012	0,695	9,088	0,997	1,679	0,990	4,348
	GPR	0,741	5,461	0,438	12,336	0,948	9,708	0,885	15,502
	Rep Tree	0,770	2,913	0,490	11,749	0,982	3,131	0,964	8,583
	M5P	0,817	2,544	0,559	10,917	0,989	2,598	0,979	6,575

La figure VII.10. Présente une comparaison détaillée entre les valeurs de débit observées et prévues. La tendance linéaire observée pour RF-ressemblait étroitement à la ligne  $y=x$ , avec un coefficient de corrélation R de 0,929. D'autre part, les tendances pour M5P, Rep Tree, GPR, ANFIS, ANN, LibSVM et SVM ont montré un R variable, compris entre 0.791 et 0,913.



**Figure. VII. 10.** Relation globale de débits observés et débits prédits pour l'ensemble des stations appliquées aux différents modèles d'intelligence artificielle

La figure VII.11. Montre que le modèle Random Forest donne des résultats plus étroitement alignés sur les données observées par rapport aux autres modèles. Bien que, le modèle GPR a présenté les mauvais résultats en termes de précision de prédiction.



**Figure. VII. 11.** Diagrammes de Taylor : Comparaison statistique des modèles (toutes les stations).

Le Random Forest est le modèle le plus performant dans cette étude, grâce à ses caractéristiques uniques, telles que la réduction du biais et de l'erreur de surajustement, la capacité à gérer les variables catégoriques et numériques et à évaluer l'importance des variables. Le réseau de neurones est également un algorithme d'apprentissage automatique avancé qui peut modéliser les relations complexes entre les entrées et les sorties, ce qui peut être plus adapté à de données complexes, telles que la production des débits d'eau naturelle. Le réseau de neurones artificiels ANN peut également traiter les données manquantes et les données

omises, ce qui peut améliorer la qualité de la prédiction. En revanche, le SVM peut être moins adapté à de données complexes ou à des relations complexes entre les entrées et les sorties et peut être plus sensible à l'erreur de surajustement en raison de sa capacité limitée à gérer les données manquantes et les données omises.

**VII.4.2. Caractéristiques des modèles d'algorithmes d'apprentissage pour la prédiction de débit dans le Nord de l'Algérie**

La section dédiée à l'exploration détaillée des divers modèles algorithmiques déployés dans le cadre de la prédiction du débit dans le nord de l'Algérie dans une région spécifique offre une analyse approfondie. Chaque modèle, sélectionné avec soin pour son adéquation au contexte hydrologique, est minutieusement décrit en termes d'architecture, de méthode d'apprentissage, de noyau utilisé, de paramètres du noyau et de paramètres spécifiques au modèle. L'inclusion de ces détails cruciaux vise à établir une base solide pour la compréhension et l'interprétation des résultats obtenus par ces modèles d'apprentissage automatique. Cette section représente une ressource essentielle pour les chercheurs et les praticiens du domaine, leur fournissant un aperçu complet des choix algorithmiques opérés dans le contexte de la modélisation du débit. Elle facilite également une analyse approfondie des performances et des implications de chaque modèle, contribuant ainsi à enrichir la compréhension globale des processus hydrologiques dans la région étudiée.

**Tableau. VII. 11.** Caractéristiques des modèles d'algorithmes d'apprentissage automatique utilisés pour la prédiction de débit dans le Nord de l'Algérie.

Modèle	Architecture	Méthode d'apprentissage	Noyau Utilisé	Paramètres du Noyau	Paramètres spécifiques du modèle
<b>Gaussian Processes Processus (GPR)</b>	Non-paramétrique	Gaussian Processus Regression	Linear Kernel: $K(x, y) = x^T y$	Length scale ( $\gamma$ ) et Noise ( $\sigma^2$ ) pour un noyau linéaire	-L 1.0 : Paramètre de l'échelle de longueur. -N 0 : Paramètre de bruit. -E 1.0 -C 250007" : Spécification du noyau. Il s'agit d'un PolyKernel avec un exposant de 1.0 et un paramètre de coût de 250007. -S 1 : Graine pour la génération de nombres aléatoires.
<b>Support Vector Machien (SVM)</b>	Optimisation minimale séquentielle pour la régression	Support vector Machien	PolyKernel	Utilisation du PolyKernel . -C 1.0: Paramètre de coût C du SVM (Support Vector Machine). -E 1.0: Exposant du noyau polynomial. -C 250007 : Paramètre de coût pour le noyau polynomial.	Options internes spécifiques au modèle SMOreg, y compris les paramètres de réglage pour la formation. Nombre maximal d'itérations : 0.001 Utilisation de la validation croisée (-V) Paramètre de tolérance (-P) : 1.0E-12 Paramètre de seuil de tolérance (-L) : 0.001 Taille du cache pour les exemples (-W) : 1-S 1 : Graine pour la génération de nombres aléatoires.

<b>LIBSVM</b>	Support Vector	LibSVM Machine à vecteurs de support (wrapper LIBSVM)	SVM Linéaire Kernel	Modèle de régression (epsilon-SVR). -Noyau linéaire. -Degré du noyau polynomial fixé à 3. -paramètre Gamma du noyau mis à 0.0. -paramètre Coef0 du noyau mis à 0.0.	Grande bibliothèque de noyau, flexibilité -paramètre Nu réglé à 0,5. -taille de la mémoire cache définie sur 40 Mo. -paramètre de coût défini sur 1,0. -paramètre Epsilon réglé à 0,001. -Epsilon pour le critère de terminaison fixé à 0,1. -Spécification du fichier modèle (non complet dans le texte fourni). - graine aléatoire définie sur 1
<b>M5P</b>	Arbre de Décision	Model M5 Tree	Modèles linéaires lissés utilisés	-Taillez l'arbre pour avoir au maximum 4 feuilles. -num-decimal-places 4 : affiche les valeurs numériques avec 4 décimales.	Pruning utilisé pour simplifier l'arbre
<b>Rep Tree</b>	Arbre de décision	Régression de l'arbre de décision	Noyau Utilisé (Kernel Used) : Non applicable (N/A)	Profondeur Max, Nombre Min Obj	-M 2 : Nombre minimum d'instances par feuille. -V 0,001 : Variance minimale pour la régression (utilisée pour l'élagage). -N 3 : Nombre minimum d'instances par feuille pour la régression (utilisé pour l'élagage). -S 1 : graine pour le brassage aléatoire des données. -L -1 : Si supérieur à 0, définit la valeur en dessous de laquelle une scission ne sera pas tentée (utilisée pour l'élagage). -I 0,0 : Nombre d'itérations de boosting et paramètre de trimming du poids (0 pour pas de trimming).
<b>Random Forest</b>	RandomTree	Ensemble Bagging	Noyau Utilisé (Kernel Used) : Non applicable (N/A)	Nombre d'arbres, profondeur, etc.	100 : Nombre de caractéristiques à prendre en compte lors de la recherche du meilleur partage. -I 100 : Nombre d'arbres (itérations) dans la forêt aléatoire. -num-slots 1 : Nombre d'emplacements d'exécution (traitement parallèle). -K 0 : Nombre de plis pour le backfitting (0 pour utiliser les données jusqu'au point actuel dans l'arbre en croissance). -M 1,0 : Nombre minimum d'instances par vantail et nombre de plis pour le retrofitting. -V 0,001 : Variance minimale pour la régression. -S 1 : graine pour le brassage aléatoire des données.
<b>Artificial Neural Networks (ANN)</b>	L'architecture est représentée par 20-20-1. Couche d'entrée : non spécifiée. Couche cachée : 20 neurones. Couche de sortie : 1 neurone	Feedforward Neural Network	Non applicable	Non applicable	Sigmoid

<b>Adaptive Neuro Fuzzy Inference System (ANFIS)</b>	Gaussian MFs (3)	Apprentissage hybride	Non applicable	Non applicable	Gaussmf, linear
--	------------------	-----------------------	----------------	----------------	-----------------

### VII.5. Modèles SVM-Hybridés

Dans cette étude, les modèles SVM autonomes et quatre modèles d'ensemble, à savoir SVM-Bagging, SVM-Dagging, SVM-RSS et SVM-RF\* ont été ainsi développés et leurs performances en termes de différentes mesures de performance pour les phases de formation et de test ont été évaluées. Lors de la phase de formation (tableau VII.12), tous les modèles d'ensemble ont largement surpassé les modèles SVM autonomes, qui ont affiché les valeurs NSE et R les plus faibles (0,67 et 0,89) et l'erreur de modélisation la plus importante (RMSE = 4,68). Le modèle SVM-Dagging a montré la meilleure performance de formation avec RMSE=2.43, NSE=0.91, et R=0.96, suivi par les modèles SVM-Bagging, SVM-RF\*, et SVM-RS.

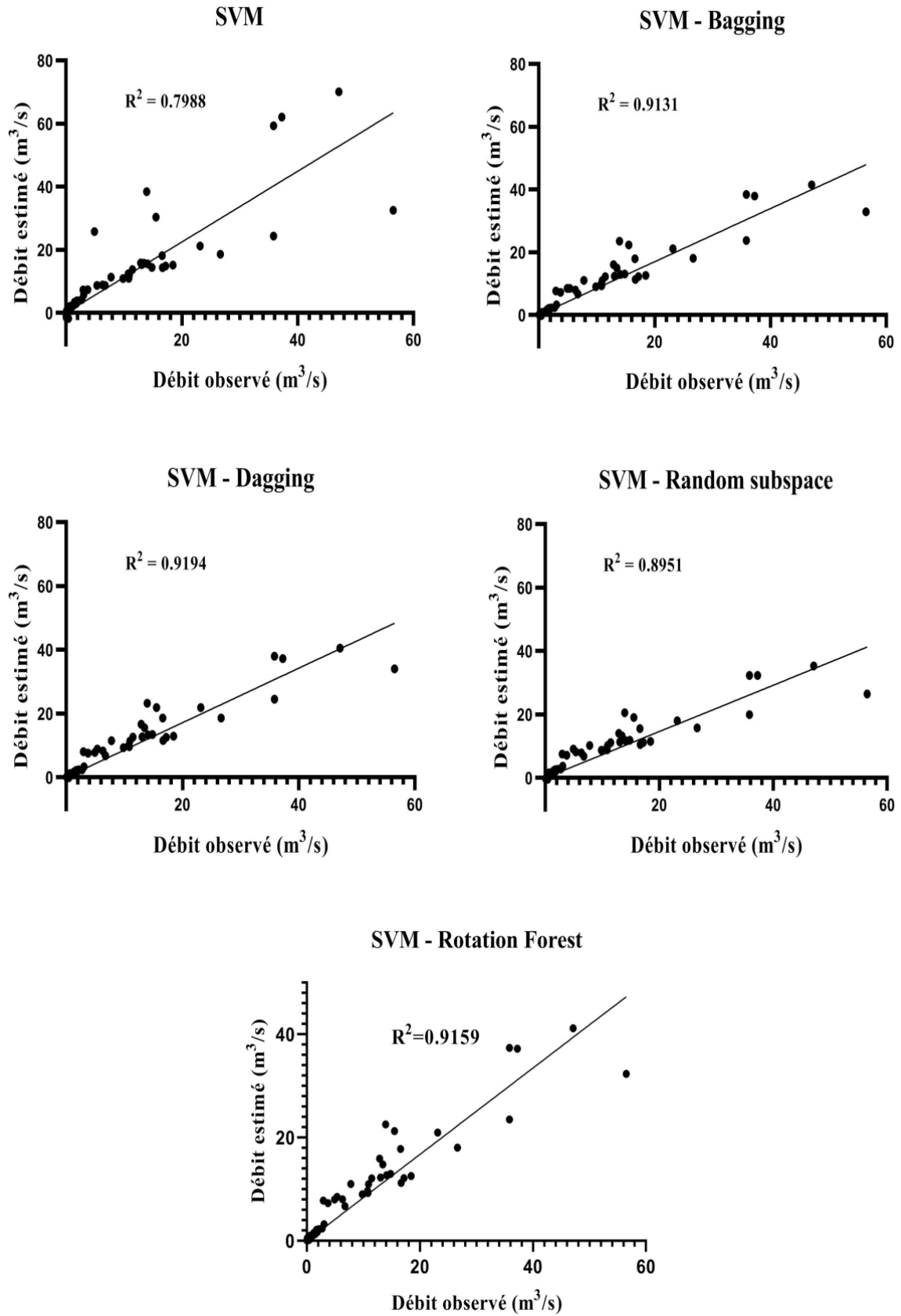
En termes de performance de validation, une fois de plus, le modèle SVM autonome a montré la performance la plus faible avec les valeurs NSE et R les plus basses (0,68 et 0,88) et l'erreur de modélisation la plus importante (RMSE= 7,71). Le modèle SVM-Dagging a montré la prédiction la plus précise avec un RMSE= 6,58, NSE=0,76 et R=0,96, suivi par les modèles SVM-Bagging, SVM-RF\* et SVM-RS.

**Tableau. VII. 12.** Performances du modèle au cours des phases de formation et de test.

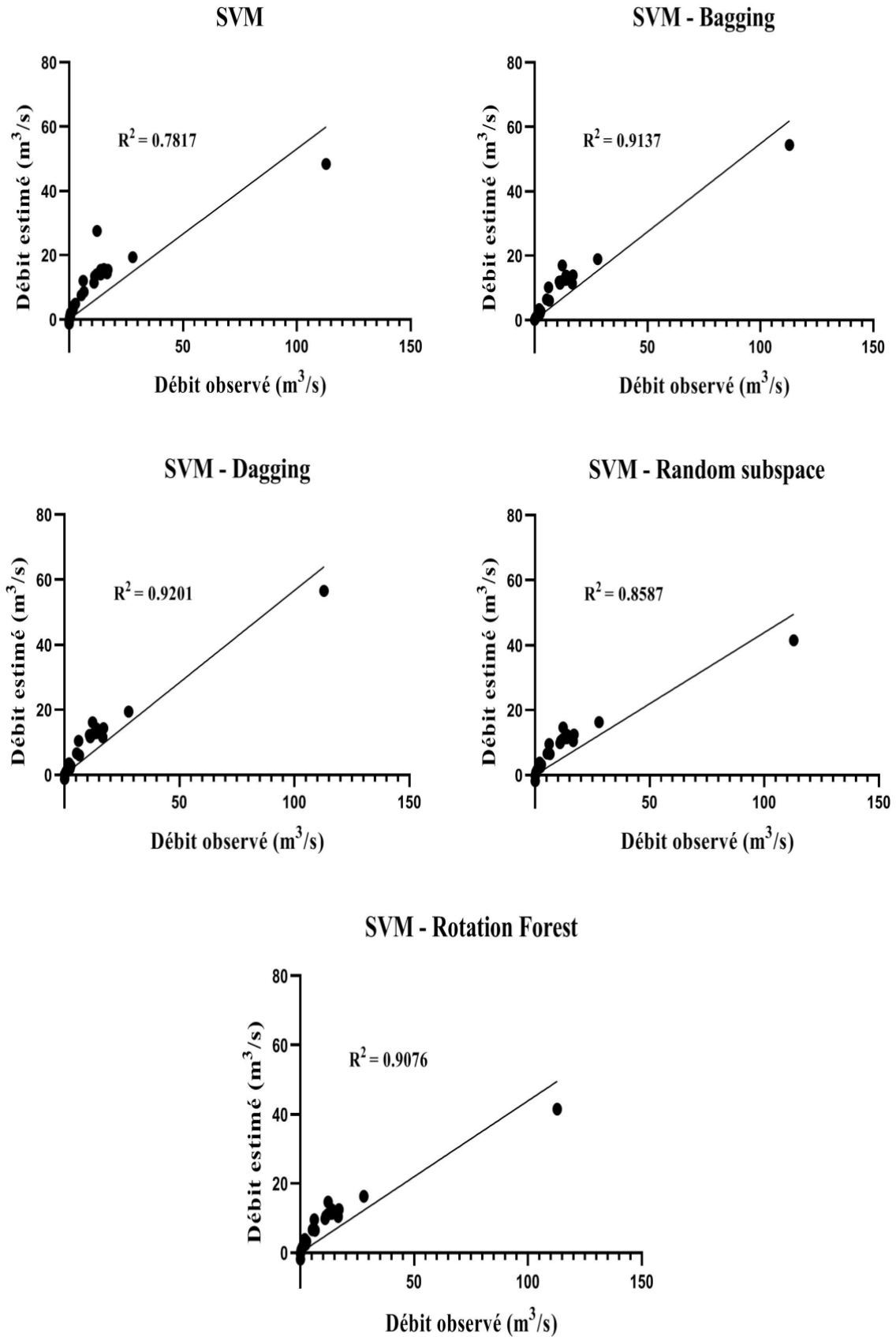
	Entraînement (70%)			Test (30%)		
	RMSE	NSE	R	RMSE	NSE	R
<b>SVM</b>	4.68	0.67	0.89	7.71	0.68	0.88
<b>SVM-Bagging</b>	2.52	0.91	0.96	6.83	0.75	0.96
<b>SVM-Dagging</b>	2.43	0.91	0.96	6.58	0.76	0.96
<b>SVM-RSS</b>	3.14	0.85	0.95	7.34	0.62	0.93
<b>SVM-RF*</b>	2.52	0.91	0.96	6.95	0.74	0.95

Les figures VII. 12 et 13 présentent une comparaison détaillée entre les valeurs de débit observées et prédites au cours des phases d'apprentissage et de test des différents modèles. La tendance linéaire observée pour le modèle SVM-Dagging ressemble beaucoup à la ligne  $y=x$ , avec un coefficient de corrélation (R) de 0,92 en termes de formation et de test. En revanche, les tendances pour SVM, SVM-Bagging, SVM-RF\* et SVM-RSS ont montré un R variable, allant de 0,92 à 0,79 pendant la phase d'entraînement et de 0,91 à 0,77 pendant la phase de test.

En consultant la littérature, on a découvert que les méthodes d'apprentissage d'ensemble ont généralement amélioré l'efficacité des modèles de base (Shah et al., 2023 ; Talukdar et al., 2022), bien que les classificateurs individuels puissent être faibles en théorie, tant qu'ils sont légèrement plus performants que les suppositions aléatoires, le classificateur final peut devenir un modèle puissant. Par conséquent, de nombreuses techniques d'apprentissage d'ensemble sont utilisées pour créer des modèles prédictifs solides et résistants, les recherches antérieures indiquant que les modèles hybrides et d'ensemble sont plus performants que les modèles individuels. Le modèle d'ensemble SVM-Dagging s'est avéré être le plus performant parmi les modèles examinés dans cette étude. Bien que les performances d'une méthode ou d'un modèle soient fortement influencées par les caractéristiques générales de la zone d'étude, qui à leur tour affectent les ensembles de données, le modèle Dagging se distingue des autres méthodes d'ensemble en raison des nombreux avantages qui lui permettent d'être exceptionnellement performant. L'un de ces avantages est sa capacité à réduire la variance en générant plusieurs sous-ensembles de données d'entraînement et en entraînant chaque sous-ensemble sur un modèle différent. Il a été démontré que la méthode Dagging offre une généralisation supérieure à d'autres méthodes, grâce à sa capacité à générer plusieurs arbres de décision qui réduisent le surajustement et améliorent les performances en matière de généralisation. L'utilisation de plusieurs arbres de décision dans la méthode Dagging garantit la diversité des modèles, ce qui est essentiel pour améliorer les performances de l'ensemble. En outre, l'inclusion d'autres arbres de décision rend la méthode Dagging plus résistante aux valeurs aberrantes et aux données bruyantes.

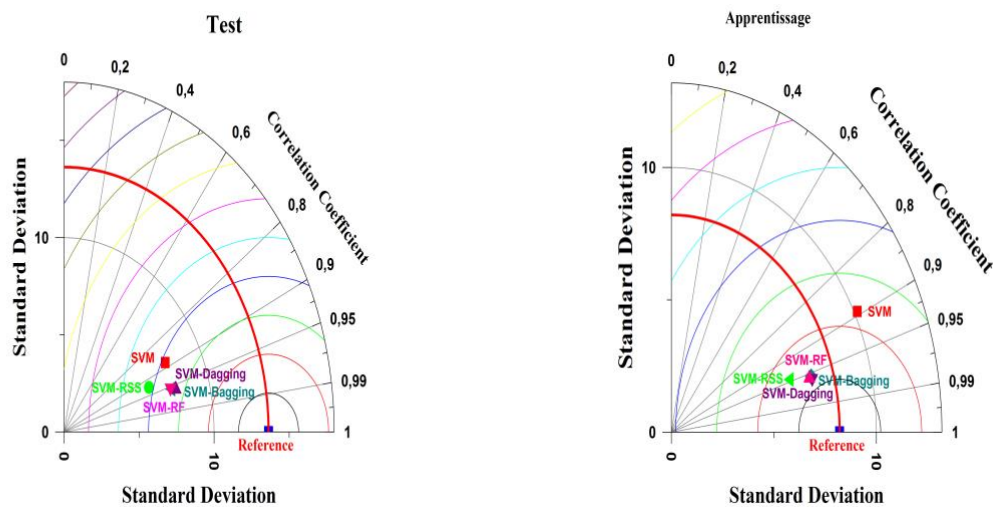


**Figure. VII. 12.** Relation débits observés et débit prédits dans le cours d'eau "Chiffa" pour l'ensemble des modèles hybridés lors de la phase d'apprentissage : a) SVM ; b) SVM-Bagging ; c) SVM-Dagging ; e) SVM- Random subspace ; f) SVM- Rotation Forest.



**Figure. VII. 13.** Relation des débits observés et débits prédits dans le cours d'eau "Chiffa" pour l'ensemble des modèles hybridés lors de la phase de test : a) SVM ; b) SVM-Bagging ; c) SVM-Dagging ; e) SVM- Random subspace ; f) SVM- Rotation Forest.

Le diagramme de Taylor a été utilisé pour représenter visuellement les mesures d'efficacité des résultats de la modélisation en traçant un ensemble des points sur un graphique polaire. Ce diagramme a démontré efficacement la variation spatiale du débit prévu par les modèles par rapport à la valeur observée au cours de la période d'essai pour la station hydrométriques des gorges de la Chiffa en amont du cours d'eau. Le diagramme de Taylor définit l'écart type (ET) entre les valeurs observées et prédites le long d'intervalles radiaux avec des racines, tandis que les valeurs R sont déterminées par les angles de direction. Le diagramme suppose que les valeurs observées sont affichées de manière indépendante et que les performances du modèle sont plus élevées lorsque les indicateurs de performance de la prédiction sont plus proches des valeurs observées. La figure VII.14 montre que le modèle SVM-Dagging montre des résultats parfaits et les débits prédis sont plus proches aux observés par rapport aux modèles SVM-Bagging, SVM-RF\* et plus particulier aux modèles SVM et SVM-RSS. Cela suggère que le modèle SVM-Dagging appliqué dans cette stations est le plus fiable et plus pertinent.



**Figure. VII. 14.** Diagrammes de Taylor montrant la comparaison statistique des cinq modèles hybridés étudiés pendant la période d'essai.

## VII.6. Conclusion

La méthode du modèle rugueux de référence (MMR) permet, entre autres, d'étendre les séries hydrométriques dans les stations mal jaugées afin de mieux dimensionner les différents ouvrages hydrauliques. En effet, le calcul des débits s'effectue à l'aide des paramètres mesurables, tels que la rugosité absolue, la section du courant d'eau, le rayon hydraulique, la pente longitudinale du cours d'eau, coefficient  $k_s$  et la viscosité cinématique du liquide en écoulement.

L'application de l'équation de Chiu a montré la bonne relation linéaire entre les vitesses d'écoulement moyennes et maximales, la valeur du paramètre d'entropie peut être considérée comme constante à n'importe quel endroit des cours d'eau naturels étudiés. Dans l'ensemble, malgré la moindre précision des résultats obtenus par la méthode de Chiu par rapport à la méthode MMR, cette dernière se distingue par des performances robustes, facilitant une estimation précise des débits.

Dans l'ensemble, l'utilisation extensive des méthodes, telles que celles présentées dans ce travail permettra sans aucun doute de remédier à l'insuffisance des mesures de débit et d'estimer de manière plus appropriée les matériaux liquides et même les apports solides transportés par ces cours d'eau naturels. Compte tenu de la facilité d'application et de la grande précision de la MMR, cette méthode pourrait être appliquée efficacement dans les cours d'eau naturels au climat semi-aride, notamment dans les stations non et mal jaugées.

L'utilisation des méthodes d'apprentissage automatique pour l'estimation des débits fluviaux a montré des résultats prometteurs dans cette étude. Les modèles de prévision de débit tels que, l'ANN, l'ANFIS, le SVM, le LibSVM, le M5P, le Rep Tree, le GPR et le Random Forest ont été utilisés pour estimer les débits des cours d'eau naturels pour cinq stations hydrométriques du nord de l'Algérie. Ces modèles ont été validés en termes de critères statistiques et d'interprétations graphiques.

Les résultats obtenus indiquent que, même en considérant les étiages observés pendant la période estivale à la station d'Oued Sebaou, le modèle ANFIS a démontré une performance exceptionnelle. Simultanément, l'ANN a manifesté les performances optimales pour la station d'Oued Chiffa.

Pour les stations (Chellif, Bellah et Mazafran), le modèle RF a également produit les meilleurs résultats. Enfin, en combinant les données de ces cinq stations hydrométriques étudiées, un modèle global d'apprentissage automatique a été développé pour toute la zone d'étude dans le but de prédire d'une façon précise et facile les débits de ces cours d'eau en Algérie. Ce modèle a été validé en termes de l'ensemble de critères statistiques et d'interprétations graphiques.

En conséquence, cette étude indique que le modèle Random Forest, ANN et ANFIS ont un grand potentiel dans les bassins versants pauvres en données et présente une meilleure alternative par rapport aux autres modèles d'apprentissage appliqués (SVM, LibSVM, M5p, Rep Tree et GPR).

L'approche de modélisation hybridée intègre une machine à vecteurs de support (SVM) avec quatre techniques d'apprentissage d'ensemble différentes, notamment le Bagging, le

Dagging, le sous-espace aléatoire et la forêt de rotation. Les modèles ont été développés et évalués à partir des données de la station de jaugeage «Amont des gorges» en tenant toujours les paramètres quantitatifs des cours d'eau. Les résultats montrent que le modèle SVM-Dagging a surpassé les autres modèles sur la base des critères statistiques, le modèle atteint un RMSE de 6,58, NSE de 0,76 et R de 0,96. Ces résultats indiquant le haut niveau de précision dans la prévision du débit d'eau. Le modèle SVM-Bagging a également présenté des résultats prometteurs avec RMSE de 6,83, NSE de 0,75 et R de 0,96. Ces résultats obtenus ont des implications importantes pour la gestion des ressources en eau dans les bassins versants pauvres en données, où des modèles prédictifs précis sont essentiels. Les modèles SVM d'ensemble proposés peuvent être utilisés pour améliorer les performances de prédiction et fournir une meilleure compréhension des processus hydrologiques. Leur utilisation se révèle particulièrement efficace dans les régions à climat semi-aride et méditerranéen, en particulier dans les stations non et mal jaugeées en Algérie. En aidant les décideurs à prendre des décisions éclairées en matière de gestion des ressources hydriques superficielles, ces modèles permettent une planification plus précise et opportune.

# **CONCLUSION GÉNÉRALE**

## **Conclusion générale**

L'estimation précise du débit d'un cours d'eau est importante pour la gestion des ressources en eau, la planification des projets hydroélectriques et la prévention des catastrophes naturelles. La prévision hydrologique représente un axe de recherche important, en hydrologie, car elle permet de comprendre les mécanismes de génération des débits et d'anticiper les phénomènes hydrologiques extrêmes, telles que les sécheresses et les inondations.

Dans cette étude, tout d'abord, nous mettons en lumière sur la possibilité d'utiliser les méthodes MMR et CHIU pour évaluer les débits des cours d'eau naturels en milieu méditerranéen et semi-aride. Les données recueillies par cinq stations hydrométriques du nord de l'Algérie ont été utilisées. Ensuite, nous explorons le potentiel et la pertinence des méthodes d'intelligence artificielle, nous utilisons huit modèles différents pour prévoir les débits fluviaux sur les séries chronologiques étudiées. Enfin, L'utilisation des modèles SVM-hybridés pour améliorer encore les prédictions des résultats obtenus.

Les résultats obtenus ont montré l'efficacité de la méthode MMR et de la méthode de Chiu dans l'estimation des débits liquides. La comparaison a révélé que la méthode MMR est à la fois facile à appliquer et parvient à atteindre une précision accrue pour chaque station étudiée. Cette méthode permet entre autres l'extension des séries hydrométriques dans les stations mal jaugées afin de mieux dimensionner les différents ouvrages hydrauliques.

Pour la méthode de Chiu, elle présente ainsi des bonnes performances pour l'estimation de ces débits, bien qu'elle soit moins précise que la méthode MMR. L'utilisation de telles méthodes permet de compenser le manque de mesure de débit et d'estimer de manière plus appropriée la charge liquide et même sédimentaire par les cours d'eau naturels.

La seconde contribution porte sur l'élaboration des modèles l'apprentissage automatique basés sur les données des jaugeages pour améliorer la précision de l'estimation des débits liquides. Huit modèles ont été utilisés ; les réseaux de neurones artificiels (ANN), système d'inférence neuro floue adaptative (ANFIS), les forêts aléatoires (Random Forest), arbre d'élague à erreur (Rep Tree), arbre M5P, la régression par processus gaussien (GPR), les machines à vecteurs de support (SVM) et la machine à vecteurs de support libérale (LibSVM). Les performances de ces modèles ont été validées en termes des critères statistiques et d'interprétations graphiques. Les résultats obtenus mettent en évidence :

- Les modèles Random Forest, ANN et ANFIS sont prometteurs pour les bassins versants avec peu des données, et offrent une alternative meilleure que d'autres modèles d'apprentissage utilisés (SVM, LibSVM, M5P, Rep Tree et GPR).
- En utilisant la combinaison des données de cinq stations, la méthode de la forêt aléatoire (Random Forest) est employée en tant que modèle d'apprentissage automatique pour prédire les débits dans toute la zone d'étude.
- Les modèles d'apprentissage automatique peuvent être utilisés avec succès pour prédire les débits liquides des cours d'eau en Algérie.
- Ces modèles peuvent aider à la prise des décisions éclairées pour la gestion de l'eau, la prévention des dommages liés aux inondations et l'amélioration de la sécurité des populations riveraines, en particulier dans les bassins mal et non jaugés.

Pour les modèles SVM-hybridés, les résultats obtenus sont très satisfaisants :

- Le modèle SVM-Dagging surpassait les autres modèles en termes de critères statistiques, ce modèle a également montré des résultats prometteurs.
- Les modèles en ensemble SVM-hybridés proposés peuvent améliorer les performances de prédiction du débit et offrir une meilleure compréhension des processus hydrologiques. Ces modèles ont un grand potentiel pour les bassins versants jaugés et non jaugés.

Dans la totalité, l'utilisation extensive des méthodes d'estimations des débits, telles que celles présentées dans ce travail permettra sans doute de pallier le manque de mesures de débit et d'estimer de manière plus appropriée les matières liquides et solides transportées par les cours d'eau naturels. Ces méthodes pourraient être appliquées efficacement dans les cours d'eau naturels à climat semi-aride et méditerranéen, en particulier dans les stations non jaugées.

Ces modèles peuvent aider les décideurs à prendre des décisions éclairées en matière de gestion de l'eau, en permettant une planification plus précise et en temps opportun des ressources superficiels, et peuvent également aider à la prévention des dommages liés aux inondations et à l'amélioration de la sécurité des populations riveraines.

Comme perspectives, nous proposons :

- Élargir l'utilisation de la méthode MMR dans l'évaluation des débits dans les oueds naturels du nord algérien, ce qui permet ainsi d'améliorer la précision des estimations hydrologiques, en particulier dans des régions où les données sont souvent insuffisantes.
- Employer des différents modèles d'intelligence artificielle pour établir des modèles globales offrant une compréhension approfondie des flux d'écoulement et estimer encore les apports liquides de la manière la plus précise possible.
- Développer des modèles hybrides novateurs permettant de déterminer les débits dans les bassins peu ou mal jaugés, tout en créant des modèles globaux couvrant l'ensemble du nord de l'Algérie, ce qui comble le déficit des données des apports liquides et solides et améliore la gestion des ressources hydriques superficielles en Algérie.

# **RÉFÉRENCE BIBLIOGRAPHIQUE**

## **Références bibliographiques**

Abdiansah, A., & Wardoyo, R. (2015). Time complexity analysis of support vector machines (SVM) in LibSVM. *International journal computer and application*, 128(3), 28-34.

Abraham, A. (2005). Artificial neural networks. *Handbook of measuring system design*.

Abroudi, A., Shokouhifar, M., & Farokhi, F. (2013). Improving the performance of artificial neural networks via instance selection and feature dimensionality reduction. *International Journal of Machine Learning and Computing*, 3(2), 176.

Achite, M., & Meddi, M. (2004). Estimation du transport solide dans le bassin-versant de l'oued Haddad (Nord-Ouest algérien). *Science et changements planétaires/Sécheresse*, 15(4), 367-373.

Achite, M., & Ouillon, S. (2007). Transport de sédiments en suspension dans un bassin versant semi-aride, Wadi Abd, Algérie (1973–1995). *Journal d'hydrologie*, 343 (3-4), 187-202.

Achour B. (2007) 'Calcul des conduites et canaux par la MMR (Conduites et canaux en charge) [Calculation of pipes and channels by the RMM (Conduits and channels in charge)]', Tome, Vol. 1, Larhyss Edition Capitale, ISBN 978-9961-9701-0-2, Biskra, p.610.

Achour, B. (2012) *Circular Pipe with Load And Free Surface*, Larhyss Edition Al Djazair.

Achour, B. (2013) 'Design of pressurized vaulted rectangular conduits using the rough model method', *Adv. Mat. Res.*, Vol. 779, pp414–419, Trans. Tech. Publications, <https://doi.org/10.4028/www.scientific.net/AMR.779-780.414>.

Achour, B. (2014) 'Computation of normal depth in trapezoidal open channel using the rough model method', *Adv. Mater. Res.*, Vols. 955–959, pp.3231–3237, Trans Tech Publications Ltd, <https://doi.org/10.4028/www.scientific.net/AMR.955-959.3231>.

Achour, B. and Bedjaoui, A. (2006a) 'Discussion. Exact solutions for normal depth problem', *J. Hydraulic. Res.*, Vol. 44, No. 5, pp.715–717.

Achour, B. and Bedjaoui, A. (2006b) 'Contribution au calcul de la profondeur normale dans un canal rectangulaire', *LARHYSS Journal*, Vol. 5, No. 5, pp.139–147, P-ISSN 1112-3680/EISSN 2521-9782.

Achour, B. and Bedjaoui, A. (2014) 'Design of a pressurized trapezoidal shaped conduit using the rough model method (part 2)', *Applied Mechanics and Materials*, Vols. 580–583, pp.1828–1841, <https://doi.org/10.4028/www.scientific.net/AMR.945-949.892>.

Achour, B. and Riabi, M. (2014) 'Design of a pressurized trapezoidal shaped conduit using the rough model method (part 1)', *Adv. Mater. Res.*, Vols. 945–949, pp.892–898, Trans Tech Publications Ltd, <https://doi.org/10.4028/www.scientific.net/AMR.945-949.892>.

Achour, B. and Sehtal, S. (2014) 'The rough model method (RMM). Application to the computation of normal depth in circular conduit', *Open Civil Engineering Journal*. Vol. 8, No. 1, pp.57–63, DOI: 10.2174/1874149501408010057.

Achour, B., Bedjaoui, A., Khattaoui, M. and Debabeche, M. (2002) 'Contribution au calcul des écoulements uniformes à surface libre et en charge (Première partie)', *LARHYSS Journal*, Vol. 1, No. 1, pp.7–36, P-ISSN 1112-3680/E-ISSN 2521-9782.

Agnihotri, A., Sahoo, A., & Diwakar, M. K. (2022). Flood prediction using hybrid ANFIS-ACO model: a case study. In *Inventive Computation and Information Technologies: Proceedings of ICICIT 2021* (pp. 169-180). Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-16-6723-7\\_13](https://doi.org/10.1007/978-981-16-6723-7_13)

Aharony, O., Seiberg, N. et Tachikawa, Y. (2013). Lire entre les lignes des théories de jauge à quatre dimensions. *Journal de physique des hautes énergies*, 2013 (8), 1-42.

Ahmadlou, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejhad, D., Shahabi, H., & Panahi, M. (2019). Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto International*, 34(11), 1252-1272. <https://doi.org/10.1080/10106049.2018.1474276>

Aizerman, A. (1964). Theoretical foundations of the potential function method in pattern recognition learning. *Automation and remote control*, 25, 821-837.

Akgün, E., & Demir, M. (2018). Modeling course achievements of elementary education teacher candidates with artificial neural networks. *International Journal of Assessment Tools in Education*, 5(3), 491-509.

Al-Abadi, A. M., & Shahid, S. (2016). Spatial mapping of artesian zone at Iraqi southern desert using a GIS-based random forest machine learning model. *Modeling Earth Systems and Environment*, 2, 1-17.

Al-Abadi, Alaa M., and Shamsuddin Shahid. "Spatial mapping of artesian zone at Iraqi southern desert using a GIS-based random forest machine learning model." *Modeling Earth Systems and Environment* 2 (2016): 1-17.

Al-Hmouz, A., Shen, J., Al-Hmouz, R., & Yan, J. (2011). Modeling and simulation of an adaptive neuro-fuzzy inference system (ANFIS) for mobile learning. *IEEE Transactions on Learning Technologies*, 5(3), 226-237.

Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms : an overview. In *Journal of physics: conference series* (Vol. 1142, p. 012012). IOP Publishing.

Amat, J. P., Le Coeur, C., Dorize, L., & Gautier, E. (2008). *Éléments de géographie physique*. Editions Bréal.

Amat, J., Paul, E., Watkins, L. R., & Maier, S. F. (2008). Activation of the ventral medial prefrontal cortex during an uncontrollable stressor reproduces both the immediate and long-term protective effects of behavioral control. *Neuroscience*, 154(4), 1178-1186.

Ammari, A. (2012). *Vulnérabilité à l'Envasement des Barrages (cas du bassin Hydrographique des Côtiers Algérois)* (Doctoral dissertation).

Ammari, A., & Remini, B. (2010). Estimation of Algerian rivers discharges based one Chiu's equation. *Arabian Journal of Geosciences*, 1(3), 59-65.

Ammari, A., & Remini, B. (2014). *Vulnérabilité à l'envasement des barrages (cas du bassin Algérois)*. Éditions universitaires européennes.

Ammari, A., Moramarco, T., & Meddi, M. (2017). A simple entropy-based method for discharges measurements in gauged and ungauged river sites: the case study of coastal Algerian rivers. *Bulletin de l'Institut Scientifique, Rabat*, 39, 35-44.

Andréassian, V., Lerat, J., Le Moine, N., & Perrin, C. (2012). Neighbors: Nature's own hydrological models. *Journal of Hydrology*, 414, 49-58.

Aoulmi, Y., Marouf, N., & Mohamed, A. (2020). The assessment of artificial neural network rainfall-runoff models under different input meteorological parameters Case study: Seybouse basin, Northeast Algeria. *Journal of Water and Land Development*. <https://doi.org/10.24425/jwld.2021.138158>

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.

Aoulmi, Y., Marouf, N., Rasouli, K., & Panahi, M. (2023). Runoff Predictions in a Semiarid Watershed by Convolutional Neural Networks Improved with Metaheuristic Algorithms and Forced with Reanalysis and Climate Data. *Journal of Hydrologic Engineering*, 28(7), 04023018.

Ardabili, S., Mosavi, A., Dehghani, M., & Várkonyi-Kóczy, A. R. (2020). Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review. In *Engineering for Sustainable Future: Selected papers of the 18th International Conference on Global Research and Education Inter-Academia-2019 18* (pp. 52-62). Springer International Publishing.

Ardilly, P. (2006). *Les techniques de sondage*. Editions Technip.

Aswad, F. M., Kareem, A. N., Khudhur, A. M., Khalaf, B. A., & Mostafa, S. A. (2021). Tree-based machine learning algorithms in the Internet of Things environment for multivariate flood status prediction. *Journal of Intelligent Systems*, 31(1), 1-14. <https://doi.org/10.1515/jisys-2021-0179>

Aswad, F. M., Kareem, A. N., Khudhur, A. M., Khalaf, B. A., & Mostafa, S. A. (2021). Tree-based machine learning algorithms in the Internet of Things environment for multivariate flood status prediction. *Journal of Intelligent Systems*, 31(1), 1-14. <https://doi.org/10.1515/jisys-2021-0179>

Bacro, J. N., & Chaouche, A. (2006). Incertitude d'estimation des pluies extrêmes du pourtour méditerranéen: illustration par les données de Marseille. *Hydrological sciences journal*, 51(3), 389-405.

Baker, D. B., Richards, R. P., Loftus, T. T., & Kramer, J. W. (2004). A new flashiness index: Characteristics and applications to midwestern rivers and streams 1. *JAWRA Journal of the American Water Resources Association*, 40(2), 503-522.

Balasubramanian, A. (2017). Soil erosion—causes and effects. Centre for Advanced Studies in Earth Science, University of Mysore, Mysore.

Balkema, AA, & De Haan, L. (1974). Durée de vie résiduelle au grand âge. *Les Annales des probabilités* , 2 (5), 792-804.

Basu, J. K., Bhattacharyya, D., & Kim, T. H. (2010). Use of artificial neural network in pattern recognition. *International journal of software engineering and its applications*, 4(2). Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, 22(5), 717-727.

Batout, S., Houichi, L., & Marouf, N. (2022). Influence of the envelope curve on the estimate of probable maximum precipitation (PMP) in the coastal region of Algeria. *Modeling Earth Systems and Environment*, 8(2), 2083-2093.

Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine learning*, 36, 105-139.

Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine learning*, 36, 105-139.

Behnood, A., Behnood, V., Gharehveran, M. M., & Alyamac, K. E. (2017). Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm. *Construction and Building Materials*, 142, 199-207.

Belhachat, F., Larbes, C., Barazane, L., & Kharzi, S. (2007, November). Commande neuro-floue d'un hacheur MPPT. In 4<sup>ème</sup> conférence internationale " Computer Integrated Manufacturing", CIP (Vol. 7, pp. 03-04).

Bennett, K. P., & Bredensteiner, E. J. (2000, June). Duality and geometry in SVM classifiers. In ICML (Vol. 2000, pp. 57-64).

Bertoni, A., Folgieri, R., & Valentini, G. (2005). Bio-molecular cancer prediction with random subspace ensembles of support vector machines. *Neurocomputing*, 63, 535-539. <https://doi.org/10.1016/j.neucom.2004.07.007>

Bhavsar, H., & Panchal, M. H. (2012). A review on support vector machine for data classification. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 1(10), 185-189.

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25, 197-227.

Blöschl, G. (Ed.). (2013). *Runoff prediction in ungauged basins: synthesis across processes, places and scales*. Cambridge University Press.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A., Parajka, J., Merz, B., ... & Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, 573(7772), 108-111. <https://doi.org/10.1038/s41586-019-1495-6>

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., & Savenije, H. (Eds.). (2013). *Runoff prediction in ungauged basins: synthesis across processes, places and scales*. Cambridge University Press. <https://doi.org/10.1002/>

Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1-2), 245-271.

BOIS P. (1971) - Une méthode de contrôle des séries chronologiques utilisées en climatologie et en hydrologie. *Laboratoire de Mécanique des Fluides*. Université de Grenoble, section hydrologie, 49 p.

BOIS P. (1986) - Contrôle des séries chronologies corrélées par l'étude du cumul des résidus. *Actes des Deuxièmes Journées Hydrologiques de l'ORSTOM (Montpellier)*, p. 89-100.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152).

Boudjemia, M. T., & Touati, A. (2016). *Etude expérimentale des écoulements à surface libre dans un convergent avec extension en marches d'escalier* (Doctoral dissertation, École Nationale Polytechnique).

Bowden, G. J., Maier, H. R., & Dandy, G. C. (2002). Optimal division of data for neural network models in water resources applications. *Water resources research*, 38(2), 2-1. <https://doi.org/10.1029/2001WR000266>

Braga, P. L., Oliveira, A. L., Ribeiro, G. H., & Meira, S. R. (2007, August). Bagging predictors for estimation of software project effort. In 2007 international joint conference on neural networks (pp. 1595-1600). IEEE.

Brazdil, P. B., Soares, C., & Da Costa, J. P. (2003). Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results. *Machine Learning*, 50, 251-277.

Breiman, L. (1996). Bagging predictors *Machine Learning* 24 (2), 123-140 (1996) 10.1023. A: 1018054314350.

Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.

BRUNET-MORET Y. (1977) - Test d'homogénéité. *Cahiers ORSTOM, série Hydrologie*, vol. XIV, n° 2, p. 119-129.

Brunet-Moret, Y. (1977). Test d'homogénéité. *Cahiers ORSTOM, série hydrologie*, 14(2), 119-128.

Bühlmann, P., & Yu, B. (2002). Analyzing bagging. *The annals of Statistics*, 30(4), 927-961.

Bunod, R., Augstburger, E., Brasnu, E., Labbe, A., & Baudouin, C. (2022). Intelligence artificielle et glaucome: une revue de la littérature. *Journal Français d'Ophtalmologie*, 45(2), 216-232.

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2), 121-167.

Bush, Vannevar. 1945. As We May Think. *The Atlantic Monthly*. July 1945.

Buyukyildiz, M., & Kumcu, S. Y. (2017). An estimation of the suspended sediment load using adaptive network based fuzzy inference system, support vector machine and artificial neural network models. *Water resources management*, 31, 1343-1359.

Capaccioli, M., & De Vaucouleurs, G. (1983). Luminosity distribution in galaxies. II-A study of accidental and systematic errors with application to NGC 3379. *The Astrophysical Journal Supplement Series*, 52, 465-486.

Carlier, M. (1972). *Hydraulique générale et appliquée* (No. 532 C3).

Carlier, P., Hannachi, H., & Mouvier, G. (1986). The chemistry of carbonyl compounds in the atmosphere—A review. *Atmospheric Environment* (1967), 20(11), 2079-2099.

Carnot, S., Clausius, R., & Kelvin, W. T. B. (1899). *The Second Law of Thermodynamics: Memoirs by Carnot, Clausius, and Thomson* (Vol. 6). Harper & brothers.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408, 189-215.

Chambon, G., Ghemmour, A. et Naaim, M. (2014). Etude expérimentale des écoulements viscoplastiques à surface libre dans un régime uniforme et stable. *Journal de mécanique des fluides*, 754, 332-364.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3), 1-27.

Chaudhry, M. H. (2008). *Open-channel flow* (Vol. 523). New York: Springer.

Chen, Y. C., & Chiu, C. L. (2002). An efficient method of discharge measurement in tidal streams. *Journal of hydrology*, 265(1-4), 212-224.

Cheng, M. Y., Huang, C. C., & Roy, A. F. V. (2013). Predicting project success in construction using an evolutionary Gaussian process inference model. *Journal of Civil Engineering and Management*, 19(sup1), S202-S211.

Chiu, C. L. (1987). Entropy and probability concepts in hydraulics. *Journal of Hydraulic Engineering*, 113(5), 583-599.

Chiu, C. L. (1988). Entropy and 2-D velocity distribution in open channels. *Journal of Hydraulic Engineering*, 114(7), 738-756.

Chiu, C. L. (1989). Velocity distribution in open channel flow. *Journal of Hydraulic Engineering*, 115(5), 576-594.

Chiu, C. L., & Chen, Y. C. (1999). Efficient methods of measuring discharge and reservoir-sediment inflow. *Risk Analysis in Dam Safety Assessment*. Water Resources Publications, Inc., Colorado, 97-116.

Chiu, C. L., & Chiou, J. D. (1986). Structure of 3-D flow in rectangular open channels. *Journal of Hydraulic Engineering*, 112(11), 1050-1067.

Chiu, C. L., & Said, C. A. A. (1995). Maximum and mean velocities and entropy in open-channel flow. *Journal of Hydraulic Engineering*, 121(1), 26-35.

Chou, J. S., Tsai, C. F., & Lu, Y. H. (2013). Project dispute prediction by hybrid machine learning techniques. *Journal of Civil Engineering and Management*, 19(4), 505-517.

Church, M. (2006). Bed material transport and the morphology of alluvial river channels. *Annu. Rev. Earth Planet. Sci.*, 34, 325-354.

Clausius, R. (1865). Sur diverses formes facilement applicables, qu'on peut donner aux équations fondamentales de la théorie mécanique de la chaleur. *Journal de mathématiques pures et appliquées*, 10, 361-400.

Colebrook, C.F. (1939) 'Turbulent flow in pipes, with particular reference to the transition between the smooth and rough pipe laws', *J. Inst. Civil Eng.*, Vol. 11, No. 4, pp.133–156.

Coles, S., Pericchi, LR et Sisson, S. (2003). Une approche entièrement probabiliste de la modélisation des précipitations extrêmes. *Journal d'hydrologie* , 273 (1-4), 35-50.

Cosandey, C., & Robinson, M. (2000). *Hydrologie continentale.*: 360 p.

Costa, J. E., Cheng, R. T., Haeni, F. P., Melcher, N., Spicer, K. R., Hayes, E., ... & Barrick, D. (2006). Use of radars to monitor stream discharge by noncontact methods. *Water Resources Research*, 42(7).

Dalkılıç, H. Y., & Hashimi, S. A. (2020). Prediction of daily streamflow using artificial neural networks (ANNs), wavelet neural networks (WNNs), and adaptive neuro-fuzzy inference system (ANFIS) models. *Water Supply*, 20(4), 1396-1408. <https://doi.org/10.2166/ws.2020.062>

Dang, C. N., Moreno-García, M. N., & De la Prieta, F. (2021). Hybrid deep learning models for sentiment analysis. *Complexity*, 2021, 1-16.

Darcy, H. (1854) 'Sur des recherches expérimentales relatives au mouvement des eaux dans les tuyaux', *Comptes rendus des séances de l'Académie des Sciences*, Vol. 38, pp.1109–1121.

Dave, V. S., & Dutta, K. (2014). Neural network based models for software effort estimation: a review. *Artificial Intelligence Review*, 42(2), 295-307.

Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and trends® in signal processing*, 7(3–4), 197-387.

Devasena, C. L. (2014). Comparative analysis of random forest, REP tree and J48 classifiers for credit risk prediction. *International Journal of Computer Applications*, 975(8887), 30-36.

Diconne, M. A. (1972). Les jaugeages des débits des cours d'eau au moulinet hydrométrique par la méthode d'intégration. *La Houille Blanche*, (1), 35-44.

Dietterich, TG (1997). Recherche en apprentissage automatique. *Revue AI* , 18 (4), 97-97.

Donelan, J. M., & Kram, R. (2000). Exploring dynamic similarity in human running using simulated reduced gravity. *Journal of Experimental Biology*, 203(16), 2405-2415.

Dramais, G., Le Coz, J., Gallavardin, A., Duby, P., Hauet, A., & Laronne, J. (2011, October). Mesures sans contact des débits de crue: avancées et perspectives. In *Ecotechs 2011* (pp. 4-p).

Dramais, G., Le Coz, J., Le Boursicaud, R., Hauet, A., & Lagouy, M. (2014). Jaugeage par radar mobile, protocole et résultats. *La Houille Blanche*, (3), 23-29.

Dubreuil, P., & Dubreuil, P. (1974). *Initiation à l'analyse hydrologique: dix exercices suivis des corrigés*. Masson.

- Ebden, M. (2015). Gaussian processes: A quick introduction. arXiv preprint arXiv:1505.02965.
- Ehteram, M., Salih, S. Q., & Yaseen, Z. M. (2020). Efficiency evaluation of reverse osmosis desalination plant using hybridized multilayer perceptron with particle swarm optimization. *Environmental Science and Pollution Research*, 27(13), 15278-15291. <https://doi.org/10.1007/s11356-020-08023-9>
- En ligne Nicolas, A. (2012). Structures des ophiolites et dynamique de la lithosphère océanique (Vol. 4). Springer Science et médias d'affaires.
- En ligne Petersen-Øverleir, A. (2006). Étape de modélisation - relations de débit affectées par l'hystérésis à l'aide de la formule de Jones et de la régression non linéaire. *Revue des sciences hydrologiques*, 51 (3), 365-388.
- FARTAS, F., MAROUF, N., & REMINI, B. (2017). QUANTIFICATION DU TRANSPORT SOLIDE EN SUSPENSION DANS LE BARRAGE DE FOUM EL GHERZA–BISKRA. ALGERIE. *Journal of Water and Environmental Sciences*, 1, 198-218.
- Ficke, A. D., Myrick, C. A., & Hansen, L. J. (2007). Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries*, 17, 581-613.
- Fisher, R. A. (1948). Conclusions fiduciaires. In *Annales de l'institut Henri Poincaré* (Vol. 10, No. 3, pp. 191-213).
- Francke, T., López-Tarazón, J. A., & Schröder, B. (2008). Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. *Hydrological Processes: An International Journal*, 22(25), 4892-4904.
- Gallart, F., Llorens, P., Latron, J., Cid, N., Rieradevall, M., & Prat, N. (2016). Validating alternative methodologies to estimate the regime of temporary rivers when flow data are unavailable. *Science of the Total Environment*, 565, 1001-1010.
- García-Pedrajas, N., & Ortiz-Boyer, D. (2008). Boosting random subspace method. *Neural Networks*, 21(9), 1344-1362.
- Ghebrehiwot, A., & Kozlov, D. (2021). Reanalysis dataset-based hydrologic predictions for ungauged basins. In *E3S Web of Conferences* (Vol. 264, p. 01001). EDP Sciences. <https://doi.org/10.1051/e3sconf/202126401001>
- Gibson, C. A., Meyer, J. L., Poff, N. L., Hay, L. E., & Georgakakos, A. (2005). Flow regime alterations under changing climate in two river basins: implications for freshwater ecosystems. *River Research and Applications*, 21(8), 849-864.
- Goldman, S. (1953). *Information theory*. Prentice-Hall, Inc., New York, N.Y. Jaynes, E. T. (1957). "Information theory and statistical mechanics I." *Physics Review*, 106, 620-630.

Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., ... & Lander, E. S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *science*, 286(5439), 531-537.

Goswami, M., O'connor, K. M., & Bhattarai, K. P. (2007). Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment. *Journal of Hydrology*, 333(2-4), 517-531.

Goula, AM, Karapantsios, TD, Achilias, DS et Adamopoulos, KG (2008). Isothermes de sorption d'eau et température de transition vitreuse de la pulpe de tomate séchée par pulvérisation. *Journal of Food Engineering* , 85 (1), 73-83.

Goula, B.T.A., et al., 2007. Estimation des pluies exceptionnelles journalières en zone tropicale: cas de la Côte d'Ivoire par comparaison des lois lognormale et de Gumbel. *Hydrological Sciences Journal*, 52 (2), 49–67.

Graf, W. H., & Altinakar, M. S. (2000). *Hydraulique fluviale: écoulement et phénomènes de transport dans les canaux à géométrie simple (Vol. 16)*. PPUR presses polytechniques.

Guo, Y., Zhang, Y., Zhang, L., & Wang, Z. (2021). Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. *Wiley Interdisciplinary Reviews: Water*, 8(1), e1487. <https://doi.org/10.1002/wat2.1487>

Hafied, y., marouf, n., bouziane, m. T., remini, b., & lubomir, s. (2019). Load sediments quantification in algerian north-west basins by ann (artificial neurons network) method. *Geoscience engineering*, 65(3), 1-17. Doi 10.35180/gse-2019-0013

Hale, R., Marshall, S., Jeppe, K., & Petti grove, V. (2014). Separating the effects of water physiochemistry and sediment contamination on *Chironomus tepperi* (Skuse) survival, growth and development: A boosted regression tree approach. *Aquatic Toxicology*, 152, 66–73. <https://doi.org/10.1016/J.AQUATOX.2014.03.014>

Henderson, J., Lemon, O., & Georgila, K. (2008). Hybrid reinforcement/supervised learning of dialogue policies from fixed data sets. *Computational Linguistics*, 34(4), 487-511.

Hoang, N. D., Pham, A. D., Nguyen, Q. L., & Pham, Q. N. (2016). Estimating compressive strength of high performance concrete with Gaussian process regression model. *Advances in civil engineering*, 2016.

Horvath, D., Brown, J. B., Marcou, G., & Varnek, A. (2014). An evolutionary optimizer of libsvm models. *Challenges*, 5(2), 450-472.

Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). *A practical guide to support vector classification*.

Huang, Nantian, Guobo Lu, and Dianguo Xu. "A permutation importance-based feature selection method for short-term electricity load forecasting using random forest." *Energies* 9.10 (2016): 767.

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics & proteomics*, 15(1), 41-51.

Huntingford, C., Jeffers, E. S., Bonsall, M. B., Christensen, H. M., Lees, T., & Yang, H. (2019). Machine learning and artificial intelligence to aid climate change research and preparedness. *Environmental Research Letters*, 14(12), 124007.

Hutter, M. (2004). *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Springer Science & Business Media.

Hwang, S. C., Park, J. C., Gotoh, H., Khayyer, A., & Kang, K. J. (2016). Numerical simulations of sloshing flows with elastic baffles by using a particle-based fluid–structure interaction analysis method. *Ocean Engineering*, 118, 227-241.

Ibrahim, K. S. M. H., Huang, Y. F., Ahmed, A. N., Koo, C. H., & El-Shafie, A. (2022). A review of the hybrid artificial intelligence and optimization modelling of hydrological streamflow forecasting. *Alexandria Engineering Journal*, 61(1), 279-303.<https://doi.org/10.1016/j.aej.2021.04.100>

Jaafari, A., Zenner, E. K., Panahi, M., & Shahabi, H. (2019). Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agricultural and forest meteorology*, 266, 198-207. <https://doi.org/10.1016/j.agrformet.2018.12.015>

Jakkula, V. (2006). Tutorial on support vector machine (svm). School of EECS, Washington State University, 37(2.5), 3.

Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.

Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing—a computational approach to learning and machine intelligence [Book Review]. *IEEE Transactions on automatic control*, 42(10), 1482-1484.

Jang, J. S., & Sun, C. T. (1995). Neuro-fuzzy modeling and control. *Proceedings of the IEEE*, 83(3), 378-406.

Jean, A. (2020). Une brève introduction à l'intelligence artificielle. *médecine/sciences*, 36(11), 1059-1067.

Kalmegh, S. (2015). Analysis of weka data mining algorithm reptime, simple cart and randomtree for classification of indian news. *International Journal of Innovative Science, Engineering & Technology*, 2(2), 438-446.

Karr, J. R. (1991). Biological integrity: a long-neglected aspect of water resource management. *Ecological applications*, 1(1), 66-84.

KEULEGAN, G.H. (1938): Laws of turbulent flow in open channels, Research paper RP 1151, *Journal of Research*, U.S. National Bureau of Standards, Vol. 21, 707-741.

Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied soft computing*, 11(2), 2664-2675.

Kim, G., Kim, Y., Lim, H., & Kim, H. (2010). An MLP-based feature subset selection for HIV-1 protease cleavage site analysis. *Artificial intelligence in medicine*, 48(2-3), 83-89.

KIRONOTO, B. A., Graf, W. H., & Reynolds. (1995). Turbulence characteristics in rough non-uniform open-channel flow. *Proceedings of the institution of civil engineers-water maritime and energy*, 112(4), 336-348.

Kohler, D. (1949). William Faulkner and the social conscience. *College English*, 11(3), 119-127.

Kotsianti, S. B., & Kanellopoulos, D. (2007). Combining bagging, boosting and dagging for classification problems. In *Knowledge-Based Intelligent Information and Engineering Systems: 11th International Conference, KES 2007, XVII Italian Workshop on Neural Networks, Vietri sul Mare, Italy, September 12-14, 2007. Proceedings, Part II 11* (pp. 493-500). Springer Berlin Heidelberg. [https://citations.springernature.com/item?doi=10.1007/978-3-540-74827-4\\_62](https://citations.springernature.com/item?doi=10.1007/978-3-540-74827-4_62)

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations. *Hydrological Sciences Journal*, 49(4).

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. Recherche théorique. *Hydrological sciences journal*, 49(4).

Krasnogorskaya, N., Belozeroва, E., Longobardi, A., & Nafikova, E. (2019). The map-correlation method for ungauged catchments streamflow prediction in the Ufa River, Russian Federation. *International Journal of Hydrology Science and Technology*, 9(6), 603-626. <https://doi.org/10.1504/IJHST.2019.103440>

Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications*. InTech, 1-18.

Kuncheva, L. I., & Rodríguez, J. J. (2007). An experimental study on rotation forest ensembles. In *Multiple Classifier Systems: 7th International Workshop, MCS 2007, Prague, Czech Republic, May 23-25, 2007. Proceedings 7* (pp. 459-468). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-72523-7\\_46](https://doi.org/10.1007/978-3-540-72523-7_46)

Latif, J., Xiao, C., Imran, A. et Tu, S. (janvier 2019). Imagerie médicale utilisant des algorithmes d'apprentissage automatique et d'apprentissage profond : une revue. En 2019 2ème conférence internationale sur les technologies de l'informatique, des mathématiques et de l'ingénierie (iCoMET) (pp. 1-5). IEEE.

Lázaro-Gredilla, M., Titsias, M. K., Verrelst, J., & Camps-Valls, G. (2013). Retrieval of biophysical parameters with heteroscedastic Gaussian processes. *IEEE Geoscience and Remote Sensing Letters*, 11(4), 838-842.

LeCun, Y., Bengio, Y. et Hinton, G. : Apprentissage en profondeur, *Nature*, 521, 436–444, <https://doi.org/10.1038/nature14539> , 2015.

Lenard, M. J., Madey, G. R., & Alam, P. (1998). The design and validation of a hybrid information system for the auditor's going concern decision. *Journal of Management Information Systems*, 14(4), 219-237.

Leopold, L. B., & Langbein, W. B. (1962). *The concept of entropy in landscape evolution* (Vol. 500). US Government Printing Office.

Li, H., Lee, Y. C., Zhou, Y. C., & Sun, J. (2011). The random subspace binary logit (RSBL) model for bankruptcy prediction. *Knowledge-Based Systems*, 24(8), 1380-1388.

Li, Z., Sun, Z., Liu, J., Dong, H., Xiong, W., Sun, L., & Zhou, H. (2022). Prediction of river sediment transport based on wavelet transform and neural network model. *Applied Sciences*, 12(2), 647. <https://doi.org/10.3390/app12020647>

Longuet-Higgins, M. S. (1953). Mass transport in water waves. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 245(903), 535-581.

Lou, H., Wang, P., Yang, S., Hao, F., Ren, X., Wang, Y., ... & Gong, T. (2020). Combining and comparing an unmanned aerial vehicle and multiple remote sensing satellites to calculate long-term river discharge in an ungauged water source region on the Tibetan Plateau. *Remote Sensing*, 12(13), 2155. <https://doi.org/10.3390/rs12132155>

Luo, Q., Wu, J., Yang, Y., Qian, J., & Wu, J. (2016). Multi-objective optimization of long-term groundwater monitoring network design using a probabilistic Pareto genetic algorithm under uncertainty. *Journal of Hydrology*, 534, 352-363. <https://doi.org/10.1016/j.jhydrol.2016.01.009>

Madvar, HR, Dehghani, M., Memarzadeh, R., Salwana, E., Mosavi, A., & Shahab, S. (2020). Dérivation d'équations optimisées pour l'estimation du coefficient de dispersion dans les cours d'eau naturels à l'aide d'algorithmes ANN hybrides avec PSO et CSO. *Accès IEEE* , 8 , 156582-156599.

Malamud, B. D., & Turcotte, D. L. (2006). The applicability of power-law frequency statistics to floods. *Journal of hydrology*, 322(1-4), 168-180.

Malavoi, J. R., & Bravard, J. P. (2010). *Éléments d'hydromorphologie fluviale*. Onema.

Malavoi, J. R., & Souchon, Y. (1996). *Dynamique fluviale et dynamique écologique*. La Houille Blanche, (6-7), 98-107.

Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, 42(10), 4621-4631.

Manguerra, HB, & Engel, BA (1998). PARAMÉTRAGE HYDROLOGIQUE DES BASSINS VERSANTS POUR LA PRÉDICTION DE L'ÉCOULEMENT À L'AIDE DE SWAT 1. *JAWRA Journal of the American Water Resources Association* , 34 (5), 1149-1162.

Marouf, N. (2012). *Etude de la qualité des eaux et transport solide dans le barrage de beni-haroun (mila) son impact sur l'environnement de la région (Doctoral dissertation, Université Mohamed Khider–Biskra)*.

Marouf, N., & Remini, B. (2011). Temporal variability in sediment concentration and hysteresis in the Wadi Kebir Rhumel Basin of Algeria. *HKIE Transactions*, 18(1), 13-21. <https://doi.org/10.1080/1023697X.2011.10668219>

Marouf, N., & Remini, B. (2019). Impact study of Beni-Haroun dam on the environmental and socio-economic elements in Kébir-Rhumel basin, Algeria. *Journal of Water and Land Development*(43).

Mathew, A., Amudha, P., & Sivakumari, S. (2021). Deep learning techniques: an overview. *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020*, 599-608.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.

Meddi, M., Toumi, S., & Assani, A. A. (2017). Application of the L-moments approach to the analysis of regional flood frequency in Northern Algeria. *International Journal of Hydrology Science and Technology*, 7(1), 77-102. <https://doi.org/10.1504/IJHST.2017.080959>

Messalti, S. (2015, March). A new neural networks MPPT controller for PV systems. In IREC2015 the sixth international renewable energy congress (pp. 1-6). IEEE.

Meylan, P. et Musy A., 1999. Hydrologie fréquentielle. Gestion et protection de la ressource en eau. Office fédéral de l'éducation de la science, no. 96.01. Bucarest: HGA

Meylan, P., Favre, A. C., & Musy, A. (2012). Predictive hydrology: a frequency analysis approach. CRC Press.

Mishra, A. K., & Coulibaly, P. (2009). Developments in hydrometric network design: A review. *Reviews of Geophysics*, 47(2). <https://doi.org/10.1029/2007RG000243>

Mitchell, Machine Learning, McGraw-Hill, 1997.

Mitchell, T. (1997). Introduction to machine learning. *Machine learning*, 7, 2-5.

Mohamed, W. N. H. W., Salleh, M. N. M., & Omar, A. H. (2012, November). A comparative study of reduced error pruning method in decision tree algorithms. In 2012 IEEE International conference on control system, computing and engineering (pp. 392-397). IEEE.

Moramarco, T., & Singh, V. P. (2001). Simple method for relating local stage and remote discharge. *Journal of Hydrologic Engineering*, 6(1), 78-81. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2001\)6:1\(78\)](https://doi.org/10.1061/(ASCE)1084-0699(2001)6:1(78))

Moramarco, T., Ammari, A., Mirauda, D., & Pascal, V. (2008). Addressing the flows monitoring in natural channels through the entropy theory. In *International Congress on Environmental Modelling and Software*. International Environmental Modelling and Software Society (iEMSs).

Moramarco, T., Saltalippi, C., & Singh, V. P. (2004). Estimation of mean velocity in natural channels based on Chiu's velocity distribution equation. *Journal of Hydrologic Engineering*, 9(1), 42-50.

Musy, A., & Higy, C. (2004). *Hydrologie : Une science de la nature* (Vol. 21). PPUR presses polytechniques.

Musy, A., & Higy, C. (2004). *Hydrologie: Une science de la nature* (Vol. 21). PPUR presses polytechniques.

Nadir, M., & Boualem, R. (2016). Study of Beni Haroun dam pollution (Algeria). *Desalination and Water Treatment*, 57(6), 2766-2774.

Naghibi, Seyed Amir, Kouros Ahmadi, and Alireza Daneshi. "Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping." *Water Resources Management* 31 (2017): 2761-2775.

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of big data*, 2(1), 1-21.

Nanson, G. C., & Knighton, A. D. (1996). Anabranching rivers: their cause, character and classification. *Earth surface processes and landforms*, 21(3), 217-239.

NF EN ISO 748 (2009), – Mesure de débit des liquides dans les canaux découverts – Méthodes d’exploration du champ des vitesses, Norme ISO, 50 p

Nguyen-Tuong, D., Seeger, M., & Peters, J. (2009). Model learning with local gaussian process regression. *Advanced Robotics*, 23(15), 2015-2034.

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. <https://doi.org/10.1016/j.jhydrol.2019.124296>

Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.

Nourani, V., Baghanam, A. H., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. *Journal of Hydrology*, 514, 358-377. <https://doi.org/10.1016/j.jhydrol.2014.03.057>

Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *science*, 313(5790), 1068-1072.

Oliveira, Sandra, et al. "Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest." *Forest Ecology and Management* 275 (2012): 117-129.

Olyaie, E., Banejad, H., Chau, K. W., & Melesse, A. M. (2015). A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. *Environmental monitoring and assessment*, 187, 1-22.

Omran, B. A., Chen, Q., & Jin, R. (2016). Comparison of data mining techniques for predicting compressive strength of environmentally friendly concrete. *Journal of Computing in Civil Engineering*, 30(6), 04016029.

Ondo, J. C., Ouarda, T. B., & Bobée, B. (1997). *Revue bibliographique des tests d'homogénéité et d'indépendance*.

Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J. (2017). Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, 48(3), 128-138.

Oueslati, O., De Girolamo, A. M., Abouabdillah, A., Kjeldsen, T. R., & Lo Porto, A. (2015). Classifying the flow regimes of Mediterranean streams using multivariate analysis. *Hydrological Processes*, 29(22), 4666-4682.

Oza, N. C., & Russell, S. (2001). Online ensemble learning (p. 2). University of California, Berkeley.

Ozcift, A., & Gulten, A. (2011). Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. *Computer methods and programs in biomedicine*, 104(3), 443-451.

Pal, M., & Deswal, S. (2010). Modelling pile capacity using Gaussian process regression. *Computers and Geotechnics*, 37(7-8), 942-947.

Peiry, J. (1989). L'utilisation du cadastre sarde de 1730 pour l'étude des rivières savoyardes: l'exemple de la vallée de l'Arve (Haute-Savoie)./The " Mapped Sarde". Use for an historical study of the Arve river (Haute-Savoie). *Géocarrefour*, 64(4), 197-203.

Perumal, M., Moramarco, T., Sahoo, B., & Barbetta, S. (2007). A methodology for discharge estimation and rating curve development at ungauged river sites. *Water Resources Research*, 43(2). <https://doi.org/10.1029/2005WR004609>

Perumal, M., Moramarco, T., Sahoo, B., & Barbetta, S. (2010). On the practical applicability of the VPMS routing method for rating curve development at ungauged river sites. *Water resources research*, 46(3). <https://doi.org/10.1029/2009WR008103>

Pham, B. T., Avand, M., Janizadeh, S., Phong, T. V., Al-Ansari, N., Ho, L. S., ... & Prakash, I. (2020). GIS based hybrid computational approaches for flash flood susceptibility assessment. *Water*, 12(3), 683.

Pham, B. T., Nguyen-Thoi, T., Qi, C., Van Phong, T., Dou, J., Ho, L. S., ... & Prakash, I. (2020). Coupling RBF neural network with ensemble learning techniques for landslide susceptibility mapping. *Catena*, 195, 104805. <https://doi.org/10.1016/j.catena.2020.104805>

Pham, B. T., Nguyen-Thoi, T., Qi, C., Van Phong, T., Dou, J., Ho, L. S., ... & Prakash, I. (2020). Coupling RBF neural network with ensemble learning techniques for landslide susceptibility mapping. *Catena*, 195, 104805. <https://doi.org/10.1016/j.catena.2020.104805>

Pham, B. T., Tien Bui, D., Prakash, I., & Dholakia, M. B. (2016). Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. *Natural Hazards*, 83, 97-127.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, 119-131.

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegard, K. L., Richter, B. D., ... & Stromberg, J. C. (1997). The natural flow regime. *BioScience*, 47(11), 769-784.

Polikar, R. (2012). Ensemble learning. *Ensemble machine learning: Methods and applications*, 1-34.

PRANDTL, L. (1926): On fully developed turbulence, *Proceedings of the 2nd International Congress of Applied Mechanics, Zurich*, 62-74.

Prat, N., Gallart, F., Von Schiller, D., Polesello, S., García-Roger, E. M., Latron, J., ... & Froebrich, J. (2014). The mirage toolbox: an integrated assessment tool for temporary streams. *River research and applications*, 30(10), 1318-1334.

Quinlan, J. R. (1987). Simplifying decision trees. *International journal of man-machine studies*, 27(3), 221-234.

Quinlan, J. R. (1992, November). Learning with continuous classes. In *5th Australian joint conference on artificial intelligence (Vol. 92, pp. 343-348)*.

Rahman, S. A., & Chakrabarty, D. (2020). Sediment transport modelling in an alluvial river with artificial neural network. *Journal of Hydrology*, 588, 125056. <https://doi.org/10.1016/j.jhydrol.2020.125056>

Rajasekaran, S., & Pai, G. V. (2003). *Neural networks, fuzzy logic and genetic algorithm: synthesis and applications (with cd)*. PHI Learning Pvt. Ltd..

Ramachandra Rao, A., & Hsieh, C. H. (1987). Maximum entropy probability distributions for flood frequency analysis. *Civil Engineering Systems*, 4(2), 67-76.

Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning (Vol. 1, p. 159)*. Cambridge, MA: MIT press.

Razavi, T., & Coulibaly, P. (2013). Streamflow prediction in ungauged basins: review of regionalization methods. *Journal of hydrologic engineering*, 18(8), 958-975. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000690](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000690)

Rey, F., Ballais, J. L., Marre, A., & Rovéra, G. (2004). Rôle de la végétation dans la protection contre l'érosion hydrique de surface. *Comptes rendus géoscience*, 336(11), 991-998.

Rezaeianzadeh, M., Tabari, H., Arabi Yazdi, A., Isik, S., & Kalin, L. (2014). Flood flow forecasting using ANN, ANFIS and regression models. *Neural Computing and Applications*, 25, 25-37.

Rodriguez, J. J., Kuncheva, L. I., & Alonso, C. J. (2006). Rotation forest: A new classifier ensemble method. *IEEE transactions on pattern analysis and machine intelligence*, 28(10), 1619-1630.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.

Rosgen, D. L. (1994). A classification of natural rivers. *Catena*, 22(3), 169-199.

Rott, N. (1990). Note on the history of the Reynolds number. *Annual review of fluid mechanics*, 22(1), 1-12.

Rules, A. F. I. T. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on systems, man, and cybernetics*, 23(3), 665.

Saha, S., Kundu, B., Paul, G. C., & Pradhan, B. (2023). Proposing an ensemble machine learning based drought vulnerability index using M5P, dagging, random sub-space and rotation forest models. *Stochastic Environmental Research and Risk Assessment*, 1-28. <https://doi.org/10.1007/s00477-023-02403-6>

Saha, S., Kundu, B., Paul, G. C., & Pradhan, B. (2023). Proposing an ensemble machine learning based drought vulnerability index using M5P, dagging, random sub-space and rotation forest models. *Stochastic Environmental Research and Risk Assessment*, 1-28. <https://doi.org/10.1007/s00477-023-02403-6>

Sahoo, A., & Ghose, D. K. (2022, June). Application of Hybrid Support Vector Machine model for Streamflow Prediction in Barak valley, India. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1032, No. 1, p. 012016). IOP Publishing. <https://doi.org/10.1088/1755-1315/1032/1/012016>

Sahoo, A., Samantaray, S., & Ghose, D. K. (2019). Stream flow forecasting in Mahanadi River Basin using artificial neural networks. *Procedia Computer Science*, 157, 168-174. <https://doi.org/10.1016/j.procs.2019.08.154>

Sahoo, B., Perumal, M., Moramarco, T. et al. (2014) 'Rating curve development at Ungauged River sites using variable parameter Muskingum discharge routing method', *Water Resour. Manage.*, Vol. 28, pp.3783–3800, <https://doi.org/10.1007/s11269-014-0709-9>

Salcedo-Sanz, S., Pérez-Aracil, J., Ascenso, G., Del Ser, J., Casillas-Pérez, D., Kadow, C., ... & Castelletti, A. (2022). Analysis, characterization, prediction and attribution of extreme atmospheric events with machine learning: a review. *arXiv preprint arXiv:2207.07580*.

Samanataray, S., & Sahoo, A. (2021). A comparative study on prediction of monthly streamflow using hybrid ANFIS-PSO approaches. *KSCE Journal of Civil Engineering*, 25(10), 4032-4043. <https://doi.org/10.1007/s12205-021-2223-y>

Samantaray, S., & Sahoo, A. (2021). Modelling response of infiltration loss toward water table depth using RBFN, RNN, ANFIS techniques. *International Journal of Knowledge-based and Intelligent Engineering Systems*, 25(2), 227-234.

Samantaray, S., & Sahoo, A. (2022). Prediction of suspended sediment concentration using hybrid SVM-WOA approaches. *Geocarto International*, 37(19), 5609-5635. <https://doi.org/10.1080/10106049.2021.1920638>

Samantaray, S., Sahoo, A., & Agnihotri, A. (2023). Prediction of Flood Discharge Using Hybrid PSO-SVM Algorithm in Barak River Basin. *MethodsX*, 10, 102060. <https://doi.org/10.1016/j.mex.2023.102060>

Samantaray, S., Sahoo, A., & Ghose, D. K. (2020). Assessment of sediment load concentration using SVM, SVM-FFA and PSR-SVM-FFA in arid watershed, India: a case study. *KSCE Journal of Civil Engineering*, 24, 1944-1957. <https://doi.org/10.1007/s12205-020-1889-x>

SÁNCHEZ-MONTOYA, M. D. M., Puntí, T., Suarez, M. L., VIDAL-ABARCA, M. D. R., Rieradevall, M., Poquet, J. M., ... & Prat, N. (2007). Concordance between ecotypes and macroinvertebrate assemblages in Mediterranean streams. *Freshwater Biology*, 52(11), 2240-2255.

Saritha, R. R., Paul, V., & Kumar, P. G. (2019). Content based image retrieval using deep learning process. *Cluster Computing*, 22, 4187-4200.

Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2(6), 420.

Sepahvand, A., Singh, B., Sihag, P., Nazari Samani, A., Ahmadi, H., & Fiz Nia, S. (2021). Assessment of the various soft computing techniques to predict sodium absorption ratio (SAR). *ISH Journal of Hydraulic Engineering*, 27(sup1), 124-135.

Shah, M., Kantawala, H., Gandhi, K., Patel, R., Patel, K. A., & Kothari, A. (2023, January). Theoretical Evaluation of Ensemble Machine Learning Techniques. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 829-837). IEEE. [10.1109/ICSSIT55814.2023.10061139](https://doi.org/10.1109/ICSSIT55814.2023.10061139)

Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., ... & Chau, K. W. (2020). Predicting standardized streamflow index for hydrological drought using machine learning models. *Engineering Applications of Computational Fluid Mechanics*, 14(1), 339-350. <https://doi.org/10.1080/19942060.2020.1715844>

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell system technical journal*, 30(1), 50-64.

Shannon, C. E., Weaver, W., & Wiener, N. (1950). The mathematical theory of communication. *Physics Today*, 3(9), 31-32.

Sisson, S. A., Pericchi, L. R., & Coles, S. G. (2006). A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean. *Stochastic Environmental Research and Risk Assessment*, 20, 296-306.

Sivapalan, M., Schaake, J., & Sapporo, J. (2003). PUB science and implementation plan. V5., online available at: [http://pub.iwmi.org/UI/Images/PUB Science Plan, 5](http://pub.iwmi.org/UI/Images/PUB_Science_Plan_5).

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., ... & Zehe, E. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological sciences journal*, 48(6), 857-880. <https://doi.org/10.1623/hysj.48.6.857.51421>

Skurichina, M., & Duin, R. P. (2002). Bagging, boosting and the random subspace method for linear classifiers. *Pattern Analysis & Applications*, 5, 121-135.

Srinivasan, D. B., & Mekala, P. (2014). Mining social networking data for classification using reptime. *International Journal of Advance Research in Computer Science and Management Studies*, 2(10).

Subramanya, K. (2008). edition 3. *Engineering hydrology*. New Delhi. Tata McGraw Hill, 155-162.

Svetnik, V., Liaw, A., Tong, C., & Wang, T. (2004). Application of Breiman's random forest to modeling structure-activity relationships of pharmaceutical molecules. In *Multiple Classifier Systems: 5th International Workshop, MCS 2004, Cagliari, Italy, June 9-11, 2004. Proceedings 5* (pp. 334-343). Springer Berlin Heidelberg.

Swain, J. B., & Patra, K. C. (2017). Streamflow estimation in ungauged catchments using regionalization techniques. *Journal of Hydrology*, 554, 420-433. <https://doi.org/10.1016/j.jhydrol.2017.08.054>

Taïbi, R. (1993). Contribution à l'étude de l'écoulement des cours d'eau de l'Algérie septentrionale (Doctoral dissertation).

Talukdar, S., Mallick, J., Sarkar, S. K., Roy, S. K., Islam, A. R. M. T., Praveen, B., ... & Sobnam, M. (2022). Novel hybrid models to enhance the efficiency of groundwater potentiality model. *Applied Water Science*, 12(4), 62. <https://doi.org/10.1007/s13201-022-01571-0>

TAMRABET, Z., Marouf, N., & Remini, B. (2019). Quantification of suspended solid transport in Endja watercourse [Dehamecha basin-Algeria]. *GeoScience Engineering*, 65(4), 71-91.

Tao, H., Al-Sulttani, A. O., Salih Ameen, A. M., Ali, Z. H., Al-Ansari, N., Salih, S. Q., & Mostafa, R. R. (2020). Training and testing data division influence on hybrid machine learning model process: application of river flow forecasting. *Complexity*, 2020, 1-22. <https://doi.org/10.1155/2020/8844367>

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of geophysical research: atmospheres*, 106(D7), 7183-7192. <https://doi.org/10.1029/2000JD900719>

Ting, K.M., & Witten, I.H. (1997). Stacking bagged and dagged models. In: *Proceeding ICML '97 Proceedings of the Fourteenth International Conference on Machine Learning*. pp. 367–375. ISBN:1-55860-486-3.

Ting, K.M., & Witten, I.H. (1997). Stacking bagged and dagged models. In: *Proceeding ICML '97 Proceedings of the Fourteenth International Conference on Machine Learning*. pp. 367–375. ISBN:1-55860-486-3.

Touaibia, B., Aidaoui, A., Gomer, D., & Achite, M. (2001). Quantification et variabilité temporelles de l'écoulement solide en zone semi-aride, de l'Algérie du Nord. *Hydrological sciences journal*, 46(1), 41-53.

Toy, T. J., Foster, G. R., & Renard, K. G. (2002). *Soil erosion: processes, prediction, measurement, and control*. John Wiley & Sons.

Träsch, M. (2019). *Caractérisation expérimentale et numérique du comportement hydrodynamique d'une hydrolienne à membrane ondulante (Doctoral dissertation, Université de Lille (2018-2021))*.

Tricart Jean. Les types de lits fluviaux. In: *L'information géographique*, volume 24, n°5, 1960. pp. 210-214; doi : <https://doi.org/10.3406/ingeo.1960.2025> [https://www.persee.fr/doc/ingeo\\_0020-0093\\_1960\\_num\\_24\\_5\\_2025](https://www.persee.fr/doc/ingeo_0020-0093_1960_num_24_5_2025)

Tsai, C. F., & Chen, M. L. (2010). Credit rating by hybrid machine learning techniques. *Applied soft computing*, 10(2), 374-380.

Tung, Y. K. (1990, December). Comparison of correction factor for dispersion coefficient based on entropy and logarithmic velocity profiles. In *Hydraulic Engineering-Proceedings of the 1990 National Conference* (pp. 463-468). Publ by ASCE.

Turing, A. M. (1950). *Mind. Mind*, 59(236), 433-460.

Turing, A. M., Girard, J. Y., Basch, J., & Blanchard, P. (1995). La machine de Turing (pp. 47-102). Editions du seuil.

Turing, Alan. 1950. Computing Machinery and Intelligence. *Mind* 49, 433 – 460.

Turner, K., & Oza, N. C. (1999, July). Decimated input ensembles for improved generalization. In *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339)* (Vol. 5, pp. 3069-3074). IEEE.

Vapnik, V. (1963). Pattern recognition using generalized portrait method. *Automation and remote control*, 24, 774-780.

Vapnik, V. (2006). Estimation of dependences based on empirical data. Springer Science & Business Media.

Vapnik, V. N. (1995). The nature of statistical learning. Theory.

Vapnik, V., Golowich, S., & Smola, A. (1996). Support vector method for function approximation, regression estimation and signal processing. *Advances in neural information processing systems*, 9.

Verniers, G. (1995). Amenagement ecologiques des berges des cours d'eau: techniques de stabilisation.

Vieira, J., Dias, F. M., & Mota, A. (2004, March). Neuro-fuzzy systems: a survey. In *5th WSEAS NNA international conference on neural networks and applications*, Udine, Italia (pp. 87-92).

Vincent, P., & Bengio, Y. (2001). K-local hyperplane and convex distance nearest neighbor algorithms. *Advances in neural information processing systems*, 14.

Voiron-Canicio, C., Killian, A., Florian, S., & Nicolas, T. (2013). Assessing possible changes in a town's buildings–Fuzzy logic and 3D simulation Applied to the city of Nice.

Von Kaarmaan, T. (1935). Some aspects of the turbulence problem. Guggenheim Aeronautical Laboratory.

Walia, N., Singh, H., & Sharma, A. (2015). ANFIS: Adaptive neuro-fuzzy inference system-a survey. *International Journal of Computer Applications*, 123(13).

Wang, Y., & Witten, I. H. (1997). Induction of model trees for predicting continuous classes. *proceedings of the poster papers of the european conference on machine learning*. Prague: University of Economics, Faculty of Informatics and Statistics.

Were, Kennedy, et al. "A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape." *Ecological Indicators* 52 (2015): 394-403.

Whitfield, PH, Burn, DH, Hannaford, J., Higgins, H., Hodgkins, GA, Marsh, T., & Looser, U. (2012). Réseaux hydrologiques de référence I. Le statut et les orientations futures potentielles des réseaux hydrologiques nationaux de référence pour la détection des tendances. *Journal des sciences hydrologiques*, 57 (8), 1562-1579.

Whitham, G. B. (1958). On the propagation of shock waves through regions of non-uniform area or flow. *Journal of Fluid Mechanics*, 4(4), 337-360.

World Economic Forum, 2016. The fourth industrial revolution: what it means, how to respond. <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/> (accessed 18 Aug 2018)

Woźniak, M., Grana, M., & Corchado, E. (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion*, 16, 3-17.

Xie, M., Zhou, J., Li, C., & Zhu, S. (2015). Long-term generation scheduling of Xiluodu and Xiangjiaba cascade hydro plants considering monthly streamflow forecasting error. *Energy Conversion and Management*, 105, 368-376. <https://doi.org/10.1016/j.enconman.2015.08.009>

Yadav, B., Gupta, P. K., Patidar, N., & Himanshu, S. K. (2020). Ensemble modelling framework for groundwater level prediction in urban areas of India. *Science of the Total Environment*, 712, 135539. <https://doi.org/10.1016/j.scitotenv.2019.135539>

Yang, S., Yang, D., Chen, J., Santisirisomboon, J., Lu, W., & Zhao, B. (2020). A physical process and machine learning combined hydrological model for daily streamflow simulations of large watersheds with limited observation data. *Journal of Hydrology*, 590, 125206.

Yaseen, Z. M., Ebtahaj, I., Bonakdari, H., Deo, R. C., Mehr, A. D., Mohtar, W. H. M. W., ... & Singh, V. P. (2017). Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model. *Journal of Hydrology*, 554, 263-276. <https://doi.org/10.1016/j.jhydrol.2017.09.007>

Yaseen, Z. M., Sulaiman, S. O., Deo, R. C., & Chau, K. W. (2019). An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, 569, 387-408. <https://doi.org/10.1016/j.jhydrol.2018.11.069>

Young, A. R. (2006). Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. *Journal of Hydrology*, 320(1-2), 155-172. <https://doi.org/10.1016/j.jhydrol.2005.07.017>

Zamoum, S., & Souag-Gamane, D. (2019). Monthly streamflow estimation in ungauged catchments of northern Algeria using regionalization of conceptual model parameters. *Arabian Journal of Geosciences*, 12, 1-14. <https://doi.org/10.1007/s12517-019-4487-9>

Zeghmar, A., Marouf, N., & Mokhtari, E. (2022). Assessment of soil erosion using the GIS-based erosion potential method in the Kebir Rhumel Watershed, Northeast Algeria. *Journal of Water and Land Development* (52).

Zeroual, A. (2009). Contribution à la modélisation des courbes de Tarage (cas de bassin versant côtier Algérois (Doctoral dissertation).

Zeyneb, T., Nadir, M., & Boualem, R. (2022). Modeling of suspended sediment concentrations by artificial neural network and adaptive neuro fuzzy interference system method—study of five largest basins in Eastern Algeria. *Water Practice and Technology*, 17(5), 1058–1081. <https://doi.org/10.2166/wpt.2022.050>.

Zhang, C., & Ma, Y. (éd.). (2012). *Apprentissage automatique d'ensemble : méthodes et applications*. Springer Science et médias d'affaires.

Zhang, D. (2000, June). Applying machine learning algorithms in software development. In *Proceedings of the 2000 Monterey workshop on modeling software system structures in a fastly moving scenario* (pp. 275-291).