

Qualité des données hydrométriques prédites par une approche machine learning dans l'évaluation de la ressource en eau au nord de la Côte d'Ivoire

Kamenan Satti Jean-Robert

Doctorant, EDP-Science et Techniques de l'Ingénieur, Institut National Polytechnique Houphouët-Boigny
Yamoussoukro, Côte d'Ivoire

Adja Miessan Germain

Enseignant-Chercheur, Hydrogéologie, Sciences de la Vie et de la Terre (SVT), École Normale Supérieure (ENS), Abidjan, Côte d'Ivoire

Youan Ta Marc

Enseignant-Chercheur, Hydrogéologie, Télédétection, Base de Données et WEBSIG, Centre Universitaire de Recherche Appliquée en Télédétection (CURAT), Université Félix Houphouët-Boigny,
Abidjan, Côte d'Ivoire

Kouassi Amani Michel

Enseignant-Chercheur, Hydrogéologie, Sciences de la Terre et des Ressources Minières (STeRMi), Laboratoire des Sciences et Techniques de l'Ingénieur, Institut National Polytechnique Félix Houphouët-Boigny (INP-HB), Yamoussoukro, Côte d'Ivoire

Njeugeut Mbiafeufu Amandine Carine

Chercheur, Centre Universitaire de Recherche Appliquée en Télédétection (CURAT), Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire

[Doi:10.19044/esj.2024.v20n36p155](https://doi.org/10.19044/esj.2024.v20n36p155)

Submitted: 19 October 2024
Accepted: 12 December 2024
Published: 31 December 2024

Copyright 2024 Author(s)
Under Creative Commons CC-BY 4.0
OPEN ACCESS

Cite As:

Jean-Robert K.S., Germain A.M., Youan Ta M., Kouassi A.M. & Njeugeut Mbiafeufu A.C. (2024). *Qualité des données hydrométriques prédites par une approche machine learning dans l'évaluation de la ressource en eau au nord de la Côte d'Ivoire*. European Scientific Journal, ESJ, 20 (36), 155. <https://doi.org/10.19044/esj.2024.v20n36p155>

Résumé

L'écoulement des rivières revêt un intérêt particulier dans la gestion rationnelle des ressources en eau. Son estimation permet d'évaluer quantitativement la disponibilité des ressources en eau, ainsi que de prévenir

des catastrophes naturelles telles que les inondations et les sécheresses. Ces estimations, généralement effectuées à l'aide de méthodes statistiques d'apprentissage automatique, sont parfois biaisées, entraînant des écarts par rapport aux observations *in situ*. Ainsi, cette étude a pour objectif d'évaluer la qualité des données de débit de la rivière Bagoé prédites par des réseaux neuronaux, afin d'aider les décideurs dans leurs projets de développement durable. Pour ce faire, deux modèles basés sur des réseaux neuronaux ont été développés pour prévoir les variations des débits mensuels de la rivière Bagoé aux instants t et $t+3$ respectivement. Les modèles sont validés en utilisant le critère de Nash (%), le coefficient de corrélation de Pearson (R), le rapport du débit maximal et le critère de robustesse. Les résultats obtenus indiquent que les modèles neuronaux développés expliquent plus de 82 % des débits mensuels de la rivière Bagoé. Ces derniers varient respectivement de 0,002 m³s⁻¹ à 300,12 m³s⁻¹, avec une moyenne de 31,10 m³s⁻¹ en calibration, et de 0,002 m³s⁻¹ à 243,32 m³s⁻¹, avec une moyenne de 31,29 m³s⁻¹ en validation. Cependant, des décalages dans des débits extrêmes ont été observés à la fois en calibration et en validation.

Mots-clés: Données de débit, Prévision, Machine Learning, Fleuve Bagoé, Côte d'Ivoire

Quality of hydrometric data predicted by a machine learning approach in water resource assessment in northern Cote d'Ivoire

Kamenan Satti Jean-Robert

Doctorant, EDP-Science et Techniques de l'Ingenieur, Institut National Polytechnique Houphouët-Boigny
Yamoussoukro, Côte d'Ivoire

Adja Miessan Germain

Enseignant-Chercheur, Hydrogéologie, Sciences de la Vie et de la Terre (SVT), École Normale Supérieure (ENS), Abidjan, Côte d'Ivoire

Youan Ta Marc

Enseignant-Chercheur, Hydrogéologie, Télédétection, Base de Données et WEBSIG, Centre Universitaire de Recherche Appliquée en Télédétection (CURAT), Université Félix Houphouët-Boigny,
Abidjan, Côte d'Ivoire

Kouassi Amani Michel

Enseignant-Chercheur, Hydrogéologie, Sciences de la Terre et des Ressources Minières (STeRMi), Laboratoire des Sciences et Techniques de l'Ingénieur, Institut National Polytechnique Félix Houphouët-Boigny (INP-HB), Yamoussoukro, Côte d'Ivoire

Njeugeut Mbiafeu Amandine Carine

Chercheur, Centre Universitaire de Recherche Appliquée en Télédétection (CURAT), Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire

Abstract

River flow is of particular interest in the rational management of water resources. Their estimation enables quantitative assessment of water resource availability, as well as anticipation of natural disasters such as floods and droughts. These estimates, generally made using statistical methods or machine learning, are sometimes biased, resulting in variations that differ from those observed in situ. This study aims to evaluate the quality of flow data predicted by neural networks for the Bagoé River to support sustainable development planning. To achieve this, two neural network-based models were developed to forecast variations in monthly flows in the Bagoé River at time t and time $t+3$ respectively. The models are validated using the Nash criterion (%), Pearson's correlation coefficient (R), the maximum flow ratio and the robustness criterion. The results obtained indicate that the neural models developed expressed more than 82% of the monthly flows of the Bagoé River. These ranged respectively from $0.002 \text{ m}^3\text{s}^{-1}$ to $300.12 \text{ m}^3\text{s}^{-1}$, with

an average of $31.10 \text{ m}^3\text{s}^{-1}$ in calibration, and $0.002 \text{ m}^3\text{s}^{-1}$ to $243.32 \text{ m}^3\text{s}^{-1}$, with an average of $31.29 \text{ m}^3\text{s}^{-1}$ in validation. However, some shifts in extreme flows were observed in both calibration and validation.

Keywords: Flow data, forecasting, Machine Learning, Bagoé River, Ivory Coast

1. Introduction

Les eaux de surface ont toujours occupé une place primordiale dans l'organisation et le fonctionnement des sociétés humaines. En Afrique particulièrement, elles sont capitales dans l'approvisionnement en eau pour la consommation et essentielles dans le développement des activités industrielles, agricoles et pastorales (Camara & Bangoura, 2017; Kohnert, 2023). Depuis ces dernières années, ces ressources en eau sont fréquemment menacées par les effets du changement climatique, ressentis essentiellement en une baisse de la pluviométrie et une chute des débits des cours d'eau pouvant entraîner parfois des risques d'évènements d'extrême hydrologique (sécheresse) (Ardoin-Bardin et al., 2005; Sané & Faye, 2015; Beguería et al., 2023; EL Yadari & Idrissi, 2023). Anticiper l'occurrence de ces catastrophes, demeure l'un des défis à l'échelle tant nationale (Adja et al., 2019 ; Kouao et al., 2022) qu'internationale (Cardinal et al., 2023). Bien évidemment, cela nécessite la connaissance des débits des cours d'eau qui est l'un des paramètres clés permettant d'évaluer quantitativement la ressource en eau (Souanef, 2015). De façon traditionnelle, les débits des cours d'eau sont déterminés par des instruments de mesure au sol, par des enquêtes et par des méthodes de calcul. Aujourd'hui avec la dégradation des stations de mesure en vue de collecter les données *in situ*, des méthodes de calcul sont couramment utilisées pour estimer ou combler les données manquantes de débits. Dans la région de la Bagoé au Nord de la Côte d'Ivoire où l'utilisation de l'eau est d'une nécessité permanente et abondante compte tenu des différentes activités anthropiques qui demandent une quantité énorme en eau à savoir l'agriculture, l'élevage, la construction des barrages agro pastoraux, les cours d'eau sont sujets à des effets du changement climatique comme révélés par les travaux de Adja (2009) et de Adja et al. (2019). Quoi qu'il existe des stations de mesure des débits sur cet espace, leur dégradation au fil du temps par manque d'entretien et aussi le coût très élevé des équipements de pointe en vue de collecter les données *in situ* constitue un handicap dans la bonne conduite d'une étude hydrologique. Dès lors, l'usage des méthodes de calcul pour estimer ou pour combler des données manquantes de débits s'avère être une alternative. Aujourd'hui avec le progrès technique et la puissance de calcul des ordinateurs de dernière génération, les algorithmes du Machine Learning sont couramment utilisés en économie, en urbanisme et

particulièrement en environnement et en hydrologie avec des résultats très satisfaisants (Jofack, 2016 ; Assoko, 2022 ; Njeugeut et *al.*, 2023). En effet, ces outils s'avèrent être performants motivés par leur propriété d'efficacité et de flexibilité. Mais la qualité des valeurs exprimées par ces derniers, reste toutefois problématique pouvant même entacher toute étude hydrologique. Ainsi, disposer une chronique de valeurs estimées de qualité acceptable, respectant les variations de celles observées sur le terrain reste l'un des défis majeurs. C'est dans cette optique, que la présente étude ambitionne évaluer la qualité des débits du fleuve Bagoé estimés par les modèles de prévision issus du Machine Learning pour aider les acteurs locaux dans la planification et la priorisation des projets d'aménagement dans la région de la Bagoé.

2. Présentation de la zone d'étude

La zone d'étude est constituée par une portion du bassin versant de la Bagoé, recoupant les départements de Boundiali et Kouto, situés au Nord de la Côte d'Ivoire (Figure 1). Elle fait partie de la plaine alluviale du fleuve Bagoé (un affluent du fleuve Niger). Le fleuve Bagoé s'y écoule dans une direction sud-nord pour se jeter dans le fleuve Bani (Assouman et *al.*, 2016) au Mali. Le climat est de type soudano-sahélien à guinéen marqué par deux saisons : une courte saison des pluies de mai à octobre et une longue saison sèche de novembre à avril (Girard, 1961 ; Assouman et *al.*, 2016). Les pluies d'avril à juin sont très faibles et espacées. Le maximum de pluie est enregistré aux mois d'août et septembre (Adja, 2009). Dans cette zone, le pic pluviométrique se signale particulièrement en août (Assouman et *al.*, 2016). La moyenne annuelle précipitée avoisine les 1400 mm. La zone d'étude connaît de fortes insolation avec une température ambiante moyenne de 26,5°C. Le relief est accidenté dans la partie ouest du bassin où s'observent de nombreux pointements rocheux ; lesquels dépassent parfois les 600 mètres. Il est nettement plus atténué et aplani dans l'Est et le Nord du bassin. La végétation est constituée de savane arborée, arbustive ou herbeuse. Elle est partiellement défrichée dans les régions les moins accidentées et les plus peuplées pour laisser place à diverses cultures, en occurrence la culture du riz pratiquée dans les bas-fonds. Les peuples de la zone d'étude sont majoritairement Sénoufo. S'y ajoutent les peuls venus des pays voisins. Ils tirent leurs revenus principalement des activités agricoles et pastorales ; lesquelles sont fortement tributaires des ressources en eau.

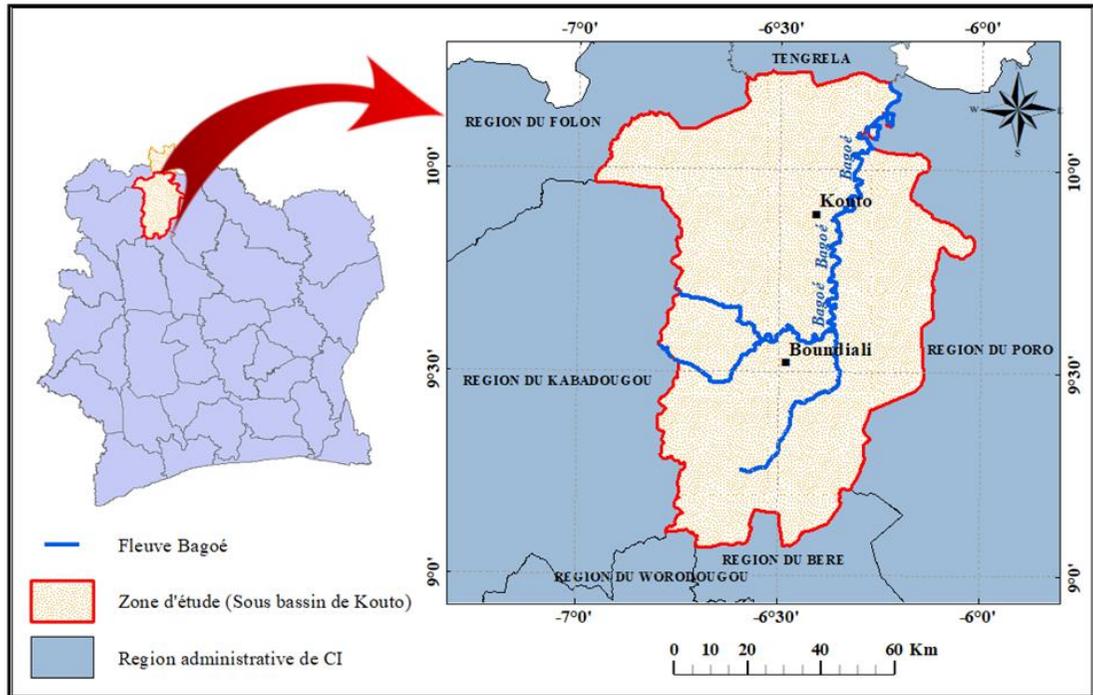


Figure 1 : Localisation de la zone d'étude

3. Matériel et méthodes

3.1. Matériel

Les données utilisées dans cette étude sont constituées essentiellement des données mensuelles de pluie, de température et de débit provenant respectivement de la station pluviométrique de Boundiali, de la station synoptique de Korhogo et de la station hydrométrique de Kouto. Ces données sont longues de 56 années (1961-2016) pour les données de pluie, de 56 années (1961-2016) pour les données de température et de 35 années (1961-1995) pour celles des débits. Elles ont été reconstituées à partir des bases de données pluviométriques de sociétés et services à savoir : la Société d'Exploitation et de Développement Aéroportuaire et Météorologique (SODEXAM), la Sous-Direction des Services de l'Hydrologie, la Direction Régionale du Ministère de l'Agriculture desdites villes, les services de Société Ivoire Coton (Ex CIDT) et de la « Climatic Research Unit (CRU) » de l'Université d'East Anglia à Norwich (Angleterre) (Mahé et al., 2005; New et al., 1998). Les données de températures ont permis d'évaluer l'évapotranspiration potentielle qui elle-même est déterminée par la méthode de Thornthwaite.

3.2. Méthodes

Cette étude a adopté une méthodologie déclinée en trois points que sont la mise en place des modèles neuronaux, la prévision des débits du fleuve Bagoé et l'évaluation de la qualité des débits estimés.

3.2.1. Mise en place des modèles à base de réseaux de neurones

- **Présentation des reseaux de neurones artificiels (RNA)**

Un réseau de neurones artificiels se définit comme un arrangement de neurones artificiels qui apprend à modéliser l'hydro système à travers des exemples bien sélectionnés au cours d'un certain nombre d'itérations finis. Selon leur architecture, plusieurs types de modèles neuronaux existent parmi lesquels le perceptron multicouche (PMC) qui est couramment utilisé pour la simulation et la prévision des débits mensuels (Koffi, 2007).

- **Construction et optimisation des modeles de reseaux de neurones**

L'approche adoptée dans cette phase a nécessité six (06) points importants dont la construction d'un jeu de donnée de qualité, la selection du vecteur d'entrée pertinent, la détermination d'une architecture du reseau adapté, le type d'apprentissage, la règle d'optimisation ainsi que les critères de validation des modèles. Le jeu de données composés des données de pluie, d'évapotranspiration et de débit du cours d'eau de la période 1961-1995, a été exprimé dans la même échelle de variation comprise entre -1 et 1 à travers la formule de l'équation 1 (Li & Liu, 2011) afin d'éviter la saturation et la mauvaise performance des modèles. L'ensemble de données obtenu est subdivisé en trois lots dont le premier lot (1961-1985) représentant 70% des données, a servi à la période d'apprentissage du modèle devant permettre au réseau d'apprendre ; le deuxième lot (1986-1990) qui concerne 15% des données, est utilisé pendant la période de test et enfin le troisième lot (1991-1995) qui concerne les 15% des données restantes, a été utilisé hors apprentissage et a servi à la validation des modèles conçus. Concernant le vecteur d'entrée pertinent des modèles, il a été sélectionné suivant l'approche combinée de la corrélation croisée (Bowden et *al.*, 2005) et de l'essai-erreur qui sont des expériences recommandées par Jofack(2016). En effet, la corrélation croisée dont sa formulation mathématique (Stollsteiner et *al.*, 2012) est donnée par l'équation (2), a permis de déterminer le décalage mensuel et l'essai-erreur a permis de tester et d'évaluer les différentes combinaisons du vecteur d'entrée. Quant à l'architecture neuronale, le modèle perceptron multicouche à une couche de sortie permettant de prédire le débit mensuel est retenu comme modèle neuronal de cette étude. Quant au nombre de neurones dans chaque couche du resseau et le choix des fonctions d'activation associés aux neurones, ils sont déterminés par l'approche expérimentale essai-erreur. Concernant l'optimisation, les modèles neuronaux conçus ont été utilisé en mode supervisé avec la règle d'optimisation Broyden-

Fletcher-Goldfarb-Shanno (BFGS) (Liu, 2024). La validation des performances des modèles est assurée par quatre (04) mesures de qualité dont le critère de Nash, le coefficient de Pearson, le rapport des débits maximaux et le critère de robustesse (Dechemi et al., 2003 ; Tarik & Dechemi, 2004 ; Koffi et al., 2014). La description et la formulation mathématiques de tous ces critères sont consignées dans le Tableau 1.

$$\hat{X} = \frac{(Y_{max}-Y_{min}) \times (X-X_{min})}{(X_{max}-X_{min})} + Y_{min} \quad (1)$$

Avec X = élément du vecteur (entrée ou sortie) à normaliser, \hat{X} = valeur normalisée de X ; X_{max} = valeur du plus grand élément du vecteur à normaliser ; X_{min} = valeur du plus petit élément du vecteur à normaliser ; Y_{max} = valeur maximale, dans le cas général elle vaut 1 ; Y_{min} = C'est la valeur minimale, dans le cas général elle vaut -1.

$$R_{xy}(t_0) = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_{i-t_0} - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_{i-t_0} - \bar{y})^2}} \quad (2)$$

Où t_0 = décalage ou retard mensuel (0 mois comme début, 1, ..., 12 mois) ; N = nombre d'observation ; \bar{x} = la moyenne de la variable d'entrée (Pluie, température, ETP et débit). \bar{y} = la moyenne de la variable de sortie (débit) ;

Tableau 1 : Description des différents critères de performances

Critères	Formules mathématiques	Critères
Nash	$\text{Nash}(\%) = (100) \times \left[1 - \frac{\sum_{i=1}^N (Qob_i - Qcal_i)^2}{\sum_{i=1}^N (Qob_i - \bar{Q})^2} \right]$ <p>(3)</p>	Le modèle est d'autant meilleur que cette valeur est proche de 100 %. Un Nash inférieur à 60 % est l'indication d'un modèle défaillant
Person	$r = \frac{\sum_{i=1}^N qob_i qcal_i}{\sqrt{\sum_{i=1}^N qob_i^2 \sum_{i=1}^N qcal_i^2}}$ <p>(Brou et al., 2017)</p> <p>(4)</p>	$R \geq 0,80$ alors très forte corrélation entre variables ; $0,5 \leq R < 0,80$ alors forte Corrélation entre variables. $R < 0,5$ alors faible corrélation entre variables
Rapport des débits Maximaux	$I_{max} = \frac{Q_{cal_{max}}}{Q_{obs_{max}}}$ <p>(Tarik et Dechemi, 2004)</p> <p>(5)</p>	Plus cet indice tend vers 1 plus la reproduction du débit maximal mensuel observé est très bonne
Robustesse	Valeur absolue de la différence de la valeur du critère de Nash-Sutcliffe en calage et en validation (Koffi et al., 2014)	Le modèle est dit robuste si la valeur de critère est inférieure ou égal à 10%

Avec Qob_i et $Qcal_i$ respectivement les débits mesurés (observés) et calculés pour les $i \in \{1, 2, \dots, N\}$; \bar{Q} est la moyenne des débits mesurés (observés). $qob_i = Qob_i - \bar{Qob}$ et $qcal_i = Qcal_i - \bar{Qcal}$.

Q_{ob_i} et Q_{cal_i} respectivement les débits mesurés (observés) et calculés pour les $i \in \{1,2, \dots, N\}$, $\overline{Q_{ob}}$ et $\overline{Q_{cal}}$ sont les moyennes respectives des débits observés et calculés ; N = nombre d'entrées

3.2.2. Estimation des débits par les réseaux de neurones

L'estimation des débits du fleuve Bagoé devant aboutir à la prévision de ce paramètre a consisté en une modélisation des débits mensuels par deux modèles à base des réseaux de neurones notés P_I et P_{II} conçus et optimisés. Tous ces modèles ont utilisé une combinaison des données de pluie, d'évapotranspiration (ETP) et de débit. Deux cadres de prévision ont été clairement définis. Il s'agit: La prédiction à « l'instant t » : Le modèle prédit le débit au moment t (mois en cours) en considérant que les valeurs pour la pluie et pour l'ETP jusqu'au moment t , ainsi que pour les débits aux moments précédents jusqu'à $t-1$, sont disponibles. Cette prévision est assurée par le modèle P_I ;

La prédiction à « l'instant $t+3$ » : Le modèle prédit le débit du 3ème prochain mois en utilisant le vecteur des entrées de la prédiction à « l'instant t ». Cette prévision est assurée par le modèle P_{II} .

3.2.3. Evaluation de la qualité des débits estimés par les modèles conçus

Cette étape est effectuée à l'aide de l'analyse des résidus calculés à cet effet. Ces derniers se définissant comme l'écart entre les débits observés et les débits calculés par les modèles (Bodian, 2011). Ensuite, une analyse statistique consistant à vérifier deux points de qualité, est effectuée. Selon Lek et *al.* (1995), le premier consiste à vérifier si les résidus calculés sont stationnaires par rapport à la variable estimée en sortie et pour le second si la moyenne des résidus est statistiquement nulle. L'examen du premier point est réalisé par une analyse des représentations graphiques des résidus calculés en fonction des débits estimés par le modèle. Ensuite la vérification du deuxième point est réalisée par une comparaison entre la distribution des résidus et la distribution de la loi normale associée.

4. Résultats

4.1. Modèles conçus et débits prédits

- Modèles conçus

À l'issue de plusieurs expériences réalisées, les paramètres les plus optimaux des modèles neuronaux développés ont été retenus pour prédire le débit du fleuve Bagoé. Ces derniers sont alors consignés dans le tableau 2.

Tableau 2 : Récapitulatif des architectures des modèles retenus

Modèles	Vecteur d'entrée	Nombre neurones	Architecture	Fonction d'activation
P _I	$ETP_t; ETP_{t-1}; \dots \dots; ETP_{t-11}$	24	27-24-1	Exponentielle - Logistique
	$P_t; P_{t-1}; P_{t-2}; P_{t-3}, \dots, P_{t-11}$			
	$ETP_t; ETP_{t-1}; \dots \dots; ETP_{t-11}$			
	$Q_{t-1}; Q_{t-2}; Q_{t-3}$			
P _{II}	$P_t; P_{t-1}; P_{t-2}; P_{t-3}, \dots, P_{t-11}$	22	27-22-1	Exponentielle - Logistique
	$ETP_t; ETP_{t-1}; \dots \dots; ETP_{t-11}$			
	$Q_{t-1}; Q_{t-2}; Q_{t-3}$			

L'analyse du tableau 2 montre que le modèle P_I possède vingt-quatre (24) neurones sur sa couche cachée et un (1) neurone sur sa couche de sortie qui estime le débit à l'instant t (mois), soit une architecture 27-24-1. Quant au modèle P_{II}, il possède vingt-deux (22) neurones sur sa couche cachée et un (1) neurone sur sa couche de sortie qui prédit le débit du fleuve du 3ème prochain mois, soit une architecture 27-22-1. La fonction exponentielle et logistique sont utilisées comme fonction d'activation respectivement des neurones de la couche cachée et du neurone de la couche de sortie du modèle P_I et P_{II}

- **Débits prédits**

Les performances des modèles développés sont résumées dans le tableau 3.

Tableau 3 : Critères de performance des modèles de simulation

Modèle		Calage			Validation			Critère de robustesse (%)
Code	Structure	Nash (%)	r	i _{max}	Nash (%)	r	i _{max}	
P _I	27-24-1	94,00	0,97	0,96	94,21	0,97	1,03	0,21
P _{II}	27-22-1	88,14	0,93	0,82	82,75	0,91	0,98	5,39

L'analyse du tableau 3 montre que les valeurs du critère de Nash calculées sont de 94,00 % pour le modèle P_I et de 88,14 % pour le modèle P_{II} en calage et respectivement de 94,21 % et de 82,75 % en validation ; les valeurs du critère de robustesse (i_{max}) calculées sont de 0,21 % pour P_I et de 5,39 % pour P_{II}. Elles sont toutes inférieures à 10% ; les valeurs du coefficient de corrélation (r) déduites, sont de 0,97 pour P_I et de 0,93 pour P_{II} en calage et respectivement de 0,97 et de 0,91 en validation ; le rapport des débits maximaux est de 0,96 pour P_I et de 0,82 pour P_{II} en calage et respectivement de 1,03 et 0,98 en validation. Toutes ces valeurs obtenues, traduisent une bonne performance des modèles de prévision développés. Le critère de robustesse de chaque modèle étant inférieur à 10%, montre que les modèles de prévision sont robustes par conséquent stables. Afin d'apprécier l'évolution

des débits, les hydrogrammes des débits prédits ont été représentés par les figures 2 et 3 en calage et par les figures 4 et 5 en validation

En calage:

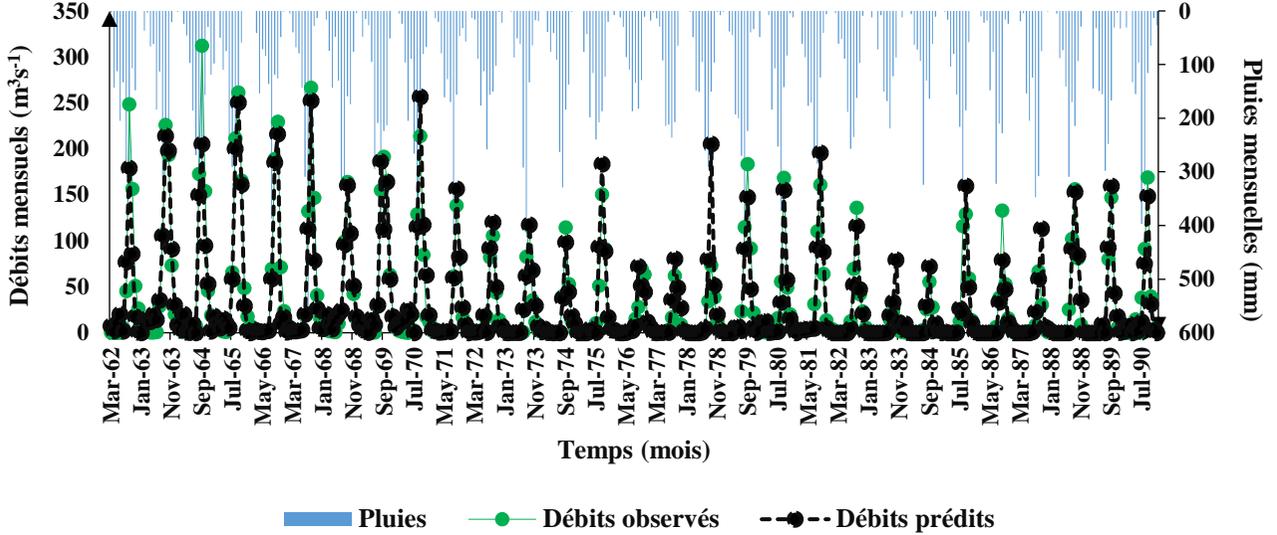


Figure 2 : Hydrogrammes des débits mesurés et prédits par le modèle P_I en calage à la station de Kouto

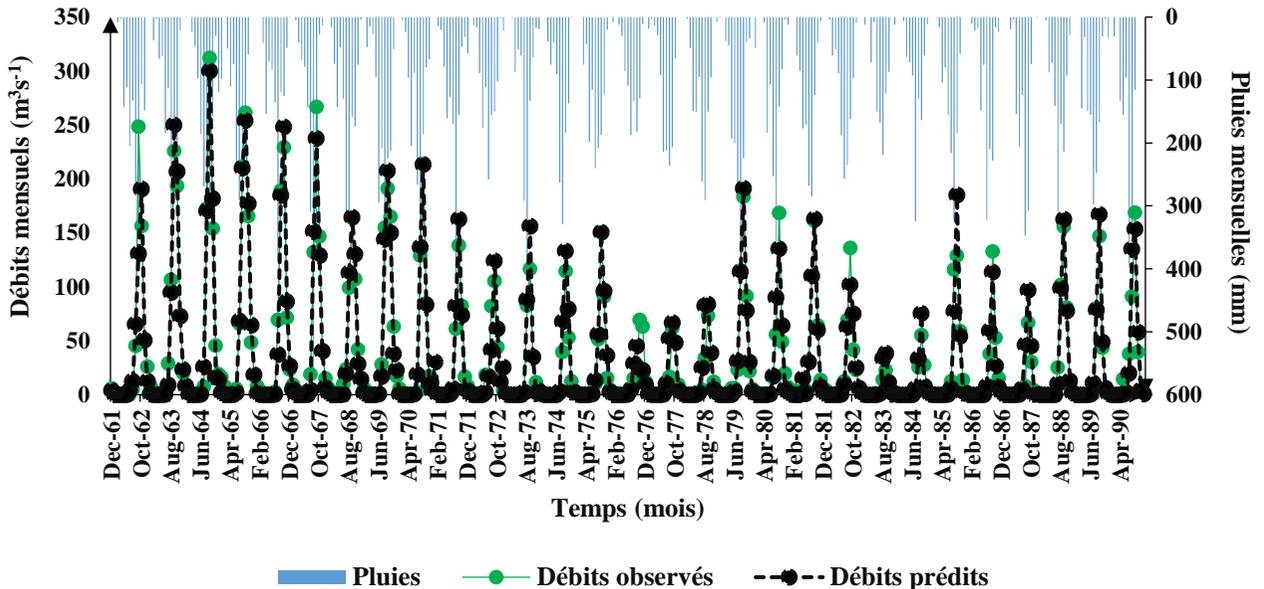


Figure 3 : Hydrogrammes des débits mesurés et prédits par le modèle P_{II} en calage à la station de Kouto

L'analyse des hydrogrammes (Figure 2 et 3) montre que tous les modèles de prévision développés reproduisent bien la dynamique des

écoulements du fleuve Bagoé à la station de Kouto en phase de calage. Une tendance de bonne synchronisation entre les débits prédits et ceux mesurés *in situ*, est observée. Cependant des décalages entre les débits mesurés et ceux prédits par ces modèles de prévision sont observés. Par exemple certains debits de crues (septembre 1962 ; septembre 1963; septembre 1964; octobre 66; septembre 70; septembre 1975; septembre 78; septembre 1979; septembre 1981; septembre 1986) sont mal estimés par les modèles conçus.

En validation

Les hydrogrammes mesurés et prédits par les modèles en calage et en validation sont illustrés par les figures 4 et 5.

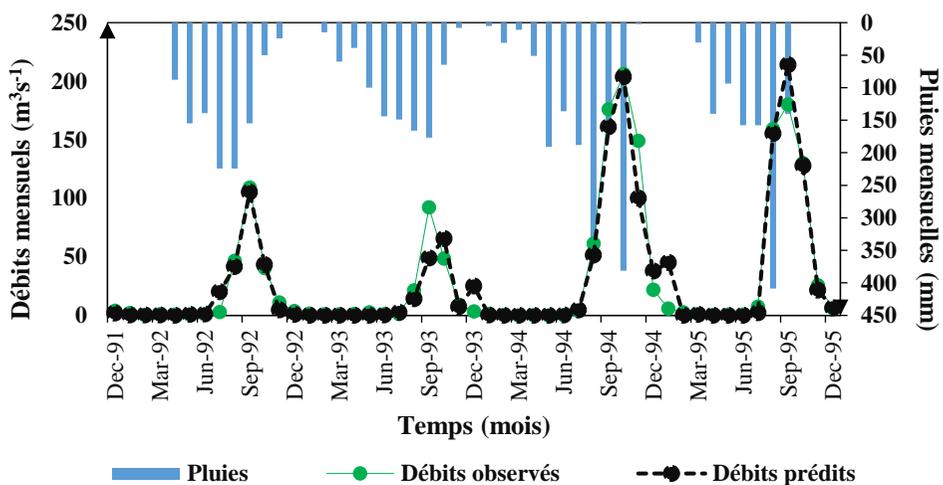


Figure 4 : Hydrogrammes des débits mesurés et prédits par le modèle P_I en validation à la station de Kouto

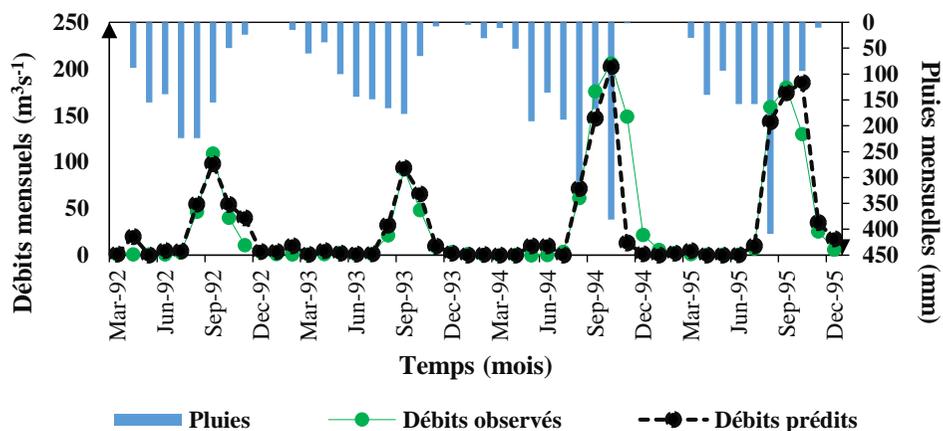


Figure 5 : Hydrogrammes des débits mesurés et prédits par le modèle P_{II} en validation à la station de Kouto

L'analyse des hydrogrammes (Figure 4 et 5) montre que tous les modèles de prévision développés reproduisent bien la dynamique des écoulements du fleuve Bagoé à la station de Kouto dans la phase de validation. En effet, une bonne synchronisation entre les débits prédits et ceux mesurés *in situ* est observée. Toutefois, cette évolution synchrone des débits laisse apparaître des décalages entre les débits mesurés et ceux prédits par ces modèles de prévision. Certains débits de crues observés dans les périodes de grandes eaux sont parfois mal estimés. Par exemple, les débits de mai 1992 ; novembre 1992 ; novembre 1994 ; décembre 1994 marquant les périodes fortes du régime hydrologique du bassin sont par endroit mal estimés. Au regard de tous ces résultats obtenus, une comparaison entre les performances des modèles de prévision montre que le modèle P_1 constitue le meilleur prédicteur des débits du fleuve Bagoé.

4-2. Qualité des données estimées par les modèles de prévision

- **Qualité des données estimées par le modèle P_1**

La description statistique des débits observés et prédits par le modèle P_1 est consignée dans les tableaux 3.

Tableau 3 : Paramètres statistiques de la prévision des débits mensuels par le modèle P_1

	Calage			Validation		
	Observés	Prédits	Résidus	Observés	Prédits	Résidus
Moyenne	31,27	31,22	0,00	31,56	30,90	0,65
Minimum	0	0,002	-55,97	0,001	0,002	-39,20
Maximum	312,17	300,12	117,87	206	216,82	49,10
Ecart type	55,81	56,78	13,60	56,59	55,10	13,56

L'analyse du tableau 3 montre que pour le modèle P_1 , en calage, les moyennes des valeurs de débits estimés et observés sont presque identiques. L'écart type des valeurs estimées est légèrement élevé à celui des valeurs observées avec une moyenne des erreurs résiduelles quasi nulle. En effet, les débits prédits par le modèle P_1 oscillent entre $0,002 \text{ m}^3\text{s}^{-1}$ et $300,12 \text{ m}^3\text{s}^{-1}$, avec une moyenne de $31,22 \text{ m}^3\text{s}^{-1}$. Le résidu moyen traduisant l'écart moyen entre les débits observés et les débits prédits est quasi nul. Ce résultat montre qu'en calage, P_1 restitue en moyenne presque le débit observé. En validation, les débits prédits par P_1 varient de $0,002 \text{ m}^3\text{s}^{-1}$ à $216,82 \text{ m}^3\text{s}^{-1}$, avec une moyenne de $30,90 \text{ m}^3\text{s}^{-1}$. L'écart résiduel moyen entre ces valeurs observées et prédites, reste dans le voisinage de la valeur nulle ($0,65 \text{ m}^3\text{s}^{-1}$). Le résidu qui traduit l'écart entre les débits observés et ceux prédits varie de $-55,97 \text{ m}^3\text{s}^{-1}$ à $117,87 \text{ m}^3\text{s}^{-1}$ en validation et de $-39,20 \text{ m}^3\text{s}^{-1}$ à $49,10 \text{ m}^3\text{s}^{-1}$ en validation. Ce résultat montre qu'en validation P_1 restitue en moyenne des valeurs qui sont dans le voisinage des valeurs de débits observés.

Les figures 6 et 7 présentent respectivement la distribution des résidus et la relation des résidus en fonction des débits prédits par P_1 .

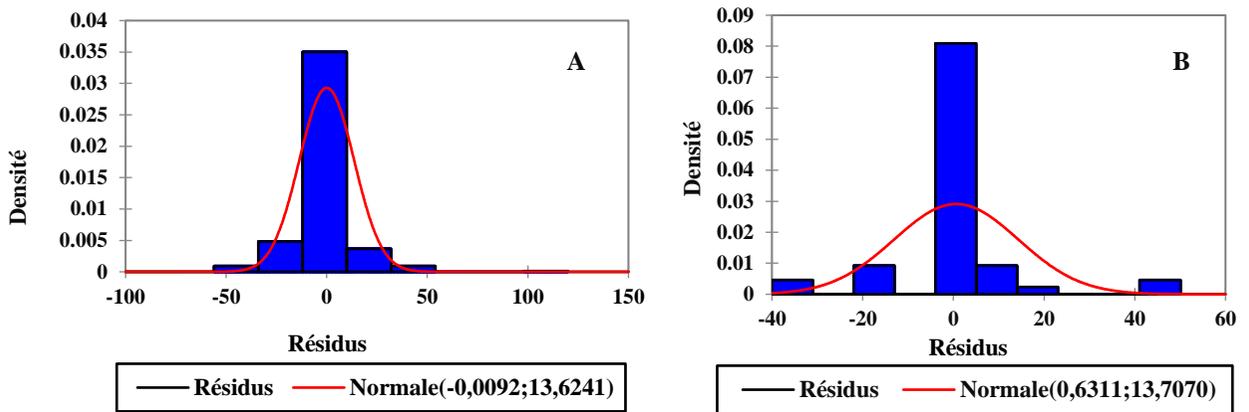


Figure 6 : Distribution des résidus pour le modèle P_1 , A : en calage ; B : en validation

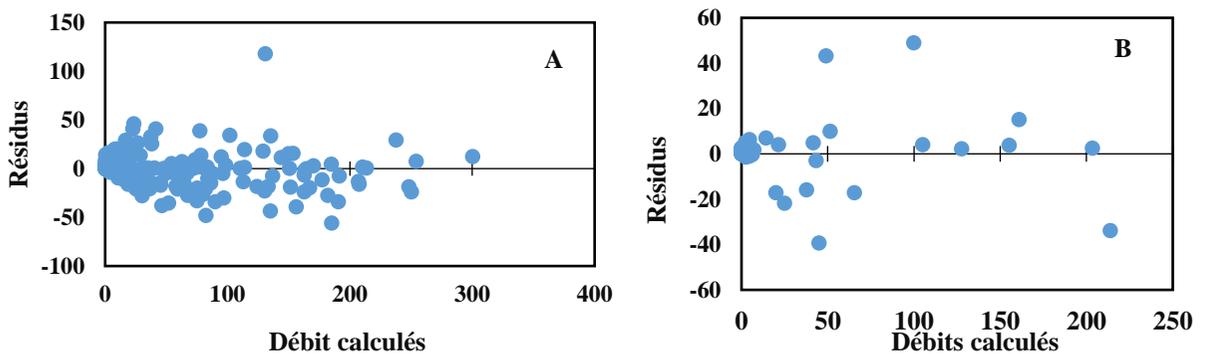


Figure 7 : Relation entre les valeurs estimées et les résidus pour le modèle P_1 , A : en calage ; B : en validation

L'analyse de ces figures montre que les résidus calculés pour le modèle P_1 ne suivent pas une distribution normale (Figure 6) en calage comme en validation. Néanmoins, on remarque un regroupement d'une très forte proportion des valeurs de résidus au niveau de la moyenne nulle (Figure 7). Aussi, la relation entre les résidus et les débits estimés du modèle fournit un bon ajustement des faibles débits qui représente la quasi-totalité des points (Figure 7). Concernant les débits forts, les erreurs sont parfois importantes. Une tendance de sous-estimation ou de surestimation de certains débits forts est observée pour ce modèle à la station de Kouto dans l'ensemble des deux phases (calage et validation).

• **Qualité des données estimées par le modèle P_{II}**

La description statistique des débits observés et prédits par le modèle P_{II}, est consignée dans les tableaux XII.

Tableau I : Paramètres statistiques de la prévision des débits mensuels par le modèle P_{II}

	Calage			Validation		
	Observés	Prédits	Résidus	Observés	Prédits	Résidus
Moyenne	31,27	30,99	0,15	31,56	31,69	-0,13
Minimum	0	0,13	-61,43	0,001	0,18	-37,32
Maximum	312,37	285,08	79,26	206	243,32	58,64
Ecart type	55,61	52,57	17,26	56,59	54,17	14,10

L'analyse du tableau XII montre qu'en calage, les moyennes des valeurs de débits estimés et observés sont presque identiques et une moyenne des erreurs résiduelles très faible (0,15). En effet les débits prédits par le modèle P_{II} oscillent entre 0,13 m³s⁻¹ et 285,08 m³s⁻¹, avec une moyenne de 30,99 m³s⁻¹. Le résidu moyen traduisant l'écart moyen entre les débits observés et les débits prédits, est estimé à 0,15 qui s'inscrit dans le voisinage de la valeur nulle.

En validation, les débits estimés, oscillent entre 0,18 m³s⁻¹ et 243,32 m³s⁻¹, avec une moyenne de 31,69 m³s⁻¹. L'écart moyen entre les débits observés et les débits prédits est estimé à -0,13 m³s⁻¹ qui tend presque vers la moyenne nulle. Le résidu qui traduit l'écart entre les débits observés et ceux prédits varie de -61,43 m³s⁻¹ à 79,26 m³s⁻¹ en calage et de -37,32 m³s⁻¹ à 58,64 m³s⁻¹ en validation. Tous ces résultats traduisent que le modèle P_{II} restitue en moyenne des valeurs qui sont dans le voisinage des valeurs de débits observés. Les figures 8 et 9 donnent respectivement la distribution des résidus et la relation des résidus en fonction des débits simulés pour le modèle P_{II}.

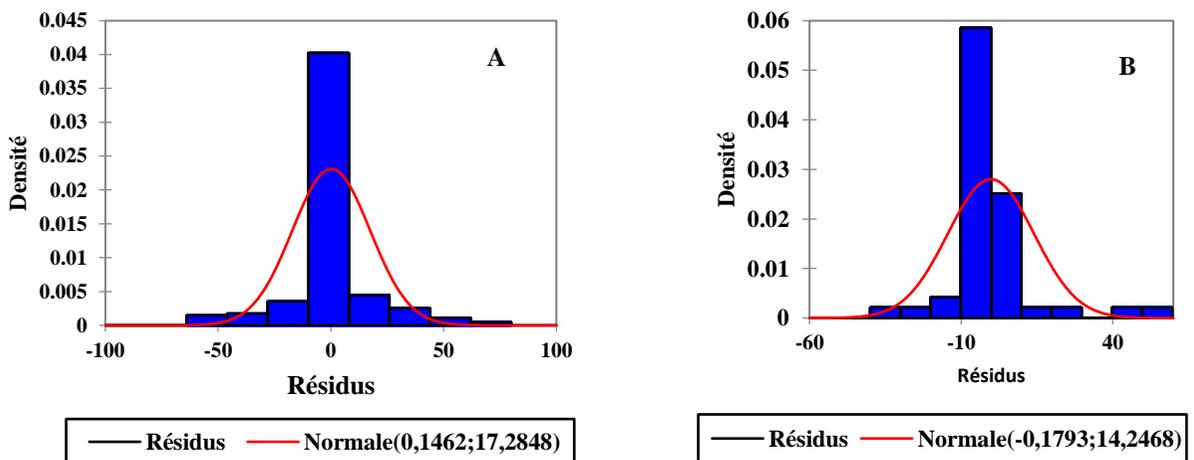


Figure 8 : Distribution des résidus pour le modèle P_{II}, A : en calage ; B : en validation

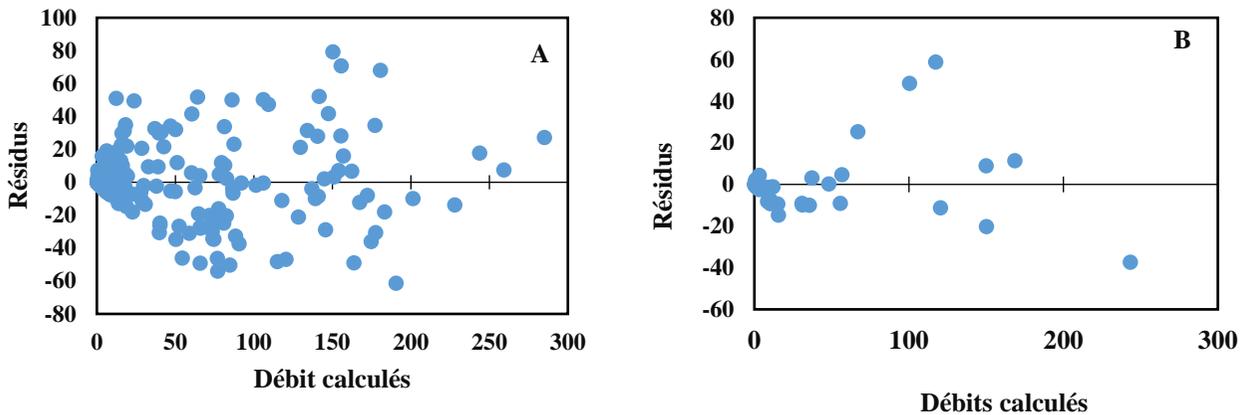


Figure 9 : Relation entre les valeurs estimées et les résidus pour le modèle P_{II} , A : en calage ; B : en validation

L'analyse de ces figures montre que la distribution des résidus pour le modèle P_{II} n'est pas une distribution normale (Figure 8) en calage comme en validation. Néanmoins, on remarque un regroupement d'une très forte proportion des valeurs de résidus au niveau de la valeur nulle (Figure 9). En outre, la relation entre les résidus et les débits prédits par le modèle P_{II} fournit un bon ajustement des débits faibles (Figure 9). Quant aux débits forts, une tendance de sous-estimation ou de surestimation est observée pour ce modèle à la station de Kouto en calage comme en validation.

Discussion

La dynamique de l'écoulement du fleuve Bagoé est évaluée par deux modèles de réseau de neurones P_I et P_{II} assurant respectivement la prévision à l'instant t et la prévision à l'instant $t+3$. Ces modèles de prévision ont estimé l'écoulement du fleuve Bagoé à plus de 93% en calage et à plus de 91 % en validation. En effet, les coefficients de corrélation de Pearson, traduisant la corrélation entre les débits observés et les débits calculés par ces derniers, varient de 0,91 à 0,97. Ces résultats obtenus concordent avec les travaux antérieurs de Koffi et *al.*(2014) ayant utilisés la même méthodologie. En effet, les travaux réalisés par ces auteurs à la station hydrométrique de Tortiya sur le Bandama Blanc dans le cadre des prévisions de débit à l'horizon de 3 mois ont donné une variation du coefficient de corrélation comprise entre 0,88 et 0,91. Concernant les hydrogrammes observés et prédits par les modèles développés, ces derniers sont synchrones et traduisent une variation similaire des débits observés *in situ*. Cependant des décalages sont observés pour des débits forts du regime hydrologique. Plusieurs raisons pourraient expliquer une telle anomalie parmi lesquelles les paramètres intrèques du modèle

neuronal. Des observations similaires ont été relevées de la part de Koffi (2007) quant à l'incapacité des modèles neuronaux à reproduire des débits extrêmes. Dans ses travaux sur le Bandama blanc au Nord de la Côte d'Ivoire, Koffi (2007) a observé l'incapacité des modèles de réseaux de neurones à reproduire certains débits du fleuve qui ne sont pas seulement le fait des seules pluies tombées sur le bassin versant. Parallèlement à cette explication, les RNA durant leur phase d'apprentissage, se familiarisent aux données. Ils extraient des singularités dans ces données. Aussi, dans cette étude de modélisation, les crues s'observent une fois dans l'année et généralement durant la période d'août ou de septembre ou d'octobre. Comme les données de débit de crue sont en nombre très insuffisant par rapport aux autres débits, les modèles RNA disposent donc peu de données de débit de crue afin d'extraire la singularité. De façon Générale, le surajustement et le sous-ajustement sont des problèmes très courants dans le domaine de l'apprentissage automatique, notamment avec les réseaux de neurones artificiels. Les différentes valeurs des erreurs résiduelles calculées dans cette étude, illustrent parfaitement la capacité des réseaux de neurones à sur estimer ou à sous-estimer les valeurs prédites. Les débits estimés par tous les modèles neuronaux conçus, oscillent respectivement entre $0,002 \text{ m}^3\text{s}^{-1}$ et $300,12 \text{ m}^3\text{s}^{-1}$, avec une moyenne de $31,10 \text{ m}^3\text{s}^{-1}$ en calage et $0,002 \text{ m}^3\text{s}^{-1}$ et $243,32 \text{ m}^3\text{s}^{-1}$, avec une moyenne de $31,29 \text{ m}^3\text{s}^{-1}$ en validation.

Conclusion

Cette étude portant sur l'étude de la qualité des valeurs estimées par les algorithmes du Machine Learning, notamment les réseaux de neurones dans l'évaluation des ressources en eau en vue d'anticiper des risques de sécheresse et d'inondation, est initiée pour tenter de faire des prévisions des débits mensuels du fleuve Bagoé à la station hydrométrique de Kouto. Elle a conduit d'abord à la construction des modèles de réseau de neurones et ensuite à leur utilisation en mode prévision de débits. Les résultats obtenus indiquent que ces modèles neuronaux développés pour les débits mensuels ont exprimé plus de 91% de ces débits. Ces derniers oscillent respectivement entre $0,002 \text{ m}^3\text{s}^{-1}$ et $300,12 \text{ m}^3\text{s}^{-1}$, avec une moyenne de $31,10 \text{ m}^3\text{s}^{-1}$ en calage et $0,002 \text{ m}^3\text{s}^{-1}$ et $243,32 \text{ m}^3\text{s}^{-1}$, avec une moyenne de $31,29 \text{ m}^3\text{s}^{-1}$ en validation. L'évolution satisfaisante des débits prédits permet d'exprimer la dynamique de l'écoulement du fleuve Bagoé. Cependant, quelques décalages sont observés au niveau des débits extrêmes en calage comme en validation. Pour les recherches futures, l'application de la présente approche méthodologique aux bassins versants de comportement hydrologique similaire peut être envisagée. Aussi, l'hybridation des réseaux de neurones de prévision avec d'autres algorithmes méta heuristiques, notamment l'algorithme des sauterelles, l'algorithme des poissons artificiels et l'algorithme des particules artificiels

pourraient corriger les anomalies observés. L'identification des hyperparamètres par des fonctions recherches de grille pourrait améliorer les performances des réseaux de neurones dans la prévision des variables hydroclimatiques.

Conflit d'intérêts : Les auteurs n'ont signalé aucun conflit d'intérêts.

Disponibilité des données : Toutes les données sont incluses dans le contenu de l'article.

Déclaration de financement : Les auteurs n'ont obtenu aucun financement pour cette recherche.

References:

1. Adja, M.,G. (2009). Etude de l'état hydrique saisonnier du bassin versant de la Bagoé dans un contexte de variabilité climatique. Départements de Boundiali et Tengréla, (milieux soudano-sahéliens au Nord-Ouest de la Côte d'Ivoire). Thèse de Doctorat de l'Université de Cocody (Abidjan), Côte d'Ivoire, 195 p.
2. Adja, M. G., Djemin, J. E., N'dri, F. O., Kouame, F. K. & Biemi, J.(2019).Contribution des modèles hydrologiques et des SIG à l'étude de l'impact de la variabilité climatique et du changement de l'occupation du sol sur les ressources en eau du bassin versant de la Volta en Côte d'Ivoire, *International Journal of Innovation and Applied Studies ISSN 2028-932*, Vol. 26, (No. 2), pp. 652-667
3. Ardoin-Bardin, S., Dezetter, A., Servat, E., Mahé, G., Paturel, J. E., Dieulin, C., & Casenave, L. (2005). Évaluation des impacts du changement climatique sur les ressources en eau d'Afrique de l'Ouest et Centrale. *Regional Hydrological Impacts of Climatic Change—Hydroclimatic Variability*, pp 194-202.
4. Assoko, A., S.,V. (2022). Conception d'outils de prévision des extrêmes hydrologiques en région tropicale ouest africaine : cas du bassin versant de la Marahoué en côte d'ivoire. Thèse de Doctorat, Institut national polytechnique Felix Houphouët-Boigny, 240p
5. Assouman, S. F., Kone M., & Koli, B. Z., (2016). Morphologie du paysage, variabilite pluviometrique et aménagement de la plaine alluviale de la bagoé : l'exemple de guinguerini au nord de la cote d'ivoire. *Revue de Géographie Tropicale et d'Environnement*, N°1, 13. pp. 5-17
6. Beguería, S., Sans, F., Antigua, I., Cakir R., Domènech, M., Haro-Monteagudo, D., Huang, P., Jódar, J., Lambán, Jiménez, L. J., & Le Coent, P. (2023). Adaptation au changement climatique dans la gestion

- des ressources en eau des Pyrénées. Mémoires scientifiques du projet, Piragua, vol. 2., 243p
7. Bowden, G. J., Dandy, G. C. & Maier, H.R.(2005). Input determination for neural network models in water resources applications: Part 1 - Background and methodology. *Journal of Hydrology*, 301(1-4), pp. 75-92, DOI: 10.1016/j.jhydrol.2004.06.021
 8. Brou, L. A., Kouassi, K. L. K., Konan, K. S., Kouadio, Z. A., Konan, K. F. et Kamagaté B. (2017). Rain-Flow Modeling Using a Multi-Layer Artificial Neural Network on the Watershed of the Cavally River (Côte d'Ivoire). *Journal of Water Resource and Protection*, Volume 09, n°12, pp 1403-1413.
 9. Camara, S., & Bangoura, A. (2017). Valorisation des ressources en eau en guinee en vu d'un developpement economique, social et durable. *Journal of Water and Environmental Sciences*, 1, pp 106-114.
 10. Cardinal, H., Delbec, M., Monier, L., & Willinger, B. (2023). La plateforme Prévi-Bièvre : Un outil de gestion dynamique et de prévision des inondations au service des riverains de la Haute Vallée de la Bièvre. *Novatech 2023: 11e Conférence internationale sur l'eau dans la ville*, 6p
 11. Dechemi, N., Benkaci, T., & Issolah, A. (2003). Modélisation des débits mensuels par les modèles conceptuels et les systèmes neuro-flous. *Revue des sciences de l'eau*, 16(4), pp. 407-424. <https://doi.org/10.7202/705515ar>
 12. El Yadari, S., & Idrissi, A. J. (2023). Changement climatique et effets de la secheresse hydroclimatique sur les ressources en eau et quelques formes d'adaptations : cas du bassin versant d'inaouene, Maroc. *Revue des Etudes Multidisciplinaires en Sciences Economiques et Sociales*, 8(2), pp. 142-151
 13. Fletcher, R.(2000). *Practical methods of optimization*. John Wiley & Sons
 14. Girard, G. (1961). *Etude du bassin de la Haute Bagoé. Résultats de la campagne 1960*, ORSTM, 82 p.
 15. Kharroubi, O., Blanpain, O., Masson, E., & Lallahem, S. (2016). Application du réseau des neurones artificiels à la prévision des débits horaires : Cas du bassin versant de l'Eure, France. *Hydrological Sciences Journal*, 61(3), pp. 541-550.
 16. Koffi, Y. B., Ahoussi, K. E., Kouassi, A. M., Kouassi, O., Kpangui, L. C. & Biemi, J. (2014). Application des réseaux de neurones formels pour la prévision des débits mensuels du Bandama blanc à la station de Tortiya (Nord de la Côte d'Ivoire). *Afrique Science : Revue Internationale Des Sciences et Technologie*, Volume 10, n°3, pp. 134-135

17. Kohnert, D. (2023). L'eau, une bénédiction et une malédiction: comment résoudre les conflits liés à l'eau en Afrique de l'Ouest? Hamburg. En ligne: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-88149-7> , 30p (Consulté le 2 Mars 2024)
18. Kouao, J.-M., Kouassi, A. M., & Kouame, K. A. (2022). Prévion de la sécheresse météorologique et agrométéorologique dans un contexte de changements climatiques en afrique de l'ouest : cas de la côte d'ivoire, *Rev. Ivoir. Sci. Technol.*, 40, pp. 279 – 307
19. Liu, H. (2024). Empowering Predictivity and Speed of hiPSC CM Assays by Machine Learning Approach [Phdthesis, Sorbonne Université]. <https://theses.hal.science/tel-04573685>, 136p
20. Lek, S., Dimopoulos, I., Derraz, M. & El Ghachtoul, Y. (1996). Modélisation de la relation pluie-débit à l'aide des réseaux de neurones artificiels. *Revue des Sciences de l'eau*, Volume 9, n°3, pp. 319-331. <https://doi.org/10.7202/705255ar>
21. Lucea, M. (2006). Modélisation dynamique par réseaux de neurones et machines à vecteurs supports : contribution à la maîtrise des émissions polluantes de véhicules automobiles. Thèse de Doctorat de l'Université PARIS 6, Paris, France, 199 p.
22. Mahé, G., Olivry J.C. & Servat E. (2005). Sensibilité des cours d'eau ouest-africains aux changements climatiques et environnementaux : extrêmes et paradoxes. Colloque AISH Scientific Assembly:
23. New, M., Hulme, M. & Jones, P. (1998). Representing 20th century space-time climate variability. II: Development of 1901–1996 monthly terrestrial climate fields. *J. Climate* 13, pp. 2217–2238
24. Njeugeut, A., C., M., Youan, Ta M., Kamenan, S., J-R., Assoma, T., V., & Jourda J., P. (2023). Cartographie Automatique des Zones Inondées et Evaluation des Dommages dans le District d'Abidjan à l'Aide de l'Imagerie Satellitaire Radar Sentinel-1 Depuis Google Earth Engine. *European Scientific Journal, ESJ*, 22, pp. 124-156. Retrieved from <https://eujournal.org/index.php/esj/article/view/17285>
25. Ouisse, K. (2013). Prévion des crues par modèle de réseau de neurones artificiels : Application au bassin versant de l'Eure (Lille 1). Thèse de Doctorat de l'Université de Lille 1, France, 210 p.
26. Sané, T., & Faye, C. (2015). Le changement climatique dans le bassin versant de la Casamance: Évolution et tendances du climat, impacts sur les ressoources en eau et stratégies d'adaptation., Atelier scientifique et du lancement de l'initiative « Casamance: un réseau scientifique au service du development en Casamance » du 15-17 juin 2015 à Ziguinchor, Sénégal, L'Harmattan, ISBN : 978-2-343-07690-4, pp 89-110

27. Souanef, N. (2015). Analyse fréquentielle des débits max de crues de L'Oued Abiod. Master, Université Mohamed Khider, Algérie, 134 p, <http://thesis.univ-biskra.dz/1556/>
28. Stollsteiner, P., Wuillemier, A., Bessière, H & Seguin, J (2012). Guide pour la détermination des indicateurs piézométriques en vue d'une gestion quantitative de la ressource. [Rapport de recherche] Brgm; Onema. ffhal-02936929ff