Document généré le 2 juin 2025 14:07

Revue des sciences de l'eau
Journal of Water Science

Modélisation des débits mensuels par les modèles conceptuels et les systèmes neuro-flous
A monthly streamflows modelling using conceptual models and neural fuzzy system

N. Dechemi, T. Benkaci et A. Issolah

Volume 16, numéro 4, 2003

URI : https://id.erudit.org/iderudit/705515ar
DOI : https://doi.org/10.7202/705515ar

Aller au sommaire du numéro

Éditeur(s)

Université du Québec - INRS-Eau, Terre et Environnement (INRS-ETE)

ISSN

0992-7158 (imprimé)
1718-8598 (numérique)

Découvrir la revue

Citer cet article

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 / Journal of Water Science*, *16*(4), 407–424. https://doi.org/10.7202/705515ar

Résumé de l'article

La modélisation pluie-débit au pas de temps mensuel, a été étudiée par le biais de quatre modèles qui appartiennent à deux catégories, les modèles conceptuels (modèles à réservoirs), et les modèles basés sur les réseaux de neurones, et la logique floue

Les modèles conceptuels mensuels utilisés sont les modèles de Thornthwaite et Arnell et le modèle GR2M, ainsi que deux modèles représentés par les réseaux de neurones à apprentissage supervisé et le modèle neuro-flou qui combine une méthode d'optimisation neuronale et une logique floue.

Une application de ces modèles a été effectuée sur le bassin de la Cheffia (Nord-Est Algérien), et a confirmé les performances du modèle basé sur la logique floue. Par sa robustesse et son pouvoir d'extrapolation non-linéaire, ce modèle a donné d'excellents résultats, et représente donc une nouvelle approche de la modélisation pluie-débit au pas de temps mensuel.

Tous droits réservés © Revue des sciences de l'eau, 2003

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

https://apropos.erudit.org/fr/usagers/politique-dutilisation/

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

https://www.erudit.org/fr/

REVUE DES SCIENCES DE L'EAU, Rev. Sci. Eau 16/3(2003) 407-424

Modélisation des débits mensuels par les modèles conceptuels et les systèmes neuro-flous

A monthly streamflows modelling using conceptual models and neural fuzzy system

N. DECHEMI*[1], T. BENKACI*[2], A. ISSOLAH[3]

Reçu le 2 juillet 2001, accepté le 18 juin 2003**.

SUMMARY

Rainfall-runoff modelling is very important for environmental issues, as well as for water management. Due to this importance, several models have been developed to describe the transformation of rainfall to runoff. From these models, we can distinguish three categories: conceptual models; physically-based models and black box models. Conceptual models are designed to approximate within their structures the general sub-processes that govern the hydrological cycle, and they are often used because of their simplicity. The physically-based models are generally distributed models, involve complex descriptions using partial derivative equations, and need some parameter calibration to be adjusted or estimated in situ. These models can not be applied on a monthly scale. In contrast, the black box models rely on linear (or nonlinear) relationships between inputs (rainfall) and outputs (runoff), and they have been widely accepted as a practical tool on different time scales.

In this paper, rainfall-runoff modelling on a monthly scale was studied using four models, from two different categories; conceptual models (reservoir models), and models based on artificial neural network and fuzzy logic. The monthly conceptual models used were the Thornthwaite-Arnell model and the GR2M model with two reservoirs. These models are regarded as mathematical models, and are of simple conception with a reduced number of parameters. In addition, these models are considered the most valid. The two other models were based on artificial neural networks and fuzzy logic, which combine neural optimization methods and fuzzy logic. These models incorporate a flexible mathematical structure that is capable of identifying complex nonlinear relationships between input and output data sets. In

1 Professeur, Laboratoire de recherche Construction et Environnement, École Nationale Polytechnique d'Alger, 10, avenue Hassen-Badi, B.P 16182, El Harrach, Alger, Algérie, tél. : (0 21) 52-53-01/03, fax : (0 21) 52-29-73, e.mail : ndechemi@caramail.com

2 Magister, département d'hydraulique, Institut national agronomique d'Alger, 22, boulevard Bougara, Alger, Algérie. B.P 16000, e.mail : benkaci.tarik@caramail.com

3 Chargé de cours, département d'hydraulique, Institut national agronomique d'Alger, 10, avenue Hassen-Badi, El Harrach, Alger, Algérie, e-mail : aliissolah@caramail.com

* Correspondance. E-mail : ndechemi@caramail.com, benkaci.tarik@caramail.com

** Les commentaires seront reçus jusqu'au 30 juin 2004.

contrast to conceptual deterministic models, these models proceed using data learning through input-output systems. Artificial neural network models have been often shown to provide a better representation of the rainfall-runoff relationships. However, it is necessary to investigate different learning methods used with these models.

There are two different learning modes (training). One is data learning (incremental training), which consists of training for each data set, where the weights and biases on the network model are updated each time an input is presented to the network, thus the error between simulated and target (observed) data is minimised for each input. The alternative to data learning is block learning (batch training). In block mode the weights and biases on the network model are updated only after the entire training set has been applied to the network.

We have tried a block learning data method, which consisted of learning from the simulation of all data sets. Thus, it evaluates the influence of this model in the streamflow forecasting in real time.

In Algeria, the droughts recorded during the previous years resulted in a reduction of surface water and in unbalanced resources that affected the phreatic underground water due to intensive exploitation. The results from evaluation studies emphasised the instability and vulnerability of surface water resources. The government has decided to carry out an emergency plan, by constructing several reservoirs and dams over the next few years in different regions of the country. However, several hydrometric gauges are disabled, so the series of hydrometric data are short or have gaps, and thus water resource evaluation has become impossible.

One of the objectives of the monthly rainfall-runoff modelling was estimating the stream flow at the mouth of the watershed, so the rainfall-runoff relationship on a monthly scale represents a solution and a reliable method for water management projects. We have selected and applied four models on data from the Cheffia watershed situated in north-eastern Algeria. The catchment of the Cheffia river includes various sub-basins, and has an area of about 575 km^2. The study was carried out on a twelve-year data set, split into a six-year calibration period, and a six-year validation period. Our research compared the models based on model characteristics, like simplicity and parameterisation, and also conceptual models were compared to parsimonious models. In addition, our research compared modelling results, based on the assessment of quantitative indices and statistics, such as the Nash criterion, the root mean squared error and a comparison of means during the calibration and validation periods.

Model results have confirmed the strong performance of the fuzzy logic based model, for two periods, and this model best stimulated streamflows. Whereas the neural network model based on block learning is unable to reproduce the high runoff values, this model can to be used for simulation of the runoff only. Because of its robustness and non-linear extrapolation power, the neuro-fuzzy logic model gave better results, so it represents a new method of rainfall-runoff modelling in monthly time steps.

Key-word: *model, rainfall-runoff, artificial neural network, supervised learning, fuzzy logic, cheffia.*

RÉSUMÉ

La modélisation pluie-débit au pas de temps mensuel, a été étudiée par le biais de quatre modèles qui appartiennent à deux catégories, les modèles conceptuels (modèles à réservoirs), et les modèles basés sur les réseaux de neurones, et la logique floue.

Les modèles conceptuels mensuels utilisés sont les modèles de Thornthwaite et Arnell et le modèle GR2M, ainsi que deux modèles représentés par les réseaux de neurones à apprentissage supervisé et le modèle neuro-flou qui combine une méthode d'optimisation neuronale et une logique floue.

Une application de ces modèles a été effectuée sur le bassin de la Cheffia (Nord-Est Algérien), et a confirmé les performances du modèle basé sur la logique floue. Par sa robustesse et son pouvoir d'extrapolation non-linéaire, ce modèle a donné d'excellents résultats, et représente donc une nouvelle approche de la modélisation pluie-débit au pas de temps mensuel.

Mots clés : *modèle, pluie-débit, réseaux de neurones artificiels, apprentissage supervisé, logique floue, cheffia.*

1 - INTRODUCTION

La modélisation pluie-débit a connu ces dernières années un développement considérable, surtout au pas de temps journalier, où les modèles conçus, généralement conceptuels, apportent de plus en plus d'éléments, qui font que le processus pluie-débit soit au mieux reproduit. Au pas de temps mensuel, l'approche est plus complexe. Cela est dû à la non-linéarité de la relation pluie-débit qui s'accroît, ainsi qu'à la difficulté de quantifier certains paramètres de la modélisation en particulier l'évapotranspiration réelle du sol (BENKACI, 2001). L'essai d'une régression linéaire pluie-débit s'avère inefficace, le coefficient de corrélation est souvent peu significatif ($R^2 < 0,5$), ajoutant à cela, le risque d'obtention des débits négatifs. Les modèles conceptuels représentent néanmoins un intérêt, selon ARNELL (1992), ces derniers sont étudiés de façon à explorer l'impact du changement climatique dans l'aménagement des ressources hydriques, ainsi que dans la prévision des débits à long terme.

D'après VANDEWIELLE *et* XU (1992), ces modèles servent à :

- compléter les séries des débits tronquées, lorsque celles des pluies et des évapotranspirations sont complètes et fiables, et les séries des débits jugées trop courtes et insuffisantes ;
- compléter l'information dans différents endroits du bassin, c'est-à-dire obtenir des débits dans des régions, ou sous-bassins non jaugés ;
- lorsqu'ils contiennent des données d'humidité, ces modèles constituent une information utile pour l'agriculture.

Cette modélisation de la relation pluie-débit au pas de temps mensuel, est bien évidemment influencée par différents problèmes, tels que les erreurs de mesures des données hydropluviométriques, la non homogénéité des pluies durant le mois, les erreurs du modèle. Grâce aux avancées scientifiques dans le domaine des systèmes non-linéaires, représentées par les réseaux de neurones artificiels et la logique floue, ces modèles constituent une nouvelle méthode d'approximation des systèmes complexes.

Ces modèles procèdent par un apprentissage de données et permettent d'obtenir des débits à partir d'une série d'entrées (pluie), effectuées pendant la

période de calage. L'avantage majeur de ces techniques se résume en leur pouvoir d'interpolation et d'extrapolation non-linéaire.

Les modèles conceptuels qui incorporent des lois physiques simples et considèrent le processus pluie-débit linéaire, ont souvent été une référence valide dans le domaine de la modélisation au pas de temps journalier (Modèle GR3j, EDIJATNO 1991). À l'échelle mensuelle, ces modèles voient leurs tâches limitées, face au problème de la non-linéarité de cette relation. Dans pareilles situations, l'intégration d'équations non-linéaires complexes dans de nouveaux modèles s'avère utile, et la transformation pluie-débit prend une nouvelle dimension dans la modélisation.

Dans notre étude, nous avons comparé quatre modèles appartenant à ces deux catégories à savoir les modèles conceptuels, et les modèles basés sur les systèmes neuro-flous.

2 - PRÉSENTATION DES MODÈLES

La transformation des pluies en débits à l'échelle mensuelle, nécessite des modèles fiables et simples, et cela vues les difficultés concernant à la fois leur conception et leur application.

Les modèles utilisés dans cette étude, sont considérés comme étant les plus robustes.

2.1 Le modèle de Thornthwaite

Ce modèle a été développé par THORNTHWAITE (1950), puis révisé et amélioré par MATHER (1981), ALLEY (1985) et ARNELL (1992) (MAKHLOUF, MICHEL, 1994), c'est un modèle conceptuel à deux réservoirs et trois paramètres.

On calcule la fraction de la pluie (Qp), qui contribue directement à la genèse des débits (figure 1) en fonction du paramètre X1($0 < X_1 < 1$).

Le réservoir d'humidité du sol (H) reçoit au début du mois, l'autre fraction de la pluie (Pluie nette) Pn :

$$Pn = P - Q_P \quad (1)$$

La capacité maximale du réservoir (H) est définie par un paramètre X_2. Lorsque la différence entre la pluie nette (Pn) et l'évapotranspiration potentielle mensuelle (E) est positive, cette quantité (Pn-E), va alimenter ce réservoir, et lorsqu'elle dépassera un seuil maximal, l'excès d'eau (DH) va directement dans le réservoir (S). Le niveau de ce réservoir augmente et devient (S') pour donner un débit (Q_S) :

$$Q_S = X_3.S' \quad (2)$$

X_3 : Troisième paramètre du modèle ($0 < X3 < 1$).

Dans le cas où l'évapotranspiration potentielle (E), est supérieure à la pluie nette (Pn), le niveau du réservoir H diminue :

$$H' = H \exp(-(E - Pn)/X_2) \quad (3)$$

Le débit total mensuel (Qt) exprimé en lame d'eau (mm) évacué à l'exutoire est :

$$Qt = Q_S + Q_P \quad (4)$$

Dans notre approche, la relation (3) a été modifiée de la manière suivante :

$$H' = H^{(-(E-Pn)/X2)} \quad (5)$$

Cette relation a été obtenue à partir des résultats des tests expérimentaux effectués sur trois bassins (BENKACI, 2001), l'objectif de ce changement est de minimiser les valeurs du réservoir (H), lorsque l'évapotranspiration (E) est supérieure à la pluie P. Ceci n'est pas éloigné de la réalité, car dans les mois où l'évapotranspiration est supérieure à la pluie, la réserve du sol est au minimum.

Les résultats obtenus sont meilleurs, et l'évolution des niveaux des réservoirs (H), (S) est conforme à la physique du phénomène étudié.

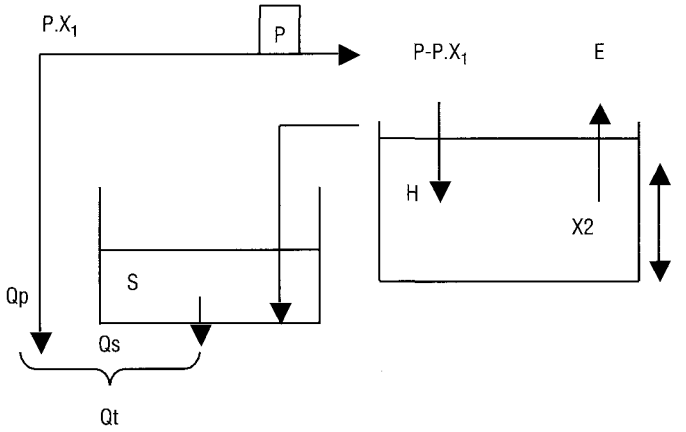

Figure 1 Architecture du modèle de Thornthwaite.
Structure of the Thornthwaite model.

2.2 Le modèle GR2M

Le modèle GR2M (MAKHLOUF 1994, MAKHLOUF et MICHEL, 1994) est un modèle à deux réservoirs (H) et (S), avec deux (2) paramètres X_1 et X_2 (figure 2).

On calcule Pn' et En' :

$$Pn' = X1 \cdot (P - [P \cdot E / (P^{1/2} + E^{1/2})^2]) \quad (6)$$

$$En' = X1 \cdot (E - [P \cdot E / (P^{1/2} + E^{1/2})^2]) \quad (7)$$

Pn' et En' serviront d'influence sur le réservoir H, et donneront respectivement H1 ou H2.

1er cas. Influence de Pn' : Pn' transforme H en H_1 :

$$H1 = (H + AV) / [1 + (HV/A)] \quad (8)$$

H_1 et H sont respectivement l'humidité du sol à la fin et au début du mois (mm) ;

A : Capacité maximale du réservoir (H) estimée à 200 mm ;

V : Tangente hyperbolique de (Pn'/A).

L'excès d'humidité sera : $Pe = Pn' + H - H_1$ (9)

Un cinquième de l'excès d'humidité (1/5 Pe) ruissellera, et participera au débit total Qt.

2e cas. Influence de En' : En' transforme H_1 en H_2 :

$$H_2 = H_1 (1 - W)/ [1 + W (1 - H_1/ A)] \quad (10)$$

W : Tangente hyperbolique de (En'/A).

Le deuxième réservoir (S) recevra la quantité (4/5 du Pe), et le niveau augmente, devient S1, libère un débit Qs :

$$Qs = X_2.S_1 \quad (11)$$

Le débit total à l'exutoire (Qt) calculé par ce modèle est :

$$Qt = Qs + 1/5\ Pe \quad (12)$$

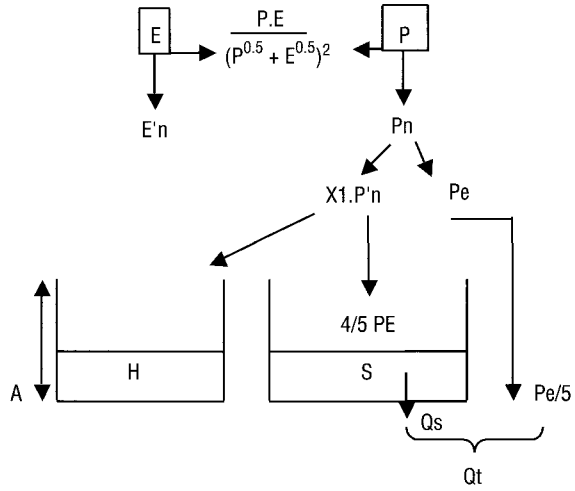

Figure 2 Architecture du modèle GR2M.
Structure of the GR2M model.

2.3 Modélisation par les réseaux de neurones artificiels

L'étude de la relation pluie-débit par les réseaux de neurones a été traitée par plusieurs hydrologues (FRENCH *et al.* 1992, SHAMSELDIN 1996, HSU *et al.* 1995, DIMOPOULOS *et al.* 1996, MINNS *et* HALL 1996, COULIBALY *et al.* 1999, SAJIKUMAR *et* THANDAVESWARA 1999).

2.3.1 Généralités sur les réseaux de neurones

Le neurone formel

Le neurone formel a été développé pour la première fois par Mc CULLOCH et PITTS (1943), ce modèle représente un produit scalaire entre un vecteur d'entrée et un vecteur poids, suivi d'une fonction seuil.

Fonctionnement mathématique général du neurone

Il est régi par trois fonctions :

a) La fonction Discriminant (fonction de base) : Elle définit l'activité du neurone, la plus utilisée est la fonction linéaire (LBF), sa forme générale est définie par :

$$\varphi(Z,W) = Z^T W \tag{13}$$

Où W représente la matrice des poids et Z les entrées qui viennent de l'extérieur ou des autres neurones vers le neurone en question.

b) La fonction dynamique linéaire : Détermine la dynamique du réseau.

c) La fonction non-dynamique non-linéaire : C'est la fonction dite d'activation, qui a pour objectif de rendre l'activité du neurone bornée. Pour cela la fonction d'activation doit être non-linéaire, continue dérivable et monotone. (YEDDOU, 1998). Parmi les fonctions les plus utilisées, on retrouve la fonction sigmoïde, ou la fonction log-sigmoïde.

Structure générale des réseaux de neurones multicouches

On distingue deux types de réseaux :

Les réseaux de neurones monocouche : tels que le Perceptron et l'Adaline, ces réseaux sont exploités essentiellement dans la classification des données.

Les réseaux de neurones multicouches : Développés par MINSKY et PAPERT (1969) (RUMELHART *et al.* 1986), ces réseaux peuvent constituer une solution aux problèmes non-linéaires, et cela par l'introduction de fonctions non-linéaires et le nombre de couches cachées.

Un réseau de neurones multicouches est formé d'un assemblage d'éléments (neurones), chacun d'eux est lié à d'autres neurones. Ceux-ci sont disposés dans une série de couches : la première couche (couche d'entrée), la (ou les) seconde(s) couche(s) intermédiaire(s) (appelées couches cachées), et la dernière appelée couche de sortie. Donc un réseau neuronal contient (n-2) couches cachées. Seules les couches d'entrées et de sorties peuvent coder les informations spécifiques au problème traité (contrairement aux couches cachées). Les neurones des couches cachées sont reliés (connectés) aussi bien par ceux de la couche d'entrée que ceux de la couche de sortie, et donc chaque neurone peut recevoir une variété d'entrées et produire une seule sortie. Il n'y a pas de connexion entre les neurones au sein d'une même couche (figure 3).

La première couche appelée couche d'entrée reçoit les informations, le rôle principal de cette couche réside dans la fonction d'identité, c'est-à-dire que le nombre d'entrées dans cette couche est égal au nombre de sorties effectuées par celle-ci. (FAUSSET, 1994, in SHAMSELDIN, 1996). La dernière couche génère les sorties (ou résultats) finales du réseau. Chaque point de connexion (appelé coefficient ou poids) entre deux neurones joue le rôle d'un synapse, l'élément principal d'interaction entre les neurones. Ces connexions ou poids synaptiques ont un rôle primordial dans le fonctionnement parallèle et adaptif des neurones. Les neurones des couches cachées transforment les entrées en sorties et cela en deux étapes :

– Pour chaque neurone, chaque entrée (dans notre cas, la pluie) est multipliée par son poids correspondant. Les ETP ne sont pas prises en considération dans ce modèle.

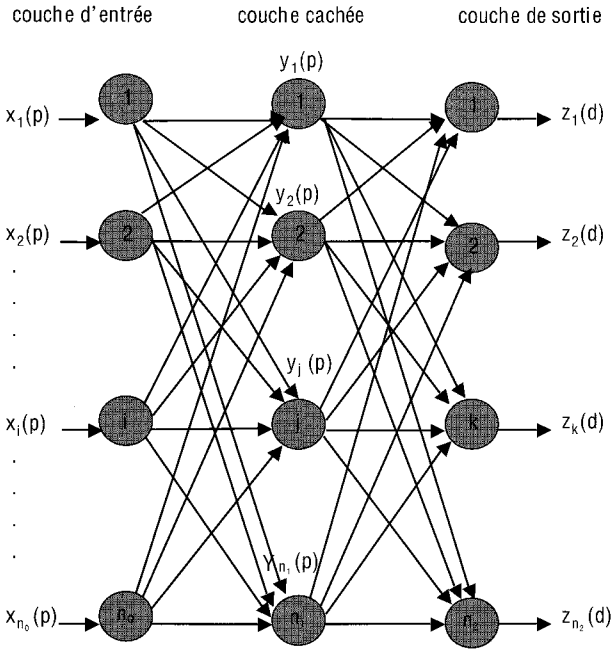

Figure 3 Structure générale d'un réseau de neurones multicouches.
Representation of a multi-layer artificial neural network.

– L'entrée du neurone est transformée en sortie (débit simulé) : chaque neurone reçoit des entrées sous forme vectorielle (X), effectue une somme pondérée (α), et génère à l'aide d'une fonction de transfert (G) un résultat réel (Y):

$$Y = G(WX + b). \tag{14}$$

$W = (w_{i1}, w_{i2} ..., w_{in})$ représente les poids du neurone (ou matrice des poids) ;

$X = (x_{i1}, x_{i2} ..., x_{in})$ sont les entrées du neurone i (ou vecteur d'entrée), b est le biais du neurone ou valeur seuil.

La fonction d'activation généralement non linéaire classique est la fonction sigmoïde inspirée par le neurone formel, définie par :

$$G(\alpha) = \frac{1}{(1 + e^{-\alpha})} \tag{15}$$

2.3.2 Apprentissage des réseaux de neurones

L'apprentissage est une procédure adaptive pour laquelle les connexions (synapses) des neurones sont ajustées face à une source d'informations.

L'apprentissage est défini comme étant n'importe quel changement opéré dans la mémoire du réseau. Ainsi cette modification affecte les poids synaptiques qui relient les neurones entre eux. (YEDDOU 1998, BENKACI 2001).

L'apprentissage a comme objectif l'amélioration des performances futures du réseau, sur la base d'une connaissance acquise au fur et à mesure des expériences passées.

Il existe plusieurs règles d'apprentissage qui peuvent être regroupées en trois catégories :

L'apprentissage supervisé (ou à partir d'exemples) : associe une réponse spécifique à chaque signal d'entrée. La procédure s'effectue progressivement, jusqu'à ce que l'erreur entre les sorties du réseau et les résultats désirés (données observées) soit minimisée.

L'apprentissage non supervisé : consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données, aucun résultat désiré n'est fourni au réseau.

L'apprentissage renforcé : représente une technique nouvelle, qui a en commun avec l'apprentissage supervisé, la présence d'un superviseur régi par une mesure de degrés de performance (un grade ou score) qui juge l'évolution de l'apprentissage du réseau après quelques itérations, et une similitude avec l'apprentissage non supervisé, puisqu'il nécessite que des variables d'entrées, sans définir les variables à la sortie du réseau.

2.3.3 Caractéristiques du modèle utilisé

Dans notre étude nous présentons une brève synthèse des caractéristiques de notre réseau, qui est un réseau neuronal à apprentissage supervisé en bloc. La fonction d'activation $f(u)$ utilisée est la fonction sigmoïde.

Algorithme de Backpropagation de Leavenberg-Marquard

C'est un algorithme d'optimisation du second ordre. Dans ce cas on utilise la dérivée seconde de la fonction objective afin d'améliorer le choix de la direction à entreprendre dans l'espace des poids, à la recherche du minimum (YEDDOU, 1998).

Pour chercher le minimum on a :

$$\Delta W = -\eta_k[\ddot{E}(k)]^{-1}\,\dot{E}(k) \tag{15}$$

$\ddot{E}$: Surface d'erreur qui est considérée quadratique.

η : taux d'apprentissage qui est positif et de petite valeur.

Soient F^L_i : les vecteurs des dérivées de l'erreur à la sortie du réseau par rapport à ces vecteurs poids w^l_i , $l = 1...L$, $i = 1...n_l$.

$$F_i^L = \sum_{p=1}^{M}\sum_{i=1}^{ni}\left(d_i^P - y_i^L\right)\left[-\frac{\partial y_i^L}{\partial W_i^l}\right] \tag{16}$$

M : représente le nombre d'exemples d'entraînement, n_l le nombre de neurones dans la couche l, y_i^L la *iéme* sortie du réseau, et d_i^p sa sortie désirée pour la p *éme* entrée.

Soit F : Vecteur rassemblant des vecteurs F_i^L.

$$\text{Pour trouver l'optimum on a : } H\Delta w = -F \tag{17}$$

H représente la matrice Hessien de la fonction erreur en sortie.

La méthode de Leavenberg-Marquard remplace le calcul du Hessien par une approximation numérique donnée par : $H = FF^T + \lambda\Omega$ (18)

H est la valeur du calcul Hessien, FF^T est une matrice positive qui sert à conditionner la matrice H. Ω représente la matrice diagonale, λ est un coefficient positif.

2.4 Modélisation par la logique floue

Depuis son introduction, la logique floue a été le sujet de plusieurs applications industrielles et scientifiques (Commande du métro au japon 1985, estimations géophysiques et biologiques BARDOSSY *et* DUCKSTEIN, 1995, estimation des imprécisions des produits phytosanitaires, FREISSINET, 1997).

2.4.1 Généralités sur les sous-ensembles flous

Un sous-ensemble flou F défini sur un univers de discours U, est caractérisé par une fonction d'appartenance $\mu_F : U \rightarrow [0,1]$. Le sous-ensemble flou F est défini par :

$$F = \{(u, \mu_F(u))/u \in U\} \quad (19)$$

La figure 4 présente un exemple de fonction d'appartenance de forme trapézoïdale définie sur un univers de discours *U*, ainsi que la fonction d'appartenance pour un ensemble classique.

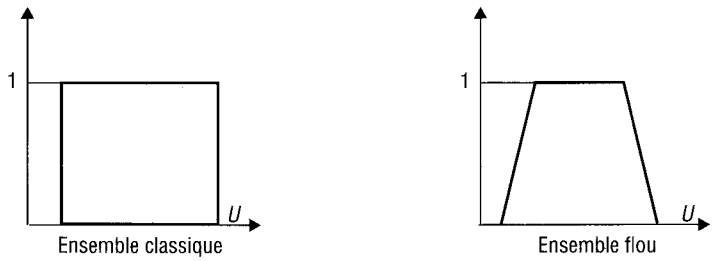

Figure 4 Fonction d'appartenance d'un ensemble précis et d'un ensemble flou.
Membership function of non-fuzzy and fuzzy logic.

2.4.2 Raisonnement et logique floue

Le raisonnement à partir de connaissances imparfaitement définies utilise la logique floue pour pallier les insuffisances de la logique classique, le raisonnement flou nécessite la définition des notions de variable linguistique, proposition floue et implication floue.

Une variable linguistique : *X* est caractérisée par un triplet (*X*,*T*(*X*),*U*), où *X* est le nom de la variable, *T*(*X*) l'ensemble des étiquettes et *U* l'univers de discours correspondant.

Exemple : X est le nom de la variable (ex : taille), T(X) l'ensemble des étiquettes correspondantes (ex : petite, grande, moyenne) et U l'univers de discours (HARTANI, 1995) (ex : U= [100, 200]).

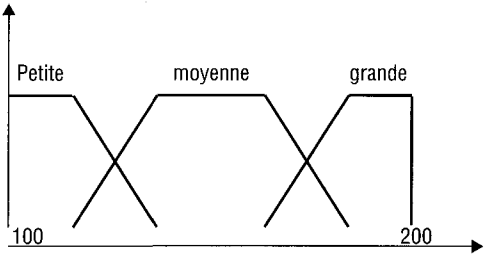

Figure 5 Étiquette pour la variable linguistique « taille ».
Classical set of « dimension » fuzzy set.

2.4.3 Modélisation par les systèmes flous

La méthodologie générale de la logique floue consiste à estimer l'imprécision d'une moyenne de résultats d'une simulation (FREISSINET, 1998).

La modélisation des systèmes flous passe par trois étapes :

– L'identification de leur structure externe : permet de déterminer les valeurs d'entrées et de sortie nécessaires et/ou suffisantes à la modélisation.

– L'identification de leur structure interne: basée sur des règles floues, cette étape permet l'identification des paramètres des sous-ensembles flous, par une méthode d'optimisation.

– L'approximation linguistique : utilise des variables linguistiques, cette opération consiste à associer à tout sous-ensemble flou une étiquette linguistique.

2.4.4 Paramétrisation de système d'inférence floue (SIF)

Un système d'inférence floue (SIF), peut être défini comme une fonction de transfert d'une partie E, d'un espace de dimension n vers une partie F d'un espace de dimension m. Il est défini comme suit :

SIF : $E \subset R^n \rightarrow F \subset R^m$

$$X = (X_1, \dots X_n) \rightarrow Y = (Y_1, \dots Y_m) \tag{20}$$

Avec $Y = SIF_P(X)$ et P = (Paramètres du SIF).

La fonction de transfert est interprétée selon une base de règles floues par une succession de trois mécanismes :

a - La fuzzification : est obtenue par l'intermédiaire d'une application associant à une observation numérique une partie floue du domaine défini par les différents symboles.

b - Processus d'inférence : permet la détermination de l'image de la composition des parties floues par l'intermédiaire de la procédure de fuzzification.

c - La déffuzzification : elle est définie comme étant une application qui associe à la partie floue obtenue par la procédure d'interface une valeur numérique correspondant à la sortie y du système flou.

2.4.5 Caractéristiques du modèle utilisé

Le modèle développé est un réseau neuro-flou, qui utilise la logique floue et une structure neuronale à apprentissage supervisé. L'identification des paramètres de la structure interne du système flou se fait par l'utilisation d'une technique numérique s'appuyant sur des algorithmes d'apprentissage hybride, de type Sugeno, qui combine la méthode des moindres carrés avec la rétropropagation du gradient (Backpropagation). L'inférence floue du modèle de Sugeno, est caractérisée par des fonctions d'appartenance à la sortie du système flou linéaires ou constantes.

La règle générale de ce modèle est :

$$\text{Si } X = A \text{ Alors } Y \text{ est } B \tag{21}$$

L'application de cette règle au modèle sera de la forme :

Si la valeur de la pluie est grande, alors le débit est élevé.

A l'entrée du système flou, les fonctions d'appartenance sont gaussiennes.

3 - LES DONNÉES UTILISÉES

L'aire d'étude est le bassin de la Cheffia (S = 575 km^2), situé dans le nord-est algérien, il est alimenté par l'oued Bouna Moussa. Les caractéristiques hydrologiques du bassin sont présentées dans le tableau 1.

Nous avons une série de données mensuelles de pluie, d'ETP et de débits observées de 1978 à 1989, soit au total un échantillon de douze ans (figure 6).

La durée de calage des modèles est de six ans (1978-1983), celle de validation va de 1984 à 1989.

Tableau 1 Caractéristiques hydrologiques du bassin.

Table 1 *Hydrological characteristics of the river basin.*

Module interannuel	804,2 mm
Q moyen	4,0 m^3/s
Q spécifique moyen	6,9 l/s/km^2
Lame écoulée	219,8 mm
C. d'écoulement	27 %

3.1 Critères de validation des modèles conceptuels mensuels

Ces critères permettent de juger de la qualité des modèles, en donnant une estimation globale et numérique de l'écart entre les résultats calculés et les données observées sur la période de calage.

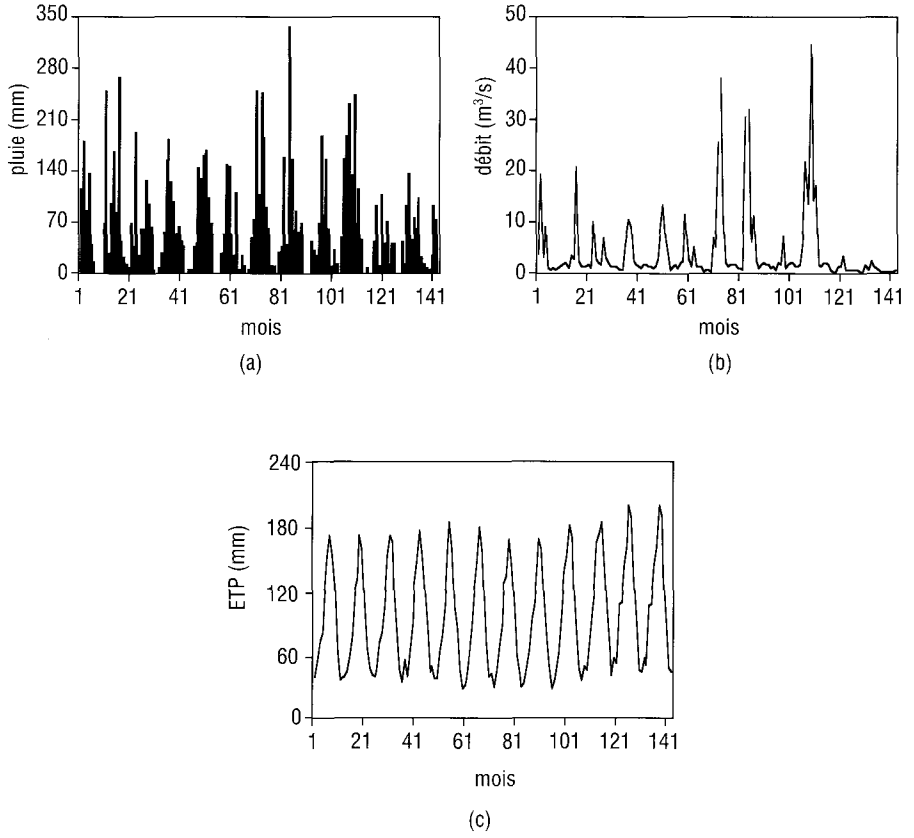

Figure 6 Évolution des variables hydrométéorologiques, (a) précipitations, (b) débits, (c) évapotranspirations.

Hydro-meteorological variables : (a) rainfall, (b) streamflow, (c) evapotranspiration.

3.1.1 Calcul de la racine carrée de l'erreur moyenne (RMSE)

Cet indice permet une évaluation combinée de la variance et du biais. Il est donné par :

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Q_i o - Q_i s)^2}{n}} \tag{22}$$

$Q_i o$ et $Q_i s$ sont les débits (m^3/s) observé et simulé pendant la période de calage;

n est le nombre de données durant la période de calage.

Le modèle est bien optimisé si la valeur du RMSE est proche de zéro.

3.1.2 Critère de Nash (T) :

Introduit par NASH et SUTCCLIFFE (1970), il est défini par :

$$T = \left(1 - \frac{\sum_{i=1}^{n} (Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n} (Q_i - \overline{Q}_i)^2} \right) \cdot 100 \qquad (23)$$

Avec :

Q_i : Débit observé durant la période de calage ;

$\hat{Q}_i$: Débit calculé par le modèle durant la période de calage ;

$\overline{Q}_i$: Moyenne des débits observés pendant la période de calage.

Le critère de Nash peut être interprété comme étant la proportion de la variance du débit observé expliqué par le modèle. Si T = 100 %, l'ajustement est parfait, par contre si T< 0, le débit calculé par le modèle est une plus mauvaise estimation que le simple débit moyen. (XANTHOULIS, 1985).

3.1.3 Moyenne biaisée (B)

Elle définit la différence entre la moyenne des débits observés et celle des débits simulés. Ce paramètre est défini par la relation suivante :

$$B = M_{obs} - M_{sim} \qquad (24)$$

M_{obs}, M_{sim} respectivement la moyenne des débits observés et celle des débits simulés durant la période de calage, lorsque {B} tend vers zéro, le modèle est sans biais.

4 – RÉSULTATS ET INTERPRÉTATIONS

Pour mieux illustrer la qualité de simulation des modèles, nous présentons les résultats observés durant les périodes de calage et de validation, et cela pour mieux analyser la robustesse des modèles ainsi que leur pouvoir prédictif.

4.1 Paramètres des modèles

Dans le tableau 2, nous présentons les valeurs des paramètres optimisés des modèles. Pour le modèle des réseaux de neurones, nous donnons le nombre de couches cachées (c), ainsi que le nombre de neurones (N) dans chaque couche. Pour le modèle de la logique floue, nous présentons seulement le nombre de règles floues (P).

D'après le tableau 3, on voit que les modèles ont généralement donné de bons résultats, les valeurs du RMSE sont faibles, et le critère de Nash oscille entre 61,3 et 95,8 %.

Tableau 2 Résultats des modèles.
Table 2 *Results of model simulations.*

Modèles	Calage			Validation		
	RMSE (m^3/s)	Nash (%)	B(m^3/s)	RMSE (m^3/s)	Nash (%)	B(m^3/s)
Thornthwaite	3,5	78,2	0,1	4,1	70,8	0,3
GR2M	3,5	78,0	– 0,5	3,7	74,3	0,7
R. neurones	2,6	61,3	0,0	7,0	35,0	1,8
L. floue	0,9	95,8	– 0,1	2,1	89,7	0,1

Tableau 3 Paramètres des modèles.
Table 3 *Models parameters.*

Modèles	Paramètres
Thornthwaite	X1= 0,16 – X2 = 200 – X3 = 0,85
GR2M	X1 = 0,8 – X2 = 0,83
R. neurones	(c1)N = 2 – (c2) N = 4 – (c3) N = 1
L. floue	nombre de règles (P) = 60

Pour les modèles conceptuels, l'optimisation des modèles a été basée sur les valeurs du RMSE et sur le critère de Nash. Ceci dit, dans certains cas, lorsque la série est plus ou moins longue, et les débits assez fluctuants, l'optimisation des modèles basée sur les valeurs du critère de Nash est généralement peu efficace. Dans ce cas on obtient seulement une moyenne de tous les débits simulés (BENKACI, 2001).

Le modèle de Thornthwaite reprend assez bien la série, les débits élevés sont simulés de façon régulière, les cycles de dessèchement et de réhumectation du sol sont correctement reproduits par le modèle (figure 7(a)).

Pour le modèle GR2M, la série modélisée coïncide avec la série observée, les débits sont généralement bien simulés, les valeurs du RMSE et du critère de Nash sont satisfaisantes, et sont respectivement 3,5 m^3/s et 78,0 %, et cela avec deux paramètres seulement.

Pour le modèle basé sur les réseaux de neurones, les résultats obtenus sont moins performants que les modèles conceptuels, le critère de Nash est moins optimisé (61,3 % dans la phase de calage), et les débits de pointe ne sont pas reproduits par le modèle. Cela s'explique par le fait que l'apprentissage effectué dans ce modèle est un apprentissage en bloc (Block Adaptive Learning) qui correspond à un apprentissage après passage de tous les exemples d'entraînement, et donc le réajustement se fait suivant la moyenne de tous les exemples.

Dans ce cas, le modèle ne prend pas en considération l'écart entre le débit observé et le débit simulé, mais seulement l'erreur entre la moyenne de tous les débits observés et celle de tous les débits simulés par le réseau de neurones. De ce fait, le volume d'apport est reproduit de façon correcte en période

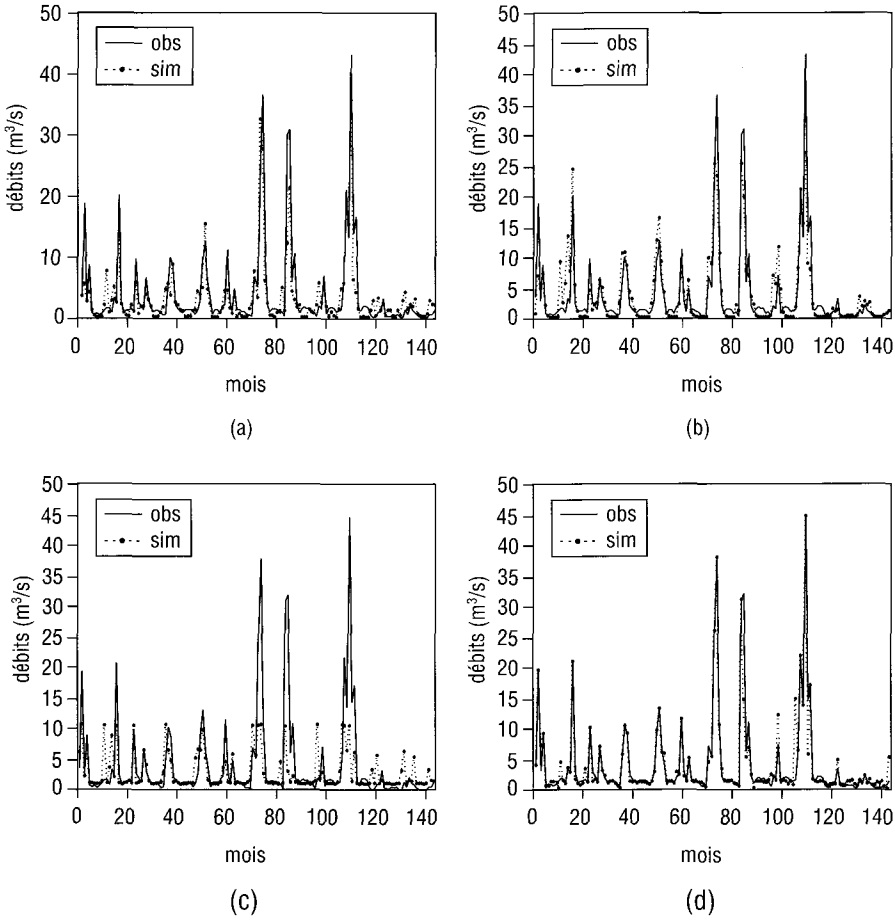

Figure 7 Résultats des modèles, modèle de Thornthwaite (a), modèle GR2M (b), Réseaux de neurones (c), Logique floue (d).

Models results : (a) Thornthwaite model, (b) GR2M model, (c) neural network and (d) Fuzzy logic.

de calage, dans la phase de validation, la moyenne des débits simulés est largement inférieure (B = 1,8 (m^3/s)). On peut conclure, que ce mode d'apprentissage est incompatible avec la modélisation des débits en temps réel.

Les résultats obtenus par la logique floue sont nettement meilleurs, le critère de Nash est de 95,8 %, la valeur du RMSE est faible (0,9 m^3/s), les débits de pointe sont régulièrement reproduits, les débits sont simulés de façon correcte, étant donné les bons résultats obtenus par le modèle.

La logique floue semble être la meilleure méthode de modélisation et d'approximation de systèmes non-linéaires, puisque les résultats obtenus sont les meilleurs.

Pour tous les modèles, on observe une détérioration de la modélisation entre la période de calage (optimisation) et celle de validation, ceci est illustré

par les valeurs du RMSE qui augmentent lorsqu'on passe à la phase de validation. Le critère de Nash se détériore et diminue pour le modèle des réseaux de neurones.

Si on fait une comparaison entre les modèles, on remarque que la logique floue a donné de meilleurs résultats pour les deux périodes. On note une valeur du critère de Nash assez élevée ; 95,8 % pour la phase de calage, et 89,7 % pour la phase de validation. Les débits maxima sont reproduits de façon régulière et dans les deux périodes.

Sur le plan de la paramétrisation, les modèles conceptuels (Thornthwaite et GR2M), par leur conception simple, intègrent des équations avec un nombre de paramètres réduit (2 ou 3 paramètres seulement), ce qui les rend plus parcimonieux. Par contre, la logique floue représente un système d'équations non-linéaires complexes, et l'optimisation du modèle nécessite un nombre de règles floues assez important.

Les modèles conceptuels, contrairement aux modèles neuro-flous, présentent des paramètres reliés à une signification physique du bassin, et peuvent servir dans l'étude de la gestion des ressources en eau. Ces modèles ont donné de bons résultats, mais accusent un déclin de reproduction de débits maxima lors de la phase de validation, ce qui nous conduit à ne pas se fier de leur pouvoir prédictif.

5 – CONCLUSION

La logique floue est une technique d'approximation relativement récente qui permet d'évaluer les imprécisions des résultats d'une simulation.

Le modèle flou utilisé dans cette étude, est une combinaison d'une méthode de simulation pluie-débit neuronale à apprentissage supervisé et une logique floue, ce qui permet de décomposer le processus complexe, en processus physiques plus simples. Son application dans le domaine de la modélisation pluie-débit a donné de meilleurs résultats que les approches conceptuelles et les réseaux de neurones, et s'avère plus fiable que les modèles conceptuels déterministes les plus robustes.

Pour le modèle neuronal, l'essai de la méthode d'apprentissage en bloc des données, n'a pas atteint les objectifs attendus, les débits sont mal simulés et les résultats confirment son incompatibilité dans la modélisation des débits en temps réel.

La reproduction de la relation complexe pluie-débit par le modèle flou, semble être la mieux adaptée. Cette méthode est dotée d'une rapidité de diagnostic et d'un pouvoir de généralisation assez importants, et peut modéliser une série de données mensuelles même jugée trop courte.

La logique floue représente donc une technique puissante d'approximation et d'optimisation, et constitue un moyen efficace dans la modélisation pluie-débit.

RÉFÉRENCES BIBLIOGRAPHIQUES

ALLEY M.W., 1985. Water balance models in one-month-ahead stream flow forecasting. *Water ressources research*, 21(4), 597- 606.

ARNELL N., 1992. Factors controlling the effects of climate change on river Flow regimes in humid temperatures environments. *J. of hydrology*, 132, 321-342.

BENKACI T., 2001. Contribution à la modélisation de la relation pluie-débit au pas de temps mensuel et journalier par les modèles conceptuels et les réseaux de neurones. Th. Magister. Institut national agronomique d'Alger.

BARDOSSY A., DUCKSTEIN L., 1995. Fuzzy rule based modelling with applications to geophysical , biological and engineering sytems. *CRC press*, p 232.

COULIBALY P., ANCTIL F., BOBEE B., 1999. Prévision hydrologique par réseaux de neurones artificiels Etat de l'art. *J. of civil engineering*, 26., 293-301.

DIMOPOULOS I., LEK S., LAUGA J., 1996. Modélisation de la relation pluie-débit par les réseaux connexionnistes et le filtre de Kalman. *J des sciences hydrologiques*, 41,177-193.

EDIJATNO. (1991). Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier. Th. Doct. univ. Louis Pasteur. Cemagref (Anthony).

FREISSINET C., 1997. Estimation des imprécisions dans la modélisation du devenir des produits phytosanitaires dans les sols ; une méthode basée sur la logique floue. Th. Doct. Univ. Grenoble, France.

FREISSINET C., 1998. Estimation des imprécisions dans la modélisation du devenir des produits phytosanitaires dans les sols ; une méthode basée sur la logique floue. *La houille blanche*, 8, 14-16.

FRENCH M.N., KRAJEWSKI W.F., CUYKENDALL R.R., 1992. Rainfall forecasting in space and time using a neural network. *J. of hydrology*,137, 1-31.

HARTANI R., 1995. Modélisation de systèmes neuroflous, contributions théoriques et applications. Th. Doct. Univ. Paris VI.

HSU K., GUPTA H., SOROOSHIAN S., 1995. Artificial neural network modelling of the rainfall - runoff process. *Water ressources research*, 31, 2517- 2530.

JANG, J.-S. R., 1993. ANFIS: Adaptive-Network-based Fuzzy Inference Systems. *IEEE, Man, and Cybernetics*, 23, 665-685.

MAKHLOUF Z., 1994. Compléments sur le modèle pluie-débit GR4J et essai d'estimation de ses paramètres. Th. Doct. Univ. Paris-sud Orsay, CEMAGREF Anthony.

MAKHLOUF Z., MICHEL C., 1994. A two parameters monthly model for french watersheds. *J. of hydrology*, 162, 299-318.

MATHER J.R., 1981. Using computed stream flow in watershed analysis. *Water ressources bulletin*, *17(3)*, 474- 482.

MINNS A., HALL J., 1996. Artificial neural network as rainfall-runoff models. *J. des sciences hydrologiques*, 41, 399-417.

NASH J.E., SUTCLIFFE J.V., 1970. River flow forecasting through conceptual models. Part 1 - A discussion of principles. *J of hydrology*, 10, 282-290.

RUMELHART D.E, HINTON G.E., WILLIAMS J., 1986. Learninig internal representation by error propagation, in Parallel Distributed Processing, Vol. 1, 318-362, MIT Press, Cambridge.

SAJIKUMAR N., THANDAVESWARA S., 1999. A non-linear rainfall-runoff model using an artificial neural network. *J of hydrology*, 216, 32-55.

SHAMSELDIN Y., 1996. Application of neural network technique to rainfall-runoff Modeling. *J. of hydrology*, 199, 272-294.

SUGENO M., 1985. Industrial applications of fuzzy control. *Elseiver sciences publications*.

VANDEWIELE G., XU, C., 1992. Methodology and comparative study monthly water balance model in Belgium. *J. of hydrology*, 134, 315-347.

XANTHOULIS D., 1985. Modélisation mathématique des débits d'étiage des cours d'eau. Th. Doct. Faculté des sciences agronomiques de Gembloux.

YEDDOU L., 1998. Étude de synthèse sur les réseaux de neurones et leur application. Th. Magister. École nationale polytechnique d'Alger.