**Modélisation de la relation pluie-débit dans les bassins sahariens à l’aide des systèmes hybrides intelligents**

Mohamed Chettih1, Khaled Chorfi1 & Kaddour Mouattah1

*1 Laboratoire de Recherche en Ressources en Eau, Sols et Environnement, Département de Génie Civil, Faculté de Technologie, Université Amar Telidji de Laghouat, B.P. 37 G ,03000 – Laghouat – Algérie.*

Email : m.chettih@lagh-univ.dz

**Résumé.**  Au cours de cette dernière décennie, l’application des systèmes hybrides intelligents dans différents domaines a montré de bonnes performances et une efficacité inégalée. Cependant, le caractère non-linéaire, le comportement non-stationnaire et la nature multi-échelle des régimes hydrologiques et plus particulièrement le régime hydrologique saharien, requiert le recours à des outils spécifiques des systèmes dynamiques non-linéaires. A ce titre, nous proposons dans cette note l’application d’outils d’analyse issus de la dynamique non-linéaire des systèmes complexes à la relation pluie-débit, et des modèles hybrides intelligents basé sur la reconstruction de l’espace des phases, et la transformée en ondelettes pour la prédiction des débits. Dans ces modèles, les séries générées à l’aide de la technique de reconstruction de l’espace de phase de Takens et à l’aide de la décomposition du signal dyadique en une succession d’approximations et de détails, constitueront la base de données au réseau de neurones et permettront ainsi de prendre en compte la dynamique du signal hydrologique et d’apprendre son évolution. Les résultats obtenus dans cette étude, ont montré la bonne performance des techniques utilisées. Les modèles proposés, possèdent un bon pouvoir prévisionnel si toutefois la chaoticité et la fractalité du processus hydrologique sont mises en évidence. Le recours à cette technique constitue une alternative pleinement justifiée, même si le signal hydrologique est souvent considéré comme un processus chaotique de faible dimension. Ces résultats encourageants ouvrent un certain nombre de perspectives, où il serait intéressant de tenter des modèles hybrides intelligents optimisant simultanément par algorithmes génétiques les paramètres des réseaux de neurones.

**Mots clefs**  Bassins Sahariens, Intelligence Arificielle, Système Hybride, Modèle Neuro-Chaotique.

1. INTRODUCTION

L’un des principaux objectifs de la modélisation d’une série temporelle est la prédiction de l’évolution future de la série à partir de celles qui ont été observées [1], [2]. La plupart des modèles proposés supposent que les séries étudiées sont stationnaires. Au cours de ces des dernières décennies, un grand nombre d’approches automatisées ou informatisées ont été mises en œuvre pour modéliser ces processus[3]. Le caractère non-linéaire, le comportement non-stationnaire et la nature multi-échelle des séries hydrologiques sont connus depuis longtemps [4], [5], [6]. L’avènement de l’informatique et l’évolution technologique de la puissance de calcul ont facilité la formulation d'approches non linéaires et permis d’avoir d’autres alternatives plus efficaces [7]. Parmi ces méthodes, les réseaux neuronaux, qui présentent une technique de calcul pour la prévision hydrologique largement décrite par Robert J. Abrahart et ses collaborateurs [8]. Les réseaux de neurones fournissent une voie de modélisation qui peut être utile quand il ya suffisamment de données pour relier par exemple les pluies aux débits et surtout où les résultats sont nécessaires en temps réel [9]. Malgré leur efficacité, les réseaux de neurones sont souvent considérés comme des solutions de boîte noire. Mais en revanche, ils peuvent fournir des résultats très utiles. Ils demeurent toutefois un sujet d’un grand intérêt pour les chercheurs qui désirent améliorer les performances de ces réseaux et étendre leur champ d’applications. Le *Soft Computing* , constitue une nouvelle technique pour la construction de systèmes hybrides intelligents capables de raisonnement et d'apprentissage dans un environnement incertain et imprécis. Cependant, la complexité des régimes hydrologiques requiert à recourir à des outils spécifiques des systèmes dynamiques non-linéaires [10],[11]. A ce titre, les applications de la dynamique non linéaire aux systèmes et aux processus hydrologiques ont reçu une attention considérable et tout à fait particulière par les chercheurs au cours de ces dernières années [12] [13]. La réalisation de résultats encourageants a permis de reconnaitre que les systèmes hybrides intelligents pourraient fournir de nouvelles perspectives et des voies alternatives vers la modélisation et la prévision des processus hydrologiques [9], [10], [14]. Nous proposons dans le cadre de l’étude des bassins sahariens un modèle hybride intelligent Neuro-chaotique pour la prédiction des débits journaliers. Les résultats obtenus seront comparés aux résultats obtenus par d’autres modèles.

**2. RECONSTRUCTION DE L’ESPACE DES PHASES**

L'espace de phase est un outil utile pour représenter l'évolution d'un système dans le temps. Il s'agit essentiellement d'un graphe ou d'un schéma dont les coordonnées représentent les variables nécessaires pour décrire complètement l'état du système à tout moment [18]. La dimension de l'espace de phase est le nombre de degrés de liberté d'un système dynamique. Dans le contexte des systèmes Hamiltoniens, c'est le nombre de paires de variables d'état. Lors de la recherche de la structure d'un système dynamique, il faut reconstruire ou plonger la série temporelle dans un espace de phase de dimension supérieure. La technique de reconstruction de l'espace de phase la plus importante est la méthode de retards (délai). La méthode est connue comme le théorème du plongement de Takens [19].La prévision à l’aide des réseaux de neurones a été déjà tentée pour des séries temporelles purement chaotiques en montrant une bonne performance [16]. Mais pour des séries hydrologiques où il pourrait s’agir d’un processus de chaos déterministe de faible dimension [17], la reconstruction de l’espace des phases permettra dans ce cas de constituer la base de données au réseau de neurones et de prendre ainsi en compte la dynamique du signal hydrologique et d’apprendre son évolution.

## **2.1 Théorème de Takens**

Supposons que l’on observe une série temporelle : , générée par un processus inconnu. Afin de comprendre la Dynamique du système sous-jacent, on se limite à l’étude de la dynamique sur l’attracteur et à partir de la série initiale on peut également générer différents signaux. A ce titre, on va chercher à reconstruire l’attracteur du système à partir de la série temporelle observée. Pour cela, on forme vecteurs à m Dimensions, appelés m-historiques, dont les composantes sont les valeurs consécutives de la série observée décalées d’un retard ou délai τ fixé :

 (1)

avec , où m est la Dimension de plongement, c’est la Dimension de l’espace des phases dans lequel l’attracteur est reconstruit. Le théorème de Takens établit qu'à condition de choisir *m* assez grand, le comportement des m-historiques imitera celui du système dynamique sous-jacent inconnu. En particulier, si la dynamique du système est chaotique, les m-historiques auront aussi un comportement chaotique. En pratique, si , où *n* est la dimension inconnue du vecteur d'état, l’attracteur ainsi reconstruit aura les mêmes propriétés topologiques que le système initial.

Si l’espace des phases est représenté en trois dimensions, cette suite de points peut montrer graphiquement l’évolution du système dans le temps. L’ensemble des trajectoires possibles constitue le portrait de phases. Celui-ci peut aider à percevoir l’attracteur du système.

La figure 1 montre le portrait de phase en trois dimensions de l'attracteur reconstruit pour la série des débits moyens journaliers. La structure assez bien définie de l’attracteur indique à priori la possibilité d’une dynamique déterministe.



**Fig. 1** : Portrait de phase à trios dimensions des débits journaliers pour τ = 3.

On voit bien toute la puissance de ce théorème, et son utilité pour l’analyse des séries temporelles. Néanmoins, un certain nombre de problèmes restent en suspens. Ainsi, le choix du décalage temporel *τ*  et de la dimension de plongement optimale *m* peut s'avérer délicat.

## **2.2 Données et Site d’étude**

Le Bassin de l'Oued Seklafa draine essentiellement les grès à dragées du Crétacé inférieur du les formations du Crétacé inférieur du synclinal d'Ed Dor et Kef El Gada ainsi que le massif de Seklafa. L'exutoire est considéré au niveau de la station hydrométrique de Seklafa codée C0104 par l'Agence Nationale des Ressources Hydrauliques. Le bassin de l’Oued Seklafa appartient au grand bassin méridional de l’Oued M’Zi, qui fait partie des bassins endoréiques sahariens qui se déverse dans Chott El Melrhir.



**Fig. 2** : Série chronologique des débits de l’Oued Seklafa

## **2.3 Choix du décalage temporel**

Pour le choix du décalage temporel τ, différentes méthodes intuitives ont été proposées. Le premier zéro de la fonction d'autocorrélation, par exemple, garantit l’indépendance entre et . Pour les débits journaliers de l’Oued Seklafa, on constate sur la figure 3, que le premier zéro est obtenu pour un décalage temporel τ = 9 jours.



**Fig. 3**: Fonction Autocorrélation des débits journaliers. Le premier zéro de la fonction autocorrélation est obtenu pour τ = 9 jours.

Une indépendance plus générale sera néanmoins garantie par le premier minimum de la fonction d'information mutuelle qui représente le degré de dépendance de deux variables au sens probabiliste. Elle est défini par la somme qui lie densité de probabilité aux distributions marginales :

où et sont respectivement les densités des lois de et

La figure 4 montre que la fonction d’information mutuelle est minimale pour τ = 3 jours.

**Fig. 4** : Fonction d’information mutuelle des débits journaliers. La fonction d’information mutuelle est minimale pour τ = 3 jours.

Pour la dimension de plongement *m* optimale, l’Algorithme des plus proches faux voisins s’avère efficace [20]. L’Algorithme se base sur l’idée que lorsqu’un objet en dimension *m* est projeté en dimension *m − 1*, certains points initialement éloignés deviennent voisins. On parle alors de *faux voisins*. Inversement, en partant de la dimension 1, le nombre de faux voisins à chaque passage à la dimension supérieure devrait diminuer, jusqu’à devenir nul lorsque la topologie de l’attracteur est entièrement dévoilée.

La figure 5 permet de trouver la dimension de plongement. Les bonnes valeurs sont celles qui ont un pourcentage de faux voisins proche de zéro. Elle révèle que le pourcentage des faux voisins est moins de 20% dans un espace de dimension égale à 2. La Dimension de plongement est optimale à partir de 6.



**Fig. 5** : Distribution des plus proches faux voisins de la série des débits journaliers.

**3. STRUCTURE DU MODELE**

L’approche proposée est basée sur l’utilisation des modèles de réseaux de neurones artificiels couplés aux informations issues de la reconstruction de l’espace de phase.

Ainsi, la série générée à l’aide de la technique de reconstruction de l’espace de phase de Takens constituera les vecteurs d’entrée dans un modèle de réseau de neurones et forme ainsi un modèle hybride neuro-chaotique (CNN).

La structure de réseau utilisée dans cette étude est le Perceptron Multicouche. Le réseau comporte une couche pour capter les entrées, une couche cachée et une couche pour émettre les sorties. Chaque couche contient des processeurs élémentaires connectés à d’autres neurones par la voie des poids (*Wij* et *Wjk*) représentatifs de la force de connexion (Fig. 6).

En hydrologie, souvent, pour l’apprentissage, l’Algorithme de rétro-propagation reste le plus utilisé. A ce titre, trois méthodes ont été testés :

- La Méthode de Levenberg-Marquardt (LM) ;

- La Méthode Resilient Backpropagation (RBP) ;

- la Méthode Scaled Conjugate Gradient (SCG).

Néanmoins, le choix de l’algorithme de la rétro-propagation pour l’apprentissage du réseau impose des fonctions dérivables. A ce titre, la fonction sigmoïde a été choisie pour fonction de transfert pour les neurones des couches cachées.



**Fig. 6** : Structure du modèle Neuro-Chaotique

**4. INDICATEUR DE PERMORMANCE**

La performance du modèle est validée par des paramètres statistiques des phases d’apprentissage, de validation et de test. Les paramètres statistiques utilisés dans cette étude sont : L’erreur moyenne des carrés ***ASE***(Average Squared Error), le coefficient d’efficacité de Nash–Sutcliffe **CE**et le coefficient de corrélation***r***. Ces paramètres sont donnés par les relations suivantes :

où est la valeur mesurée du débit, est le débit calculé par le modèle, est le débit moyen mesuré est le débit moyen simulé et *N* est le nombre de données.

## **5. PREVISION HYDROLOGIQUE**

Sur la base des résultats obtenus par l’analyse chaotique de la série de débits, nous avons pris la dimension de plongement *m = 6,* et le retard *τ = 3*.

La base de données contient les valeurs de pluies et les valeurs de débits générées à l’aide de la méthode de reconstruction de l’espace de phase (Fig. 6). La base de données a été subdivisée en deux ensembles :

* Un ensemble pour le calage du modèle, c’est la phase d’apprentissage ;
* Un autre ensemble pour la validation du modèle sur un jeu de données n’ayant pas participé au calage du modèle.

En effet, la phase de test a été effectuée sur l’ensemble de données.

En fonction des algorithmes d’apprentissage choisis, les résultats obtenus pour la phase d’apprentissage sont résumés dans le tableau ci-dessous :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Algorithmes d’apprentissage  | Nombre de neurones | ASE  | CE | *r* |
| LM | 8 | 3.1747 | 0.9042 | **0.9509** |
| RBP | 5 | 5.6442 | 0.7635 | **0.7664** |
| SCG | 3 | 8.1348 | 0.6898 | **0.6899** |

Tableau 1. Paramètres du modele pour differents algorithmes d’apprentissage.

Le modèle le plus performant est obtenu à l’aide de l’Algorithme de Levenberg-Marquard pour un nombre de neurones de 8 dans la couche cachée. Le tableau 2 résume les principaux résultats pour les différentes phases.

|  |  |  |  |
| --- | --- | --- | --- |
| Phases | ASE  | CE | *r* |
| Apprentissage | 3.1747 | 0.9042 | ***0.9509*** |
| Validation | 8.6936 | 0.4825 | ***0.7309*** |
| Test |  5.5238 | 0.7893 | ***0.8903*** |

Tableau 2. Paramètres statistiques du modele CNN

La figure 7 illustre les résultats entre débits observés et simulés. Les résultats obtenus montrent une bonne concordance expliquée par des coefficients assez élevés et une erreur moyenne des carrés acceptable.



Fig. 7 : Comparaison entre débits observés et débits simulés pour les phases d’apprentissage, de validation et de Test.

Pour évaluer la performance du modèle Neuro-chaotique (CNN), une comparaison a été réalisée avec le modèle classique de Régression Linéaire Multiple (RLM), le modèle de Réseau de Neurones Artificiels (ANN) et le modèle d’inférence neuro-floue (ANFIS). L’application de ces modèles a été faite sur les mêmes jeux de données. Le tableau 3 montre les résultats obtenus par ces différents modèles pour la phase de test.

Ces résultats montrent la bonne performance du modèle Neuro-Chaotique (CNN) qui dépasse celle des autres modèles. Cette performance peut s’expliquer par le fait que si le système est chaotique, son espace de phase constitue un outil puissant pour représenter son évolution dans le temps, et par conséquent une meilleure prévision.

|  |  |  |  |
| --- | --- | --- | --- |
| Models  | ASE  | CE | *r* |
| MLR | 6.9800 | 0.7300  | ***0.8600***  |
| ANN | 6.8867 | 0.7364 | ***0.8583*** |
| ANFIS |  6.3381 | 0.7577  | ***0.8956***  |
| CNN |  5.5238 | 0.7893 | ***0.8903*** |

Tableau 3. Résultats obtenus par les modèles MLR, ANN, ANFIS, et CNN.

**6. CONCLUSION**

Les résultats obtenus dans cette étude, ont montré l’efficacité du modèle hybride Neuro-Chaotique dans la modélisation de la relation pluie-débit pour la prévision des débits. Le modèle Neuro-Chaotique possède un bon pouvoir prévisionnel si toutefois sa chaocité est mise en évidence. Le recours à cette méthode hybride constitue une alternative pleinement justifiée, même si le signal hydrologique est considéré souvent comme un processus chaotique de faible dimension. Ces résultats encourageants ouvrent un certain nombre de perspectives, où il serait intéressant de tenter des modèles hybrides intelligents en couplant les transformées en ondelettes aux systèmes Neuro-Chaotiques prenant en compte la longue dépendance et l’effet multi-échelles et en optimisant simultanément par algorithme génétique les paramètres des réseaux de neurones artificiels.

Dans cet article, un algorithme génétique continu a été appliqué à l’interprétation automatique d’un essai de pompage et à l’identification des conductivités hydrauliques de deux problèmes d’écoulements permanents en milieux anisotropes et hétérogènes saturés. Les résultats de l’interprétation de l’essai de pompage ont montrés une bonne concordance entre valeurs mesurées et calculées. Les modèles directs simulant des écoulements dans des milieux poreux ont été inversés à l'aide d'un couplage (algorithme génétique continu/ modèle aux éléments finis) afin d'identifier les paramètres physiques relatifs a ces deux modèles. Les paramètres physiques ont été identifiés avec succès et cela grâce à la capacité de l'algorithme génétique continu à localiser l'optimum global. La méthodologie métaheuristique utilisée s’avère un bon outil pour le calage automatique des modèles numériques en hydrogéologie. Ce travail nécessite, bien entendu, l’extension à des cas réels où les mesures sont très limitées et sujettes à des imperfections.

**REFERENCES**

1. H. I. D. Abarbanel, R. Brown, and J.B. Hadtke, “Prediction in chaotic nonlinear systems. Methods for time series with broadband Fourier spectra,” *Physics Rev.* A, 41(4), 1990, pp. 1782-1807.
2. H. I. D. Abarbanel, “Analysis of observed chaotic data,” Eds. Springer-Verlag, New York Inc. 1996, p. 272.
3. H. Kentz and T. Schreiber, “Nonlinear time series analysis,” Eds. Cambridge University Press. 2003, p. 369.
4. J. Amorocho, “Mesures of the linearity of hydrologic systems”, *Journal of Geophysical Research*, 68(8), 1963, pp. 2237-2249.
5. D. Labat, R. Ababou and A. Mangin, “Rainfall-runoff relation for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses”. *Jour. Hydrol*. 238, 2000, pp 149-178.
6. D. Schertzer and S. Lovejoy, “Multifractal simulations and analysis of clouds by multiplicative processes”. *Atmospheric research*, 21, 1988, pp. 337–361.
7. B. Sivakumar and R. Berndtsson, ERNDTSSON R. “Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting”. Eds. World Scientific Publishing, N.J. USA, 2010, p. 519.
8. R. J. Abrahart, P. E. Kneale, L. M. See, Neural networks for hydrological modelling. Taylor & Francis, London, UK, 2004, p. 304.
9. P. C. Nayak, K .P. Sudheer and S. K. Jain, “Rainfall-Runoff modeling through hybrid intelligent system,” *Water Resources Research,* vol. 43 w07415, 2007, pp. 1-17.
10. G. Tayfur, “Soft computing in water resources engineering : Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms,” Eds. WIT, 2012, p. 288.
11. B. Sivakumar, R. Berndtsson, J. Olsoon, K. Jinno, “Evidence of chaos in the rainfall-runoff process,” *Hydrological Sciences Journal*, 46(1), 2001, p. 131-145.
12. B. Sivakumar, “Chaos theory in hydrology: important issues and interpretations. *Journal of Hydrology*, 227, 2000, p. 1-20.
13. M. Siek, “Predecting Storm Surges. Chaos, computational intelligence, Data assimilation, ensembles”. CRC Press. Eds. Taylor & Francis, 2011, p. 213.
14. S. Velickov, “Nonlinear Dynamics and Chaos with Applications to Hydrodynamics and Hydrological Modelling,” Eds. Taylor & Francis, London, UK, 2006, p. 396.
15. A. V. Gromov and A. N. Shulga, “Chaotic time series prediction with employment of ant colony optimization,” *Expert Systems with Applications*, 39, 2012, pp. 8474-8478.
16. Y. Lisheng, H. Yigang, D. Xueping and L. Zhaoquan, “Adaptive Chaotic Prediction Algorithm of RBF Neural Network Filtering Model based on Phase Space Reconstruction,” *Journal of computers,* vol. 8 (6), 2013, pp. 1449-1455.
17. M. Chettih, W. Tadj and M.B. Taouti, “ Contribution of nonlinear dynamics and related techniques for the characterization of hydrological systems in semi-arid climate,”  3rd International Conference Water–Climate’2014: SYNERGIES NORTH – SOUTH, [October 21, 22 & 23, Hammamet, Tunisia, 2014].
18. N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, “Geometry from a Time Series,” *Physical Review Letters*, Vol. 45, 1980, pp. 712-716.
19. F. Takens, “Detecting strange attractors in turbulence,” *Springer Lecture Notes in Mathematics,* vol. 898, 1981, pp. 366–381.
20. R. Brown , P. Bryant, et H.D.I. Abarbanel. Computing the lyapunov spectrum of a dynamical system from an observed time series. *Phys. Rev. A*, 43(6), 1991, pp. 2787-2806.
21. J.P. Eckmann, S. O. Kamphorst and D. Ruelle, “Recurrence plots of dynamical systems,” *EPL* (Europhysics Letters) *4* (9), 1987, pp. 973-977.
22. C. L. Webber and N. Marwan, “Recurrence Quantification Analysis: Theory and Best Practices (Understanding Complex Systems),” Ed. Springer, 2014, p. 421.
23. A. Metin, “Nonlinear Biomedical Signal Processing: Dynamics Analysis and Modeling. Vol. II, IEEE Press. Inc. New York, 2001, p. 341.
24. A.Wolf, J. B. Swift, H. L. Swinney and A. Vastano, “Determining Lyapunov exponents from a time series,” Physica D 16, 1985, pp. 285 -317.
25. M. Rosenstein, J. Collins, C. J. De Luca, “A practical method for calculating largest Lyapunov exponents from small data sets,” Physica, D 65, 1992, pp. 117-134.
26. P. Grassberger and I. Procaccia, “Measuring the strangeness of strange attractors,” Physica D 9, 1983, pp. 189-208.