



MODELISATION DE LA CONSOMMATION EN EAU POTABLE DANS LES CAPITALES AFRICAINES AU SUD DU SAHARA: APPLICATION DES RESEAUX DE NEURONES FORMELS A LA VILLE DE YAMOOUSSOUKRO, CAPITALE POLITIQUE DE LA COTE D'IVOIRE

Yao Blaise KOFFI¹
Kouassi Ernest AHOUSI²
Amani Michel KOUASSI³
Kpangui Loukou Christophe⁴
Jean BIEMI⁵

ABSTRACT

Face à la croissance démographique que va engendrer le transfert effectif de la capitale politique et administrative de la Côte d'Ivoire à Yamoussoukro, il faut craindre une congestion, voire une inadéquation des équipements dans le secteur de l'approvisionnement en eau potable de cette localité. Pallier à ce problème passe par des prévisions de la consommation de cette ressource en fonction de la taille de la population grandissante. Le recours à des modèles capables de traiter la non-linéarité de cette relation s'avère donc nécessaire. C'est dans cette optique que cette étude présentant une modélisation de la consommation au pas de temps mensuel basée sur l'utilisation des réseaux de neurones est réalisée. Des modèles neuronaux types Perceptron Multicouches (PMC) bouclés dirigés sur la consommation observée au temps mensuel, à deux (2) couches ont été développés. Une Analyse en Composantes Principales Normée (ACPN) a été réalisée sur les différentes données démographiques (population), hydroclimatiques (pluie et température) et de consommation d'eau potable de la ville pour faire ressortir les corrélations afin de sélectionner les variables explicatives les plus pertinentes. Ces modèles neuronaux ont été calés avec l'algorithme

¹ Université de Cocody ; UFR des Sciences de la Terre et des Ressources Minières ; 22 BP 582 Abidjan 22, KOFFI Yao Blaise, 22 BP 801 Abidjan 22, **E-mail:** yaomonie@yahoo.fr

² Université de Cocody ; UFR des Sciences de la Terre et des Ressources Minières ; 22 BP 582 Abidjan 22

³ Institut National Polytechnique Félix Houphouët-Boigny (INP-HB); Département des Sciences de la Terre et des Ressources Minières (STeRMi) ; BP 1093 Yamoussoukro

⁴ Université de Cocody ; UFR des Sciences de la Terre et des Ressources Minières ; 22 BP 582 Abidjan 22

⁵ Université de Cocody ; UFR des Sciences de la Terre et des Ressources Minières ; 22 BP 582 Abidjan 22

de la rétro-propagation avec le principe de la validation croisée pour améliorer leurs performances de généralisation. La fonction d'activation choisie est la variante logistique du type sigmoïde. L'application des réseaux de neurone pour l'estimation de la consommation en eau potable de la ville de Yamoussoukro a permis d'obtenir d'assez bons résultats. En effet, les coefficients de corrélation entre les consommations prédites et celles mesurées varient de 0,63 à 0,97. On note donc que les performances obtenues sont relativement bonnes et pourraient être améliorées en utilisant une base de données plus importante. Le choix de variables explicatives de bonne qualité pourrait en outre permettre d'obtenir de meilleurs résultats.

Key Words: Perceptron multicouche, Optimisation, Préviation, Apprentissage, Consommation, Eau potable.

INTRODUCTION

Sur l'ensemble du District d'Abidjan, bien que les ressources de la nappe d'Abidjan soient assez suffisantes pour satisfaire les besoins en eau potable des populations, il y a cependant un déficit journalier d'environ 150 000 m³ (EGEP, 2009). Cette situation est causée par la vétusté des équipements d'approvisionnement en eau potable et de l'inadéquation entre la démographie galopante et le manque de financement de ces infrastructures. Comme conséquence, on assiste à de fréquentes perturbations au niveau de la distribution d'eau dans certains quartiers populaires (Yopougon et Abobo) de la mégapole d'Abidjan. Mais, cette situation n'est pas le propre de la seule ville d'Abidjan, on observe également des perturbations dans la distribution de l'eau potable à l'intérieur du pays notamment dans certaines villes proches de Yamoussoukro (de Hiré, Sinfra, Oumé, etc.). Il est alors à craindre que cette situation ne se transpose à Yamoussoukro à cause du transfert effectif de la capitale politique avec son corolaire d'augmentation de la population. En prévision du fait que ce transfert pourrait perturber la distribution de l'eau potable, il est donc impérieux d'estimer, ou mieux de prévoir les besoins en eau mensuels des populations afin d'en assurer un approvisionnement durable. Mais, partant du fait que la plupart des méthodes statistiques conventionnelles de prévision sont basées sur des modèles linéaires et que le succès pratique de ces approches est limité par leur linéarité (Lek et al, 1996), il y a lieu de proposer d'autres méthodes de modélisation. Parmi les plus prometteuses figurent les approches basées sur les réseaux de neurones. C'est donc pour mettre en place un modèle neuronal capable d'effectuer des prévisions de la consommation mensuelle en eau potable de la ville de Yamoussoukro dans le but de permettre aux décideurs d'assurer un accès permanent et durable en eau des populations que cette étude a été menée.

MATERIEL ET METHODES

Site d'étude

La ville de Yamoussoukro, capitale politique de la Côte d'Ivoire, est située au centre du pays dans la région des Lacs à environ 248 km au Nord-Ouest d'Abidjan. Elle est localisée entre 6°35' et

6°90' de latitude Nord et 4°80' et 5°30' de longitude Ouest (Figure-1). Elle est située sur un relief plat recouvert de la savane pré-forestière parsemée plus ou moins densément de petits arbres et entrecoupées de bosquets et flots forestiers. Des galeries forestières occupent les bas-fonds le long des marigots. Sur les sols sableux drainés, apparaissent des peuplements de rôniers. Le substratum rocheux de cette localité est essentiellement caractérisé par des granitoïdes éburnéens et des formations birrimiennes du socle précambrien. Quant au climat, il est de type équatorial caractérisé par quatre (04) saisons. Au niveau de l'hydrographie, cette région est drainée par le fleuve Bandama et ses affluents. Cette ville fait partie du District de la même dénomination avec une densité de 79,8 hab. /km² et un taux d'accroissement de 2,6% observé sur la période de 1988 à 1998 (INS, 2009). A Yamoussoukro, 80% des ménages sont connectés au réseau d'eau potable de la Société de Distribution d'Eau de la Côte d'Ivoire (SODECI) (Seidl, 2006).

Présentation des données utilisées

Les données utilisées dans cet article couvrent la période janvier 1995 à décembre 2008 en faisant abstraction aux données aberrantes de l'année 2000. Ces données sont entre autres : la consommation en eau potable de la ville de Yamoussoukro fournies par la Direction Régionale de la SODECI de Yamoussoukro ; la population de la ville de Yamoussoukro estimée aux divers horizons de la période d'étude à partir du taux de croissance, lui-même calculé à l'aide de certaines valeurs de population fournies par l'Institut National de la Statistique (INS) ; la pluviométrie et la température obtenues auprès de la Société d'Exploitation et de Développement Aéroportuaire, Aéronautique et Météorologique (SODEXAM). Toutes ces données ont été observées au pas de temps mensuel.

METHODES

Prétraitement des données pour la sélection des variables pertinentes

L'Analyse en Composantes Principales normée (ACPn) est une méthode statistique linéaire de traitement de données par la description d'un tableau de variables. Chaque entrée supplémentaire dans un réseau ajoute une dimension supplémentaire à l'espace dans lequel résident les données. L'ACPn a été utilisée pour la réduction de la dimension de l'espace des variables et l'élimination des variables non pertinentes. L'analyse est dite normée parce que les teneurs des variables sont centrées sur une moyenne et réduites. Le nuage de points obtenu est projeté sur un sous espace de dimension inférieure. Cette réduction du nombre de facteurs se traduit par la recherche des composantes principales qui sont des axes ayant absorbé le maximum de variance totale c'est-à-dire d'informations se rapportant à ces axes. Le premier axe est celui qui exprime le plus fort pourcentage de la variance totale du nuage de points. Le deuxième axe perpendiculaire au premier est celui qui exprime la plus grande partie de la variance résiduelle. L'ACPn appliqué aux données d'une zone n'est valable que lorsque les plans factoriels donnent plus de 70% de l'information (variance totale exprimée). Plusieurs variables qui portent dans une certaine mesure la même information sont fortement corrélées. En effet, deux variables sont corrélées lorsque leur coefficient

de corrélation est supérieur ou égal à 0,75 (Diei et al, 2004). L'utilisation d'un nombre minimal de ces variables corrélées est suffisante pour la modélisation avec les réseaux de neurones formels ; car, un nombre trop important de variables explicatives dans ce type de modèle augmentera la taille des paramètres à optimiser et la performance du réseau s'en trouvera altérée (Dreyfus et al, 2004).

Aperçu sur les réseaux de neurones formels et organisation des bases de données

Un Réseau de Neurones peut se définir comme un maillage de plusieurs neurones, généralement organisés en couches (Dreyfus et al, 2004 ; Koffi, 2007). De manière générale, le sens de transfert de l'information dans un réseau est défini par la nature des connexions qui sont directes ou récurrentes. Selon le chemin suivi par l'information dans le réseau on peut classer les Réseaux de Neurones en deux grandes catégories : les réseaux "feed-forward" et les réseaux "feed-back" (Dreyfus et al, 2004; Koffi, 2007). Les perceptrons multicouches (PMC) utilisés dans cette étude sont à deux (2) couches : une couche cachée et une couche de sortie (**Figure-2**).

La fonction neurone logistique a été choisie pour cette étude ; et les modèles neuronaux développés dans cette étude ont été entraînés avec la méthode de l'apprentissage supervisé en s'appuyant sur l'algorithme de rétro-propagation. Les modèles développés sont uniquement des Perceptrons Multicouches Bouclés Dirigés de prévision. Quatre modèles de ce type, différents du point de leurs architectures sont développés dans cet article. En ce qui concerne la répartition des données, elles ont été subdivisées en trois sous ensembles distinctes : 50% des données pour le calage du modèle, 25% pour arrêter l'apprentissage lorsque l'erreur sur cet ensemble dit de validation croisée atteint un minimum et 25% pour tester et évaluer la performance du réseau afin de le valider. Les données utilisées dans cette étude ont été auparavant normées afin d'être comprise entre 0 et 1 pour faciliter la convergence des modèles pendant leur apprentissage.

Critères de performance des modèles

La performance d'un modèle issu d'une méthode d'apprentissage s'évalue par sa capacité de prédiction ou de généralisation. En effet, les critères de performance mesurent la qualité ou la confiance que l'on peut accorder aux résultats d'une prévision en vue même, dans un cadre légal, d'une certification (Besse, 2005). Pour évaluer la qualité prévisionnelle des modèles développés dans cette étude, le coefficient de corrélation, la racine carrée de l'erreur quadratique moyenne et la moyenne biaisée ont été utilisés.

❖ Coefficient de corrélation « R »

Le coefficient de corrélation lui-même est utilisé pour mesurer la corrélation linéaire entre la consommation réelle ou mesurée et la consommation calculée. Sa formule est :

$$R = \frac{\sum_i (X_i - \bar{X}) \times (Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2 \times \sum_i (Y_i - \bar{Y})^2}}$$

Avec X_i = consommation mesurée ; Y_i = consommation calculée ; \bar{X} = moyenne des consommations mesurées ; \bar{Y} = moyenne des consommations calculées

Selon Koffi et *al* (2006) par convention la relation est parfaite si $R = 1$; très forte si $R > 0,8$; forte si R se situe entre 0,5 et 0,8 ; d'intensité moyenne si R se situe entre 0,2 et 0,5 ; faible si R se situe entre 0 et 0,2 ; nulle si $R = 0$. De manière générale si R est inférieure à 0,7 alors la relation est sujette à problème (Anctil et *al*, 2004).

❖ **Racine carrée de l'erreur quadratique moyenne « RMSE »**

La racine carrée de l'erreur quadratique moyenne est utilisée comme la mesure de la performance globale du modèle. Cet indice permet une évaluation combinée de la variance et du biais. Le modèle est bien optimisé si la valeur de RMSE est proche de zéro, ce qui tend vers une parfaite prévision. Sa formulation mathématique est donnée par la relation suivante (Koffi, 2007) :

$$RMSE = \sqrt{\frac{\sum_i (X_i - X_{cal})^2}{N}}$$

Avec X_i et X_{cal} qui sont les consommations observées et calculées ; N est le nombre de données.

❖ **Moyenne biaisée « B » (Dechemi et *al*)**

Ce critère définit la différence entre la moyenne des consommations observées et celle des consommations calculées (ou prédites). Lorsque la moyenne biaisée « B » tend vers zéro, les résultats du modèle sont sans biais. Ce paramètre est défini par la relation suivante :

$$B = |C_{moy_Obs} - C_{moy_Est}|$$

Avec : C_{moy_Obs} , la moyenne des consommations observées et C_{moy_Est} , la moyenne des consommations estimées.

RESULTATS

Variables explicatives pertinentes

L'Analyse en Composantes Principales Normée a permis de réduire la dimension de l'espace des variables et d'éliminer les variables explicatives non pertinentes. Les variables traitées sont les suivantes : la population, la consommation du mois précédent, le cumul des consommations des deux derniers mois et le cumul des consommations des trois derniers mois, la consommation du mois en cours, la pluie et la température. L'étude du tableau des valeurs propres, du pourcentage de variance propre et du pourcentage de variance exprimée cumulée (Tableau-1) issu de l'ACP a permis de rechercher les principaux facteurs à l'origine de l'évolution des variables. Ainsi les trois premiers facteurs F1, F2 et F3 expliquent à eux seuls 92,58% de la variance totale exprimée. Cela signifie que ces trois facteurs donnent une quantité d'informations suffisante pour être prise en compte et mener une étude.

L'analyse du tableau des coordonnées des variables (Tableau-2) montre que le facteur F1 est déterminé par la population, la consommation du mois précédent, le cumul des consommations des deux derniers mois et le cumul des consommations des trois derniers mois et la consommation du

mois en cours. Ce facteur exprime la variation des séries temporelles. Quant au facteur F2, il est déterminé par la pluie. Ce facteur exprime la variation pluviométrique. Le facteur F3, lui est déterminé par la température, il exprime la variation climatique.

En dehors de la consommation qui est la variable à expliquer, le facteur F1 contrôle quatre autres variables. L'analyse de la matrice de corrélation (Tableau-3) montre que ces quatre variables (population au mois (m), consommation du mois (m-1), cumul des consommations des deux mois précédant le mois (m), cumul des consommations des trois mois précédant le mois (m)) sont fortement corrélées entre elles. Ce qui veut dire qu'elles apportent la même information (on parle alors de redondance). Dans un premier temps, la population a été conservée; car, elle est la variable explicative dont les coefficients de corrélation avec les trois autres variables sont les plus bas (0,78 ; 0,81 et 0,83). Ensuite, en ce qui concerne ces trois dernières variables explicatives relatives aux consommations des mois précédents, les coefficients de corrélation sont très élevés ($\geq 0,92$). De manière arbitraire, la consommation (m-1) a été conservée et elle a servi à faire le bouclage du modèle. En définitive les variables explicatives retenues pour cette étude sont : la population, la pluie (pour le facteur F2), la température (pour le facteur F3) et la consommation (m-1) (pour le facteur F1).

Présentation de l'architecture des modèles et de leurs performances

Les architectures des modèles neuronaux de prévision obtenus sont résumées dans le tableau-4. Il s'agit des modèles neuronaux dont les performances sont les plus optimales. Ils présentent un nombre minimal de neurones cachés obtenus à la suite de la validation croisée.

Le tableau-5 résume la performance des quatre modèles retenus dans cet article. A l'analyse, on note que les modèles (3 9 1) et (4 8 1) sont ceux qui retiennent l'attention vu qu'ils présentent des coefficients de corrélation (R) supérieurs ou proches de 0,90 pour toutes les phases, ce qui n'est pas le cas des deux autres modèles (2 8 1) et (3 5 1) qui sont moins performants pour la prévision de la consommation mensuelle en eau potable de la population de la ville de Yamoussoukro.

Pour les deux modèles (3 9 1) et (4 8 1), la corrélation entre les consommations calculées (ou prédites) et celles observées est assez forte ; ces deux modèles sont donc très performants. La racine carrée de l'erreur quadratique moyenne (RMSE) de la consommation normée a été aussi analysée surtout en phase de calage. Il faut rappeler que le modèle est bien optimisé si la valeur de RMSE est proche de zéro. Dans cette phase de calage, le coefficient RMSE des différents modèles (3 9 1) et (4 8 1) sont les plus petits respectivement 0,027 et 0,029 contre 0,036 et 0,034 respectivement pour les modèles (2 8 1) et (3 5 1). Compte tenu de ces résultats, la suite de cette étude ne s'intéressera qu'aux modèles (3 9 1) et (4 8 1) qui apparaissent comme les plus performants pour la prévision de la consommation en eau potable de la population de la ville de Yamoussoukro. Comme la qualité d'un modèle est basée sur l'ajustement de ce modèle à la réalité, les paragraphes suivants n'exposeront que les performances des modèles (3 9 1) et (4 8 1) obtenues pendant la phase de Test. Sur cette base, les valeurs du coefficient de corrélation (R) des modèles

retenus sont respectivement de 0,90 et de 0,92. On remarque que le coefficient de corrélation R obtenu avec le modèle (4 8 1) est supérieur à celui obtenus avec le modèle (3 9 1). Les valeurs du paramètre B consigné permet également de comparer les deux modèles au regard de l'étude des erreurs. En observant les valeurs du paramètre B obtenus en phase de Test des deux modèles, les performances du modèle (4 8 1) sont encore meilleures que celle de modèle (3 9 1). L'analyse des graphes de la figure 3 indique que les modèles (3 9 1) et (4 8 1) prévoient bien les consommations en eau de la ville de Yamoussoukro avec cependant quelques valeurs qui sont surestimées ou sous-estimées. En effet, on a pu observer dans cette étude que c'est seulement 14% et 9% des consommations mensuelles en eau des populations de Yamoussoukro, respectivement, que les modèles (3 9 1) et (4 8 1) n'arrivent pas à prévoir correctement.

DISCUSSION

Parmi les modèles développés, le modèle (2 8 1) avec comme variables explicatives la consommation du mois précédent et la population est le moins performant. Cette mauvaise performance ($R \leq 0,71$) pourrait s'expliquer par le fait que les variables explicatives considérées (la consommation du mois précédent et la population de la ville de Yamoussoukro) ne suffisent pas pour exprimer la variable à expliquer (la consommation du mois suivant). La mauvaise performance de ce modèle (2 8 1) pourrait être aussi due à la fiabilité des données utilisées en entrée. En effet, selon Parizeau (2004) les modèles neuronaux sont influencés par la qualité des données. Dans la présente étude, la variable « population » utilisée, à l'instar de toutes les autres données démographiques, est calculée au pas de temps annuel. Cela impose d'attribuer la même valeur de la population à chacun des douze (12) mois de l'année ; car, il est impossible de déterminer une valeur exacte du nombre de la population de la ville de Yamoussoukro pour chaque mois. Le deuxième modèle (3 5 1) développé avec en entrée les variables explicatives du premier modèle (2 8 1) et la pluie est légèrement plus performant que le modèle (2 8 1). En effet, l'ajout de la pluie en entrée a apporté de légères améliorations, mais de manière globale les performances ne sont pas très satisfaisantes et le coefficient de corrélation de Pearson demeure toujours très proche de 0,7. On en déduit que l'ajout de la pluie n'a pas apporté suffisamment d'informations significatives au modèle afin de lui permettre de faire une bonne prévision de la consommation en eau de la ville de Yamoussoukro. La pluie pourrait donc ne pas influencer significativement la consommation en eau de cette localité. On pourrait noter que dans cette ville, l'eau de pluie ne constitue pas un substitut à l'eau potable comme c'est le cas dans d'autres localités de l'intérieur du pays, surtout en milieu rural. Le troisième modèle (3 9 1) a été obtenu en remplaçant la pluie par la température dans le deuxième modèle (3 5 1). Dans ce modèle, les performances en phase de calage et de validation se sont nettement améliorées. Cette amélioration de la capacité des modèles neuronaux à prédire la consommation se ressent au niveau des coefficients de corrélation R (respectivement 0,94 et 0,93 pour le calage et la validation) et au niveau des graphes de comparaison entre les consommations calculées ou prédites et celles observées. La température apporterait donc beaucoup plus d'informations aux modèles neuronaux développés dans cette étude

que la pluie pour la prévision de la consommation en eau potable de la ville de Yamoussoukro. On peut en déduire que cette variable climatique induit un effet positif à la consommation en eau potable de la ville. En effet, en période de chaleur, les populations transpirent beaucoup et boivent alors assez d'eau pour compenser cette eau perdue. Le quatrième et dernier modèle (4 8 1) a pour variables explicatives l'ensemble des variables d'entrée de l'étude à savoir la population, la pluie, la température et la consommation. Comme précédemment remarqué, la pluie apporte encore de légères améliorations au modèle de prévision de la consommation en eau potable. Ainsi les coefficients de corrélation de Pearson passent de 0,94 à 0,97 en Apprentissage ; de 0,93 à 0,96 en Validation et de 0,9 à 0,92 en phase de Test. En somme l'ajout de la température améliore significativement les performances des modèles tandis la pluie n'apporte qu'une relative amélioration. Le choix des variables d'entrée influence beaucoup la performance des réseaux de neurones. D'autres études qui prévoyaient la même variable à expliquer mais avec des variables explicatives différentes ont donné de bien meilleurs résultats. En effet Guhl et Brémond (2000) ont construit des modèles de prévision de la demande en eau potable au pas de temps horaire. Ils se sont affranchis des données météorologiques ou climatiques du fait des incertitudes du modèle météorologique et de la difficulté à établir des liens entre ces données et la consommation. Les données utilisées étaient plutôt le mois de l'année, le jour de la semaine, l'heure de la journée et le fait que le jour considéré, corresponde à une période de vacances scolaires (oui ou non). Les résultats obtenus sont très satisfaisants jusqu'à un horizon de précision de 10 heures ($R \geq 0,9$). Les performances obtenues dans cette présente étude, pourraient donc s'expliquer par le choix des variables explicatives dont les liens avec la consommation ne seraient pas assez forts.

CONCLUSION

Il ressort de cette étude que l'utilisation de la population et la consommation d'une part et de la population, la consommation et la pluie d'autre part en entrée des Perceptrons multicouches n'ont pas conduit à de bonnes estimations de la consommation d'eau de la ville de Yamoussoukro. Mais en ce qui concerne les modèles (3 9 1) et (4 8 1), ils donnent des résultats plus satisfaisants. Le modèle (4 8 1) fournit de meilleurs résultats que le (3 9 1) pour toutes les phases. Ainsi l'ajout de la température aux variables explicatives du premier modèle (2 8 1) que sont la population et la consommation a donné le modèle (3 9 1). Et cela a permis d'obtenir de bien meilleures performances. En outre l'ajout de la pluie aux variables explicatives du modèle (3 9 1) améliore davantage les performances du modèle qui est obtenu (4 8 1) en phase de calage mais surtout sur un ensemble de données différent de celui de la base d'apprentissage (test). En définitive, de ces deux derniers modèles, le modèle (4 8 1) a été retenu ; car, il est celui qui est le moins complexe avec 8 neurones cachés déterminés à la suite du principe de la validation croisée induisant un arrêt prématuré. Bien que les performances de nos modèles soient acceptables dans l'ensemble, les perceptrons multicouches bouclés dirigés avec une seule couche cachée pourraient être plus performants et constituer un excellent outil de prévision de la consommation d'eau potable au pas de temps mensuel sur une base de données plus importante et de meilleure qualité.

Remerciements

Les différents auteurs de cet article remercient très sincèrement l’UFR STRM de l’Université de Cocody, pour avoir initié cette étude. Ils remercient également les différents instructeurs qui ont participé à l’amélioration du présent article.

Figure-1. Localisation et plan de la ville de Yamoussoukro

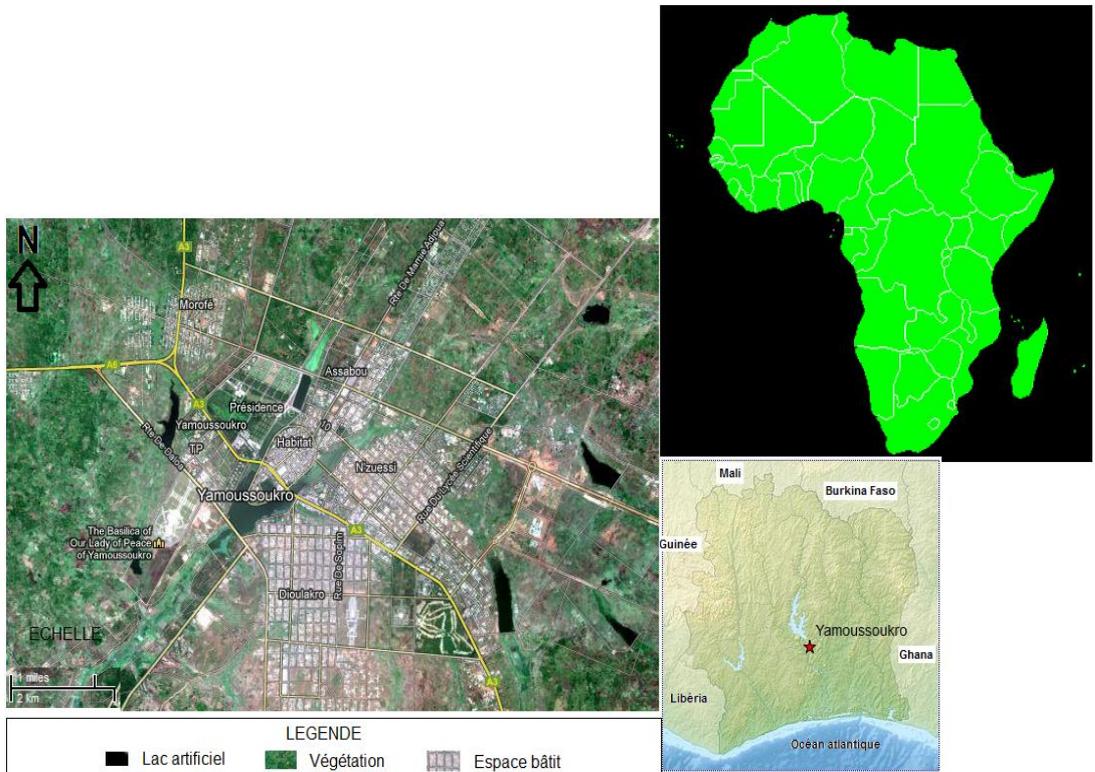


Figure-2. Aperçu d’un exemple de Perceptron Multicouche de prévision à deux couches

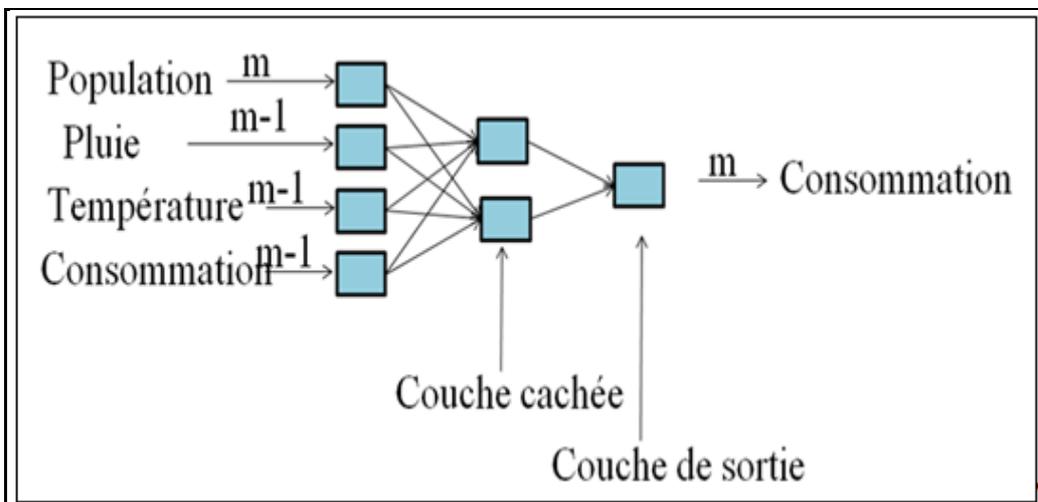
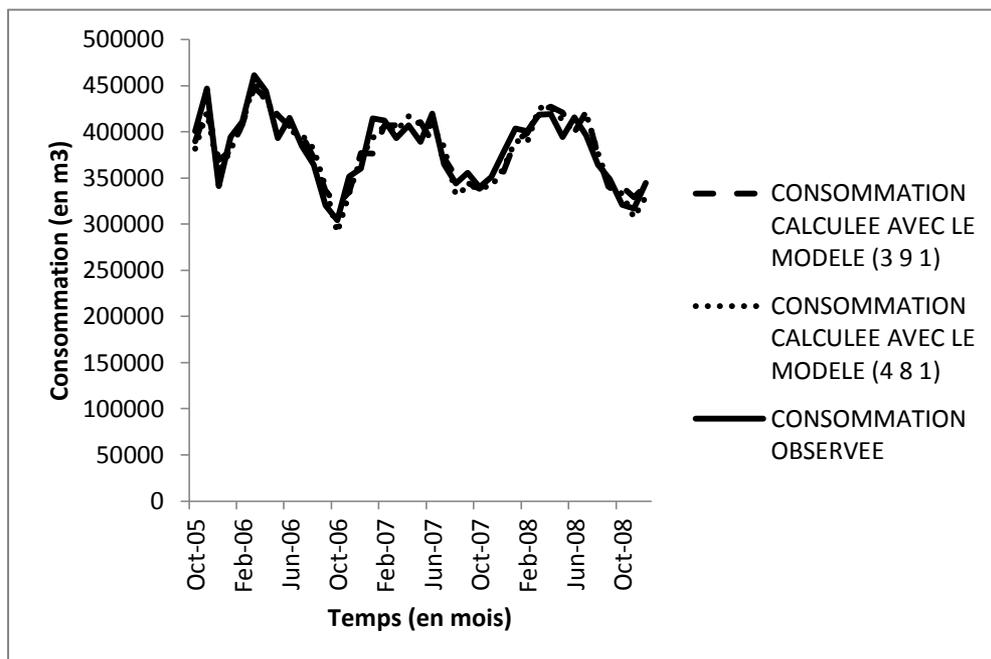


Figure-3. Comparaison des consommations observées et calculées (prédites) en phase de test des modèles (3 9 1) et modèle (4 8 1)**Tableau-1.** Récapitulatif des valeurs propres, % de variance propre et % de variance cumulée

	F1	F2	F3	F4	F5	F6	F7
Val. Propre	4,492	1,031	0,958	0,249	0,196	0,062	0,012
% variance	64,171	14,725	13,685	3,563	2,793	0,886	0,176
% cumulé	64,171	78,896	92,581	96,145	98,938	99,824	100,000

Tableau-2. Résumé des coordonnées des variables

(Pop : population au mois (m); P_{moy} : pluie moyenne mensuelle au mois (m); T_{moy} :

	F1	F2	F3	F4	F5	F6	F7
POP	0,882	-0,106	-0,145	0,429	-0,062	0,045	0,005
P_MOY	0,076	0,987	0,111	0,064	0,062	0,017	0,003
T_MOY	0,269	-0,112	0,954	0,040	-0,052	-0,006	0,000
C_M_-1	0,958	0,019	-0,027	-0,197	-0,078	0,186	-0,024
Cu_M_-2	0,976	0,057	-0,064	-0,129	-0,108	-0,062	0,085
Cu_M_-3	0,970	0,103	-0,085	-0,056	-0,115	-0,144	-0,066
C	0,908	-0,138	0,038	-0,028	0,392	-0,019	-0,001

température moyenne mensuelle au mois (m); C_{m-1} : consommation du mois (m-1); Cu_{m-2} : cumul des consommations des deux mois précédant le mois (m); Cu_{m-3} : cumul des consommations des trois mois précédant le mois (m); C : consommation du mois (m).

Tableau-3. Matrice de corrélation entre variables analysées

	POP	P_MOY	T_MOY	C_M_-1	Cu_M_-2	Cu_M_-3	C
POP	1	-0,03	0,13	0,775	0,813	0,833	0,772
P_MOY	-0,03	1	0,015	0,074	0,108	0,153	-0,04
T_MOY	0,13	0,015	1	0,224	0,196	0,172	0,274
C_M_-1	0,775	0,074	0,224	1	0,959	0,929	0,838
Cu_M_-2	0,813	0,108	0,196	0,959	1	0,981	0,838
Cu_M_-3	0,833	0,153	0,172	0,929	0,981	1	0,822
C	0,772	-0,04	0,274	0,838	0,838	0,822	1

Tableau-4. Récapitulatif des architectures des modèles de neurones retenues pour l'étude

Architecture du modèle	Variables explicatives	Nombre de neurones cachés	Variable à expliquer
(2 8 1)	Population et consommation	8	Consommation en eau
(3 5 1)	Population, pluie et consommation	5	Consommation en eau
(3 9 1)	Population, température et consommation	9	Consommation en eau
(4 8 1)	Population, pluie, température et consommation	8	Consommation en eau

Tableau-5. Performance des différents modèles avec la consommation réelle (R) et la consommation normée (RMSE)

Architecture des modèles	RMSE			R		
	Calage	Validation	Test	Calage	Validation	Test
(2 8 1)	0,036	0,044	0,045	0,66	0,71	0,63
(3 5 1)	0,034	0,046	0,039	0,69	0,70	0,70
(3 9 1)	0,029	0,043	0,041	0,94	0,93	0,90
(4 8 1)	0,027	0,042	0,044	0,97	0,96	0,92

RÉFÉRENCES

- Anctil F., Michel C., Perrin C., Andréassian V. (2004)** "A soil moisture index as an auxiliary ANN input for stream flow forecasting". Journal of Hydrology, Vol. 286, pp. 155–167
- Ashu J., Sanaga S. (2006)** "Integrated approach to model decomposed flow hydrograph using artificial neural network and conceptual techniques" Journal of Hydrology. Vol. 317, pp 291-306
- Besse P. (2005)** "Data mining II. Modélisation Statistique & Apprentissage. Publications du laboratoire de statistique et probabilités". Université Paul Sabatier, Toulouse, 103p
- Coulibaly P., Anctil F. and Bobée B. (2000)** "Daily reservoir inflow forecasting using artificial neural networks with stopped training approach". Journal of Hydrology, Vol. 230, pp. 244–257.

- Coulibaly P., Anctil F. and Bobée B. (1999)** “Prévision hydrologique par réseaux de neurones artificiels : Etat de l’art.” Revue canadienne de génie civil, Vol. 26, pp. 293-304.
- Dechemi N., Tarik B. A. and Issolah A. (2003)** “Modélisation des débits mensuels par les modèles conceptuels et les systèmes neuro-flous.” Revue des Sciences de l’Eau. Vol. 16 n°4, pp. 407-424.
- Diei B., Inza D., Goula B. T. A., Savane I., Biémi J. (2004)** “Analyse statistique des paramètres influençant la productivité des forages d’eau en milieu cristallin et cristallophyllien dans la région d’Aboisso (sud-est de la Côte d’Ivoire).” Revue Sud sciences et technologie, N°13, pp 22-30
- Dreyfus, G., Martinez, J. M., Samuelides, M., Gordon, M. B, Badran, F., Thiria, S., et Herault, L. (2004)** “Réseaux de Neurones” Méthodologie et application. 2^{ème} édition, Groupe Eyrolles, 374 p.
- Florent G. and Bernard B. (2000)** “Optimisation du fonctionnement des réseaux d’eau potable.” Prise en compte de l’aspect stochastique de la demande. Revue Ingénieries – EAT, N° 23 pp 15-23
- Koffi Y. B. (2007)** “Modélisation pluie-débit en région tropicale humide : Application des réseaux de neurones sur quatre stations hydrométriques du Bandama Blanc (Bada, Marabadiassa, Tortiya, Bou), Nord de la Côte d’Ivoire.” Thèse de doctorat d’état, Université de Cocody, 237p
- Koffi Y. B., Lasm T., Ayrat P. A., Anne J., Kouassi A. M., Assidjo E. and Biémi J. (2006)** “Optimization of Multi-Layers Perceptrons Models with Algorithms of First and Second Order. Application to the Modelling of Rainfall-Rainoff Relation in Bandama Blanc Catchment (North of Ivory Coast).” European Journal of Scientific Research, Vol. 17 N°3, pp. 313-328.
- 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. Vol. 9 N°3, pp. 319-331.
- ONEP (2009)** “Etats Généraux de l’eau potable en Cote d’Ivoire, Yamoussoukro.” Rapport de synthèse, 9p
- Parizeau M. (2004)** “Réseaux de Neurones.” GIF-21140 et GIF-64326. Université Laval, 115p.
- Schmitt A., Le Blanc B., Corsini M. M., Lafond C., Bruzek J. (2001)** “Les réseaux de neurones artificiels, un outil de traitement de données prometteur pour l’anthropologie.” Bull. et Mém. de la société d’Anthropologie de Paris, n.s, t. 13,2001, 1-2, pp. 143-150
- Seidl M. (2006)** “Enjeux et pratiques de l’assainissement en Afrique sub-saharienne. 17èmes Journées Scientifiques de l’Environnement : le Citoyen, la Ville et l’environnement” Collection HAL Archives Ouvertes, pp. 4-6.
- Thuillier D. (1997)** “Principes et applications des réseaux de neurones : deux illustrations sur l’habitat au Maroc.” Revue Région et Développement, n°5, 28p