



CONCEPTION DE MODÈLES STATISTIQUES A VARIABLES HYDROCHIMIQUES POUR LA PRÉDICTION DE LA CONDUCTIVITÉ ÉLECTRIQUE DES EAUX SOUTERRAINES

**KOUASSI A.M.¹, MAMADOU A.²,
AHOUSSE K.E.³, BIEMI J.³**

¹Institut National Polytechnique Félix Houphouët-Boigny (INP-HB) ; Département des Sciences de la Terre et des Ressources Minières (STeRMi), Laboratoire du Génie Civil, des Géosciences et des Sciences Géographiques, BP 1093 Yamoussoukro
Côte d'Ivoire

²Institut National Polytechnique Félix Houphouët-Boigny (INP-HB) de Yamoussoukro, Ecole Supérieure des Mines et de Géologie (ESMG), Bp: 1093 Yamoussoukro
Côte d'Ivoire

³Université Félix Houphouët-Boigny d'Abidjan-Cocody, Unité de Formation et de Recherche des Sciences de la Terre et des Ressources Minières (UFR-STRM), Laboratoire des Sciences et Techniques de l'Eau et de l'Environnement (LSTEE) ; 22 BP 582 Abidjan 22, Côte d'Ivoire

michel.a_kouassi@yahoo.fr

RESUME

Cette étude a porté sur l'ensemble des régions du N'zi, Iffou et Moronou (ex-N'zi-Comoé), situé au Centre-Est de la Côte d'Ivoire. L'objectif de l'étude est de concevoir des modèles statistiques à partir des régressions linéaires multiples, capables de simuler la conductivité électrique des eaux souterraines à partir de paramètres physico-chimiques de celles-ci. L'étude a été réalisée à partir d'un échantillon de 193 forages d'eau. Les paramètres physiques utilisés concernent la température, le pH, la turbidité et la conductivité électrique. Quant aux paramètres chimiques, ils concernent les cations majeurs et les anions majeurs. La méthodologie a consisté, d'une part à appliquer une analyse en composantes principales normées (ACPN) afin d'identifier les variables explicatives pertinentes et expressives de la conductivité électrique, et d'autre part, à déterminer les coefficients des régressions linéaires et les erreurs-types associées à partir d'un calage automatique. Les résultats de l'ACPN montrent que les variables les plus pertinentes et expressives pour la modélisation de la conductivité électrique des eaux souterraines sont les paramètres chimiques.

Ainsi, ces deux paramètres ont été utilisés comme variables explicatives pour la conception d'un premier modèle. Aussi, pour mettre en évidence les impacts des activités anthropiques sur la qualité des eaux souterraines, un deuxième modèle a-t-il été envisagé par intégration de la température et du pH dans le premier modèle. Les valeurs du coefficient de variation sur les coefficients de régression des variables explicatives sont relativement faibles (inférieures à 10%) en ce qui concerne les paramètres chimiques quel que soit le modèle. Elles sont suivies par le pH (23%) puis la température (511%).

Mots-clés: Conductivité électrique, ACPN; Modélisation statistique, Calage, Côte d'Ivoire.

ABSTRACT

The study takes place on the regions of N'zi-Iffou-Moronou (ex-N'zi-Comoé), located in East-Central of Ivory Coast with an area of 19,560 km². The objective of this study is to develop statistical models using multiple linear regressions able to simulate the electrical conductivity of groundwater from physical and chemical parameters of water. Sample from 193 boreholes were used for this study. The physical parameters are temperature, pH, turbidity and electrical conductivity. Chemical parameters selected are the major cations and major anions. The methodology consisted on the one hand, a normalized principal component analysis (NPCA) to identify the relevant variables, on the second hand, the model calibration to determine the coefficients of linear regressions and standard errors associated. The results of the normalized principal component analysis show that the most relevant and expressive modeling of the electrical conductivity of groundwater variables are the sum of the major cations and the sum of the major anions. These two parameters have been used as variables for designing a first model. Also, to show the impacts of anthropogenic activities on groundwater quality, a second model was proposed by integrating the temperature and pH in the first model. The values of the coefficient of variation on the coefficients of regression of the explanatory variables are relatively low (lower than 10%) with regard to the chemical parameters whatever the model. They are followed by the pH (23%) then the temperature (511%).

Keywords: Electrical conductivity, NPCA, Statistical modeling, calibration, Ivory Coast.

INTRODUCTION

La détermination de la qualité des eaux des forages se fait par l'interprétation des résultats des analyses physico-chimiques effectuées sur les échantillons

d'eau prélevés dans ces forages. Il paraît important d'étudier la fiabilité des données d'analyses physico-chimiques dès lors que les résultats des interprétations en dépendent. Pour vérifier cette fiabilité, un certain nombre de moyens sont utilisés. Il s'agit de la balance ionique et de la conductivité électrique. La balance ionique, méthode généralement utilisée, rend compte de la fiabilité des analyses physico-chimiques effectuées au laboratoire permettant ainsi de déceler d'éventuelles erreurs commises lors du prélèvement ou du dosage des échantillons. Pour ce qui concerne la conductivité électrique, il s'agit d'une part, de comparer la conductivité électrique mesurée *in situ* à celle mesurée au laboratoire et d'autre part, de comparer la conductivité mesurée *in situ* et celle simulée à partir de modèles. En effet, de nombreuses études réalisées sur la caractérisation des eaux ont révélé l'existence d'une étroite relation entre la conductivité électrique et la minéralisation totale (Achour et Guergazi, 2002; Zourhi et Carlier, 2002; Zouhri, 2003; Tidjani et al., 2006; Kouassi et al., 2012a). On peut aussi comparer les résultats des approches basées d'une part sur la balance ionique et d'autre part sur la conductivité électrique. D'après les travaux de Kouassi et al. (2013), les résultats de la balance ionique au seuil de 10% montrent que seuls les échantillons de Bondoukou et de Guiglo-Duekoué ne sont pas fiables. Les résultats de la simulation de la conductivité électrique ont confirmé la fiabilité des échantillons de Tiébissou-Didiévi, ex-N'zi-Comoé, Agboville et Tiassalé et la non fiabilité de l'échantillon de Guiglo-Duekoué. Cependant, l'échantillon de Bondoukou, déclaré non fiable par la balance ionique, s'est avéré fiable au regard des résultats de la simulation de la conductivité électrique. En émettant l'hypothèse que la balance ionique est la méthode de référence dans l'analyse de la fiabilité des résultats de dosages chimiques, il se pose la question de la performance du modèle utilisé par Kouassi et al. (2013). Il est question de rechercher des modèles plus performants et capables d'apporter des solutions proches voire meilleures que la balance ionique. Dans cette problématique, un regard pourrait être tourné vers les modèles statistiques. Alors, on pourrait se demander «quel est l'apport des modèles statistiques à la modélisation de la conductivité électrique des eaux souterraines dans l'optique de l'étude de la fiabilité des résultats d'analyses physico-chimiques?». En effet, des paramètres tels que la température de l'eau sont de plus en plus modélisés à partir de l'approche statistique (Belanger et al., 2005; Valé, 2006; etc.). Cette approche statistique peut regrouper des méthodes telles la régression linéaire, la régression non-linéaire, les modèles autorégressifs (ARMA), etc. La régression linéaire multiple a fait l'objet de plusieurs études (Belanger et al., 2005 ; Valé, 2006 ; etc.). Ces raisons font de la régression linéaire multiple, l'approche privilégiée dans cette étude.

L'objectif général de cette étude est de concevoir des modèles statistiques au moyen des régressions linéaires multiples capables de prédire la conductivité électrique des eaux souterraines à partir de paramètres physico-chimiques. Il s'agit spécifiquement d'identifier des variables explicatives pertinentes et

expressives de la conductivité électrique des eaux souterraines et de déterminer les coefficients de régression et erreurs-types des différentes équations. L'atteinte de cet objectif permettra d'apporter des solutions à notre problématique, c'est-à-dire proposer une nouvelle approche appropriée de modélisation de la conductivité électrique des eaux souterraines dans le but de développer des outils de simulation qui seront utiles à la gestion intégrée et durable des ressources en eau.

La méthodologie a consisté, d'une part à appliquer une analyse en composantes principales normées (ACPN) afin d'identifier les variables explicatives pertinentes et expressives de la conductivité électrique, et d'autre part, à déterminer les coefficients des régressions linéaires et les erreurs-types associées à partir d'un calage automatique.

PRÉSENTATION DE LA ZONE D'ÉTUDE

L'ex-région du N'zi-Comoé est aujourd'hui divisée en trois (3) régions à savoir la région du N'zi, la région du Iffou et la région du Moronou. Elle est située au Centre-Est de la Côte d'Ivoire et est comprise entre les longitudes 3°40' et 4°55' Ouest et les latitudes 6°20' et 8°10' Nord (Figure 1). Elle occupe une superficie totale de 19 560 km². Les chefs-lieux des régions du N'zi, du Iffou et du Moronou sont respectivement, les villes de Dimbokro, Daoukro et Bongouanou. Le relief de la zone d'étude est monotone et constitué d'un plateau légèrement incliné de direction Nord-Sud, avec une altitude moyenne d'environ 150 m. Les collines de la chaîne Baoulé correspondent au faciès volcano-sédimentaire, principalement les roches vertes. La végétation est constituée de lambeaux de forêt mésophile, et de larges mailles de savanes séparées par des forêts-galeries. Il existe à côté de ces végétations naturelles, des végétations liées à l'action anthropique (mosaïques forêt-cultures, mosaïques savanes-cultures, cultures). Le climat est de type tropical humide avec une alternance de 4 saisons dont deux (2) saisons pluvieuses et deux (2) saisons sèches. Le régime climatique est caractérisé par des précipitations moyennes mensuelles variant entre 11 mm et 198 mm. La hauteur pluviométrique interannuelle (1950-2000) est de 1155 mm.

L'ex-région du N'zi-Comoé appartient au domaine paléoprotérozoïque encore appelé domaine baoulé-mossi. Du point de vue lithologique, la zone d'étude est recouverte par un ensemble complexe de formations géologiques constituées essentiellement de roches ignées, métamorphiques et volcano-sédimentaires (Daouda, 1998) (Figure 2). On distingue les roches éburnéennes et les roches birimiennes. Les roches éburnéennes sont constituées essentiellement de granitoïdes notamment les granites à deux micas, les granites à biotite homogènes et hétérogènes, et les granodiorites. Les formations birimiennes sont essentiellement des métasédiments et des roches vertes (métavulcanites). Les métasédiments sont constitués en majeure partie de grès et de schistes. Les

roches vertes sont constituées de métabasalte, métaandésite, métadolérite, métagabbro, amphibolite.

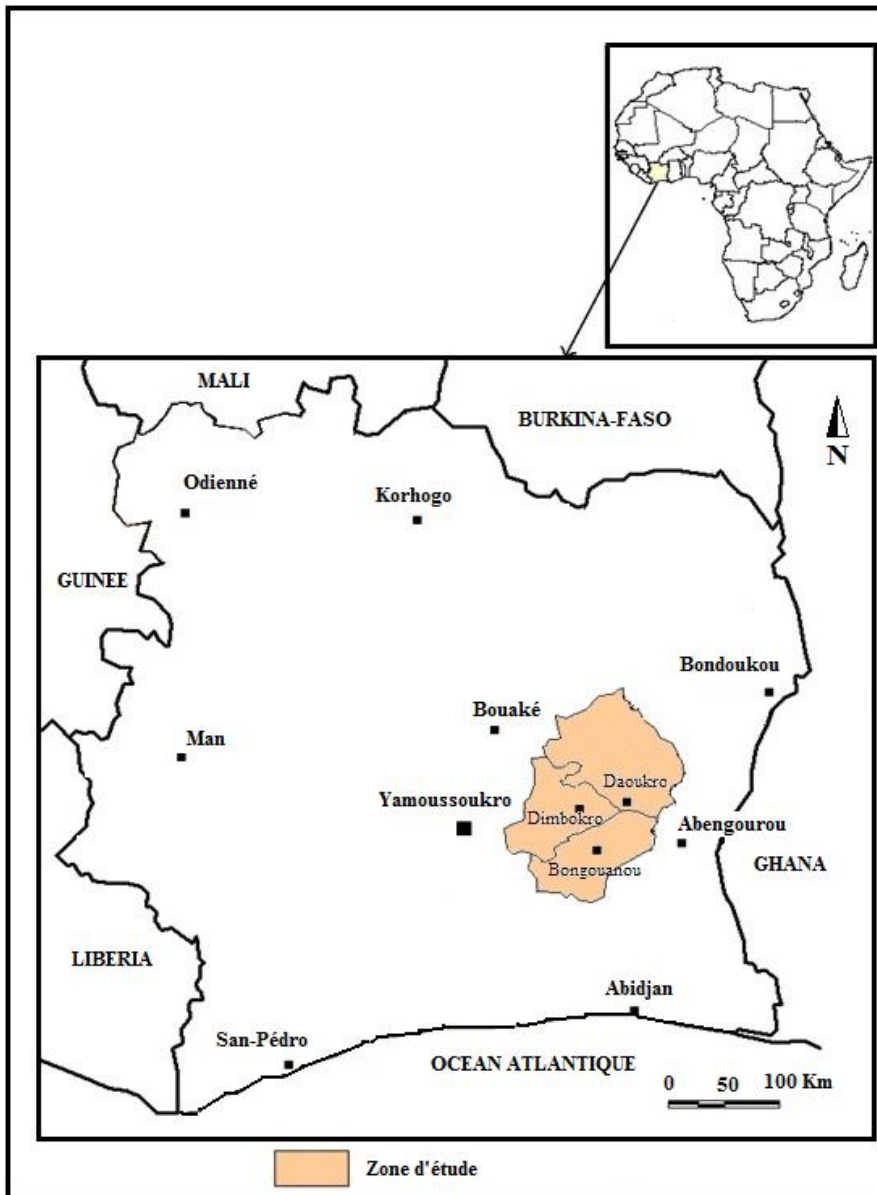
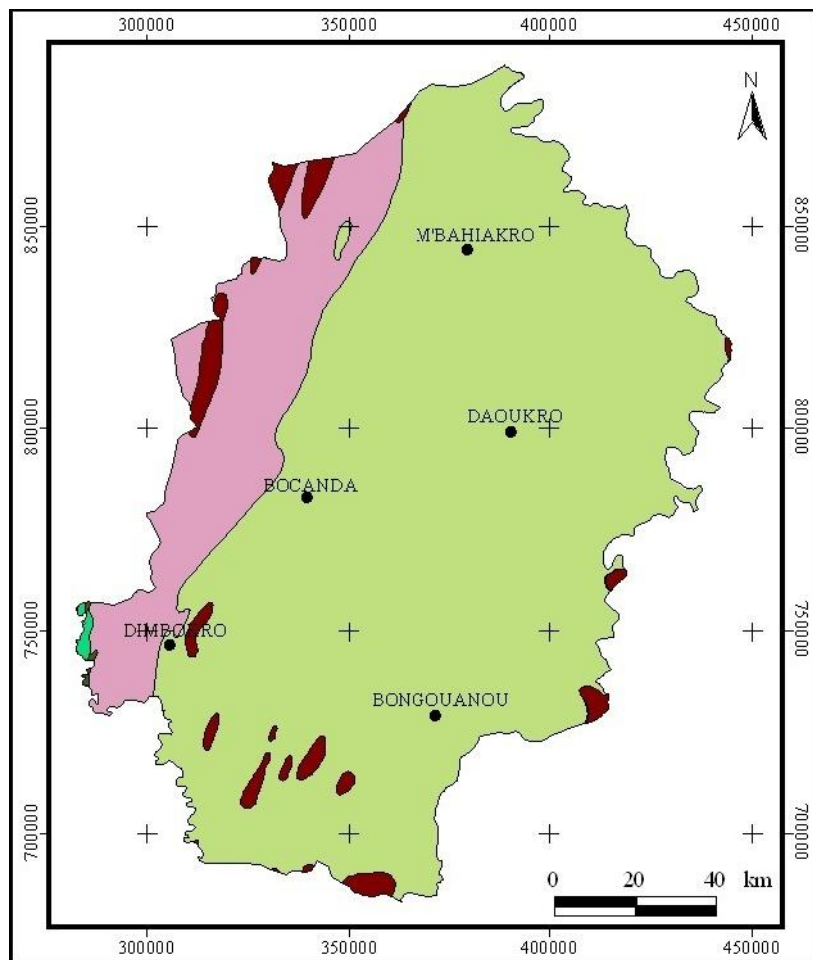


Figure 1 : Présentation de l'ex-région du N'zi-Comoé (Centre-Est de la Côte d'Ivoire)



LEGENDE

- | | |
|---|--|
|  Schistes, Grès, Conglomérats |  Granites récents |
|  Volcano-sédiments
(Laves acides ou basiques) |  Granites, Migmatites, Gneiss |

Figure 2 : Carte géologique de l'ex-région du N'zi-Comoé (Centre-Est de la Côte d'Ivoire)

MATÉRIEL ET MÉTHODES

Matériel

L'étude a été réalisée à partir d'un échantillon de 193 forages d'eau. Les données concernent les résultats des analyses des paramètres physico-chimiques réalisées sur ces forages. Les paramètres physiques concernent la

température, le pH, la turbidité et la conductivité électrique. Les données de conductivité électrique ont été corrigées. Les paramètres chimiques sélectionnés sont d'une part, les cations majeurs (le calcium, le magnésium, le sodium, le potassium) et d'autre part, les anions majeurs (le bicarbonate, le chlore, le nitrate, le sulfate). Les paramètres physiques ont été mesurés *in situ*. Les échantillons d'eau destinés aux analyses chimiques ont été recueillis dans des bouteilles en polyéthylène pour être analysés au Laboratoire d'Analyse d'Abengourou (LPA) (Côte d'Ivoire). Ces analyses ont été réalisées à l'aide d'un spectrophotomètre DR 2010 de type HACH. Ces données ont été recueillies lors du projet « Don Japonais », exécuté au cours de l'année 2000. Elles sont issues de la consultation des fiches techniques d'analyse de la qualité de l'eau des forages mises à notre disposition par la Direction de l'Office National de l'Eau Potable (ONEP) de Yamoussoukro. Les données utilisées pour ce travail proviennent de différents forages. La répartition spatiale des données est d'autant plus importante vue qu'elle rassure quant à la généralisation des modèles à définir. La figure 3 montre la répartition des forages dont les données ont fait l'objet de la présente étude.

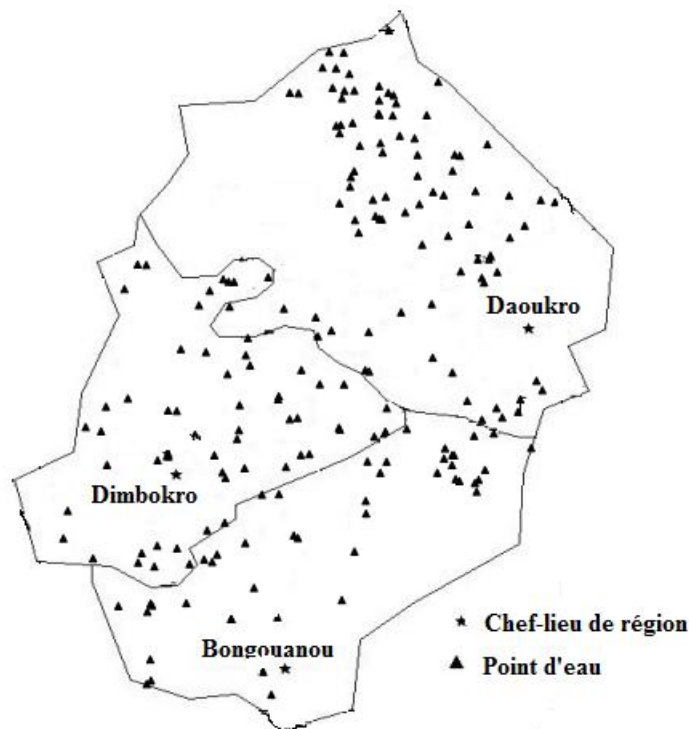


Figure 3: Carte de répartition des forages étudiés dans l'ex-région du N'Zi-Comoé

Le tableau 1 présente les caractéristiques statistiques descriptives des variables physico-chimiques utilisées.

Tableau 1: caractéristiques statistiques des variables physico-chimiques

	Min	Max	Moyenne	Ecart-type	Coef. de variation
Somme des cations	1,22	12,59	3,49	1,72	0,49
Somme des anions	0,74	12,00	3,64	1,84	0,51
Température	21,40	31,60	27,18	1,37	0,05
pH	5,12	8,21	6,58	0,48	0,07
Turbidité	0,30	3,50	0,80	0,37	0,47
Conductivité électrique	106,00	1535,00	391,40	200,53	0,51

L'analyse du tableau 1 révèle que les eaux souterraines de la zone d'étude ont des températures variant entre 21,40 et 31,60 °C, avec une moyenne de 27,18°C. Le pH est compris entre 5,12 et 8,21, avec une moyenne est de 6,58. La turbidité varie entre 0,33 et 3,49 NTU, avec une moyenne de 0,80 NTU. La conductivité électrique oscille entre 106 et 1535 µs/cm, avec une moyenne de 391,4 µs/cm. La somme des cations varie de 1,22 à 12,59 méq/l, avec une moyenne de 3,49 méq/l. La somme des anions varie de 0,74 à 12,00 méq/l, avec une moyenne de 1,84 méq/l. Les coefficients de variation des paramètres physico-chimiques des données varient de 0,51 à 0,05). La température et le pH sont homogènes ($C_v < 0,25$). Cependant, les autres éléments sont dans l'ensemble hétérogènes ($C_v > 0,25$).

Méthodologie relative à la conception des modèles

Identification des variables explicatives des modèles

Il s'agit de rechercher les variables explicatives les plus pertinentes et expressives de la conductivité électrique. La méthode de l'analyse en composantes principales normées (ACP_N) apporte une réponse pratique à ce problème. Cette méthode a été déjà utilisée avec succès par d'autres auteurs dans le domaine de l'hydrochimie (Fournier, 2006; Lasm et al., 2008; Ahoussi et al., 2010; Kouassi et al., 2011). C'est une méthode statistique multidimensionnelle qui permet de synthétiser un ensemble de données tout en minimisant la perte de l'information. Elle permet de classer un nombre important de données, afin d'en extraire les facteurs principaux qui sont à l'origine de l'évolution simultanée des variables et de leurs relations réciproques. Les données de base sont des variables quantitatives, continues, homogènes ou non, a priori corrélées entre elles deux à deux. L'ACP_N permet

donc de définir de nouvelles variables, artificielles, et des représentations graphiques permettant de visualiser les relations entre variables, ainsi que l'existence éventuelle de groupes d'individus et de groupes de variables. L'ACPN a été effectuée sur 193 échantillons d'eau. Les variables physico-chimiques prises en compte pendant cette analyse sont au nombre de six (6) (la conductivité électrique, la température, la turbidité, le pH, la somme des cations majeurs et la somme des anions majeurs). L'ACPN a permis de mettre en évidence les variables qui ont une influence majeure sur la conductivité électrique. Aussi, la régression linéaire multiple présente-elle certains défauts (Valé, 2006). Ainsi, un nombre trop important de variables explicatives par rapport au nombre d'individus la rend invalide. D'où la nécessité de réduire le nombre de paramètres en appliquant l'ACPN aux données de base. L'interprétation des résultats de l'ACPN s'est appuyée sur les valeurs propres, la matrice de corrélation et le cercle de communauté. L'ACPN appliquée aux données d'une zone n'est valable que lorsque les plans factoriels donnent plus de 70% de l'information (variance totale exprimée). Une valeur propre est choisie lorsqu'elle est supérieure ou égale à 1 sachant que la somme des valeurs propres est égale au nombre de variables utilisées pour la réalisation de l'ACPN. Elle a été réalisée à partir du logiciel Statistica 8.0.

Processus de modélisation

Au cours de cette analyse, il est question d'expliquer un paramètre fondamental qui est la conductivité électrique des eaux souterraines. Ce dernier s'exprime à travers des paramètres physico-chimiques qui sont, de ce fait, les variables explicatives. Selon Grolier et Riss (1997) in Kouadio et al. (2010), il en découle que la relation mathématique qui lie une variable expliquée (la conductivité électrique) à des variables explicatives (paramètres étudiés) est une régression linéaire multiple. C'est le principe d'analyse lorsque, dans une série statistique à p dimensions, une relation est établie entre l'une des variables et l'ensemble des autres variables. L'équation de la conductivité électrique en fonction des paramètres physico-chimiques se présentera comme suit (équation 1):

$$Y = C_1X_1 + C_2X_2 + \dots + C_nX_n + C_0 \quad (1)$$

C_0 : constante ;

Y : variable expliquée ;

X_i : variables explicatives ;

C_i (1 i n): coefficients de pondération des variables X_i .

En effet, Y est un vecteur de valeurs observées de conductivité électrique, X_i est une matrice de variables indépendantes ou explicatives, C_i est un vecteur de paramètres ou coefficients de régression à estimer, et C_0 est un vecteur des

résidus ou des perturbations aléatoires. La régression linéaire estime le vecteur C_i comme la solution des moindres carrés (Benyahya et al., 2007; Daigle et al., 2010):

$$C_i = (X_i^T X_i)^{-1} X_i^T Y \quad (2)$$

avec X^T la transposée de X .

La régression multiple est une variante de la méthode de régression simple qui peut aider à faire face à la colinéarité en choisissant itérativement les variables avec la plus grande valeur explicative. Une régression multiple ascendante commence avec aucune variable, ou un sous-ensemble des variables disponibles, et ajoute la variable la plus significative (celui avec la plus faible valeur de p , associée à des statistiques estimées F) à chaque étape du modèle. Une régression pas à pas descendante commence avec toutes les variables disponibles et supprime la variable la moins importante à chaque étape. C'est la première forme de régression multiple qui a été utilisée.

L'estimation des paramètres d'un modèle constitue l'une des étapes les plus pertinentes de son développement. Selon Hingray et al. (2012), les méthodes d'estimation sont multiples ; ceci résulte du fait que la nature des paramètres est variée et qu'il n'y a aucune méthode d'estimation globalement satisfaisante. L'estimation implique de faire des choix pour établir une stratégie de calibration. Les stratégies sont bien entendu fonctions de l'objectif visé par le modèle en tenant compte de sa pertinence et de sa performance. Le calage des modèles s'est appuyé sur le principe du «split-sample test» qui consiste à faire le calage sur les deux tiers (2/3) de l'échantillon des données disponibles (Koffi et al., 2007; Kouassi et al., 2012b). L'estimation des coefficients de pondération des variables retenues a été réalisée par calage automatique effectué avec le logiciel Excel 2013. Elle a consisté à ajuster les valeurs numériques attribuées aux paramètres des modèles pour reproduire au mieux la réponse observée. Dans la présentation des résultats du calage, il est important d'associer à chacun des coefficients de régression, l'erreur-type associée, qui est un indicateur qui peut être assimilé à l'écart-type. En fait, l'erreur-type est au coefficient de régression ce que l'écart-type est à la moyenne d'une variable. Elle consiste donc une mesure de la variabilité du coefficient de régression. Ainsi, si plusieurs régressions étaient effectuées sur autant de sous-échantillons tirés d'un même échantillon principal, la valeur des paramètres de régression ainsi obtenus différerait d'une fois à l'autre. La «robustesse» d'un coefficient donné sera d'autant plus grande que sa variation autour de la valeur la plus probable, soit le coefficient lui-même, est faible. C'est précisément ce que mesure l'erreur-type du coefficient.

RÉSULTATS ET DISCUSSIONS

Résultats de l'identification des variables explicatives par l'ACP

Les résultats de l'Analyse en composantes principales normées (ACP) sont présentés sous forme de tableaux ou graphiques permettant de synthétiser les informations et de faciliter les interprétations. La liaison existant entre toutes les variables prises deux à deux et les coefficients de corrélation entre ces différentes variables sont donnés par la matrice de corrélation (Tableau 2). En effet, la conductivité électrique présente de très forte corrélation avec la somme des cations (0,910) et la somme des anions (0,911). Ce qui traduit une influence majeure de la somme des cations et celle des anions (variables explicatives) sur la conductivité électrique qui est la variable expliquée. Des courbes de régression linéaire simple ont été utilisées pour apprécier la qualité des corrélations entre la conductivité électrique et les ions majeurs (Figures 4 et 5). En outre, les valeurs du coefficient de corrélation entre la conductivité électrique et les paramètres tels que le pH, la turbidité et la température demeurent faibles et sont respectivement de $R=0,248$, $R=-0,172$ et $R=-0,121$. Ces faibles valeurs du coefficient de corrélation montrent que ces paramètres n'interviennent pas dans la détermination de la conductivité électrique des eaux.

Tableau 2. Matrice de corrélation des variables physico-chimiques

	T	pH	Turbidité	CE	Cations majeurs	Anions majeurs
T	1					
pH	0	1				
Turbidité	0,106	-0,056	1			
CE	-0,121	0,270	-0,172	1		
Cations majeurs	-0,115	0,248	-0,127	0,910	1	
Anions majeurs	-0,027	0,360	-0,145	0,911	0,873	1

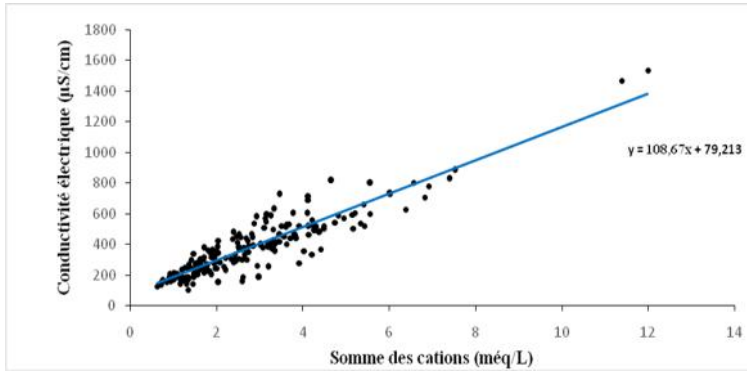


Figure 4: Conductivité électrique en fonction de la somme des cations majeurs

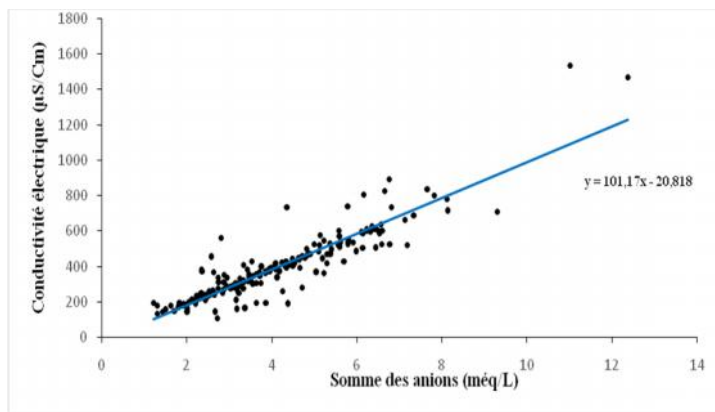


Figure 5: Conductivité électrique en fonction de la somme des anions majeurs

Le tableau 3 présente les valeurs propres, les variances exprimées pour chaque facteur et leur cumul. Le facteur F1, avec une variance exprimée de 49,59%, est le facteur le plus important. Ensuite, viennent les facteurs F2, F3 et F4 avec respectivement 17,91%, 15,09% et 14,25% de la variance exprimée. Les facteurs retenus sont ceux dont les valeurs propres sont supérieures ou égales à 1. A cela s'ajoute le critère selon lequel l'ensemble des facteurs retenus doit avoir une variance supérieure ou égale à 70%. Au regard de ces différentes conditions, les trois premiers facteurs peuvent être retenus. En effet, les valeurs propres de ces facteurs varient entre 0,905 (F3) et 2,975 (F1). Aussi, la variance cumulée est estimée à 82,57%. Les axes factoriels retenus pour cette analyse statistique sont représentatifs de la variance de l'ensemble des données. Les plans factoriels F1-F2 et F1-F3 représentent donc respectivement des variances cumulées égales à 67,50% et 64,67%. Au vu de ces pourcentages exprimés, il peut être considéré que les mécanismes qui contrôlent la variation de la

conductivité électrique des eaux souterraines sont largement contenus dans les trois premiers facteurs. De ce fait, l'analyse s'est portée sur les plans factoriels F1-F2 et F1-F3 qui mettent en évidence les tendances générales.

Tableau 3 : Matrice des valeurs propres et des variances

Facteurs	Valeur propre	Variance (%)	Cumul valeur propre (%)	Cumul variance (%)
Facteur 1	2,975	49,59	2,975	49,59
Facteur 2	1,075	17,91	4,05	67,5
Facteur 3	0,905	15,08	4,954	82,57

La figure 6 montre les résultats de l'ACPN dans le plan factoriel F1-F2. Le facteur F1 est défini par un regroupement situé dans sa partie négative et constitué par la conductivité électrique (CE), la somme des cations majeurs (SC) et la somme des anions majeurs (SA). Le facteur 2 est marqué par un regroupement composé de la température et de la turbidité dans sa partie positive. L'analyse dans le plan factoriel F1-F3 (Figure 7) montre le même regroupement pour le plan factoriel F1. Le facteur F3 est défini par deux regroupements. Le premier regroupement, composé de la température et du pH, est situé dans sa partie positive. Le deuxième regroupement est constitué par la turbidité et est situé dans sa partie négative.

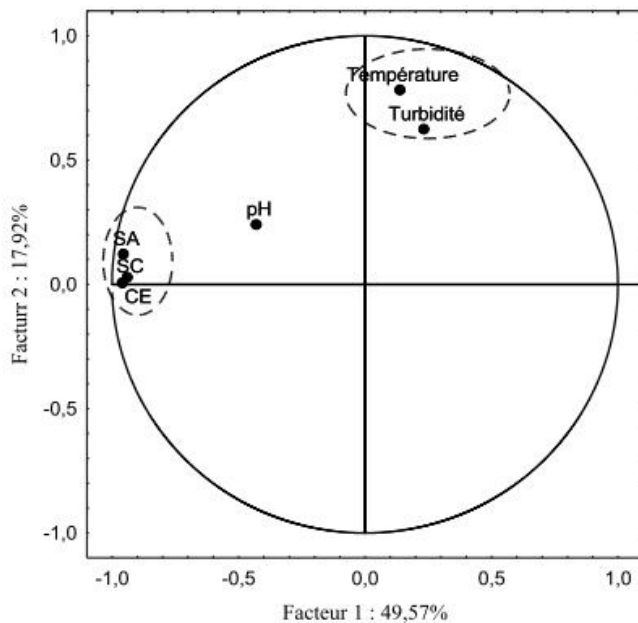


Figure 6: Espace des variables statistiques du plan F1-F2

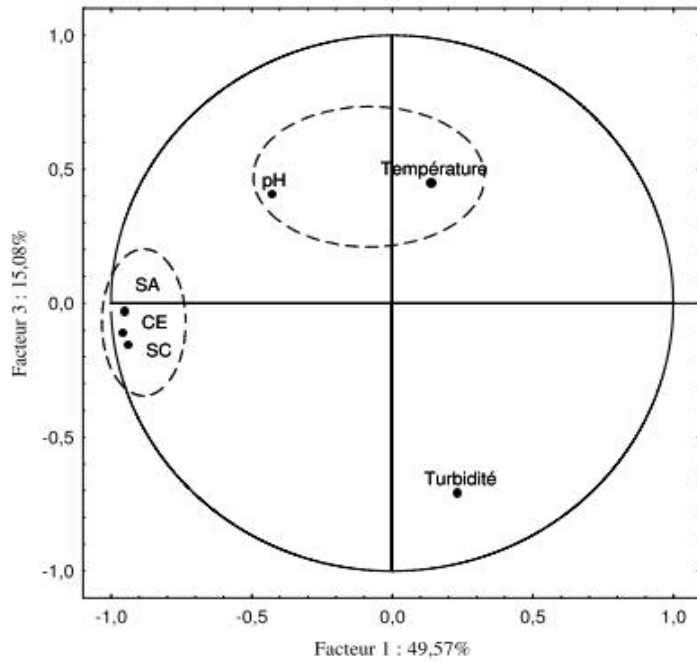


Figure 7: Espace des variables statistiques dans le plan F1-F3

Les résultats de l'ACP indiquent que les paramètres physico-chimiques étudiés n'influencent pas tous la conductivité électrique. Ceux qui influencent le plus la conductivité électrique sont la somme des cations majeurs (SC) et la somme des anions majeurs (SA). En effet, la conductivité électrique est un indicateur de la minéralisation globale (Zouhri, 2003). Il en ressort que la conductivité électrique est influencée par les paramètres chimiques de l'eau qui se résument dans ce cas précis en la somme des cations majeurs et la somme des anions majeurs. Le rapprochement du pH de la conductivité électrique dans le plan factoriel F1-F2, montre que la conductivité électrique est influencée par le pH. Cependant, l'analyse de la projection des variables dans le plan factoriel F1-F3 élimine complètement le pH des variables pouvant influencer majoritairement la conductivité électrique (figure 7). Ce qui explique la faible corrélation ($R=0,270$) entre ces deux paramètres. En outre, les résultats mettent en évidence aucune relation significative entre d'une part, la conductivité électrique et d'autre part, la température et la turbidité. Au regard de tout ce qui précède, deux modèles ont été proposés. Le premier a pour variables explicatives les ions majeurs (somme des cations majeurs et somme des anions majeurs). Pour mettre en évidence l'effet des activités anthropiques ou de pollution sur la qualité des eaux souterraines, un deuxième modèle a été envisagé par ajout de la température et du pH aux deux premières variables explicatives.

Résultats de calage des modèles

Le tableau 4 donne les coefficients de la régression linéaire multiple et les erreurs-types sur les différents coefficients du premier modèle. La somme des cations majeurs et celle des anions majeurs ont des coefficients de valeurs positives qui sont respectivement de 63,614 et de 51,331. Quant à la constante, elle est négative (-2,947). Les erreurs-types associées aux différentes variables explicatives restent relativement faibles (5,108%-5,399%). L'erreur-type associée à la constante (10,424%) est presque le double de celles des variables explicatives. Les valeurs du coefficient de variation (rapport de l'erreur-type au coefficient de régression) sur les coefficients de régression des variables sont respectivement de 8,487% pour la somme des cations et 9,951% pour la somme des anions. Ces résultats traduisent une homogénéité des coefficients de régression relatifs à la somme des cations et à la somme des anions. Ces résultats montrent l'importance des différentes variables dans la modélisation de la conductivité électrique.

Tableau 4 : Coefficients de régression et erreurs-types du modèle 1

Variables	Coefficients	Erreurs-types (%)	Coefficients de variation (%)
Constante	-2,947	10,424	353,715
Somme des cations (SC)	63,614	5,399	8,487
Somme des anions (SA)	51,331	5,108	9,951

L'équation du modèle 1 est la suivante (équation 3):

$$CE = 63,614 \times \sum_{i=1}^n \text{cations} + 51,331 \times \sum_{i=1}^m \text{anions} - 2,947 \quad (3)$$

CE: conductivité électrique simulée ($\mu\text{S}/\text{cm}$) ;

cations: somme des cations majeurs (még/l);

anions: somme des anions majeurs (még/l) ;

n: nombre de cations majeurs ;

m: nombre d'anions majeurs.

Les coefficients de régression obtenus suite à l'établissement du modèle 2 et les erreurs-types sur ceux-ci sont consignés dans le tableau 5. Les coefficients de régression relatifs à la somme des cations majeurs (60,337) et ceux des anions majeurs (58,345) varient dans le même ordre de grandeur et sont positifs. La température a un coefficient de régression positif inférieur à l'unité (0,953); cela explique une faible influence de la température sur la conductivité électrique. En outre, le pH affiche un coefficient négatif (-51,866) impliquant une diminution de la conductivité avec l'augmentation du pH. La constante a une

valeur relativement très élevée (299). Les erreurs-types associées aux paramètres chimiques sont dans le même ordre et se rapproches de celles obtenues au niveau du modèle 1. Les erreurs-types des paramètres physiques (température et pH) sont respectivement de 4,870% et 12,392%. L'erreur-type de la constante est très élevée (155,027%), ce qui démontre une instabilité de la constante au sein de la série. Les valeurs du coefficient de variation (C.V.) sur les coefficients de régression des variables explicatives fluctuent entre 8,829 et 511%. La somme des cations et la somme des anions ont des valeurs de coefficient de variation relativement faibles (inférieures à 10%). Ces résultats traduisent une homogénéité des coefficients de régression relatifs aux sommes des cations et des anions et à un degré moindre le pH. La variable la plus hétérogène est la température (511,018%). Il ressort donc que les meilleurs indicateurs de la variation de la conductivité électrique sont la somme des cations majeurs et la somme des anions majeurs.

Tableau 5 : Coefficients de régression et erreurs-types du modèle 2

Variables	Coefficients	Erreur-type (%)	Coefficients de variation (%)
Constante	298,950	155,027	51,857
Somme des cations (SC)	60,337	5,327	8,829
Somme des anions (SA)	58,345	5,224	8,954
Température	0,953	4,870	511,018
pH	-51,866	12,392	23,892

L'équation du modèle 2 est la suivante (équation 4):

$$CE = 60,337 \times \sum_{i=1}^n cations + 58,345 \times \sum_{i=1}^m anions + 0,953 \times T - 51,866 \times pH + 298,950 \tag{4}$$

- CE : conductivité électrique simulée (μS/cm) ;
- cations : somme des cations majeurs (méq/l);
- anions : somme des anions majeurs (méq/l) ;
- n: nombre de cations majeurs ;
- m : nombre d'anions majeurs ;
- T : température de l'eau (°C) ;
- pH : pH de l'eau.

CONCLUSION

Les résultats de l'ACPN ont abouti à l'identification des variables pertinentes et expressives pour la modélisation de la conductivité électrique des eaux souterraines. Ces variables sont la somme des cations majeurs et la somme des anions majeurs. En effet, celles-ci présentent une forte corrélation (0,91) avec la conductivité électrique et une bonne homogénéité et appartiennent au même regroupement factoriel dans le cercle de communauté. Pour mettre en évidence les impacts des activités anthropiques sur la qualité des eaux souterraines, un deuxième modèle a été développé par intégration de la température et du pH comme variables explicatives en plus des ions majeurs. Les erreurs-types associées aux coefficients de régression des différents modèles et les coefficients de variation ont été évalués. Les erreurs-types varient entre 5 et 10 pour le premier modèle et oscillent entre 4 et 12 pour les variables explicatives en ce qui concerne le modèle 2. Les variables explicatives ont en général des erreurs-types plus faibles que celles de la constante quel que soit le modèle. Les valeurs du coefficient de variation (C.V.) sur les coefficients de régression des variables explicatives sont relativement faibles (inférieures à 10%) en ce qui concerne la somme des cations majeurs et la somme des anions majeurs. Il ressort donc ces variables constituent les meilleurs indicateurs de la variation de la conductivité électrique des eaux souterraines. Elles sont suivies par le pH puis la température. Les résultats obtenus montrent l'importance des différentes variables explicatives dans la modélisation de la conductivité électrique. Ces différents résultats traduisent des relations étroites entre la variable expliquée (conductivité électrique) et les variables explicatives (somme des cations, somme des anions, pH, température).

REFERENCES BIBLIOGRAPHIQUES

- ACHOUR S., GUERGAZI, S. (2002). Incidence de la minéralisation des eaux algériennes sur la réactivité de composés organiques vis-à-vis du chlore, *Revue des Sciences de l'Eau*, Vol. 15, n°3, 641-660.
- AHOUSSE K.E., SORO N., KOUASSI A.M., SORO G., KOFFI Y.B., ZADE S.P. (2010). Application des méthodes d'analyses statistiques multivariées à l'étude de l'origine des métaux lourds (Cu^{2+} , Mn^{2+} , Zn^{2+} et Pb^{2+}) dans les eaux des nappes phréatiques de la ville d'Abidjan, *International Journal of Biological and Chemical Sciences*, Vol. 4, n°5, 1753-1765.
- BELANGER M., EL-JABI N., CAISSIE D., ASHKAR F., RIBI J.M. (2005). Estimation de la température de l'eau en rivière en utilisant les réseaux de neurones et la régression linéaire multiple, *Revue des Sciences de l'Eau*, Vol. 18, n°3, 403-421.
- BENYAHYA L., CAISSIE D., MERTES L.A. K., O'DANIEL S. J., WOESSNER W.W., THOMAS S.A. (2007). A review of statistical water

- temperature models, Canadian Water Resources Journal / Revue canadienne des ressources hydriques, Vol. 32, n°3, 179-193.
- DAIGLE A., ST-HILAIRE A., PETERS D. and BAIRD D. (2010). Multivariate modeling of water temperature in the Okanagan watershed, Canadian Water Resources, Journal-Revue canadienne des ressources hydriques, Vol. 35, n°3, 237-258.
- DAOUDA Y.B. (1998). Lithostratigraphie et pétrographie des formations birimiennes de Toumodi-Fêtékro (Côte d'Ivoire). Implication pour l'évolution crustale du paléoprotérozoïque du craton ouest-africain, Thèse de Doctorat de l'Université d'Orléans, France, 190p.
- FOURNIER M. (2006). Identification du transport et de la vulnérabilité du karst de la craie. Application de nouveaux outils statistiques d'analyse des données au système de Hannebot (Seine-Marine, France), Thèse de Doctorat d'hydrogéologie karstique à l'Université de Rouen, 228p.
- HINGRAY B., PICOUET C., MUSY A. (2012). Hydrologie: une science pour l'ingénieur. Edition Presse Polytechniques et Universitaires, Romandes 2^{ème} édition, 600 p.
- KOFFI Y.B., LASM T., AYRAL P.A., JOHANNET A., KOUASSI A.M., ASSIDJO E., BIEMI J. (2007). Optimisation des modèles Perceptrons Multicouches avec les algorithmes de premier et de deuxième ordre. Application à la modélisation de la relation pluie-débit du Bandama Blanc, Nord de la Côte d'Ivoire, European Journal of Scientific Research, Vol. 17, n°3, 313-328.
- KOUADIO K.E., SORO N., SAVANÉ I. (2010). Stratégie d'optimisation de la profondeur des forages en contexte de socle : Application à la région du Denguélé, Nord-Ouest de la Côte d'Ivoire, Revue des sciences de l'eau, Vol. 23, n°1, 1-15.
- KOUASSI A.M., OKAINGNI J.C., BAKA D., LASM T., KOUAMÉ K.F., BIÉMI J. (2012a). Application des méthodes statistiques et géostatistiques à l'étude de la conductivité électrique des eaux souterraines de la région du N'zi-Comoé (Centre-Est de la Côte d'Ivoire), International Journal of Biological and Chemical Sciences, Vol. 6, n°2, 897-912.
- KOUASSI A.M., N'GUESSAN B.T.M., KOUAME K.F., KOUAME K.A., OKAINGNI J.C., BIEMI J. (2012b). Application de la méthode des simulations croisées à l'analyse de tendances dans les relations pluie-débit à partir du modèle GR2M: cas du bassin versant du N'Zi-Bandama (Côte d'Ivoire), Comptes Rendus Géoscience, Vol. 344, n°5, 288-296.
- KOUASSI A.M., MAMADOU A., AHOUSI K.E., BIEMI J. (2013). Simulation de la conductivité électrique des eaux souterraines en relation avec leurs propriétés géologiques: cas de la Côte d'Ivoire, Revue Ivoirienne des Sciences et Technologie, 21&22, 138-166.
- KOUASSI A.M., YAO K.A., AKE A.Y., BIEMI J. (2011). Application des méthodes d'analyses statistiques multivariées à l'étude de la minéralisation des eaux des aquifères fissurés de la zone Guiglo-Duekoué (Ouest de la Côte

- d'Ivoire), International Journal of Biological and Chemical Sciences, Vol. 5, n°5, 2155-2169.
- LASM T., YAO K.T., OGA M.S., KOUAMÉ K.F., JOURDA P.R., KOUADIO K.E., BAKA D. (2008). Analysis of the physico-chemical characteristics of groundwater in proterozoic land region of the Tiassale area (Southern Côte d'Ivoire), European Journal of Scientific Research, Vol. 20, n°3, 526-543.
- MDIKER N., ACHEB A.E., MANDOUR A., YOUNSI A., MALIKI S.E., OUTEYEB B.K. (2009). Contribution à l'étude de la salinisation de la nappe côtière de sahel El Haouzia région d'El Jadida au Maroc, Afrique Science, Vol. 5, n°2, 232-245.
- TIDJANI A.E.B., YEBDRI D., ROTH J.C., DERRICHE Z. (2006). Exploration des séries chronologiques d'analyse de la qualité des eaux de surface dans le bassin de la Tafna (Algérie), Revue des Sciences de l'Eau, Vol. 19, n°4, 315-324.
- VALE M.M. (2006). Quantification et prédiction de la minéralisation nette de l'azote du sol in situ, sous divers pédoclimats et systèmes de culture français, Thèse de Doctorat de l'Institut National Polytechnique de Toulouse, 209p.
- ZOURHI L. (2003). Étude de la salinité des eaux souterraines dans la partie occidentale du Maroc, bassin du Rharb, Journal of Environmental Hydrology, Vol.11, Paper 9, 1-8.
- ZOURHI L., CARLIER E. (2002). Caractérisation hydrochimique d'une nappe côtière, Maroc, Journal of Environmental Hydrology, Vol. 9, Paper 4, 1-7.