

## Cartographie de l'occupation de sol des Hauts Plateaux de l'Ouest Cameroun par réseaux de neurones appliqués à une image LANDSAT 8 OLI

### [ Mapping land cover in Western Cameroon Highlands using neural networks applied to a Landsat 8 OLI image ]

V.-C. Jofack Sokeng<sup>1</sup>, F. K. Kouamé<sup>1</sup>, H. Dibi N'da<sup>1</sup>, B. Tankoano<sup>2</sup>, L. Akpa You<sup>1</sup>, and B. Ngounou Ngatcha<sup>3</sup>

<sup>1</sup>Centre Universitaire de Recherche et d'Application en Télédétection (CURAT),  
University Félix Houphouët Boigny of Cocody,  
Abidjan, Côte d'Ivoire

<sup>2</sup>Institut du Développement Rural,  
Université Polytechnique de Bobo Dioulasso,  
Bobo Dioulasso, Burkina Faso

<sup>3</sup>Département des Sciences de la Terre,  
Université de Ngaoundéré,  
Ngaoundéré, Cameroon

Copyright © 2016 ISSR Journals. This is an open access article distributed under the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**ABSTRACT:** Mapping land cover in complex regions like Western Cameroon Highlands is highly difficult. Most widely used algorithms are not easily implementable because of strong similarities observed in reflectance of different land cover units. This study evaluates the contribution of neural networks to the classification of LANDSAT 8 OLI images in order to achieve a better land cover map in this region. Image processing techniques (calculation of indices, principal components analysis, and color compositions) and a field survey allowed to discriminate and select trainings and validation sites of the main land cover units. Then, a network with 14 neurons in the input layer and 8 neurons in the output layer corresponding to different land cover classes was constructed. External and internal network parameters were experimentally selected for classification. The resulting map was finally validated with an overall accuracy of 90, 08% and a Kappa equal to 0.88. Eight land cover units have been identified. These are degraded forests, savannahs, bare soil and localities, water, wetlands, crops and burned areas. Finally, comparison with the maximum likelihood method has shown the superiority of neural networks with an overall accuracy difference of around 8%.

**KEYWORDS:** Neural Networks, Back-propagation algorithm, land cover, Landsat OLI, Western Cameroon Highlands.

**RESUME:** La cartographie de l'occupation de sol dans des régions très accidentées comme les Hauts Plateaux de l'Ouest Cameroun est très complexe. Les algorithmes couramment utilisés y sont difficiles à implémenter à cause des fortes ressemblances observées dans les réflectances des différentes unités d'occupation de sol. Cette étude évalue la contribution des réseaux de neurones à la classification des images LANDSAT 8 OLI afin de réaliser une meilleure cartographie de l'occupation de sol de cette région. Les techniques traitements d'images (calcul des indices, analyses en composantes principales, compositions colorées) et une mission de terrains ont permis de discriminer et de sélectionner les parcelles d'entraînements et de contrôles des principales unités d'occupation à cartographier. Ensuite, un réseau avec 14 neurones dans la couche d'entrée et 8 neurones dans la couche de sortie correspondant aux classes d'occupation de sol a été construit. Les paramètres externes et internes des réseaux ont été choisis de façon expérimentale pour la classification. La carte

obtenue a enfin été validé avec une précision globale de 90, 08% et un Kappa de 0,88. Huit principales unités d'occupation de sol ont été mises en évidence. Ce sont forêt dégradée, les savanes, les sols nus et localités, l'eau, les zones humides, les cultures et les surfaces brûlées. Enfin, la comparaison avec la méthode de maximum de vraisemblance a montré la supériorité des réseaux de neurones avec une différence de précision globale autour de 8%.

**MOTS-CLEFS:** Réseaux de neurones, rétropropagation d'erreur, occupation de sol, LANDSAT OLI, Hauts Plateaux de l'Ouest Cameroun.

## **1 INTRODUCTION**

La carte d'occupation du sol est un des outils indispensables pour la planification du développement et la gestion du territoire. Elle fournit des informations quantitatives nécessaires pour une bonne description du territoire. Les classes d'occupation de sol sont d'importantes unités analytiques qui permettent d'établir un lien quantitatif entre les activités humaines, leurs répercussions sur l'environnement, et la dimension géographique [1]. La cartographie de l'occupation du sol constitue un enjeu très important pour beaucoup de travaux de recherche et pour de nombreux projets et applications opérationnelles au niveau mondial. Dans la littérature, les cartes d'occupation du sol sont largement utilisées dans les études diachroniques de la couverture du sol [2] [3] [4], et comme critère très indispensable pour la cartographie des ressources [5] [6] et des risques naturels [7] [8]. Cependant, dans les régions difficiles d'accès (cas des Hauts Plateaux de l'Ouest Cameroun), cartographier l'occupation de sol avec précision est un véritable défi. Elle nécessite l'intégration de nouvelles méthodes et l'utilisation d'outils complémentaires aux méthodes conventionnelles de terrain [9].

Les progrès réalisés dans l'utilisation des données à référence spatiale ont mise en évidence l'utilité de la télédétection comme outil indispensable dans la cartographie de l'occupation du sol. En effet, par comparaison aux méthodes conventionnelles de terrain, elle permet de couvrir de vastes superficies rapidement et à peu de frais [10] [11]. De plus, la diversité des capteurs optiques, radars et des photographies aériennes disponibles, fait de la télédétection, un outil redoutable pour inventorier, analyser et classifier les différentes unités d'occupation du sol [3].

Classifier une image, c'est identifier la classe à laquelle appartient chacun de ses pixels. L'objectif est de créer une carte thématique qui représente un regroupement de classes associé aux différents éléments de l'image analysée. Les techniques de classification d'images satellites ont été largement diversifiées au fil des ans. [12] les regroupe en 2 grandes catégories : (i) les méthodes dites paramétriques (méthode de maximum de vraisemblance et des champs de Markov) qui associent à la signature spectrale, une distribution statistique connue (loi normale ou multi-normale) ; et (ii), les méthodes non paramétriques (le minimum de distance, la méthode parallépipédique, la distance de Mahalanobis, le K-plus proche voisin, l'arbre de décision, le support vaste machine et les réseaux de neurones) qui, contrairement aux premières, n'exploitent aucune distribution statistique, mais plutôt la distance spectrale. De toutes ces méthodes, la méthode de maximum de vraisemblance est la plus populaire et la plus utilisée [13] [14] [6]. Cependant, cette méthode regroupe les pixels sur le seul critère spectral estimé par la probabilité d'appartenance [15]. Par ailleurs, elle suppose que la distribution des valeurs numériques des pixels est normale (gaussienne) pour chaque classe. Toutefois, cette condition n'est pas toujours honorée puisque pour certaines classes, la distribution peut être asymétrique, bimodale ou même aléatoire surtout quand il s'agit de la classification d'un milieu naturel. De même, il est très difficile d'appliquer ces méthodes dans des environnements complexes sur le plan spectral où plusieurs facteurs (par exemple, des compositions de terrain, les irrégularités du terrain, des mosaïques dans les unités d'occupation de sol) affectent les signatures spectrales. Dans de tels environnements, très commun aux Hauts Plateaux de l'Ouest Cameroun, la production de la carte d'occupation de sol basée sur le seul critère spectral est difficile [16]. L'exploitation d'autres sources d'information devient obligatoire [9] [17] [18].

Plusieurs études [16] [19] [10] [20] [21] [11] utilisent les réseaux de neurones pour faciliter l'intégration des données environnementale (morphologiques, géologiques, données climatologiques) dans la classification d'occupation de sol. Ces réseaux, inspirés du système biologique, est un ensemble de neurones interconnectés et associés en couches. Ils sont formés d'une couche d'entrée, d'une couche cachée et d'une couche de sortie. Ils ont pour objectif, de prédire les sorties d'un système ou le comportement d'un phénomène à partir des entrées connues et définies par l'utilisateur. Contrairement aux méthodes précédentes, ces modèles n'exigent pas que la signature spectrale s'ajuste à une distribution paramétrable. Ils intègrent des informations de différentes sources et s'adaptent progressivement aux propriétés de la signature au fur et à mesure qu'on leur fournit des exemples connus [9] [17].

Cette étude applique des réseaux de neurones pour une meilleure cartographie de l'occupation de sol dans les Hauts Plateaux de l'Ouest du Cameroun. Elle présente le principe, la méthodologie et les avantages d'application de l'algorithme du perceptron multicouche à rétropropagation d'erreur pour la classification d'une image LANDSAT 8 OLI, et compare les résultats obtenus à la méthode de maximum de vraisemblance.

## 2 MATÉRIEL ET MÉTHODES

### 2.1 SITE D'ÉTUDE ET MATÉRIEL UTILISÉ

Située entre 09°50' et 10°50' de longitude Est et 5°25' et 6°15' de latitude Nord, le secteur étudié appartient aux Hauts Plateaux de l'Ouest – Cameroun (HPOC). Il a une superficie de 9000 Km<sup>2</sup> et densité de population de 128,5 habitants au Km<sup>2</sup>. Son relief est très accidenté avec des plateaux et une série de massifs (mont Léfo - 2550 m, Bamboutos - 2700 m, Mbam - 2335 m...) et Bandja. La végétation rencontrée est une savane péri-forestière arborée ou arbustive en mosaïque avec des massifs forestiers et de petites galeries forestières de plantes hydrophiles et hydrophobes. Les sols très fertiles, entraîne le développement des cultures annuelles et pérennes quel que soit le milieu. L'hydrographie, marquée de nombreuses chutes, cascades et drains rectilignes, favorise une constante humidité du sol, et explique la persistance quasi-annuelle du couvert végétal. Le climat est de type tropical humide d'altitude, avec une pluviométrie moyenne annuelle de 1800 mm.

Pour réaliser la carte d'occupation de sol de cette région, deux scènes (186/056 et 187/056) de LANDSAT 8 OLI ont été téléchargées, mosaïquées et exploitées. Ce sont des images acquises le 02 Février 2014 (saison sèche) avec 11 bandes spectrales, une résolution spatiale moyenne de 30 m et qui discriminent les principales grandes unités d'occupation de sol du site d'étude (**figure 1**). Des données SRTM de février 2000 ont permis d'extraire les indicateurs de terrain (altitude et pentes) jugés pertinents et en relation avec l'occupation de sol. Les types de sols ont été extraits des cartes pédologiques (feuilles Douala Est et Ouest) disponibles au 1 / 200 000. Toutes ces données ont été géo-référencées et corrigées géométriquement sur la base des cartes topographiques existantes.

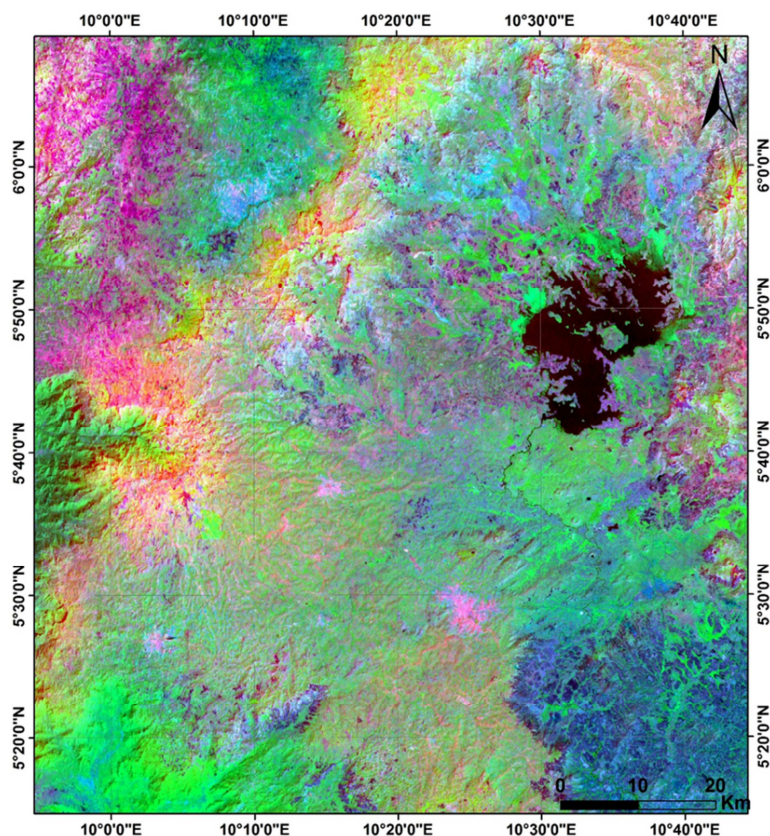


Fig. 1. Composition colorée OLI 4-5-7 rehaussant les différentes unités d'occupation de sol dans le site d'étude

## 2.2 MÉTHODOLOGIE

La réalisation de la carte d'occupation de sol des HPOC s'est déroulée en 3 étapes : (i) la discrimination des différents types d'occupation de sol à l'aide des traitements d'images et des missions de terrains ; (ii) la classification dirigée avec l'algorithme de réseaux de neurones multicouche par rétro-propagation d'erreurs ; (iii) la validation de la carte d'occupation de sol et l'évaluation de sa précision.

### 2.2.1 DISCRIMINATION DES TYPES D'OCCUPATION DE SOL

Cette phase a permis de produire des images satellites en composition colorées qui facilitent une bonne discrimination spectrale des types d'occupation de sol. Ainsi, ont été combinées plusieurs techniques de traitements notamment le calcul d'indices (NDVI, indice d'humidité, l'indice de brillance des sols), et les Analyse en Composante Principale Sélectives (ACPS). L'analyse des bandes de LANDSAT OLI et des néo-canaux issus de traitements d'images a permis de rehausser le contraste entre les bandes et de mettre en évidence les types d'occupation de sol que sont la forêt dégradée, les savanes, les sols nus et localités, l'eau, les zones humides, les cultures et les brulis. L'interprétation de ces images fausses couleurs a été fondée sur la teinte, la forme, et l'agencement spatial des objets avec leur voisinage. Les unités d'occupation de sol discriminées sur ces compositions colorées ont été par la suite identifiées et décrites sur le terrain. 450 points GPS collectés sur les différentes unités ont été utilisés dans la délimitation des sites d'entraînements (70%) et de contrôle (30%).

### 2.2.2 CLASSIFICATION DIRIGEE PAR L'ALGORITHME DE RESEAUX DE NEURONES MULTICOUCHE A RETRO-PROPAGATION D'ERREURS

#### PRINCIPE DES RESEAUX DE NEURONES ARTIFICIELS

Le neurone artificiel s'identifie au neurone humain et est défini par des entrées ( $E_i$  et  $S_j$ ), des poids de connexion, une fonction de transfert et une sortie. Le signal  $E_i$ , se trouvant à l'entrée du neurone  $j$  est multiplié par le poids  $W_{ij}$  qui représente le poids de connexion du neurone  $j$  vers le neurone  $i$ . La somme pondérée  $S_j$  des signaux d'entrées  $E_i$  ( $1 \leq i \leq n$ ) du neurone  $i$ , correspond aux signaux de sorties des neurones de la couche amont auxquels ce neurone est connecté ; la fonction de transfert  $f$  convertit la somme pondérée des entrées  $S_j$  en sortie  $S$  (équation 1).

$$S = f \left( \sum_i^n W_{ij} \cdot E(i) \right) \quad (1)$$

Dans un réseau de neurones, les neurones sont organisés en couches successives (figure 2). Chaque neurone d'une couche reçoit des signaux de la couche antérieure et transmet le résultat aux neurones de la couche postérieure en suivant un sens de propagation de l'entrée vers la sortie [22]. Avec cette configuration, la couche d'entrée reçoit les variables d'entrée et la couche de sortie fournit les résultats [23]. Les autres couches qui se trouvent entre la couche d'entrée et la couche de sortie sont appelées couches cachées. Le nombre de niveaux cachés et le nombre de neurones par niveaux représentent les paramètres architecturaux du réseau.

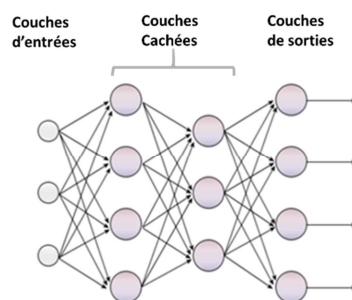


Fig. 2. Un réseau multicouche

Les réseaux de neurones sont des outils de modélisation numérique dont l'objectif est de prédire les sorties d'un système ou le comportement d'un phénomène à partir des entrées connues et définies par l'utilisateur [10]. Cette prédiction est réalisée en construisant au cours d'une phase d'apprentissage, un modèle non linéaire entre des couples entrées et sorties. Sachant que les poids précisent le lien appartenant à deux niveaux successifs, leurs valeurs sont ajustées et affinées

continuellement tout au long de la phase d'apprentissage. Pendant cette phase, des couples entrées – sorties sont fournis au réseau. Dans un premier temps, les poids sont fixés aléatoirement pour permettre au réseau de calculer ses propres sorties à partir des entrées déjà fournies [24]. Puis, ils sont alors corrigés de manière à minimiser la différence entre les sorties ainsi calculées et les sorties désirées. Cette phase de minimisation correspond à l'apprentissage ; elle est primordiale à l'efficacité du réseau [19].

Plusieurs algorithmes d'apprentissage existent. Pour des perspectives de cartographie thématique, l'algorithme d'apprentissage le plus utilisé en télédétection est la rétro propagation du gradient d'erreur [23]. Il est très adapté aux réseaux de neurones multicouches. Il permet d'estimer la correction à faire pour un poids  $W_{ij}$  défini entre deux neurones  $i$  et  $j$ . Cette correction permet de minimiser l'erreur au carré  $E_p$  entre la sortie calculée par le réseau  $S_{c,j}$  et la sortie désirée  $S_{d,j}$  du neurone  $j$ . L'erreur  $E_p$  est calculée pour chaque pixel ou valeur  $p$  appartenant au groupe d'apprentissage (équation 2).

$$E_p = \frac{1}{2} \sum_j (S_{d,j} - S_{c,j})^2 \quad (2)$$

Pour minimiser l'erreur au cours de la phase d'apprentissage, la correction du poids est proportionnelle à la dérivée de l'erreur  $E_p$  par rapport au poids (équation 3).

$$\Delta W_{ij}(n + 1) = [\eta * \delta_j * S_i] + [\alpha * \Delta W_{ij}(n)] \text{ avec } \Delta W_{ij} \approx -\frac{\partial E_p}{\partial W_{ij}} \quad (3)$$

- $\delta_j = (S_{d,j} - S_{c,j}) * f'(E_j)$  si (j) appartient au niveau de sortie du réseau
- $\delta_j = f'(E_j) * \sum_k \delta_k * W_k$  si (j) appartient à un niveau caché du réseau
- $f'(E_j)$  est la dérivée de la fonction d'activation  $f$  par rapport à l'entrée  $E_j$  du neurone  $j$ .
- $\eta$  est le coefficient de proportionnalité qui représente le taux d'apprentissage
- $n$ : indicateur du nombre d'itérations de la phase d'apprentissage
- $\alpha$ : une constante appelée momentum qui reflète l'effet du changement du poids dans l'itération précédente.

Cette équation représente la règle généralisée du delta. Cette règle permet de calculer pour chaque vecteur de pixel, la correction  $\Delta w_{ij}$  à apporter au poids  $w_{ij}$  entre deux neurones  $i$  et  $j$ . Le principe de base de la procédure utilisée pour l'apprentissage nécessite des corrections faites par de petites valeurs afin de garder une bonne décroissance de la fonction de minimisation. En effet, la réussite de l'apprentissage est proportionnelle à la valeur choisie pour le taux d'apprentissage  $\eta$ . Les valeurs du taux d'apprentissage  $\eta$  et du momentum  $\alpha$  sont spécifiques pour chaque application.

De façon générale, l'utilisation des réseaux de neurones en classification d'image nécessite la précision d'un certain nombre de paramètres reliés à l'architecture (le nombre de couches cachées, le nombre de neurones par couches, et la fonction de transfert) et à la phase d'apprentissage (le taux d'apprentissage, le momentum, et le nombre d'itérations).

#### CARTOGRAPHIE DE L'OCCUPATION DE SOL PAR RÉSEAUX DE NEURONES

Les réseaux de neurones utilisés dans cette étude, ont été implémentés dans **ENVI**. Ce logiciel de traitement d'images facilite plusieurs tâches liées à la création, l'entraînement et la simulation des réseaux de neurones pour des fins de classification d'images. Plusieurs modifications ont été apportées au programme source de ce module (**Neural Net**) afin de l'adapter aux caractéristiques spécifiques de cette étude.

Tout d'abord, l'objectif a été de définir le nombre de neurones à allouer à la couche d'entrée, la couche de sortie et aux couches cachées. A cet effet, 14 neurones ont été attribués à la couche d'entrée. Ils correspondent à 6 bandes de LANDSAT OLI (OLI 2 à 7), aux néo-canaux issus des traitements d'images (NDVI, IB, NDWI), aux composantes principales CP<sub>123</sub> de l'ACP<sub>123</sub> et CP<sub>167</sub> de l'ACP<sub>67</sub>, aux type de sol et aux paramètres morphométriques de terrain (Altitude et pente). L'utilisation d'un seul neurone pour chaque type d'information [9] réduit considérablement la taille du réseau et accélère le processus d'apprentissage et de classification. Ensuite, *Huit* (8) neurones ont été définis pour la couche de sortie. Ils représentent les 8 unités d'occupation de sol identifiés sur le terrain à savoir, les forêts, les savanes, les cultures, les zones brûlées, l'eau, les zones humides, les sols nus et les zones habitées. 29 neurones ont été affectés aux couches cachées sur la base du théorème de Kolmogorov [25]. Le réseau ainsi construit, a été entraîné par l'algorithme de rétropropagation d'erreur. Cet algorithme a permis d'ajuster et d'affiner les valeurs des poids tout au long de l'apprentissage. La correction des poids au cours de l'entraînement ne concerne que les pixels appartenant au groupe d'apprentissage. Au cours de cette phase, les poids du

réseau sont corrigés de manière à minimiser l'erreur au carré entre la réponse calculée par le réseau et la réponse attendue. Généralement, cette erreur calculée sur le groupe d'apprentissage diminue continuellement au cours de l'entraînement [9].

Plusieurs paramètres ont été définis au cours de cet apprentissage. Ces paramètres sont liés au fonctionnement interne de l'algorithme de rétropropagation. Ce sont entre autre : le nombre de couches cachées, le taux d'apprentissage, le momentum, la fonction d'activation, le seuil d'arrêt et le nombre d'itérations. Le **taux d'apprentissage** permet de calculer les changements à apporter aux valeurs de poids après chaque itération. Sa valeur affecte considérablement la performance de l'algorithme d'apprentissage. Il varie entre 0 et 1. Le **momentum** a été ajouté à la règle de Delta [26] pour éviter tout risque d'oscillation de l'erreur pendant l'apprentissage. Il varie aussi de 0 à 1 et mesure l'effet d'une itération antérieure sur l'itération en cours [19]. La **fonction d'activation** souvent sigmoïde ou tangente hyperbolique, normalise les valeurs attribuées aux neurones de sorties. Elle intervient entre le niveau d'entrée et le niveau caché d'une part, et entre le niveau caché et le niveau de sortie d'autre part. D'après [19] l'apprentissage du réseau s'arrête lorsque l'erreur dépasse une valeur optimale établie a priori (**seuil d'arrêt**) ou lorsque le **nombre d'itération** est atteint. Le **seuil d'arrêt** varie de 0 à 1 et le **nombre d'itérations** est défini par l'utilisateur. L'influence de ces paramètres sur la performance de la classification a aussi été étudiée de façon expérimentale en comparant des valeurs faibles à des valeurs fortes. Cela a permis de choisir la configuration idéale du réseau (**figure 3**) pour la cartographie de l'occupation du sol de la région d'étude.

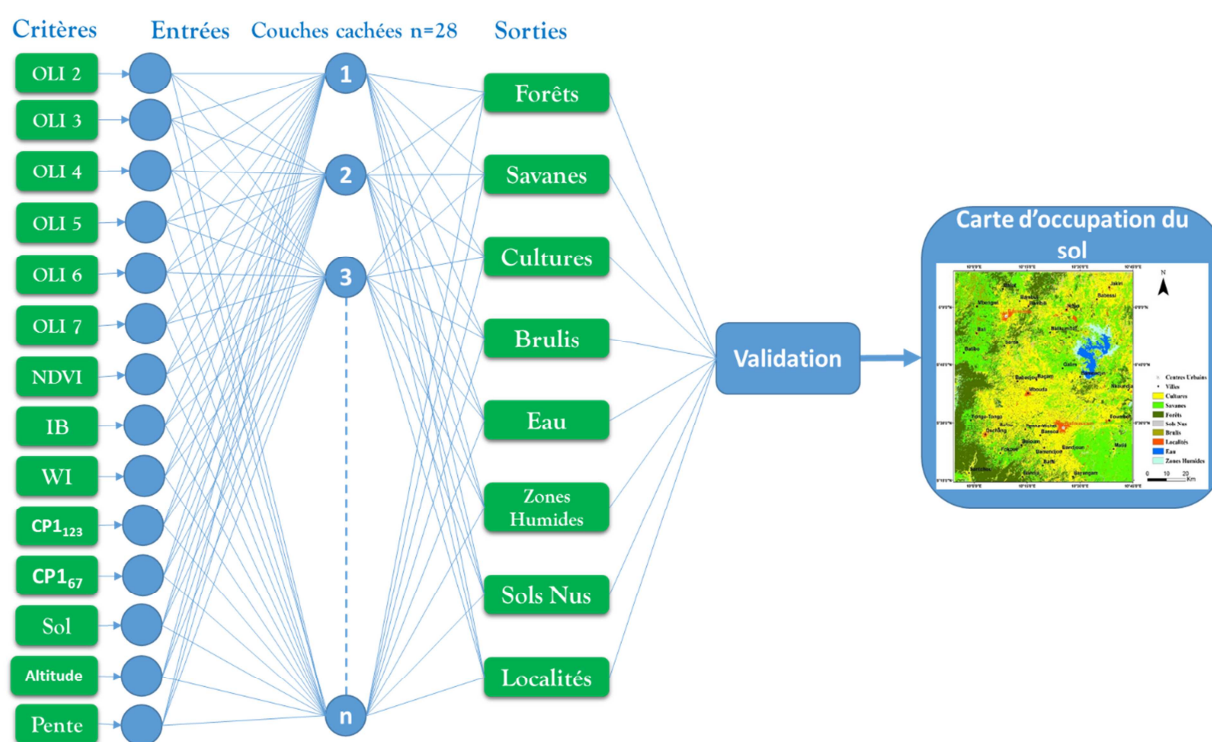


Fig. 3. Architecture du réseau construit pour la cartographie de l'occupation de sol

### 2.3 MESURE DE LA PERFORMANCE DU RÉSEAU ET VALIDATION

L'évaluation de la performance du réseau à la classification d'images s'est réalisée deux étapes. La première étape a consisté à établir la matrice d'erreur et à calculer la précision globale ainsi que l'indice Kappa de la classification neuronale obtenue sur la base des données de contrôle. Ces données sont constituées d'un ensemble de pixels n'ayant pas servi à l'apprentissage. Cette étape est appliquée pour chaque réseau construit et permet de choisir le modèle et les paramètres d'apprentissage offrant la meilleure précision de classification. La deuxième étape a consisté à mettre en évidence l'apport du réseau implémenté en comparant ses performances à celles de la méthode de maximum de vraisemblance, plus souvent utilisée. À l'issue de cette validation, la carte d'occupation du sol des Hauts Plateaux de l'Ouest du Cameroun a été obtenue.

### 3 RÉSULTATS

#### 3.1 INFLUENCE DES PARAMÈTRES DU RÉSEAU

Un réseau multicouche a été construit avec 14 neurones dans la couche d'entrée, et 8 correspondant aux unités d'occupation de sol dans la couche de sortie. Au niveau des couches cachées, 28 neurones ont été défini sur la base du théorème de Kolmogorov. Ces neurones ont été répartis en plusieurs couches cachées. Ce modèle perceptron a été par la suite entraîné par l'algorithme de rétropropagation d'erreur qui dépend de plusieurs paramètres. Différentes valeurs de ces paramètres ont aussi été étudiées. Une fois le réseau entraîné, ses performances ont été évaluées sur la base des sites de contrôle définis sur le terrain. Le **tableau 1** montre l'influence de ces paramètres internes (taux d'apprentissage, le momentum, le nombre d'itération, de seuil d'arrêt) et externe (nombre de couches cachées) sur la précision globale du réseau et de la classification.

**Tableau 1. Influence des paramètres du réseau sur la précision de la classification. CA: Couches Cachées; TA: Taux d'Apprentissage; MO: Momentum; NI: Nombre d'Itérations; AS: Seuil d'arrêt; TF: Fonction de Transfert**

Essai	CA	TA	MO	NI	SA	TF	Précision
1	1	0,01	0,9	1000	0,1	sigmoïde	89,26
2	2	0,01	0,9	1000	0,1	sigmoïde	79,87
3	3	0,01	0,9	1000	0,1	sigmoïde	74,22
4	4	0,01	0,9	1000	0,1	sigmoïde	64,21
5	1	0,05	0,9	1000	0,1	sigmoïde	81,67
6	1	0,1	0,9	1000	0,1	sigmoïde	76,33
7	1	0,15	0,9	1000	0,1	sigmoïde	68,01
8	1	0,3	0,9	1000	0,1	sigmoïde	53,38
9	1	0,5	0,9	1000	0,1	sigmoïde	45,16
10	1	0,8	0,9	1000	0,1	sigmoïde	31,58
11	1	1	0,9	1000	0,1	sigmoïde	17,33
12	1	0,01	0,1	1000	0,1	sigmoïde	65,14
13	1	0,01	0,3	1000	0,1	sigmoïde	71,67
14	1	0,01	0,5	1000	0,1	sigmoïde	74,21
15	1	0,01	0,7	1000	0,1	sigmoïde	78,62
16	1	0,01	0,9	500	0,1	sigmoïde	78,02
17	1	0,01	0,9	800	0,1	sigmoïde	80,00
18	1	0,01	0,9	1500	0,1	sigmoïde	82,16
19	1	0,01	0,9	2000	0,1	sigmoïde	79,97
20	1	0,01	0,9	4000	0,1	sigmoïde	61,63
21	1	0,01	0,9	1000	0,2	sigmoïde	85,11
22	1	0,01	0,9	1000	0,3	sigmoïde	81,97
23	1	0,01	0,9	1000	0,4	sigmoïde	88,05
24	1	0,01	0,9	1000	0,5	sigmoïde	81,89
25	1	0,01	0,9	1000	0,1	tangente	61,99

D'après ce tableau, il est clair que les précisions de classification dépendent de l'architecture des modèles neuronaux et des paramètres choisis à l'apprentissage. Le nombre de couches cachées a un effet minime sur la performance du réseau car ne fait varier la performance que 3 à 4%. L'augmentation du nombre de couche cachée (de 1 à 4) n'améliore pas la précision de la classification, mais augmente plutôt le temps de calcul. De même, une augmentation (de 0 vers 1) du taux d'apprentissage diminue de 82,76 à 61,97%, la précision de la classification en entraînant des oscillations de l'erreur pendant l'apprentissage. Cependant, si l'on augmente plutôt le momentum de 0 vers 1 et le nombre d'itération de 500 vers 4000, cela a un effet positif sur la précision de classification. Mais avec valeurs très fortes (4000), la précision diminue. En effet, lorsque la phase d'entraînement devient trop longue, le réseau peu perdre son pouvoir de généralisation. Enfin, avec des valeurs

identiques de ces paramètres, une meilleure précision a été notée pour des faibles seuils d'arrêt et une fonction de transfert de type sigmoïde.

### 3.2 CARTE D'OCCUPATION DU SOL DU SECTEUR D'ÉTUDE

Pour la suite des travaux, 1000 itérations, un momentum de 0,88, un seuil d'arrêt de 0,15, un taux d'apprentissage fixé à 0,03, une fonction de transfert sigmoïde et une seule couche cachée ont été choisies afin d'obtenir la carte d'occupation de sol (**figure 4**) du site étudié dans les Hauts Plateaux de l'Ouest du Cameroun. L'analyse statistique de la carte obtenue montre que la zone étudiée est dominée par les cultures (24%) et les savanes (16%). Les forêts n'occupent que 11%. La forte proportion (13%) occupée par les localités atteste de la densité de population de la zone étudiée.

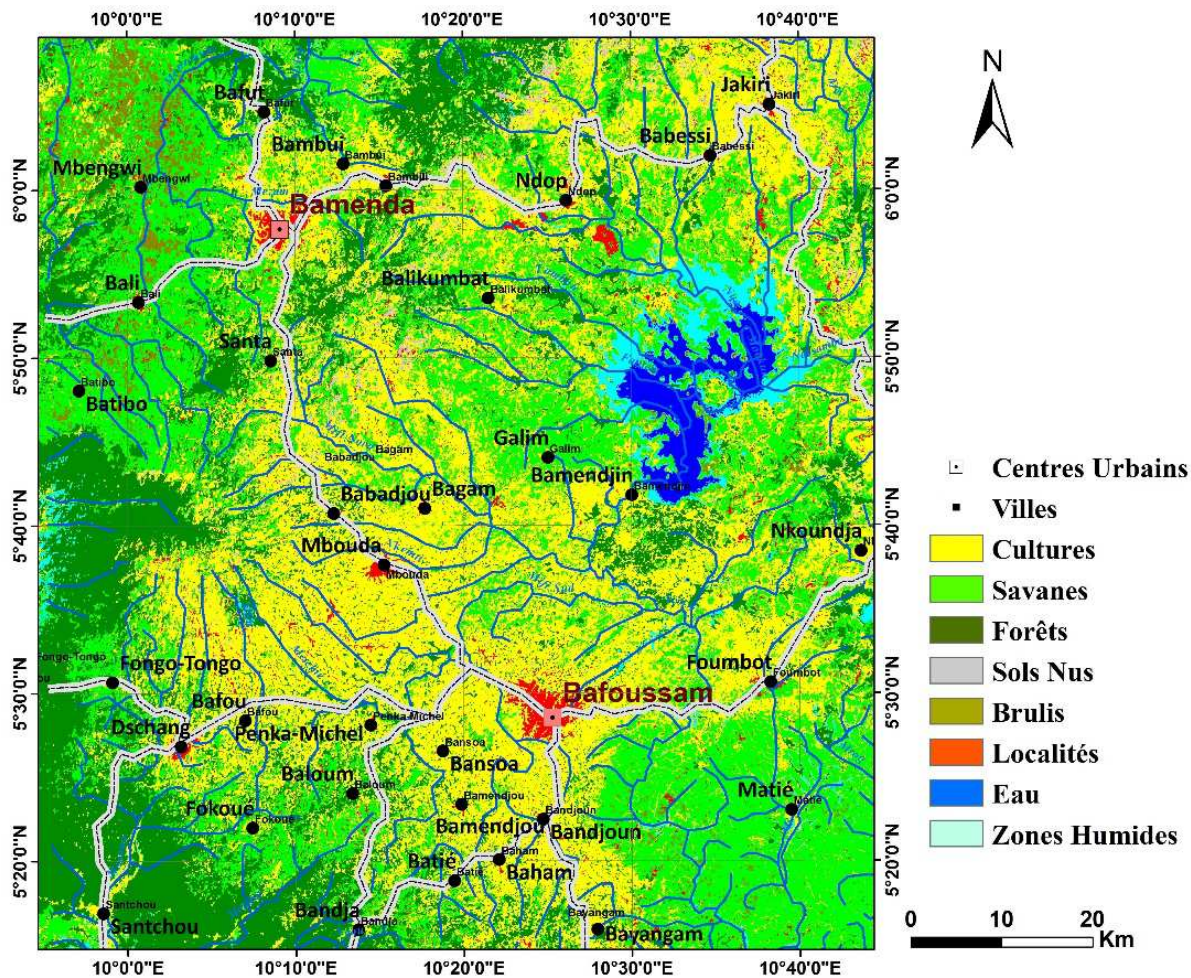


Fig. 4. Carte d'occupation de sol du site d'étude obtenue par réseaux de neurones

La validation statistique donne une précision globale de 90,08%. Le coefficient de Kappa est de 0,88. La classification est donc jugée satisfaisante. Afin de mettre en évidence l'apport des réseaux de neurones, une comparaison a été faite avec le maximum de vraisemblance qui est de loin, la méthode de classification plus utilisée dans la littérature. Cette comparaison a concerné à fois les cartes d'occupation de sols obtenus (**figure 5**), la matrice de confusion (**tableaux 2 et 3**) et les précisions qui en découlent. La comparaison des résultats obtenus montre une supériorité des réseaux de neurones dans la classification de la plupart des unités d'occupation de sol. Les réseaux de neurones donnent des valeurs supérieures de précision globale et de Kappa par rapport au maximum de vraisemblance. Pour certaines classes thématiques, la discrimination effectuée par les réseaux de neurones est très bonne alors que pour d'autres, il y a des confusions à cause des réponses radiométriques qui sont proches. L'analyse de cette matrice montre que certaines classes thématiques telles que l'eau, les zones humides et la forêt ont des précisions globales satisfaisantes qui dépassent 90%. En revanche, les classes



savanes, cultures, sols nus et localités, ont des pixels qui présentent une faible séparabilité si bien qu'il existe des confusions non négligeables entre elles.

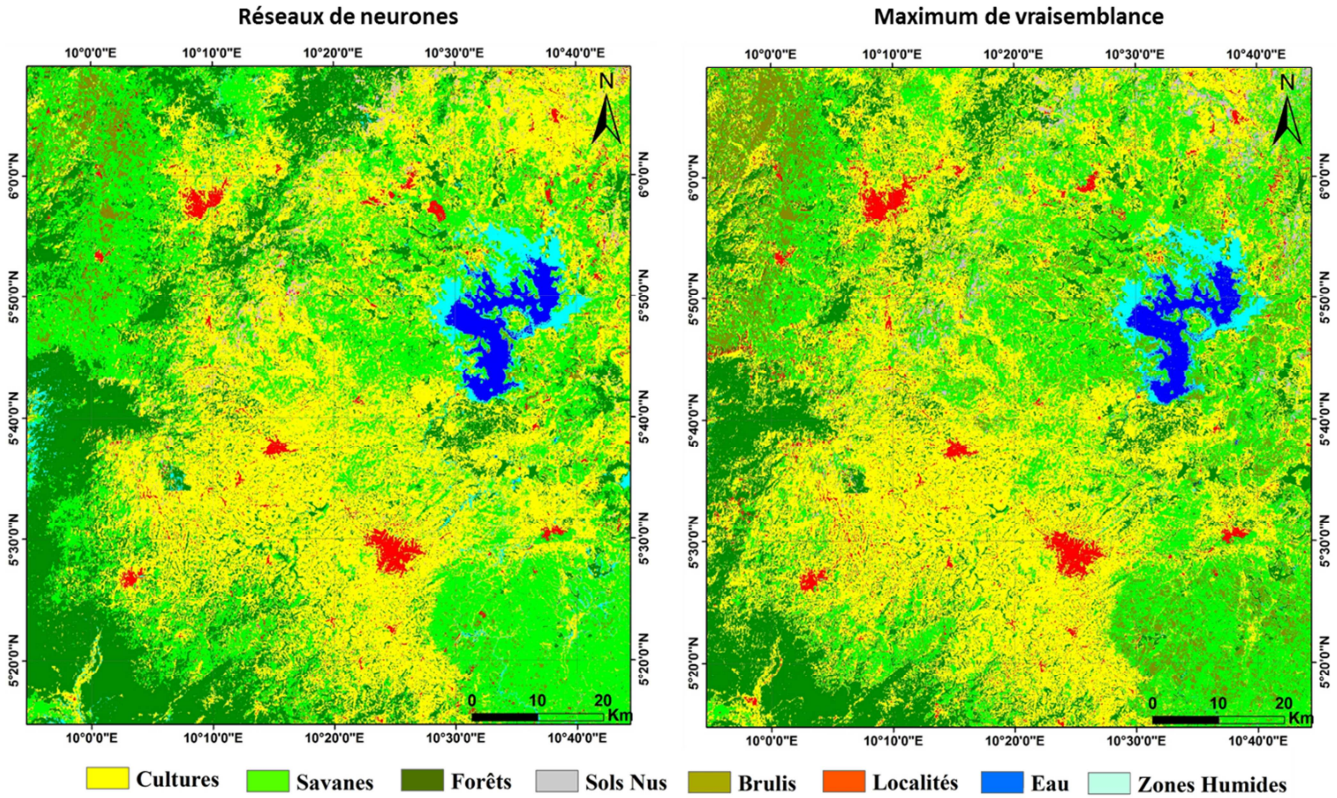


Fig. 5. Comparaison des cartes d'occupation de sol obtenues par réseaux de neurones et par maximum de vraisemblance

La courbe d'erreur constitue aussi un critère important dans la validation de la classification par réseau de neurones. En effet, l'objectif de la méthode de rétropropagation d'erreur étant de minimiser l'erreur entre la sortie calculée et celle désirée au cours de l'apprentissage, cette courbe devrait avoir une tendance décroissante et sans oscillations. Sur la courbe (figure 6) obtenue, la tendance décroissante est observée bien que quelques oscillations s'y observent ; l'allure attendue est obtenue ; preuve que la classification s'est déroulée avec succès.

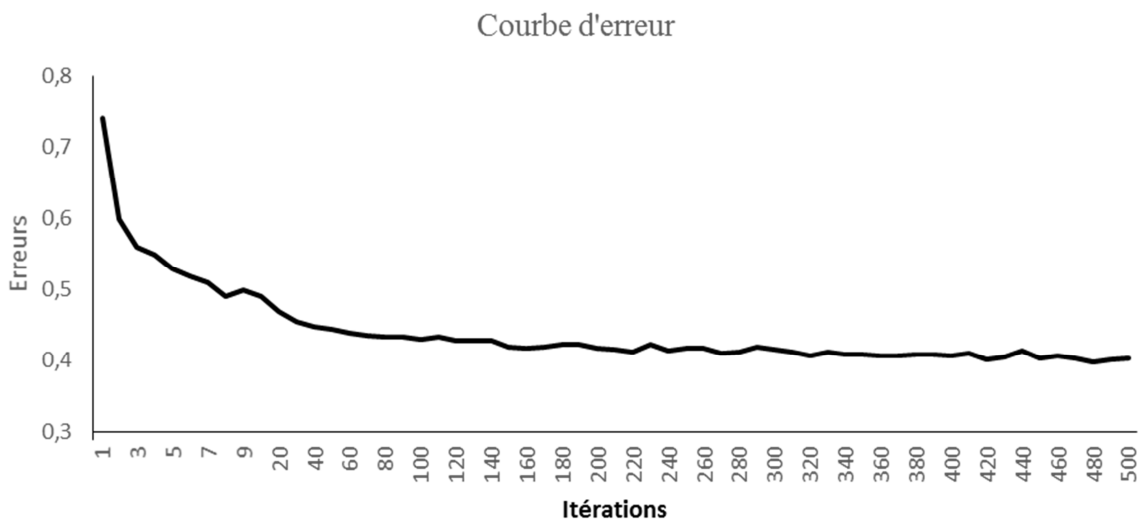


Fig. 6. Courbe de décroissance de l'erreur au cours de l'apprentissage du réseau de neurone

Tableau 2. Matrice de confusion issue de la classification par réseaux de neurones

Classes	Culture	Savane	Forets	Sols	Brulis	Villes	Eau	ZH	Total
Culture	<b>86.52</b>	16.12	1.05	14.41	15.38	3.30	0.00	0.00	23.28
Savanes	11.46	<b>82.80</b>	0.00	2.62	29.02	1.06	0.00	0.00	22.63
Forets	0.60	0.11	<b>98.95</b>	0.00	0.00	0.30	0.53	0.04	14.75
Sols Nus	0.05	0.00	0.00	<b>82.53</b>	0.23	0.00	0.00	0.00	0.92
Brulis	0.00	0.67	0.00	0.00	<b>55.38</b>	0.00	0.00	0.00	2.45
Localités	1.35	0.27	0.00	0.44	0.00	<b>95.34</b>	0.00	0.00	17.11
Eau	0.00	0.00	0.00	0.00	0.00	0.00	<b>99.01</b>	0.00	9.04
Zones Hum.	0.02	0.02	0.00	0.00	0.00	0.00	0.46	<b>99.96</b>	9.82
Total	100.00	100.00	100.00	100.0	100.0	100.0	100.0	100.0	100.0

Précision globale = 90,08

Kappa = 0,88

Tableau 3. Matrice de confusion issue de la classification par Maximum de vraisemblance

Class	Culture	Savanes	Foret	Sols	Brulis	Villes	Eau	ZH	Total
Cultures	<b>79.24</b>	18.05	1.87	0.00	7.51	5.40	0.00	0.00	21.56
Savanes	17.17	<b>80.20</b>	10.00	11.35	49.71	0.25	0.06	0.00	24.73
Forets	1.31	0.00	<b>88.13</b>	0.00	0.00	0.00	0.35	0.00	14.64
Sols	0.00	0.00	0.00	<b>88.65</b>	1.85	0.00	0.00	0.00	1.05
Brulis	0.00	1.62	0.00	0.00	<b>40.92</b>	0.05	0.00	0.00	2.07
Localités	2.29	0.13	0.00	0.00	0.00	<b>94.30</b>	0.00	0.00	17.08
Eau	0.00	0.00	0.00	0.00	0.00	0.00	<b>98.07</b>	3.94	8.47
Zones Hum.	0.00	0.00	0.00	0.00	0.00	0.00	1.52	<b>96.06</b>	10.40
Total	100.00	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Précision globale = 81,53

Kappa = 0,81

#### 4 DISCUSSION

Les réseaux de neurones ont permis de cartographier avec succès l'occupation de sol des hauts plateaux de l'ouest Cameroun. La précision obtenue est de 90,08% avec une valeur de Kappa de 0,88. La matrice de confusion a révélé de fortes confusions entre les classes cultures, savanes, et surfaces brûlées. En effet, les images (LANDSAT OLI) traitées ont été acquises au cours de la saison sèche. C'est une période pendant laquelle la plupart de la population prépare les champs pour les semis. Ainsi, des zones souvent occupées par des savanes, cultures et jachères, subissent l'effet des feux. La cohabitation entre zones de passage de feux et savanes et/ou culture, justifie cette confusion. D'autre part, le recouvrement quasi similaire et l'écologie de ces types de végétation, entraîne aussi une confusion entre cultures (très souvent du maïs, palmiers, raphia et du thé) et savanes (herbeuses, arbustives).

L'influence de plusieurs paramètres du réseau a été mise en évidence. À l'issue des différentes expérimentations réalisées, il ressort que la précision de la classification par réseaux de neurones dépend des valeurs des paramètres internes du réseau notamment la fonction d'activation, le momentum, le taux d'apprentissage, le momentum et le nombre d'itérations. Ce constat a aussi été effectué par [9] et par [19] qui ont étudié l'influence des paramètres internes du réseau sur la précision de la classification des images ETM+ de la région de Géorgie aux USA. La tendance décroissante de la courbe d'erreur obtenue atteste d'autre part que la classification a réussi. Ce critère a aussi été utilisé par [23] qui ont appliqués les réseaux de neurones pour classifier des images IKONOS de la région de Tehran en Iran. Les résultats ainsi obtenus concordent avec plusieurs études [16] [17] [19] [10] [20] [21] [11].

La comparaison des cartes d'occupation de sol, des matrices de confusion et des précisions globales entre les réseaux de neurones et le maximum de vraisemblance montre une supériorité des réseaux de neurones dans la classification de la plupart des unités d'occupation de sol. Les réseaux de neurones donnent des valeurs supérieures de précision globale et de Kappa par rapport au maximum de vraisemblance. De même, leur rendu cartographique est aussi meilleur. Ces résultats s'accordent à ceux obtenus par [27] et [23] qui ont respectivement démontré la supériorité des réseaux de neurones (par rapport au maximum de vraisemblance) dans la classification des milieux hétérogènes, complexes et où les classes d'occupation de sol sont difficilement séparables sur le seul critère spectral. Cependant, les limites de cette méthode

résident dans le nombre important de paramètres d'apprentissage et de la lenteur de son exécution. Toutefois, cette méthode a permis d'obtenir une cartographie plus fine de l'occupation de sol qui servira de critère dans la cartographie des ressources et des risques naturels de la zone.

## 5 CONCLUSION

Cette étude avait pour principal objectif de montrer l'apport des réseaux de neurones dans la cartographie de l'occupation du sol des hauts plateaux de l'Ouest du Cameroun. En effet, dans cette région, réaliser la carte d'occupation de sol était complexe étant donné le relief très accidenté et les ressemblances spectrales dans les principales unités d'occupation de sol. Les traitements des images satellitaires par les réseaux de neurones couplés aux données de terrain et à l'intégration des paramètres environnementaux, ont permis d'obtenir une bonne précision dans la classification. Huit unités d'occupation ont été mises en évidence et cartographiées avec une précision de 90,08% et un Kappa de 0,88. La comparaison de la carte d'occupation de sol obtenue par réseaux de neurones à celle issue de l'algorithme de maximum de vraisemblance a montré la supériorité des réseaux de neurones à la cartographie de l'occupation de sol. Cette carte pourrait être indispensable pour la planification du développement et la gestion du territoire dans cette région.

## REMERCIEMENTS

Les auteurs adressent leurs sincères remerciements à la Fondation Internationale pour la Science (IFS) qui leur a offert les moyens logistiques pour mener à bien les missions de terrain de collecte des sites d'entraînement et de contrôle

## REFERENCES

- [1] OSS, «Cartographie de l'occupation du sol et Spécifications techniques.», *REP-SAHEL Amélioration de la Résilience des Populations Sahéliennes aux Mutations Environnementales*, p. 36, 2013.
- [2] K. F. Kouamé, M. Bernier, D. L. Goné, M. B. Saley, R. Lefébvre, N. Soro et A. Koudou, «Intégration de données géospatiales dans un modèle hydrologique distribué pour la simulation des écoulements des eaux en milieu tropical humide de Côte d'Ivoire (Afrique de l'Ouest),» *Revue Télédétection*, vol. 7, n° 11-2-3-4, pp. 217-235, 2007.
- [3] S. Soulama, A. Kadeba, B. M. Nacoulma, S. Traoré, Y. Bachmann et A. Thiombiano, «Impact des activités anthropiques sur la dynamique de la végétation de la réserve partielle de faune de Pama et de ses périphéries (sud-est du Burkina Faso) dans un contexte de variabilité climatique,» *Journal of Applied Biosciences*, vol. 87, pp. 8047-8064, 2015.
- [4] H. Abdallah et H. Souilmi, «Analyse des changements d'occupation du sol et des transformations socio-économiques au cours du XXème siècle dans les marges arides du Tell oriental (Dorsale centrale, environs d'Es Sraïssif, Tunisie),» *Physi-Géo*, vol. 9, 2015.
- [5] T. V. Assoma, B. Adiaffi et A. Koudou, «Contribution of Remote Sensing, Multicriteria Analysis and GIS for Mapping Recharge Areas of the Coastal Aquifers in Southeast Ivory Coast.,» *American Journal of Scientific Research*, vol. 80, pp. 112-125, 2012.
- [6] T. D. Soro, B. D. Kouakou, E. A. Kouassi, G. Soro, A. M. Kouassi, K. E. Kouadio, M. S. Oga Yéi et N. Soro, «Hydroclimatologie et dynamique de l'occupation du sol du bassin versant du Haut Bandama à Tortiya (Nord de la Côte d'Ivoire),» *Vertigo*, vol. 13, 2013.
- [7] W. S. Saunders, J. G. Beban et M. Kilvington, «Risk-based approach to land use planning,» *GNS Science Miscellaneous Series 67*, p. 97p, 2013.
- [8] O. Leumbe Leumbe, D. Bitom, L. Mamdem et A. L. Tiki, «Cartographie des zones à risques d'inondation en zone soudano-sahélienne : cas de Maga et ses environs dans la région de l'Extrême Nord Cameroun,» *Afrique Sciences*, vol. 11, n° 13, 2015.
- [9] G. Hosni, Utilisation des réseaux de neurones pour la cartographie des zones humides à partir d'une série temporelle d'images RADARSAT 1, Thèse de Doctorat de l'Université du Québec, 2002, p. 230.
- [10] R. Pradham, M. P. Pradham, A. Bhusan, K. Pradham et M. K. Ghose, «Land-cover Classification and Mapping for Eastern Himalayan State Sikkim.,» *Journal of Computing*, vol. 2, n° 13, pp. 166-170, 2010.
- [11] S. Yousefi, S. Mirzaeeb, M. Tazehc, H. Pourghasemia et H. Karimid, «Comparison of different algorithms for land use mapping in dry climate using satellite images: a case study of the Central regions of Iran.,» *Desert*, vol. 20, n° 11, pp. 1-10, 2015.
- [12] S. Idbraim, «Méthodes d'extraction de l'information spatiale et de classification en imagerie de télédétection : applications à la cartographie thématique de la région d'Agadir (Maroc),» *Thèse de doctorat de l'Université de Toulouse*, p. 276, 2009.

- [13] M. Koné, A. Aman, Y. C. Adou, L. Coulibaly et K. E. N'guessan, «Suivi diachronique par télédétection spatiale de la couverture ligneuse en milieu de savane soudanienne en Côte d'Ivoire.,» *Revue Télédétection*, vol. 7, n° 11-2-3-4, pp. 433-446, 2007.
- [14] G. W. Tappoda et J. M. Fotsing, «Quantification de l'évolution du couvert végétal dans la réserve forestière de Laf-Madjam au nord du Cameroun par télédétection satellitale,» *Sécheresse*, vol. 21, n° 13, pp. 169 - 178, 2010.
- [15] F. Abdullah, L. Alqurashi et K. Lalit, «Investigating the Use of Remote Sensing and GIS Techniques to Detect Land Use and Land Cover Change: A Review.,» *Advances in Remote Sensing*, vol. 32, n° 123, pp. 193-204, 2013.
- [16] J. F. Mas, «Mapping land use/cover in a tropical coastal area using satellite sensor data, GIS and artificial neural networks.,» *Estuaire, Coastal and Shelf Science*, vol. 59, pp. 219-230, 2004.
- [17] H. Abderrahim, Apport de l'analyse de texture et des réseaux de neurones dans la cartographie de l'utilisation du sol en milieu urbain (cas de ville de Sherbrooke)., Mémoire de Master de l'Université de Sherbrooke, 2006, p. 72.
- [18] F. Pacifici, M. Chini et W. J. Emery, «A neural network approach using multi-scale textual metrics from very high-resolution panchromatic imagery for urban land-use classification,» *Remote Sensing of Environment*, vol. 113, pp. 1276-1292, 2009.
- [19] L. Zhou et X. Yang, «Use of neural networks for land cover classification from remotely sensed imagery.,» *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, vol. 37, pp. 575-578, 2008.
- [20] B. W. Szuster, Q. Chen et M. Borger, «A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones.,» *Applied Geography*, vol. 31, pp. 525-532, 2011.
- [21] G. Li, D. Lu, E. Moran et S. Hetrick, «Land-cover classification in a moist tropical region of Brazil with Landsat Thematic Mapper imagery.,» *International Journal of Remote Sensing*, vol. 32, n° 123, pp. 8207-8230, 2012.
- [22] B. Pradham et S. Lee, «Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling,» *Environmental, Modelling Software*, vol. 25, pp. 747 - 759, 2010b.
- [23] S. Ojaghi, H. Ebadi et F. F. Ahmadi, «Using artificial neural network for classification of high resolution remotely sensed images and assessment of its performance compared with statistical methods,» *American Journal of Engineering, Technology and Society*, vol. 2, n° 11, pp. 1-8, 2015.
- [24] S. Lee, K. Y. Song, Y. S. Kim et S. Park, «Regional groundwater productivity potential mapping using a geographical information system (GIS) based artificial neural network model,» *Hydrogeology Journal*, vol. 20, pp. 1511 - 1527, 2012.
- [25] C. M. Bishop, «Neural Networks for pattern recognition,» *Oxford University press*, p. 482, 1995.
- [26] D. E. Rumelhart , G. E. Hintont et R. J. Williams, «Learning representations by back-propagating errors.,» *Nature*, vol. 323, pp. 533-536, 1986.
- [27] M. R. Mustapha, H. S. Lim et M. Z. Mat Jafri, «Comparison of Neural Network and Maximum Likelihood Approaches in Image Classification,» *Journal of Applied Sciences*, vol. 10, pp. 2847-2854, 2010.
- [28] B. Pradham, S. Lee et M. F. Buchroithner, «A GIS-based back-propagation neural network model and its cross-application and validation for landslide susceptibility analyses,» *Computer, Environment and Urban Systems*, vol. 34, n° 13, pp. 216 - 235, 2010a.