
CARTOGRAPHIE DES CUIRASSES DANS LES FORMATIONS VOLCANO- SÉDIMENTAIRES DE LA ZONE D'ANIKRO- KADIOKRO (CÔTE D'IVOIRE) À L'AIDE DE LA THÉORIE DES FONCTIONS DE CROYANCE

Jean-Claude OKAINGNI¹, Koffi Fernand KOUAMÉ², Arnaud MARTIN³

¹ Institut National Polytechnique Félix Houphouët Boigny (INP-HB) de Yamoussoukro, Côte d'Ivoire ; BP 1093 Yamoussoukro, Côte d'Ivoire ; Téléphone : (225) 30 64 66 64 / 22 44 51 67 ; télécopie : (225) 30 64 04 06 / 22 44 32 93 ; Centre Universitaire de Recherche et d'Application en Télédétection (CURAT) / Université de Cocody, Abidjan, Côte d'Ivoire ; Téléphone : (225) 22 44 52 70 / 01 04 76 69 ; télécopie : (225) 22 44 52 70, Courriel : okaingni_jc@yahoo.fr

² Centre Universitaire de Recherche et d'Application en Télédétection (CURAT) / Unité de Formation et de Recherche des Sciences de la Terre et des Ressources Minières, Université de Cocody-Abidjan-Côte d'Ivoire, Côte d'Ivoire ; Téléphone : (225) 22 44 52 70 / 07 93 17 78 ; télécopie : (225) 22 44 52 70 ; Courriel : kouamef@yahoo.fr

³ Ecole Nationale Supérieure des Ingénieurs des Etudes et Techniques d'Armement (ENSIETA) du laboratoire Extraction et Exploitation de l'Information en Environnements Incertains (E3 I2 - EA3876). Téléphone : (33) 2 98 34 88 84 ; télécopie : (33) 2 98 34 87 50, Courriel : arnaud.martin@ensieta.fr

Soumis le 16 octobre 2008, accepté le 2 novembre 2009 - © Revue Télédétection, 2010, vol. 9, n° 1, p. 19-32

Résumé

Dans ce papier, on se propose d'appliquer la théorie des fonctions de croyance (ou théorie de Dempster-Shafer ou théorie de l'évidence) pour cartographier les cuirasses. L'approche présentée consiste à réaliser une classification à l'aide de la théorie des fonctions de croyance par fusion d'informations, liées aux cuirasses sur des images issues de néo-canaux obtenus à partir des canaux du satellite Landsat 7. Notre contribution réside dans l'utilisation de la méthode des **k** plus proches voisins et d'un coefficient d'affaiblissement pour prendre en compte, pour chaque pixel à classer, les incertitudes et imprécisions liées à la présence de cuirasses sur les images. La méthode a permis de cartographier les cuirasses sur la zone d'étude avec un taux de réussite supérieur à 90 %. Elle pourrait être utilisée, avec des aménagements appropriés, à d'autres fins de cartographie d'éléments géologiques.

Mots-clés : cartographie, cuirasses, images satellitaires, théorie des fonctions de croyance, Dempster-Shafer.

MAPPING BREASTPLATES IN VOLCANO-SEDIMENTARY AREA OF ANIKRO-KADIOKRO (IVORY COAST) USING THE DEMPSTER-SHAFER THEORY OF EVIDENCE

Abstract

In this paper, we propose to apply the theory of belief functions (or the evidence theory of Dempster-Shafer) to carry out mapping breastplates. The proposed approach consists in aggregating information extracted from satellite Landsat 7 images by using the belief functions theory.

Our contribution is based on the using of the classifier of **k**-nearest neighbors and a discounting coefficient in order to take into account uncertainty and imprecision of breastplates on satellite images. The methodology gives satisfactory results in our study area.

Keywords: cartography, breastplates, satellite imagery, theory of belief functions, Dempster-Shafer.

INTRODUCTION

La cuirasse est la croûte superficielle épaisse (allant jusqu'à plusieurs mètres) fortement durcie par des précipitations de fer et d'aluminium se formant surtout en climat intertropical à saisons sèches bien marquées (Foucault et Raoult, 2000).

Sa mise en place entraîne une dégradation du sol, en l'occurrence un appauvrissement et un durcissement de celui-ci (Dya, 1993). Les sols concernés ne sont plus alors cultivables. Ainsi, les populations de ces régions se déplacent vers d'autres zones ou d'autres régions à la recherche de terres cultivables. Cela entraînant souvent des conflits fonciers avec les autochtones.

Par ailleurs, les cuirasses, sous certaines conditions mécaniques et géotechniques, sont utilisées comme matériaux de réalisation d'ouvrages (constructions, remblais, etc.). D'où l'intérêt de leur cartographie afin de connaître leur position spatiale.

Des anciens travaux et des visites de terrain effectués ont permis de constater la présence de cuirasses dans le département de Toumodi (moyenne Côte d'Ivoire) (Maignien, 1954 ; Grandin et Delvigne, 1969 ; Yao-Kouamé, 2008). Toutefois, leurs continuité et extension restent à ce jour très peu étudiées.

Pour remédier à cette insuffisance, nous proposons une approche par Télédétection, à partir des images satellitaires, fondée sur une classification supervisée, réalisée à l'aide de la théorie des fonctions de croyance.

Les images satellitaires permettent de couvrir des zones relativement grandes. À partir de notre zone d'étude, elles pourront servir à extrapoler les résultats sur des zones plus grandes couvertes par celles-ci. Par ailleurs, les images satellitaires, possédant des caractéristiques spatiale (résolution spatiale) et spectrale (résolution spectrale) diversifiées, seront fusionnées à l'aide de la théorie des fonctions de croyance.

L'objectif principal de l'étude est de cartographier les cuirasses à l'aide de la théorie des fonctions de croyance. Il s'agit de façon spécifique de déterminer le cadre de discernement, les fonctions de croyance et les critères de décision, d'écrire des algorithmes et des codes de programmation en langage Matlab pour la réalisation de la cartographie.

Ce papier, qui se propose de rendre compte des travaux réalisés, présente d'abord la zone d'étude, puis les fonctions de croyance, ensuite le matériel et la méthodologie utilisés et enfin les résultats obtenus.

1. ZONE D'ÉTUDE

La zone d'Anikro-Kadiokro est située en Côte d'Ivoire, à environ 200 km d'Abidjan (capitale économique) entre Toumodi et Yamoussoukro (capitale politique). Elle est délimitée par les longitudes 5°05' et 5°10' Ouest et les latitudes 6°35' et 6°40' Nord avec une superficie d'environ 8409 ha (fig. 1).

Le climat y est de type équatorial à quatre saisons (2 saisons sèches et 2 saisons des pluies) dont la plus grande saison sèche de novembre à mi-mars est favorable à la mise en place des cuirasses (Yao-Kouamé, 2008). La végétation est dominée par une savane arborée, des îlots forestiers sur les plateaux et des forêts galeries le long des cours d'eau. Quant au réseau hydrographique, il est constitué de trois cours d'eau (Brahiba, Bomba et Bohiro) et de leurs affluents (fig. 1).

Sa géologie est composée de formations volcano-sédimentaires à prédominance basique générant des sols brunifiés avec présence d'oxyde (Yacé, 1976 in Yao-Kouamé, 2008) et de formations conglomératiques à galets de tufs (Yao-Kouamé, 2008). Selon Yao-Kouamé (2008), cela confère à la zone d'étude un milieu physique et géologique favorable à la mise en place des cuirasses.

2. THÉORIE DES FONCTIONS DE CROYANCE

La théorie des fonctions de croyance est appliquée dans plusieurs domaines notamment dans le traitement d'images satellitaires pour la réalisation de classification (Piegon et al, (2000), Milisavljevic et al, (2003), Corgne (2004)).

2.1. Principe de base

Le principe de base provient des travaux de Dempster (1967) et Shafer (1976).

Soit $\Omega = \{C_1, \dots, C_N\}$, l'ensemble des N classes possibles pour une observation \mathbf{x} , appelé cadre de discernement. La théorie des fonctions de croyance est fondée sur la manipulation de fonctions de masse définies sur l'ensemble puissance de Ω , noté 2^Ω , l'ensemble des 2^N disjonctions de Ω , au lieu de se restreindre à Ω comme le ferait la théorie des probabilités.

On définit alors pour une source S , une fonction de masse de croyance m de 2^Ω à valeurs dans $[0,1]$ vérifiant les conditions suivantes (équation 1) :

$$\sum_{A \in 2^\Omega} m(A) = 1 \quad (\text{équation 1})$$

$$m(\emptyset) = 0$$

où \emptyset est l'ensemble vide. $m(A)$ quantifie la croyance que l'on accorde au fait que la classe cherchée appartienne au sous ensemble A de Ω (et à aucun autre sous ensemble de A).

Les sous ensembles A tels que $m(A) > 0$ sont appelés éléments focaux de m .

Cas particuliers

- Si $m(\Omega) = 1$, alors la source S est dans l'ignorance totale.
- Si $m(A) = 1$ (A une disjonction d'éléments de Ω), alors la source S a une connaissance imprécise ; elle ne croit qu'en A .
- Si $m(C) = 1$ (C est un élément singleton de Ω), alors la source S a une connaissance précise, elle croit pleinement en C .
- Si $m(A) = s$ et $m(\Omega) = 1-s$, alors la source S a une connaissance incertaine et imprécise ; elle croit en partie en A , mais rien de plus.

Deux fonctions de masse de croyance m_1 et m_2 représentant les informations respectives de deux sources différentes S_1 et S_2 , peuvent être combinées selon la règle de Dempster (1967) (équation 2) :

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) * m_2(C)}{1 - K}, \quad \forall A \in 2^\Omega \setminus \{\emptyset\}; \quad \forall B, C \in 2^\Omega; \quad (\text{équation 2})$$

K (équation 3) est appelé l'inconsistance de la fusion et peut être interprété comme une mesure de conflit. Il correspond à la masse de l'ensemble vide. On a :

$$K = \sum_{B \cap C = \emptyset} m_1(B) * m_2(C), \quad \forall B, C \in 2^\Omega; \quad (\text{équation 3})$$

Si $K=1$, la combinaison entre les sources d'informations est impossible avec la règle de consensus normalisée de l'équation 2. Cela signifie que les sources sont en conflit (totalement en conflit).

Notons que la combinaison de Dempster est aussi appelée somme orthogonale et notée: $m = m_1 \oplus m_2$. Elle est commutative, associative et non idempotente.

Après combinaison, il reste à prendre une décision quant à l'élément le plus crédible de Ω , qui représente la classe attribuée à l'observation \mathcal{X} .

Plusieurs règles de décision sont alors possibles (Yager, 1986 ; Bracker, 1996 ; Sudano, 2002 ; Corgne, 2004), mais le maximum de probabilité « pignistique » reste le compromis le plus utilisé. La transformation pignistique introduite par Smets (1990) permet de répartir uniformément la masse associée à un sous ensemble de Ω sur chacun de ses éléments (équations 4) :

$$BetP(C_i) = \sum_{C_j \in A \subseteq \Omega_i} \frac{m(A)}{|A|(1 - m(\phi))}, \quad \forall C_i \in \Omega \quad i = 1, \dots, N \quad (\text{équation 4})$$

$BetP(C_i)$ est la probabilité pignistique associée à la classe C_i et $|A|$ est le cardinal de A . Avec la règle du maximum de probabilité pignistique, la décision va alors vers l'élément de Ω dont cette valeur est la plus grande (équation 5) :

$$C^* = Arg \left\{ \max_{C_i \in \Omega} [BetP(C_i)] \right\} \quad (\text{équation 5})$$

C^* correspond à la classe dont la probabilité pignistique est la plus grande.

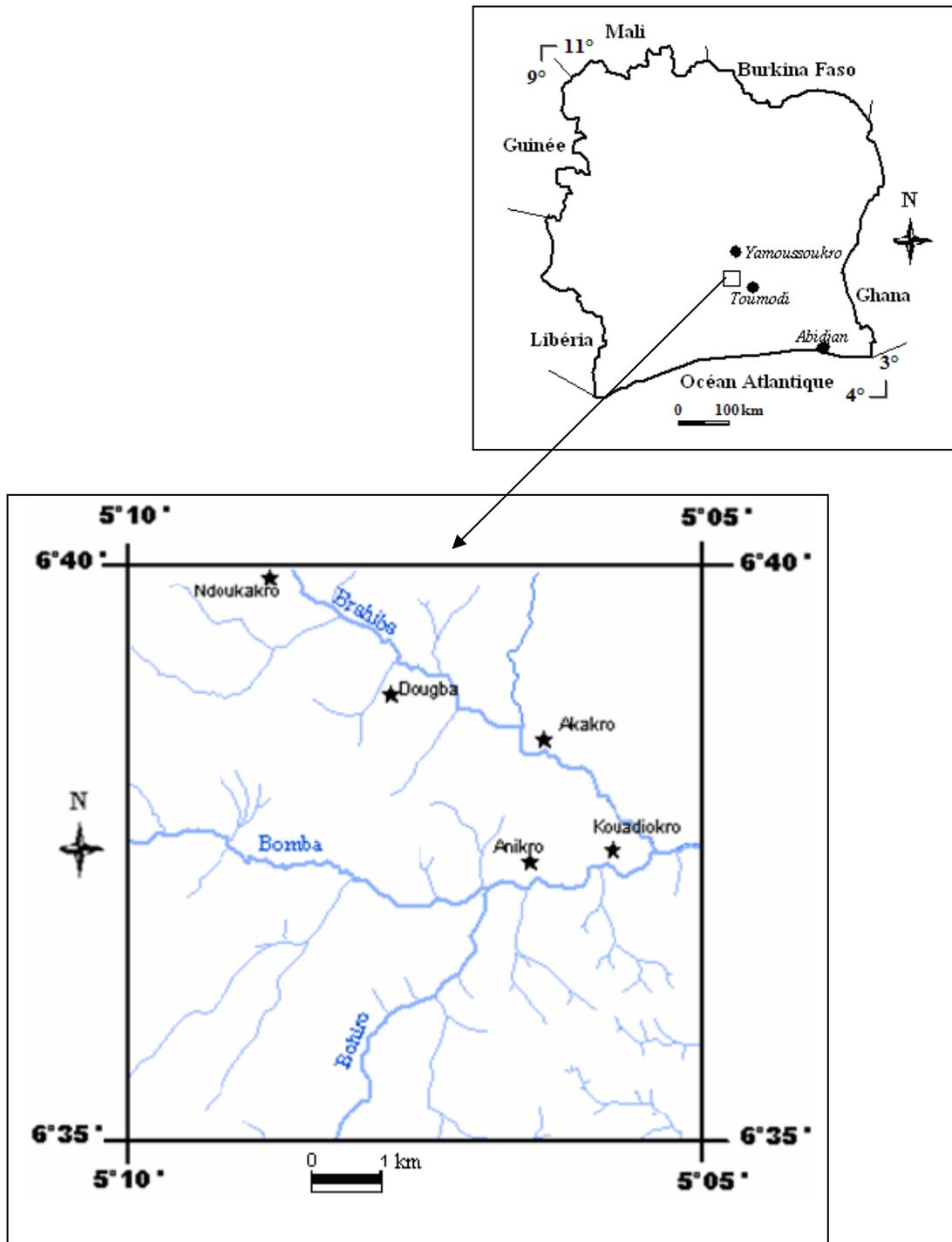


Figure 1 : carte de situation de la zone d'Anikro-Kadiokro.

2.2. Architecture d'une fusion d'informations par la théorie des fonctions de croyance

L'architecture d'une fusion d'informations par la théorie des fonctions de croyance peut être caractérisée par une structure à quatre étapes de réalisation que sont la modélisation, l'estimation, la combinaison et la décision (Martin, 2005). L'étape de modélisation prend en compte les imperfections liées aux informations. Afin de considérer la fiabilité de la source, Shafer (1976) a introduit la notion d'affaiblissement des fonctions de masse de croyance à l'aide de coefficient réel α compris entre zéro (0) et un (1). Ainsi, pour une fonction de masse de croyance m initialement définie sur le cadre de discernement Ω , la nouvelle fonction de masse de croyance m' modélisant les imperfections se met sous la forme de l'équation 6 :

$$\begin{cases} m'(A) = \alpha * m(A) & \forall A \in 2^\Omega \\ m'(\Omega) = 1 - \alpha * (1 - m(\Omega)) \end{cases} \quad (\text{équation 6})$$

α est le coefficient d'affaiblissement lié à la source ayant produit l'information.

Dans le cas où $\alpha=0$, la source concernée n'est pas fiable et alors nous nous trouvons dans un cas d'ignorance totale où toute la masse est sur Ω .

La cartographie des cuirasses dans la zone d'étude par la théorie des fonctions de croyance nécessite l'utilisation d'une méthodologie cohérente à partir de matériel et méthodes qui font l'objet de la section suivante.

3. MATÉRIEL ET MÉTHODES

3.1. Matériel

Le matériel utilisé dans le cadre de nos travaux est composé de données et de matériel informatique et de terrain.

Les données sont constituées de cartes (topographique, géologique et pédologique), de données de terrain et d'une image satellitaire couvrant la zone d'étude.

La carte topographique, coupure Gagnoa 4b (IGN, 1964) au 1 / 50 000, a permis la situation de la zone d'étude. Pour la détermination des formations géologiques et des unités morpho-pédologiques, nous avons eu recours à la carte géologique (ENSTP, 1995) au 1 / 50 000 et à la carte des unités morpho-pédologiques du degré carré de Gagnoa (DCGTx, 1993) au 1 / 100 000.

Les données de terrain sont constituées par les coordonnées géographiques de 30 sites cuirassés et 15 sites non cuirassés. Elles ont servi à caractériser les pixels échantillonnés dans les phases d'apprentissage et de programmation informatique.

En outre, nous avons utilisé des images extraites de l'image satellitaire ETM^+ de Landsat 7, scène 196/56 du 29 janvier 2002, obtenue le 03 août 2006 à 10h30 sur le site : <http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp>. Il s'agit spécifiquement des images extraites des canaux du visible (ETM^+1 , ETM^+2 , ETM^+3), du proche infrarouge (ETM^+4) et du moyen infrarouge (ETM^+5 , ETM^+7). Ce choix est guidé par la résolution spatiale (30 m. x 30 m.) qui permet, d'une part de discriminer les surfaces végétalisées, les sols nus, les routes, les pistes et les cours d'eau, et d'autre part d'avoir une précision acceptable dans la localisation de phénomènes (fracturation des formations géologiques, etc.).

Par ailleurs, à partir des travaux de Mainguet (1976) et Scanvic et Pion (1996), il ressort que les bandes spectrales vert et proche-infrarouge peuvent contribuer à caractériser les formations cuirassées.

Nous avons utilisé, comme matériel informatique et de terrain pour la réalisation des travaux, le logiciel de traitement d'images satellitaires ENVI 4.1, le logiciel Matlab 7.0 pour l'élaboration du programme de classification par la TFC et un récepteur GPS IV Garmin pour le positionnement des zones échantillonnées.

3.2. Méthodes

La méthodologie mise en œuvre s'articule autour de quatre (4) étapes. Dans un premier temps, des traitements d'images satellitaires sont effectués à partir de calcul d'indices et de compositions colorées mettant en évidence les sols nus et les affleurements, lieux de prédilection des cuirasses. Puis une phase d'apprentissage est effectuée pour la détermination des fonctions de masse de la théorie des fonctions de croyance. Ensuite, des algorithmes et des codes de programmation dans le langage Matlab ont été écrits pour réaliser une classification par la théorie des fonctions de croyance. Enfin, une validation de la classification a été faite par une analyse statistique.

3.2.1. Traitements

Trois images ont été produites à partir du calcul de l'indice de végétation normalisé ($NDVI$), de l'indice de cuirasse (IC) et d'une composition colorée RVB.

Le résultat du $NDVI$ est une image présentant un gradient croissant d'activité végétale allant du noir signifiant l'absence de couverture, au blanc qui rend compte d'une activité chlorophyllienne très élevée. Par conséquent, il est utilisé pour discriminer les sols nus et les surfaces végétalisées. La formule de calcul est présentée par l'équation 7 (Collet et al, 2001) :

$$NDVI = \frac{ETM^+4 - ETM^+3}{ETM^+4 + ETM^+3} \quad (\text{équation 7})$$

L'indice de cuirasse (Pouchin, 2001) permet de dissocier les couvertures végétalisées des étendues minérales (notamment des cuirasses). Les surfaces végétalisées et aquatiques apparaissent en noir alors que les surfaces minéralisées sont codées en gris clair ou en blanc. L'indice de cuirasse a pour formule (équation 8) :

$$IC = 3 * (ETM^+ 2) - (ETM^+ 3) - 100 \quad (\text{équation 8})$$

Plusieurs *scenarii* de compositions colorées (RVB) ont été réalisés en attribuant successivement les couleurs primaires rouge, vert et bleu à chacun des canaux ETM^+2 , ETM^+4 et ETM^+7 . Celle ayant fourni une bonne discrimination des entités de l'image a été obtenue en attribuant la couleur rouge au canal ETM^+4 (**R**), la couleur verte au canal ETM^+7 (**V**) et la couleur bleue au canal ETM^+2 (**B**). Elle est baptisée **CC** dans la suite.

3. 2. 2. Théorie des fonctions de croyance

La théorie des fonctions de croyance introduite par Dempster (1967) et Shafer (1976) offre un cadre mathématique rigoureux pour la classification et la fusion d'informations incertaines et imprécises. Nous pouvons distinguer quatre étapes (Martin, 2005) : la modélisation du cadre de discernement, l'estimation des fonctions de masse, la combinaison des fonctions de masse et la prise de décision.

- Le cadre de discernement noté Ω est l'ensemble de toutes les classes envisagées dans une problématique de classification. L'objet d'étude étant uniquement la cuirasse, la classification pour sa mise en évidence va permettre de dire si un pixel représente la cuirasse ou pas. Ainsi, nous avons modélisé le cadre de discernement en deux classes C_1 (zone peu ou fortement cuirassée) et C_2 (zone non cuirassée). Par conséquent, on a :

$$\Omega = \{C_1, C_2\} \text{ et } 2^\Omega = \{\emptyset, C_1, C_2, \Omega\}$$

Deux ensembles d'apprentissage ont été utilisés pour la détermination des classes C_1 et C_2 .

Une zone est jugée cuirassée sur le terrain si la superficie de cuirasse observée est supérieure à la résolution spatiale considérée (30m. x 30m.). Ainsi, trente zones cuirassées et quinze zones non cuirassées ont été déterminées. Les zones non cuirassées étant composées de cours d'eau, d'habitats, de zones végétalisées, etc.

- Plusieurs fonctions de masse ont été proposées dans la littérature et leur choix relève de la modélisation (Shafer, 1976 ; Yager, 1987 ; Denœux, 1995 ; Lefèvre et al, 1999).

En considérant des sources d'informations dynamiques (k plus proches voisins du pixel à classer) et le cadre de discernement susmentionné, nous définissons les fonctions de masse dans un monde clos, c'est-à-dire où tous les cas envisageables sont supposés connus dans Ω . Nous avons alors utilisé, pour l'estimation des fonctions de masse, un modèle de distance, en l'occurrence la règle des k plus proches voisins, fondée sur les travaux de Denœux (1995).

La règle des k plus proches voisins crédibiliste permet de classer un pixel x . Elle est appliquée sur l'espace à trois dimensions formé des images **NDVI**, **IC** et **CC**, cela en vue de prendre en compte simultanément les imperfections éventuelles liées à chacune des sources. Mais avant, la définition de régions d'intérêt qui sont des pixels d'apprentissage, est indispensable. Elle permet ensuite, une classification des autres pixels avec un degré de croyance obtenu par l'utilisation de coefficient d'affaiblissement pour prendre en compte les imperfections liées aux informations.

Ainsi, à partir des coordonnées géographiques des ensembles d'apprentissage susmentionnés, des pixels ont été échantillonnés sur les images **NDVI**, **IC** et **CC** à l'aide du logiciel ENVI 4.1 pour déterminer les pixels d'apprentissage caractérisant les classes C_1 et C_2 . Pour un pixel choisi, on lui associe ses valeurs correspondantes sur les images **NDVI**, **IC** et **CC**. Ce qui donne un triplet de coordonnées dans le repère (**NDVI**, **IC**, **CC**).

Soit $\mathcal{X} = \{x_{iq} = (x_{iq}^1, x_{iq}^2, x_{iq}^3) \mid i=1, 2; q=1, \dots, N_i\}$, où q est l'indice du pixel, et N_i le nombre de pixels échantillonnés par classe}, l'ensemble des pixels appartenant aux ensembles d'apprentissage pour lesquels les labels C_1 et C_2 sont attribués.

Classer un pixel x de l'image, signifie qu'il faut lui attribuer une classe, par fusion d'informations à partir de ses k plus proches voisins (en terme de distance) dans \mathcal{X} .

Soit ϕ_x , les k plus proches voisins, en terme de distance, du pixel x dans \mathcal{X} . Et soit ϕ_x^i , les plus proches voisins de x dans ϕ_x appartenant à C_i , $i = 1, 2$.

La fonction de masse associée à ϕ_x^i , pour l'appartenance du pixel x à la classe C_i , est déduite des travaux de Denœux (1995), par les formules suivantes (équations 9 et 10) :

Pour $i = 1, 2$

$$m^{x_i} : 2^\Omega \rightarrow [0,1]$$

Si $\phi_x^i \neq \emptyset$

$$m^{x_i}(C_i) = 1 - \prod_{x^t \in \phi_x^i} (1 - \alpha_0 e^{-\gamma_i d_t}) ;$$

$$m^{x_i}(C_r) = 0, r \neq i ;$$

$$m^{x_i}(\Omega) = \prod_{x^t \in \phi_x^i} (1 - \alpha_0 e^{-\gamma_i d_t}) ;$$

(équation 9)

avec α_0 le coefficient d'affaiblissement affecté aux fonctions de masse.

d_t la distance euclidienne entre x et x^t est donnée par :

$$x = (x^1, x^2, x^3) \text{ et } x^t = (x^{t1}, x^{t2}, x^{t3}) \text{ dans le repère } (NDVI, IC, CC);$$

$$d_t = ((x^1 - x^{t1})^2 + (x^2 - x^{t2})^2 + (x^3 - x^{t3})^2)^{1/2}$$

γ_i est l'inverse de la distance moyenne entre les pixels échantillons de la classe C_i ;

Sinon (Si $\phi_x^i = \emptyset$)

$$m^{x_i}(C_r) = 0, r = 1, 2$$

$$m^{x_i}(\Omega) = 1$$

(équation 10)

- La combinaison est réalisée à partir des fonctions de masse (m^{x_1} et m^{x_2}) associées aux deux sous ensembles ϕ_x^1 et ϕ_x^2 en utilisant la somme orthogonale de Dempster. La fonction de masse résultante, à partir des k plus proches voisins de x dans ϕ_x , est définie de la manière suivante (équation 11) :

$$m^x : 2^\Omega \rightarrow [0,1]$$

$$m^x(C_1) = \frac{m^{x_1}(C_1) * m^{x_2}(\Omega)}{H}$$

$$m^x(C_2) = \frac{m^{x_2}(C_2) * m^{x_1}(\Omega)}{H}$$

$$m^x(\Omega) = \frac{m^{x_1}(\Omega) * m^{x_2}(\Omega)}{H}$$

(équation 11)

Où $H = m^{x_1}(C_1) * m^{x_2}(\Omega) + m^{x_2}(C_2) * m^{x_1}(\Omega) + m^{x_1}(\Omega) * m^{x_2}(\Omega)$

- Plusieurs critères de prise de décision existent. La probabilité pignistique introduite par Smets (1993) est généralement employée comme compromis. Elle est définie ici par (équation 12) :

$$BetP^x : \Omega \rightarrow [0,1]$$

$$BetP^x(C_i) = m^x(C_i) + \frac{m^x(\Omega)}{3}, i = 1, 2 ;$$

(équation 12)

Le critère de décision, d'appartenance du pixel x à une classe, qui a été retenu est donc celui du maximum de probabilité pignistique. Il est défini par (équation 13) :

$$C^* = \max BetP^x(C_i), i = 1, 2 \quad (\text{équation 13})$$

C^* est la plus grande valeur des probabilités pignistiques associées aux classes C_1 et C_2 . Par conséquent, la classe attribuée au pixel x est la classe ayant la valeur C^* .

3. 2. 3. Algorithmes et codes

Les algorithmes élaborés et les codes de programmation du langage Matlab utilisés sont inspirés des travaux de Denœux (1995).

Nous avons écrit des fonctions avec des codes de Matlab pour déterminer d'abord les paramètres γ_i ($i=1, 2$) (équation 9) et les k plus proches voisins d'un pixel à classer, puis les fonctions de masse simples et combinées et enfin la classe à attribuer au pixel à classer. Celles-ci ont été intégrées dans un programme principal qui a permis de classer les pixels suivant deux classes C_1 , et C_2 .

Plusieurs *scenarii* suivant k (nombre de plus proches voisins considéré) et α_0 (coefficient d'affaiblissement considéré) ont été effectués sous forme d'apprentissage pour la réalisation de la classification finale. Nous avons fait varier k ($k = 2 ; 3 ; \dots ; 15$) et α_0 ($\alpha_0 = 1 ; 0,95 ; 0,90 ; \dots ; 0,50$)

3. 2. 4. Validation de la classification

La validation de la classification obtenue a été faite par une analyse statistique qui consiste à effectuer une analyse des résultats de bonne classification des classes C_1 et C_2 et des erreurs produites. Pour ce faire, nous avons calculé, à partir de deux types d'échantillons (échantillons d'estimation « échantillon I » et échantillons de vérification « échantillon II »), une matrice de confusion, pour obtenir des indices de performances, et une matrice de comparaison pour tester l'influence des pixels échantillons sur la classification.

Le nombre de classes à réaliser étant deux, nous considérons des matrices d'ordre 2.

Soient M_{cf} et M_{cp} , les matrices respectivement de confusion et de comparaison.

Pour i et j variant de 1 à 2, $M_{cf}(i,j)$ est le nombre de pixels de la classe C_i ayant été affectés à la classe C_j , après classification. Il peut être exprimé en pourcentage en le divisant par le nombre de pixels de la classe C_i .

En particulier, $M_{cf}(i,i)$ est le nombre de pixels de la classe C_i ayant effectivement été affectés à celle-ci, après classification.

Nous obtenons alors les indices de performance suivant (équations 14 et 15) (Martin *et al*, 2006) :

$$GCR_i = \frac{M_{cf}(i,i)}{N_i}, i = 1, 2 ; \quad (\text{équation 14})$$

$$ECR_i = \frac{1}{2} \left(\sum_{j=1, j \neq i}^2 \frac{M_{cf}(i,j)}{N_i} + \sum_{j=1, j \neq i}^2 \frac{M_{cf}(j,i)}{N_j} \right), i = 1, 2 ; \quad (\text{équation 15})$$

avec N_i : nombre de pixels de la classe C_i .

GCR_i : taux de pixels bien classés de la classe C_i

ECR_i : taux de pixels mal classés de la classe C_i

Pour i et j variant de 1 à 2, $M_{cp}(i,j)$ est le nombre de pixels appartenant à la fois aux classes C_i de la classification 1 (réalisée avec l'échantillon I) et C_j de la classification 2 (réalisée avec l'échantillon II). Il peut être exprimé en pourcentage en le divisant par le nombre total de pixels.

$M_{cp}(i, i)$ est, en particulier, le nombre de pixels classés à la fois par les classifications 1 et 2 comme appartenant à la classe C_i . En définissant l'expression suivante (équation 16) :

$$Sensibilité = \left(1 - \frac{M_{cp}(1,1) + M_{cp}(2,2)}{M_{cp}(1,1) + M_{cp}(2,2) + M_{cp}(1,2) + M_{cp}(2,1)} \right) * 100 ; \quad (\text{équation 16})$$

Alors, la classification sera jugée insensible aux échantillons I et II si le paramètre *Sensibilité* tend vers 0%.

La mise en œuvre de la méthodologie susmentionnée a permis d'obtenir des résultats. La présentation de ces résultats fait l'objet de la section suivante.

4. RÉSULTATS

4. 1. Traitements

4. 1. 1. Indice de Végétation Normalisé

Afin de discriminer le sol nu et la végétation, nous avons réalisé le NDVI (figure2) à partir des images des canaux ETM^+3 et ETM^+4 .

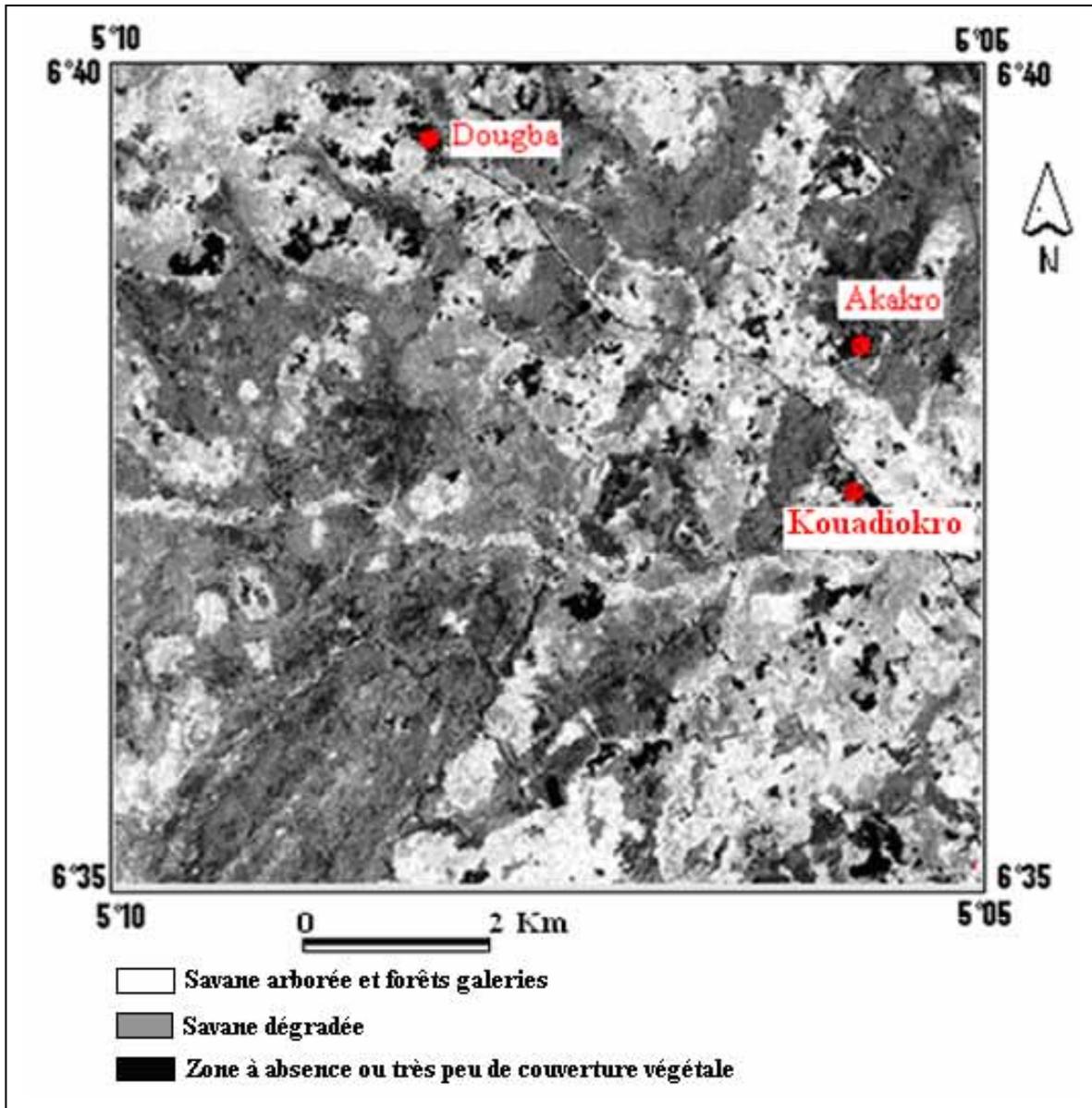


Figure 2 : Carte dérivée de l'Image du NDVI d'Anikro-Kadiokro

Ainsi, nous avons obtenu du NDVI et des visites de terrain trois entités :

- les zones de couleur grise très claire à blanchâtre marquent une activité chlorophyllienne très élevée. Elles correspondent à la savane arborée et aux forêts galeries ;
- les zones à couleur grise à gris-foncée sont révélatrices d'une très faible activité chlorophyllienne. Elles caractérisent la savane dégradée ;
- les zones de teinte noire à gris-foncée révèlent une absence ou très peu de couverture végétale. Par conséquent, elles pourraient représenter les sols bâtis à sols nus qui sont les zones de prédilection des cuirasses.

4.1. 2. Indice de Cuirasse

Nous avons réalisé, à partir des images des canaux ETM^+3 et ETM^+2 , un calcul d'indice de cuirasse de la zone d'étude. Ce qui fournit l'image ci après (fig. 3) :

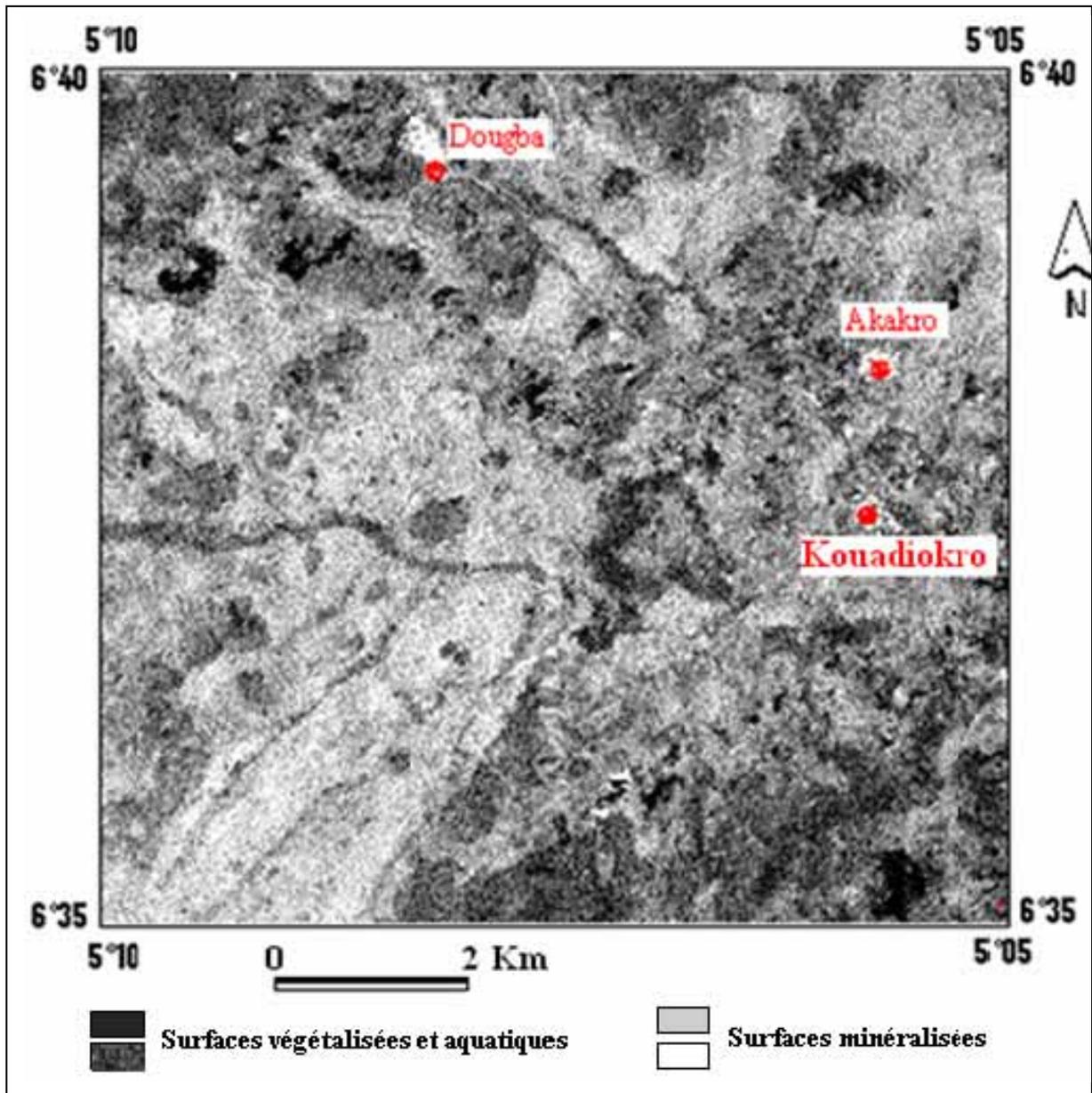


Figure 3 : Carte dérivée de l'Image de l'indice de cuirasse d'Anikro-Kadiokro

Les surfaces végétalisées et aquatiques apparaissent en noir alors que les surfaces minéralisées sont codées en gris clair ou en blanc (Pouchin, 2001).

4.1. 3. Composition colorée

Afin de discriminer les différentes entités de l'image de la zone, des compositions colorées ont été réalisées. Celle discriminant au mieux les différentes entités est représentée ci après (fig. 4) :

L'image de la composition colorée comparée à des informations de terrain indique que :

- la couleur rouge correspond à la savane arborée et aux forêts galeries ;
- la couleur verte caractérise les zones d'affleurement avec présence d'eau ;
- la couleur violette marque la savane arborée dégradée ;
- la couleur bleue met en évidence les sols à très faible couverture végétale et les habitats.

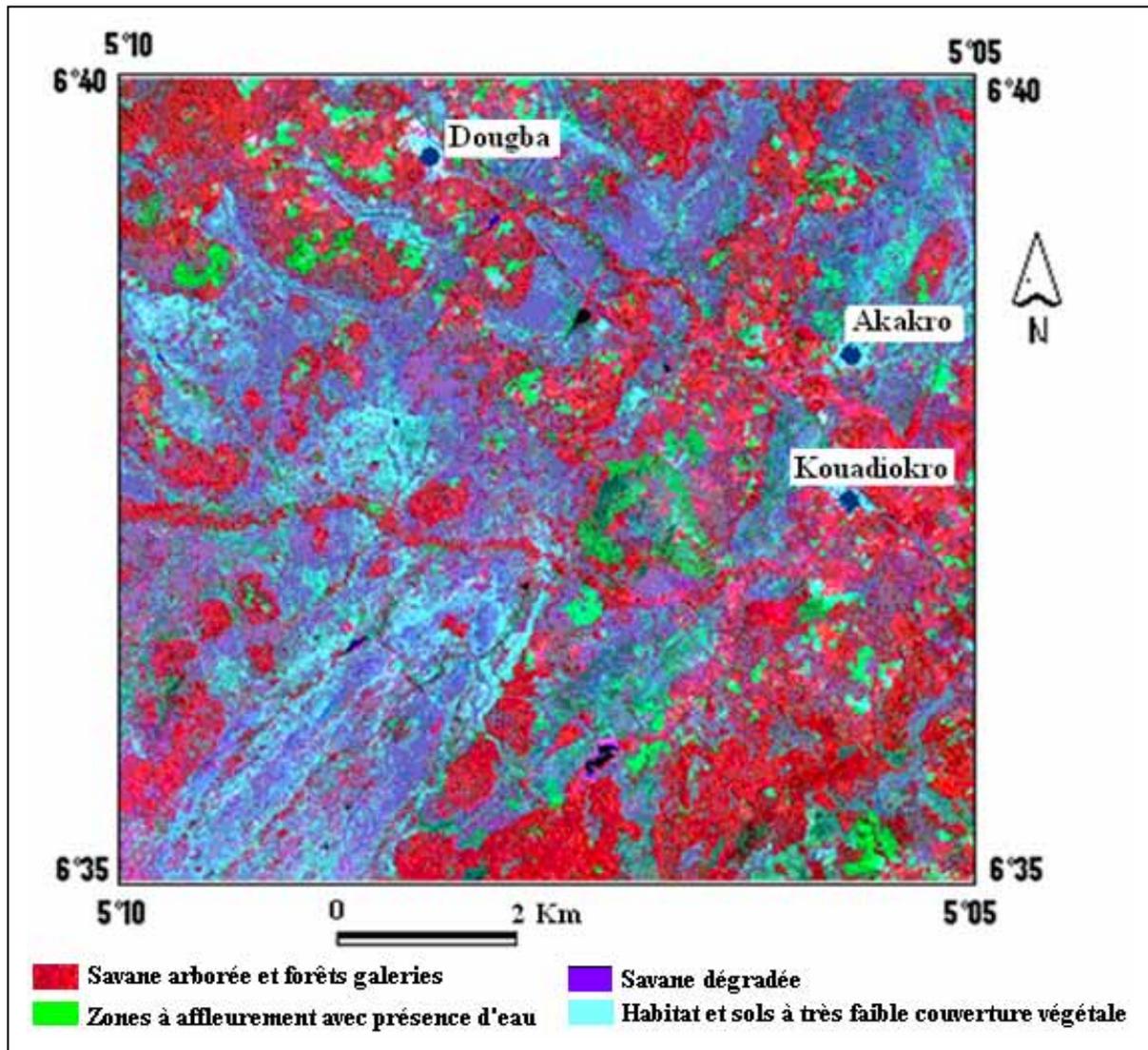


Figure 4 : Carte dérivée de la composition colorée (ETM^{+4} (rouge), ETM^{+7} (vert), ETM^{+2} (bleu)) d'Anikro-Kadiokro

4.2. Classification par la théorie des fonctions de croyance

Afin de réaliser la fusion des informations liées au cuirassement par la théorie des fonctions de croyance, nous avons écrit des codes de fonctions et un programme associés dans lequel nous avons fait varier k (nombre des plus proches voisins considérés) et α_0 (coefficient d'affaiblissement des fonctions de masse). Nous avons obtenu trois matrices de confusion, suivant les valeurs de k et α_0 (tableaux 1, 2 et 3), des classes C_1 et C_2 dont les coefficients sont exprimés en pourcentage, pour évaluer la qualité de la classification par la théorie des fonctions de croyance avec les échantillons I et II.

Tableau 1 : matrice de confusion (pour $k = 2 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8$ et $\alpha_0 = 0,95 ; 0,90 ; 0,85 ; 0,80 ; 0,75$)

Réalité	Décision	
	Classe C_1	Classe C_2
Classe C_1	54,49	45,51
Classe C_2	10,87	89,13

Tableau 2 : matrice de confusion (pour $k = 9 ; 10 ; 11$ et $\alpha_0 = 0,70 ; 0,65$)

Réalité	Décision	
	Classe C_1	Classe C_2
Classe C_1	74,47	25,53
Classe C_2	5,79	94,21

Tableau 3 : matrice de confusion (pour $k = 9$ et $\alpha_0 = 0,60 ; 0,55 ; 0,50$)

Réalité	Décision	
	Classe C_1	Classe C_2
Classe C_1	95,26	4,74
Classe C_2	2,37	97,63

Ainsi, à partir des tableaux 1, 2 et 3 et des équations (14) et (15), nous avons les indices de performance dans le tableau ci après (tableau 4) :

Tableau 4 : synthèse des indices de performance

	Tableau 1	Tableau 2	Tableau 3
GCR_1	54,49	74,47	95,26
GCR_2	89,13	94,21	97,63
$ECR_1=ECR_2$	28,19	15,66	3,55

Nous constatons, à l'analyse du tableau 4, que la meilleure performance est réalisée pour le coefficient d'affaiblissement $\alpha_0 = 0,60$ et le nombre de plus proche voisins $k = 9$. Cela nous donne l'image descriptive de la zone d'Anikro-Kadiokro (figure 5) suivante :

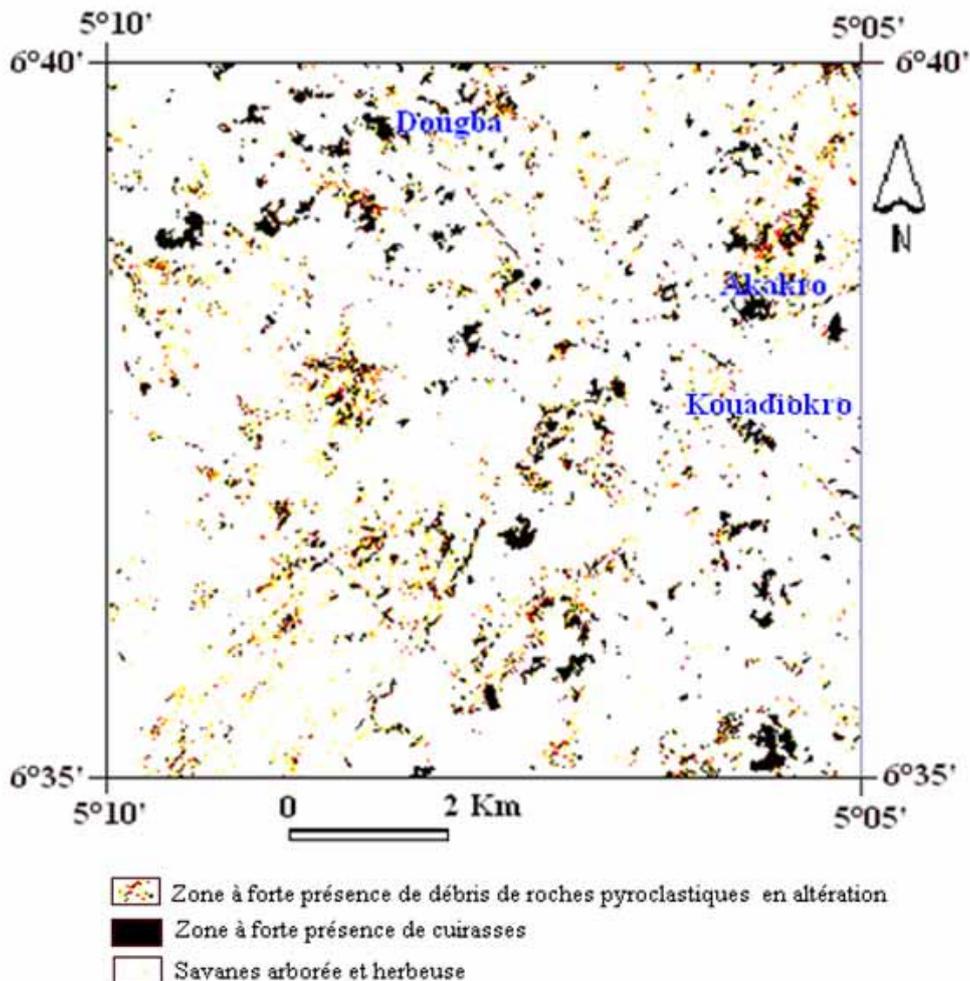


Figure 5 : Carte dérivée de la classification par la théorie des fonctions de croyance d'Anikro-Kadiokro (pour $k = 9$ et $\alpha_0 = 0,60$)

Cette figure montre, en plus des zones à forte présence de cuirasse et de savanes arborée et herbeuse, une zone en mosaïque de couleurs jaune et rouge. Après une visite de terrain, nous pouvons dire que cette mosaïque de couleurs correspond à une présence de débris de roches volcaniques (rhyodacites, dacites, tufs, etc.) à différents degrés d'altération corroborant un début de cuirassement.

Afin de tester la sensibilité de la classification aux échantillons choisis, nous avons réalisé une autre image présentant la classification avec uniquement les pixels échantillons II. Cette image a révélé dans sa généralité les mêmes entités que celles de la figure 5. Pour mieux apprécier les différences éventuelles entre les pixels, nous avons réalisé une matrice dite de comparaison pour comparer pixel par pixel les deux images. Il se résume dans le tableau ci après (tableau 5) :

Tableau 5 : matrice de comparaison

		Classification 2		Total classif 1
		Pixels cuirassés	Pixels non cuirassés	
Classification 1	Pixels cuirassés	19337	5146	24483
	Pixels non cuirassés	3069	89331	92400
	Total classif 2	22406	94477	116883

L'analyse de la matrice de comparaison (tableau5) révèle que les deux classifications effectuées ne sont pas identiques. Elle montre des différences remarquables sur l'appartenance de certains pixels à une classe donnée. En effet, nous remarquons que 4,4 % (5146 pixels) du nombre total de pixels (116883) ont été classés « cuirasse » par la classification 1 tandis que la classification 2 les classe « non cuirasse ». De plus, 2,62 % (3069 pixels) du nombre total des pixels ont été classés « cuirasse » par la classification 2 tandis que la classification 1 les classe « non cuirasse ».

Ces différences observées peuvent être amputées au choix des échantillons du fait que les images datent de 2002 ; et les superficies ayant guidé au choix de ces échantillons peuvent ne plus être les mêmes. Cependant, la *Sensibilité* = 7,02 % obtenue (équation16) est inférieure à 10 % ; alors nous pouvons dire que le programme de classification proposé est très peu sensible à la variation d'échantillons.

5. CONCLUSION

Dans cet article, nous avons proposé une application de la Théorie des Fonctions de croyance à la cartographie des cuirasses. L'approche présentée consiste à fusionner les informations des images issues de l'indice de végétation normalisé (*NDVI*), de l'indice de cuirasse (*IC*) et de la composition colorée RVB (*ETM⁺4 (R)*, *ETM⁺7 (V)* et *ETM⁺2 (B)*).

Ces informations ont été modélisées par des fonctions de croyance fondées sur un modèle de distance, en l'occurrence la règle des *k* plus proches voisins.

L'originalité de cette étude réside dans l'utilisation de la théorie des fonctions de croyance pour cartographier les cuirasses à partir d'images satellitaires.

La qualité de la classification, par la théorie des fonctions de croyance à cet effet, est fonction d'une part de la représentativité des échantillons utilisés et d'autre part des paramètres *k* (nombre de plus proches voisins) et α_0 (coefficient d'affaiblissement). Dans notre cas, nous avons utilisé *k* = 9 et α_0 = 0,60. Cela a donné un taux de pixels bien classés supérieur à 90 % et une *Sensibilité*, aux échantillons utilisés, inférieure à 10 %. La méthodologie appliquée pourrait être utilisée, avec des aménagements appropriés, à d'autres fins de cartographie d'éléments géologiques.

RÉFÉRENCES

- Bracker, H. (1996). Utilisation de la théorie de Dempster / Shafer pour la classification d'images satellitaires à l'aide de données multi-sources et multi-temporelles. Thèse de l'Université de Rennes I, 178 p.
- Collet, C. et Caloz, R. (2001). *Précis de Télédétection : traitements numériques d'image de télédétection*. Vol.3. pp 229-236.
- Corgne, S. (2004). Modélisation prédictive de l'occupation des sols en contexte agricole intensif : Application à la couverture hivernale des sols en Bretagne. Thèse de l'Université de Rennes II, Haute-Bretagne, 226 p.
- DCGTx, (1993). Carte des unités morpho-pédologiques du degré carré de Gagnoa au 1 / 100 000.
- Dempster, A. (1967). "Upper and Lower probabilities Induced by Multivalued Mapping". *Ann. Math. Statist.*, pp 325-339.
- Denœux, T. (1995). "A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory", *IEEE Transactions on Systems, Man and Cybernetics*, 25 (5). pp 804-813.
- Dya, C. S. (1993). *Connaissance des cuirasses au Burkina Faso : Géomorphologie et Utilisation*. Berichte des Sonderforschungsbereichs 268, Bd. 1, Frankfurt a. M., pp 117-131.

- ENSTP (1995). Carte géologique de la feuille Gagnoa 4b.
- Foucault, A. et Raoult, J-F. (2000). *Dictionnaire de géologie*. Les éditions MASSON, 5^e édition, 380 p.
- Grandin, G., Delvigne, J. (1969). Les cuirasses de la région birmanne volcano-sédimentaire de Toumodi : jalons de l'histoire morphologique de la Côte d'Ivoire. *Comptes Rendus de l'Académie des Sciences, Paris, série D, tome 269, n°16*, pp 1477-1479.
- IGN (1964). Carte topographique de l'Afrique de l'ouest au 1 / 50 000, feuille Gagnoa 4b.
- Lefèvre, E., Vannoorenberghe, P., Colot, O. (1999). Using information criteria in Dempster- Shafer's basic belief assignment. 8th International Conference on Fuzzy Systems, FUZZ-IEEE'99, pp 173-178.
- Maignien, R. (1954). Différents processus du cuirassement en A.O.F. Extrait des comptes rendus de la 2^e Conférence Inter africaine des sols, Léopoldville, document n°116, pp 1469-1486.
- Mainguet, M. (1976). Quelques aspects de la détection et de l'insertion spatiale des phénomènes de cuirassement sur les images aériennes et satellites. Extrait des comptes rendus de la table ronde du CEGET sur la géomorphologie des reliefs cuirassés dans les pays tropicaux chauds et humides. Vol.33, pp 209-229.
- Martin, A. (2005). La fusion d'informations, Rapport Technique, ENSIETA, 109 p.
- Martin, A., Laanaya, H., Arnold-Bos, A. (2006). "Evaluation for Uncertain Image Classification and Segmentation", *Pattern Recognition*, Vol.39, issue 11, pp 1987-1995.
- Milisavjevic, N., Bloch, I., Van Den Broek, S. and Acheroy, M. (2003). "Improving mine recognition through processing and Dempster-Shafer fusion of ground-penetrating radar data". *Pattern Recognition*, vol 36, pp. 1233-1250.
- Piegon, L., Solaiman, B., Toutin, T. and Thomson, K.P.B. (2000). "Dempster-Shafer theory for multi-satellite remotely sensed observations". *Sensor Fusion: Architectures, Algorithms and Application IV*, Orlando, USA, pp. 228-236.
- Pouchin, T. (2001). Filtres et Indices. Disponible auprès de : <http://tpouchin.dub.fr/chapitre/filtres.htm>. Consulté en Avril 2007.
- Scanvic, J. Y., Pion, J. C. (1996). « La télédétection en géologie : applications thématiques ». *Précis de télédétection*, vol. 2. Presse de l'université du Québec, Sainte Foy, Canada, pp 16-56.
- Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press, New Jersey, pp 99-107.
- Smets, P. (1993). "Belief functions : the disjunctive rule of combination and the generalized Bayesian theorem". *International Journal of Approximate Reasoning* n° 9, pp 1-35.
- Smets, P. (1990). "Constructing the pignistic probability function in a context of uncertainty". *Uncertainty in Artificial Intelligence*. Elsevier Science Publishers, Vol. 5, pp 29-39.
- Sudano, J. (2002). Inverse Pignistic Probability Transforms. Proceedings of the 5th International Conference of Information Fusion, (Fusion 2002), Annapolis, MD, USA, pp 1277-2183.
- Yacé, I. (1976). Le volcanisme éburnéen dans les parties centrale et méridionale de la chaîne précambrienne de FETTEKRO en Côte d'Ivoire. Thèse de Doctorat d'Etat, Université de Cocody-Abidjan, Côte d'Ivoire, 373 p.
- Yager, R. (1987). "On the Dempster-Shafer Framework and new combinaison rules". *Informations sciences*, 41: pp 93-137.
- Yager, R. (1986). "A general approach to decision making with evidential knowledge". In *Uncertainty in Artificial Intelligence*, Kanal L. et Lemmer J. Eds, Elsevier Science Publishers, pp 317-327.
- Yao-Kouamé, A. (2008). Etude des sols brunifiés dérivés des matériaux du complexe volcano-sédimentaire de Toumodi en moyenne Côte d'Ivoire. Thèse de Doctorat d'Etat, Université de Cocody-Abidjan, Côte d'Ivoire, 210 p.