Cartographie des écosystèmes et paramètres biophysiques satellitaires pour l'étude des flux hydriques sur le continent africain

Armel Kaptue

▶ **To cite this version:**

Armel Kaptue. Cartographie des écosystèmes et paramètres biophysiques satellitaires pour l'étude des flux hydriques sur le continent africain. Hydrologie. Université Paul Sabatier - Toulouse III, 2010. Français. NNT : . tel-00547510

HAL Id: tel-00547510

https://theses.hal.science/tel-00547510

Submitted on 16 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Discipline ou spécialité :

Télédétection et Physique de la biosphère

Présentée et soutenue par :
Armel Thibaut KAPTUE TCHUENTE

le : mercredi 10 novembre 2010

Titre :

Cartographie des écosytèmes et paramètres biophysiques pour l'étude des flux hydriques sur le continent africain

---

JURY
Jean-Philippe GASTELLU-ETCHEGORRY (Président)
Agnès BEGUE (Examinatrice)
Steven DE JONG (Examinatrice)
Catherine MERING (Examinatrice)
Catherine OTTLE (Examinatrice)

Ecole doctorale :
Sciences de l'Univers, de l'Environnement et de l'Espace (SDU2E)

Unité de recherche :
Météo France/CNRM/GMME/VEGEO

Directeur(s) de Thèse :

Jean-Louis ROUJEAN

Rapporteurs :
Eric LAMBIN
Eric SERVAT

Remerciements

Il est très difficile de remercier sur une seule page toutes les personnes qui m'ont soutenu/supporté de façon directe ou indirecte pendant ces années passées à effectuer cette thèse.

Tout d'abord, je ne trouve pas les mots pour exprimer ma gratitude à mon superviseur Jean-Louis Roujean pour m'avoir fait confiance en acceptant de me diriger (alors que l'on ne s'était jamais rencontré). Il ne m'a pas seulement introduit dans le monde de la recherche, mais il m'a aussi soutenu depuis le début de ce travail en me prodiguant des conseils et en me montrant la voie à suivre. Je suis également très reconnaissant envers Aaron Boone pour son rôle de co-superviseur durant ma thèse.

Effectuer ma thèse au Centre National de Recherches Météorologiques a été l'une de mes expériences les plus excitantes. Les personnes rencontrées et les expériences que j'ai eues continuent à m'influencer à la fois personnellement et académiquement. Tout ceci n'aurait pas été possible sans les personnes qui m'ont initié dans le domaine de la télédétection. Je pense notamment à Mathieu Kientga du 2iE (Ouagadougou), Bernard Lortic de l'IRD (Adis-Abeba). Un merci également va à Christophe François de l'université Paris-Sud pour m'avoir donné les pistes nécessaires pour la recherche d'une bourse doctorale.

Plusieurs contacts ont permis l'évolution de mon travail. Je tiens à remercier particulièrement les membres de mon comité de thèse Agnès Bégué du CIRAD (Montpellier), Steven De Jong de l'Université d'Utrecht, Catherine Mering de l'université Paris Diderot et Laurent Kergoat du LMTG (Toulouse) pour les discussions/critiques intéressantes effectuées lors des différentes réunions de mon comité de thèse. Je remercie également Jean-François Mahfouf du CNRM pour son indéfectible soutien.

J'ai également bénéficié au CNRM du support technique et de discussions de plusieurs collègues dont je ne pourrai dresser une liste exhaustive ici ; néanmoins je voudrai remercier spécialement Stéphanie Faroux, non seulement pour m'avoir accepté dans son bureau, mais encore d'avoir pu me supporter au cours de ces trois dernières années.

Finalement cette dissertation n'aurait pas été possible sans le support et la compréhension des personnes les plus importantes dans ma vie : ma famille.

Table des matières

Introduction ____ *1*

I. Contexte scientifique ____ *11*

1.1 Objectifs scientifiques ____ **11**

1.2 Etat de l'art ____ **14**

1.3 Caractéristiques du continent Africain ____ **22**

1.4 Principales données satellitales utilisées ____ **33**

1.5 Description des besoins en météorologie ____ **39**

II. Forçage physiographique sur la zone 'AMMA' ____ *43*

2.1 Introduction ____ **43**

2.2 Choix et prétraitements des données ____ **45**

2.3 Méthode de classification supervisée ____ **51**

2.4 Estimation des variables biophysiques de surface ____ **55**

2.5 Evaluation de la nouvelle physiographie ____ **59**

2.6 Résumé et conclusion ____ **68**

III. Le produit ECOCLIMAP-II sur l'Afrique ____ *71*

3.1 Introduction ____ **71**

3.2 Prétraitements des données d'observations de la terre ____ **75**

3.3 Classification non supervisée ____ **77**

3.4 Validation des produits ____ **90**

3.5 Résumé et conclusion ____ **113**

IV. Sensibilité de la modélisation des flux hydriques dans ISBA au forçage physiographique ____ *117*

4.1 Introduction ____ **117**

4.2 Le modèle de surface ISBA standard ____ **119**

4.3 Données et méthodologie ____ **122**

4.4 Distribution spatiale des différences à l'échelle régionale ____ **124**

4.5 Résultats ____ **125**

4.6 Discussion et conclusion ____ **134**

Conclusion ____ *137*

Bibliographie ____ *143*

Liste alphabétique des sigles ____ *159*

Liste alphabétique des symboles ____ *163*
Liste des tableaux ____ *167*
Annexe A : Systèmes spatiaux d'acquisition d'images ____ *173*
A.1 Bref historique ____ **173**
A.2 Termes et concepts de base ____ **174**
Annexe B : Evaluation d'une classification ____ *183*
Annexe C : Interpolation ____ *185*
Annexe D : Bilans hydriques et énergétiques dans ISBA ____ *187*
D.1 Bilan énergétique ____ **187**
D.2 Bilan hydrique ____ **188**
Annexe E : Production scientifique ____ *191*

« La survie, c'est la capacité à affronter les difficultés de circonstance et à les surmonter », Nelson Mandela, Prix Nobel de Paix, 1993.

Introduction

Les différents changements planétaires (ou changement global), induits dans notre environnement, comprennent : les changements climatiques, les changements survenant dans l'atmosphère, les changements d'occupation et d'utilisation du sol, les changements des cycles bio-géochimiques (Goudie et Cuff, 2002). La décennie 1990-2000 a été celle de la mise en œuvre de grands programmes internationaux de recherche sur l'environnement qui ont permis de comprendre toute la complexité du système de l'environnement terrestre. Celui-ci se caractérise notamment par une imbrication forte des échelles spatiales et temporelles, du local au global, qui, entre autre, relie entre eux les différents problèmes d'environnement : changement climatique, ressources en eau, biodiversité, pollution de l'air, des sols, des océans.

L'International Geosphere Biosphere Program (IGBP) est un programme, mis en place dès 1986, dont l'objectif est d'étudier le changement global, et en particulier les changements dûs aux actions anthropiques. Il permet : (i) d'observer et de documenter les changements du système terrestre, (ii) de comprendre pourquoi ces changements apparaissent, (iii) d'apporter des connaissances supplémentaires qui permettent de valider scientifiquement les appréciations sur les questions du changement global environnemental, et (iv) d'améliorer les prévisions futures du changement global. Pour faire face aux conséquences non seulement environnementales mais socio-économiques induites par les changements planétaires, d'autres programmes internationaux à l'échelle globale ont également vu le jour. Le Global Climate Observing System (GCOS) vise à long terme la mise sur pied d'un système opérationnel capable de produire des observations compréhensibles requises pour la surveillance du système climatique. Le World Climate Research Programme (WCRP) vise à améliorer les prévisions climatiques et la connaissance de l'action anthropique sur le climat. Le Global Energy and Water Cycle Experiment (GEWEX) a pour mission l'observation, la surveillance et la modélisation du cycle hydrologique et des

fluxd'énergie dans l'atmosphère terrestre et à la surface. Le Global Monitoring for Environment and Security (GMES) qui a pour objectif principal de surveiller et de mieux comprendre notre environnement.

Ayant depuis longtemps cessé d'être seulement une curiosité scientifique, les changements climatiques représentent une menace pour la croissance et le développement durable en Afrique et un obstacle à l'accomplissement d'ici 2015 des Objectifs du Millénaire pour le Développement (OMD), notamment à cause de :

(i) l'expansion de certaines maladies telles que le choléra, le paludisme, l'onchocercose, la méningite (McMichael et al., 2006 ; Molesworth et al., 2003 ; Patz et al., 2005). Cependant, il n'existe pas un consensus au sein de la communauté scientifique entre les changements climatiques et l'expansion des maladies. Alors que certaines études avaient associé la transmission du paludisme à des anomalies de température maximales dans la région de l'Afrique de l'Est (e.g. Githeko et Ndegwa, 2001), Small et al. (2003) recommandent l'analyse des variables autres que climatiques telles que la résistance aux médicaments, la migration humaine et le statut immunitaire afin de pouvoir établir une tendance à long terme ;

(ii) la fragilisation des systèmes sociaux par la migration des populations qui peuvent induire dans le cas des populations nomades leur sédentarisation et par conséquent la perte de leur identité culturelle, de leurs institutions coutumières. Ces migrations ont eu à engendrer par le passé des conflits politiques tel que celui de la guerre civile au Darfour (à l'ouest du Soudan) où la combinaison de la désertification et de plusieurs décennies de sécheresse a contraint des bergers ou populations nomades à migrer vers de meilleures terres appartenant à des cultivateurs ou populations sédentaires (Nordas et Gleditsch, 2007). Ceci a incité plusieurs chercheurs à mener des études afin de déterminer le lien entre la genèse des conflits et la dégradation des conditions environnementales (e.g.

Burke et al., 2009 ; Kahl, 2006). Cependant, ces études n'ont pu aboutir à des conclusions irréfutables et ont donc fait l'objet de contestations (Benjaminsen, 2009 ; Sutton et al., 2010).

(iii) l'érosion de la biodiversité des écosystèmes à travers l'extinction de certaines espèces fauniques (Simmons et al., 2005), floristiques (Mclean et al., 2005), marines qui peuvent être d'origine animale tels que les récifs coralliens ou d'origine végétale tels que les mangroves (Nelleman et Concoran, 2006). Cependant la compréhension des effets liés au changements climatiques sur la biodiversité ne sont pas encore suffisamment connus (Thomas et al., 2004) ; d'où la bipolarisation du débat ; tandis que les preneurs de décisions ont initialement travaillé sur l'hypothèse selon laquelle la désertification est une fonction de la planification des sols et du surpâturage, les écologistes suggèrent que les changements de la végétation sont des réponses aux fluctuations naturelles de la pluviométrie et que les bas niveaux sont une réponse adaptative à la variabilité climatique (Lovett et al., 2005) ;

(iv) la surdépendance de la production agricole africaine à plus de 95% de la pluviométrie car le rendement est sensible à la variabilité temporelle du début et de l'arrêt de la pluviométrie ou de la période de croissance. Ce qui accroît les risques d'insécurité alimentaire car en 2002 déjà, la moitié de la population de la planète souffrant de malnutrition résidait sur le continent africain (WHO, 2002). Etant à la base du PIB dans la majorité des pays africains, la production agricole est en outre essentielle pour la subsistance et le bien-être des populations (FAO, 2008 ; Nelson et al., 2009) ;

(v) la fluctuation des ressources hydriques qui affecte inéluctablement la disponibilité en eau douce. Trois-quarts des pays africains sont dans des zones où de de légères variations de la pluviométrie peuvent provoquer une baisse drastique du niveau des rivières (De Wit and Stankiewicz, 2006). Ce qui accentue ainsi le

problème d'accès à l'eau potable qui est peut être le défi environnemental le plus préoccupant auquel l'Afrique doit faire face. Déjà en l'an 2000, environ 40% des populations africaines n'avait pas accès à l'eau potable (Vörösmarty et al., 2005). En outre, alors que la consommation des besoins en eau en Afrique ne fait que s'accroître, les ressources du continent ne font que baisser suite à plusieurs séries de sécheresse et aux changements dans les modes de gestion des terres.

Ainsi, force est de constater que les conséquences dûes aux changements climatiques ne sont/seront pas les mêmes pour tous. Bien que l'Afrique génère moins de 4 % des émissions de gaz à effet de serre (GES) produites dans le monde (à titre d'exemple, un citoyen africain génère 13 fois moins de GES qu'un nord-américain), c'est paradoxalement le continent le plus vulnérable aux effets des changements climatiques (Stern, 2007). C'est dans ce contexte que s'inscrivent des initiatives spécifiques ayant une durée limitée dans le temps comme AMESD (African Monitoring of the Environment for Sustainable Developpement) et AMMA (African Analysis Multidisciplinary Analysis).

AMESD (www.amesd.org) est un programme de coopération internationale ayant pour objectifs l'amélioration de : (i) la gestion intégrée des ressources en eau, (ii) la gestion des cultures et des pâturages, (iii) la gestion des ressources marines et des zones côtières, (iv) la gestion des ressources agricoles et environnementales et (v) la protection de l'habitat naturel et la lutte contre la dégradation des sols. AMESD a débuté en 2007 et s'achèvera en 2012. Le projet AMMA (http://www.amma-international.org), a débuté en 2001 et vise à améliorer notre connaissance et notre compréhension des variabilités intra-saisonnière et inter-annuelle de la mousson africaine en prenant en compte tous les facteurs, qu'ils soient d'origine naturelle ou anthropique (Redelsperger et al., 2006). Le projet est motivé par la forte variabilité des précipitations associées à ce système de mousson, et par ses conséquences sur la sécurité alimentaire, les ressources en eau et la santé des populations.

Afin de répondre aux multiples échelles qui caractérisent la mousson africaine, le programme AMMA a été structuré autour de 4 échelles spatiales dont les frontières entre-elles sont assez perméables :

(i) l'échelle globale à laquelle la mousson interagit avec le reste du globe ;

(ii) l'échelle régionale qui s'étend jusqu'à quelques centaines de kilomètres ;

(iii) l'échelle du paysage s'étendant de la centaine de mètres à quelques dizaines de kilomètres ;

(iv) l'échelle locale d'une étendue du centimètre à la centaine de mètres.

Un des enjeux majeurs de la surveillance du changement climatique est le fonctionnement des surfaces continentales qui comprend le suivi de l'occupation du sol, c.-à-d. l'évolution urbaine, l'extension des cultures ou encore la désertification au Sahel. La surface étant en lien étroit avec l'atmosphère, de grandes évolutions de couverture des sols peuvent induire localement un changement de climat (e.g., Taylor et al., 2002). Ce suivi de l'occupation des sols passe donc par la caractérisation de l'hétérogénéité de la surface. Initialement étudiée en écologie (distribution d'une espèce vivante), l'hétérogénéité de la surface est devenue un centre d'intérêt de diverses disciplines telles que l'hydrologie (structure spatiale d'un bassin versant), géographie (distribution spatiale de la démographie), épidémiologie (répartition spatiale des vecteurs d'une maladie), la télédétection (variabilité spectrale). Bien que le concept d'hétérogénéité soit une notion complexe à définir de façon univoque (Kolasa et Rollo, 1991), l'état de l'art la décrit selon deux approches : l'approche structurelle et l'approche fonctionnelle (e.g. Gomez et al., 2004). Selon l'approche structurelle, un phénomène est dit hétérogène si une ou plusieurs variables le représentant varient dans l'espace (e.g. une classe d'occupation des sols). Par contre, selon l'approche fonctionnelle, un phénomène est dit hétérogène si l'intensité du processus sous-jacent varie dans l'espace en réponse à des variations structurelles du système (e.g. la densité de la végétation au sein d'une classe d'occupation des sols).

Dans cette étude, la caractérisation de l'hétérogénéité de surface se rapporte à l'identification des descripteurs structuraux et fonctionnels des surfaces et des couverts de même que leurs variations temporelles (saisonnière, pluri-annuelle) à diverses échelles (locale, paysagère, régionale, continentale, globale). L'hétérogénéité de la surface peut être quantifiée à l'aide d'une classification d'occupation des sols. C'est dans cette optique que la première carte globale d'occupation du sol à 8 km de résolution spatiale a été produite (Defries et al., 1995). La réalisation de cette carte a été effectuée à l'aide d'images acquises par le capteur Advanced Very High Resolution Radiometer (AVHRR) à bord de la série des plates-formes météorologiques de la National Oceanic and Atmospheric Administration (NOAA). Ces données ont par la suite permis la production des cartes à 1 km de résolution, notamment la carte IGBP du service géologique américain (Loveland et al., 2000), la carte de l'université du Maryland (Hansen et al., 2000) et le produit ECOCLIMAP-I de Météo France (Masson et al., 2003). L'avènement récent des capteurs de dernière génération (VEGETATION sur les satellites SPOT4 et SPOT5, MERIS sur ENVISAT, MODIS sur TERRA et AQUA) a permis d'accroître le nombre d'initiatives pour la cartographie de l'occupation du sol du continent africain avec notamment le produit GLC2000 à 1 km de résolution élaboré par le Joint Research Centre de l'UE (Mayaux et al. 2004), la carte GLOBCOVER à 300 m de résolution réalisée par l'ESA -agence spatiale européenne- (Arino et al. 2008) et la carte annuelle MODIS de la NASA –agence spatiale américaine- à 500 m et à 1 km de résolution (Friedl et al., 2002 ; Friedl et al., 2010). Les classes d'occupation des sols ainsi produites, ont été utilisées lors de la description de la surface en termes de fraction de végétation (*veg*) (Zeng et al., 2000), des albédos de la surface et de la végétation (Houldcroft et al., 2009), d'indice foliaire ou leaf area index (LAI) et de fraction de rayonnement actif photosynthétiquement absorbé par la canopée ou fraction of absorbed photosynthetically active radiation by the canopy (fAPAR) (Yang et al., 2006) à l'échelle du continent africain. Bien que toutes ces initiatives aient été réalisées afin de pouvoir déterminer la distribution spatiale et géographique des différents types d'occupation de sols et quantifier les variables biophysiques, on note de grandes disparités entre ces différents

produits à l'échelle du continent africain ; d'où la nécessite de générer des produits qui seront par la suite utilisés comme standard. McCallum et al. (2006) ont mis en évidence la faible concordance existant entre les produits IGBP-Dis, UMD, MODIS et GLC2000 dans un domaine de 5°x5° au sein du bassin du Congo tandis que Garrigues et al. (2008) ont montré des inconsistances spatio-temporelles des données de LAI, par classe d'occupation des sols entre ECOCLIMAP-I et MODIS, sur un transect latitudinal du continent. Ces disparités peuvent être considérées comme des informations complémentaires induites par chaque produit et ensuite combinées afin de générer des produits pour des applications spécifiques (Jung et al., 2006). L'utilisation des systèmes d'observations de la terre pour le fonctionnement des surfaces continentales et pour des fin de surveillance et de gestion de l'environnement est donc indispensable.

Les données récentes de télédétection à faible résolution spatiale et à haute répétitivité temporelle permettent de mieux caractériser l'hétérogénéité de surface sur de vastes domaines tel que l'Afrique. Cette caractérisation s'effectue par le développement de méthodes appropriées de classification de l'occupation des sols et d'estimation des variables biophysiques. Les résultats ainsi obtenus devront permettre d'étudier l'impact de la physiographie du continent africain dans les modèles biogéoclimatiques et d'avoir une meilleure compréhension de la variabilité climatique en Afrique. L'objectif de cette thèse vise une meilleure caractérisation de l'hétérogénéité de surface du continent africain. Pour cela, nous nous sommes focalisés sur l'identification des classes d'occupation des sols ainsi que sur l'estimation de variables biophysiques à l'échelle continentale. La faisabilité de cette étude a d'abord été testée sur la région ouest-africaine dans le cadre du programme AMMA avant d'être étendue à l'intégralité du continent africain. Des tests de sensibilité de la nouvelle paramétrisation de la surface en Afrique de l'Ouest ont été effectués à l'aide du schéma de surface ISBA (Interactions entre le Sol, la Biosphère et l'Atmosphère) (Noilhan et Mahfouf, 1996). ISBA, permet de simuler les échanges d'eau et d'énergie entre le sol, la végétation et les basses couches de l'atmosphère.

Plan du manuscrit

Cette étude se décompose en 4 chapitres. Le chapitre 1 décrit le contexte scientifique ainsi que les motivations de cette étude. Le chapitre 2 présente une nouvelle paramétrisation de la surface à utiliser dans les schémas de surface pour l'Afrique de l'Ouest, comprenant une carte d'occupation des sols (obtenue par classification supervisée) et des séries temporelles de LAI, d'albédo et de *veg*. Le chapitre 3 présente la nouvelle base de données ECOCLIMAP (ou ECOCLIMAP-II dans tout ce qui va suivre) mise à jour sur l'ensemble du continent africain. La discrimination des classes d'occupation des sols a été effectuée à l'aide d'une classification hybride automatique basée sur la Transformée de Fourier et le classificateur des 'k plus proches voisins'. Toujours au chapitre 3, une inter-comparaison détaillée des différentes bases de données d'occupation des sols et des variables biophysiques existantes à l'échelle continentale (y compris la nouvelle paramétrisation ECOCLIMAP-II) a été effectuée afin d'illustrer les similarités et différences des produits. Le chapitre 4 présente les formalismes de base du modèle de surface ISBA et l'impact du nouveau forçage physiographique élaboré au chapitre 2 dans ledit modèle.

« Nous n'héritons pas de la terre de nos ancêtres, nous l'empruntons à nos enfants », A. de Saint-Exupéry, Terre des hommes, 1939.

I. Contexte scientifique

1.1 Objectifs scientifiques

Le climat contemporain africain est en général similaire à celui qui prévalait il y a 2000 ans, mais avec l'apparition alternée des épisodes plus arides ou plus humides durant les deux derniers siècles (Nicholson, 2001). Une période humide s'est par exemple installée du IX^{ème} au XIV^{ème} siècle en Afrique du Nord, du XVI^{ème} au XVIII^{ème} siècle au Sahel et également dans les autres parties de l'Afrique vers la moitié du XVIII^{ème} siècle. A la fin du XIX^{ème} siècle, des conditions anormalement sèches voient le jour et persistent jusqu'au début du XIX^{ème} siècle. Après le retour à des conditions relativement humides vers la moitié du XIXème siècle, l'an 1895 marque le déclin mondial de la pluviométrie dans les tropiques. Ainsi donc, le XX^{ème} siècle a commencé par une période d'aridité similaire, à celle de 100 ans auparavant, quoique probablement moins extrême. Ceci a abouti à des sécheresses sévères et généralisées essentiellement dans les années 1910 à l'exception des parties orientales et tempérées. Les années 1920 et 1930 ont marqué le retour à des conditions favorables en Afrique de l'ouest pendant que des conditions arides perduraient dans les parties australes et méridionales de l'Afrique. Des fluctuations plus extrêmes ont apparu dans les années 1950 par une augmentation remarquable de la pluviométrie jamais observée depuis les années 1870. La décennie 1970-1980 marque à nouveau une aridification qui aura des conséquences néfastes sur les conditions de vie des populations. Et de nouveau une phase humide depuis le début des années 1990.

Les trois composantes du système terrestre que sont l'atmosphère, l'océan et la biosphère continentale interagissent étroitement et déterminent le climat et son évolution. La biosphère continentale joue un rôle majeur dans le cycle du carbone et les processus de surface (Goudie et Cuff, 2002). Afin de modéliser et prédire le climat global, il est donc crucial de mieux décrire les processus de surface (cycle de l'eau, physiologie des plantes, cycle du carbone) dans les modèles de surface (LSM). Les LSM cherchent à produire des

simulations réalistes des transferts de masse, et d'énergie entre le sol, la végétation et l'atmosphère. Le calcul des échanges entre la surface et l'atmosphère dans les LSM, aussi bien que la prévision de l'impact potentiel du changement d'occupation des sols sur le climat est compliqué par le fait que le caractère de la surface continentale est spatialement variable. La couverture végétale, les types de terrain, les texture et humidité des sols, l'étendue des aires urbaines contribuent toutes à l'hétérogénéité de la surface. La large étendue des échelles de variabilité dans les surfaces continentales est évidente dans les cartes globales de végétation produites à partir d'images acquises dans le visible et le proche infrarouge. L'échelle de ces hétérogénéités peut être plus petit, et dans certains cas beaucoup plus, que les caractéristiques des points de grille utilisés dans la plus part des LSM utilisés dans les études climatiques (de l'ordre des dizaines de km). Presque tous les LSM emploient des techniques pour essayer de prendre en compte l'hétérogénéité au sein d'une grille.

La technique la plus couramment utilisée est l'approche dite 'par tuiles' selon laquelle le paysage est représenté comme une mosaïque de quatre types principaux de surface : la mer, les plans d'eau internes, les surfaces bâties et les surfaces naturelles (figure 1.1). Les flux de carbone, d'eau et d'énergie sont calculés séparément pour chaque tuile (terre, mer, eau et bâti) en utilisant pour chacune d'entre elles des paramètres inhérents de surface et une spécification des proportions des types fonctionnels. Ensuite les flux sont moyennés en affectant à chaque tuile un poids identique à la proportion de la surface qu'elle recouvre. La performance de l'approche par tuiles a fait l'objet de plusieurs études (Avissar et Pielke, 1989 ; Pielke et al., 1997 ; Salmun et Molod, 2006). La communauté scientifique a développé une panoplie de modèles de surface ou Land Surface Model (LSM) durant les dernières décennies (e.g. Bonan et al., 2003 ; Dai et al., 2003 ; Krinner et al., 2005 ; Noilhan et Mahfouf, 1996 ; Saux-Picart et al., 2009). Pour une revue compréhensive sur le développement des LSM, le lecteur pourra se référer aux travaux de Pitman (2003).

Dans les modèles de surface ou LSM, les propriétés de la surface sont décrites à l'aide des variables biophysiques telles que : l'indice foliaire ou Leaf Area Index (LAI), l'albédo de

surface, la longueur de rugosité, la fraction de rayonnement photosynthétiquement absorbé par la végétation ou Fraction of Absorbed Photosynthetically Active Radiation (fAPAR), la fraction de couverture végétale (*veg*), la profondeur racinaire, la rugosité, l'émissivité, la phénologie (cycle de vie de la plante), etc (e.g. Gason et al., 2004 ; Jin et Liang, 2006 ; Maignan et al., 2008, Roujean et Lacaze, 2002 ; Shepard et al., 2001 ; Zeng, 2001).

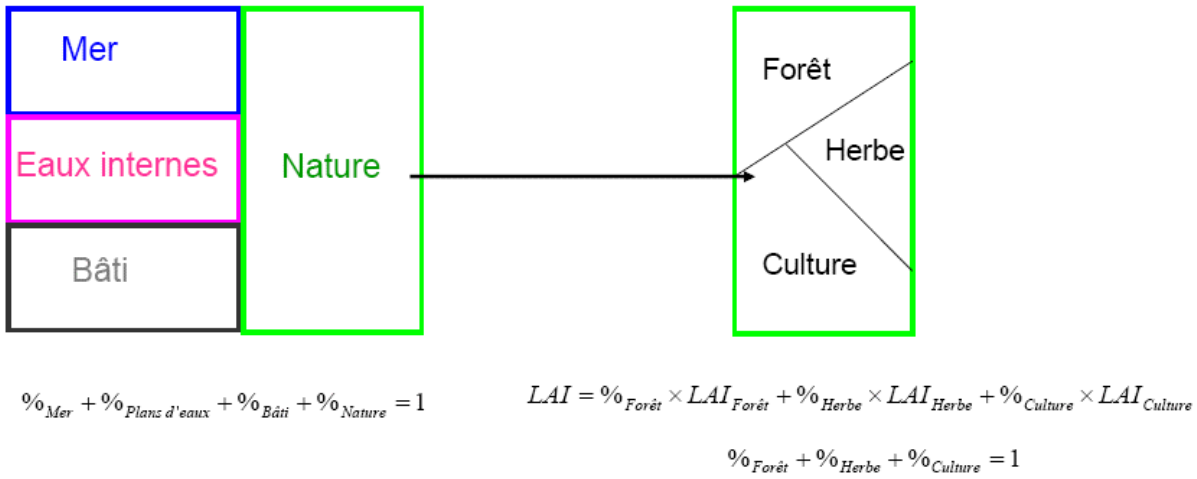

Figure I.1 Approche par tuiles dans les LSM

L'objectif général de cette thèse est de caractériser l'hétérogénéité de surface du continent africain. Il s'agit ici d'identifier les descripteurs structuraux (les classes d'occupation des sols) et les descripteurs fonctionnels (les variables biophysiques de surface) des surfaces continentales africaines. De manière spécifique, il s'agit :

(i) de cartographier l'occupation des sols en Afrique ;

(ii) d'y associer des variables biophysiques de surface ;

(iii) de montrer l'impact de la physiographie sur les bilans hydriques et énergétiques.

Parmi les variables biophysiques de surface, nous allons nous focaliser uniquement sur le LAI, le *veg*, et l'albédo car ce sont les paramètres de premier ordre des modèles de surface. L'albédo, fraction de l'énergie réfléchie du rayonnement solaire incident, module la quantité d'énergie disponible à la surface (Schaaf et al., 2002). Le LAI se définit comme la moitié de la surface foliaire par unité horizontale de sol (Stenberg, 2006). C'est un indicateur de la phénologie utilisée dans l'estimation des échanges de flux (CO_2, eau) entre la surface

et l'atmosphère. Le LAI complète l'information fournie par le *veg* qui spécifie le pourcentage de la couverture végétale (Roujean et Lacaze, 2002). Dans les modèles, cette estimation dépend généralement des types d'occupation de sols (Bonan et al., 2003 ; Dai et al., 2003 ; Krinner et al., 2005 ; Noilhan et Mahfouf, 1996). Ainsi il est crucial de disposer à la fois d'une information fiable sur l'occupation des sols et d'une représentation précise des variables biophysiques. Dans cette finalité, la télédétection spatiale est devenue un outil incontournable car elle permet d'obtenir des observations cohérentes et répétées sur de larges domaines (cf. Annexe A).

1.2 Etat de l'art

1.2.1 Cartographie de l'occupation des sols

La définition de l'occupation des sols est fondamentale parce que dans plusieurs classifications et nomenclatures existantes, elle se confond à celle de l'utilisation des sols. L'occupation des sols est le couvert biophysique observé à la surface de le terre tandis que l'utilisation des sols se caractérise par les activités anthropiques entreprises sur un type d'occupation des sols afin de le modifier ou de le maintenir (Di Grégorio et Jansen, 2000). Ainsi l'occupation des sols est un facteur déterminant de l'utilisation des sols et donc des valeurs des actions anthropiques sur l'environnement. En effet, des actions anthropiques telles que la création des clairières dans des zones forestières à des fins agricoles affecte la distribution de l'occupation des sols qui altère à son tour le fonctionnement du système climatique (Sellers et al., 1996). Au même moment le climat est le principal facteur responsable de la distribution de la végétation naturelle et donc de l'occupation des sols en l'absence d'actions anthropiques. L'importance grandissante de la planification et de la gestion environnementale a donc conduit à l'émergence de divers efforts nationaux, régionaux, continentaux et globaux de cartographie d'occupation des sols par des gouvernements et des organisations indépendantes afin de supporter le développement durable et la gestion intégrée des ressources.

Après la production au milieu des années 1980 des premières cartes globales d'occupation des sols par compilation d'atlas et de cartes existantes (Mathews, 1983 ; Olson et al., 1983 ; Wilson et Henderson-Sellers, 1985), l'imagerie satellitaire est devenue une technologie idéale pour la classification de l'occupation des sols sur de larges domaines (Cihlar, 2000). La grande disponibilité des observations spatialement étendues acquises par le capteur Advanced Very High Resolution Radiometer (AVHRR) – mis en service dès la fin des années 1970 - à bord de la série des plates-formes météorologiques de la National Oceanic and Atmospheric Administration (NOAA) a incité la communauté scientifique à développer des méthodes de caractérisation de l'occupation des sols sur de larges domaines (Tucker et al., 1985) et éventuellement à produire des cartes globales d'occupation des sols (Defries et al.,1995). Grâce à la capacité potentielle du capteur AVHRR d'observer de façon systématique diverses échelles et à la possibilité actuelle de pouvoir consulter les observations archivées de l'instant présent à des décennies antérieures, les premiers produits globaux d'occupation des sols à 1 km de résolution ont vu le jour. Loveland et al. (2000) ont développé le produit International Geosphere-Biosphere Programme Data and Information System (IGBP DISCover) continent par continent en utilisant des données composites mensuelles élaborées à partir des observations AVHRR acquises entre avril 1992 et mars 1993. Ils ont également utilisé lors de la phase d'interprétation des données auxiliaires telles que les cartes disponibles d'occupation et d'utilisation des sols, l'information topographique et climatique. Avec ces mêmes données spectrales, Hansen et al. (2000) ont produit la carte de l'Université du Maryland (UMD). La carte résultante montre une bonne séparabilité en 14 classes globales (5 classes de forêt, 2 classes de savane arborée, 2 classes de savane arbustive, 5 autres classes bien développées). Masson et al. (2003) ont généré la carte d'occupation du Centre National de Recherches Météorologiques de Météo France (ECOCLIMAP-I) en combinant plusieurs cartes globales d'occupation des sols telles que IGBP DISCover et UMD, mais également la carte CLC1993 d'occupation des sols des pays de la communauté européenne ou European Community (EC., 1993). En outre, l'information climatique issue de la carte globale de Koppe

et De Long (1958) et de la carte FIRS de la communauté européenne (EC, 1995) a aussi été prise en compte.

En raison de l'émergence de capteurs de nouvelle génération (qui possèdent forcément une meilleure résolution spatiale, spectrale, temporelle et radiométrique que l'AVHRR), un effort important a été effectué en vue d'améliorer l'information sur la distribution spatiale et les proprités des classes d'occupation des sols. Le Joint Research Center a produit la carte GLC2000 à partir des données majoritairement acquises par le capteur SPOT/VEGETATION au cours de l'année 2000 (Bartholomé et Belward, 2005). L'Agence Spatiale Européenne (ESA) a élaboré le produit GLOBCOVER en utilsant les données du capteur MERIS embarqué sur la plate-forme ENVISAT (Arino et al. 2008). Depuis le début des années 2000, la NASA génère une carte annuelle d'occupation des sols à partir des observations MODIS (Friedl et al., 2010). Le projet AFRICOVER de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO, 2004), a permis de cartographier l'occupation des sols de plusieurs pays africains à une résolution spatiale de 30 m (FAO, 2004). Les types d'occupation des sols ont été identifiés par interprétation visuelle des données LANDSAT par des experts. De ce fait, ces cartes nationales présentent un niveau de détail très élevé mais souffrent d'une large inconhérence à cause de l'hétérogénéité des dates d'acquisition et d'interprétation d'une scène à l'autre. La reconnaissance d'un comportement phénologique par l'analyse des séries temporelles d'images satellitaires ainsi que leur transcription en une carte thématique appropriée qui ont été entreprises lors des études mentionnées ci-dessus font référence à des techniques de classification (Jain et al., 1999 ; Lu et We,g, 2007 ; Nakache et Confais, 2004)

Les décisions critiques lors de ces différentes initiatives de cartographie sont liées au choix :

(i) des données utilisées en entrée notamment leur qualité et leurs caractéristiques (résolution, disponibilité) (Cihlar, 2000). La qualité des données est fonction des prétraitements initialement appliqués aux valeurs brutes afin de réduire les incertitudes en effectuant des corrections radiométriques (étalonnage, détection

des nuages/aérosols, normalisation directionnelle) et des corrections géométriques (géoréférencement, rééchantillonnage, mosaïquage). La résolution temporelle est une considération critique lors du choix des données appropriées à cause du changement d'occupation des sols dans le temps tandis que la résolution spatiale est fonction de l'échelle du domaine d'intérêt. La disponibilité ou archivage définit la période d'intérêt des données tandis que les différentes résolutions limitent l'exactitude de l'information pouvant être extraite.

(ii) des données auxiliaires (Lu et Weng, 2007) telles que le masque terre-mer, l'information texturale, topographique, climatique, démographique, etc.

(iii) de la période de couverture temporelle qui est une composante critique lors du choix des données d'entrée (Xie et al., 2008) car l'occupation des sols change dans le temps ;

(iv) d'une nomenclature adéquate pour l'usage des données de télédétection. La nomenclature d'Anderson et al. (1976) qui agrège les cultures et herbacées au niveau I ne produira pas le même nombre de classes que la nomenclature de la FAO (Di Gregorio et Jansen, 2000) qui diffère ces deux classes au même niveau ;

(v) la méthodologie de classification à cause de leur imperfectibilité (Landgrebe, 2003). En effet, il n'existe pas de méthode de classification qui puisse être utilisée de façon universelle pour tout type d'usage. Son choix est dicté par les besoins de l'utilisateur, la résolution spatiale des données d'entrée, les différentes sources de données et la nomenclature. Ainsi donc, la réponse à la question de savoir quelle approche de classification est meilleure pour une étude spécifique n'est pas aisée à fournir.

Très utilisée dans diverses applications telles que le traitement d'images, les statistiques, la biologie, mais également les sciences humaines, la classification est une opération qui a pour but de ranger les objets (individus ou variables) en se basant sur leur similarité en un nombre limité de groupes (Nakache et Confais, 2004). Chaque groupe, appelé *cluster*, comprend des objets -des pixels dans notre cas- ayant des caractéristiques similaires et

sépare des objets ayant des caractéristiques différentes (i.e. homogénéité interne et hétérogénéité externe). En accord avec Jain et al. (1999), quelque soit la méthode de classification appropriée choisie, celle-ci peut être caractérisée à l'aide de trois propriétés :

(i) le type d'algorithme pouvant être automatique ou non, et dans ce cas on parle de classification non supervisée (e.g. Loveland et al., 2000) ou supervisée (e.g. Hansen et al., 2000) respectivement. Il serait cependant fastidieux de croire en l'existence d'une classification absolument automatique, i.e. sans la moindre intervention humaine dans le processus de classification ;

(ii) la structure de la classification qui peut être hiérarchique tel que la nomenclature Land Cover Classification System (LCCS) prescrite par la FAO (Di Gregorio et Jansen, 2000) ou par partition telle que la nomenclature IGBP du service géologique américain (USGS) (Loveland et Belward, 1997);

(iii) la structure des classes qui peut être rigide (chaque comportement est assigné à une classe et une seule) (e.g. Arino et al., 2008) ou floue (chaque comportement a un degré variable d'appartenance dans une classe) (e.g. Shalan et al., 2003).

La propagation d'incertitudes et d'erreurs dans la chaîne de traitement des images est un facteur important qui influence l'exactitude de la classification. L'identification du maillon faible dans la chaîne et la réduction des incertitudes sont primordiales pour l'amélioration de l'exactitude de la classification. A cet effet, les techniques de classification, de plus en plus perfectionnées, ne cessent de voir le jour afin de déterminer les différentes variétés structurelles du paysage avec la meilleure fiabilité possible. On peut citer les classifications basées sur les réseaux de neurones (Berberoglu et al., 2000), les systèmes experts (Stefanov et al., 2001), l'approche contextuelle (De Jong et al., 2001), les ensembles flous (Shalan et al., 2003), les algorithmes basés sur la connaissance tels que le 'raisonnement évidentiel' (e.g. Sun et al., 2008) et de plus en plus on note l'apparition des combinaisons de plusieurs classificateurs (Huang et Lee, 2004 ; Lo et Choi, 2004). En ce qui nous concerne, nous ferons appel ici aux modèles non-empiétants que sont les partitions et les hiérarchies de parties car ici les objets ou pixels sont décrits par des variables quantitatives. Le lecteur

pourra se référer aux travaux de Landgrebe (2003), Lu et Weng (2008) ou Tso et Mather (2001), pour une taxonomie approfondie des spécificités des méthodes de classification d'images.

1.2.2 Estimation des variables biophysiques de surface

La modélisation des processus de surface passe par différentes étapes : identification des mécanismes, formalisation, étalonnage et validation. A l'exception de la formalisation, il est nécessaire de pouvoir disposer d'une série d'observations répétitives pour la réalisation de ces étapes. La télédétection offre donc un potentiel important car elle permet d'acquérir une quantité importante d'information sur la surface terrestre à diverses échelles spatiales et temporelles variées. Cependant, les observations satellitales ne produisent pas directement les variables d'état médiatrices encore appelées variables biophysiques qui décrivent la structure du couvert, les propriétés optiques des éléments, ou des caractéristiques dérivées utiles au bilan radiatif dans les LSM. Ce qui a conduit la communauté scientifique à mettre au point une panoplie de méthodes permettant d'estimer les caractéristiques biophysiques de la surface à partir des observations satellitales. Ces méthodes peuvent être classifiées en deux groupes : les méthodes statistiques et les méthodes physiques (Baret et Buis, 2007 ; Liang, 2004).

Les méthodes statistiques sont principalement basées sur des relations fonctionnelles empiriques avec une diversité d'indices de végétation (IV) (i.e., combinaison algébrique des réflectances de surface) qui fournissent une information sur la structure et l'état de la végétation. Ces relations ont été spécifiquement calibrées pour des types de végétation distincts en utilisant des meures in situ et des réflectances mesurées (Fernandes et al., 2003 ; Houldcroft et al., 2009 ; Los et al., 2000 ; Noilhan et Mahfouf, 1996 ; Masson et al., 2003). Parmi les IV les plus communément utilisés, on peut par exemple citer le ratio vegetation index (RVI) (Knipling, 1970), le normalized difference vegetation index (NDVI) (Rouse et al., 1973), le soil adjusted vegetation index (SAVI) (Huete, 1988), l'atmospherically resistant vegetation index (ARVI) (Kaufman et Tanré, 1992), et le global environment-

monitoring index (GEMI) (Pinty et Verstraete, 1992). De tous ces indices, le NDVI, défini comme étant la différence entre le proche infrarouge et le rouge divisé par la somme des deux (Rouse et al., 1973), est sans aucun doute le plus usité car il s'aménuise de la sensibilité à des facteurs exogènes tels que le sol, l'atmosphère et les effets angulaires ; facteurs essentiels lors de la discrimination des classes d'occupation des sols. En outre le NDVI répond favorablement aux changements de la quantité de biomasse verte (Hill et Donald, 2003), du contenu en chlorophylle (Dawson et al., 2003), à la présence de feu (Gabban et al., 2006), et à la variabilité climatique (Anyamba et al., 2002 ; Philippon et al., 2007). Les méthodes statistiques supposent que les variations des réflectances de surface sont dues aux variations d'une seule variable biophysique, et négligent l'impact des variations des conditions atmosphériques, des illuminations d'éclairement et de visée, des propriétés du sol, etc. De telles méthodes sont relativement faciles à implémenter et peuvent fournir des résultats optimaux lorsqu'elles sont appliquées à des échelles locales d'un site de validation, où des d'intensives mesures in situ ont été effectuées.

Les modèles physiques constituent une alternative aux méthodes statistiques et elles sont généralement basées sur l'usage des modèles de transfert radiatif par résolution numérique de l'équation du transfert radiatif afin d'établir une relation entre les variables biophysiques et les réflectances de surface pour des comportements donnés de réflectance du sol, des conditions d'éclairement et de visée, des propriétés optiques de la végétation et d'autres paramètres. Les variables biophysiques sont soit issues des simulations des modèles de transfert radiatif de la canopée (e.g. Myneni et al., 1999), soit basées sur des techniques d'inversion des modèles de transfert radiatif simulant les variables d'intérêt à partir des caractéristiques structurelles de la canopée (e.g. Baret et al., 2007). Les tables de correspondance (Knyazikhin et al., 1998) et les réseaux de neurones (Bacour et al., 2006) sont les principales techniques d'inversion. Pour cette approche, des vérités terrain sont principalement utilisées pour calibrer les paramètres des modèles. Les majeurs inconvénients de la méthode sont (i) la collecte laborieuse de certains paramètres du modèle

tels que l'information sur la structure de la canopée, et (ii) la non unicité de la solution de l'équation du transfert radiatif.

Indépendamment de l'approche utilisée, trois sources majeures d'incertitudes affectent l'estimation des variables biophysiques à partir des surfaces de réflectance (Garrigues et al., 2008 ; Yang et al., 2006). A cause de la nature mal-posée du problème d'inversion, l'estimation des variables biophysiques de surface à partir des réflectances de surface est instable, i.e. de légères variations de la réflectance de surface peuvent induire de larges variations des variables biophysiques estimées. La seconde source d'incertitude est liée à la manière avec laquelle l'architecture de la canopée est représentée dans l'algorithme d'estimation des variables biophysiques (e.g. prise en compte ou non de du phénomène de groupage). La dernière source d'incertitude concerne l'applicabilité de l'algorithme à une gamme de type de végétation et de conditions environnementales. D'ailleurs, le nombre des classes d'occupation des sols peut être assez faible pour représenter la variabilité globale de la structure de la végétation.

Afin d'éliminer les résidus atmosphériques, la contamination par les nuages et éventuellement les effets d'illumination et de visée (Schaaf et al., 2002 ; Roujean et al., 1992), des étapes de pré- ou post- traitement peuvent être appliquées pour la fusion de l'information disparate et complémentaire (tant sur le plan spatial que spectral) produite par les capteurs multi-spectraux au niveau du pixel (Pohl et Van Genderen, 1998), notamment:

(i) les combinaisons arithmétiques qui permettent d'améliorer le contraste à l'aide des additions/multiplications ou de détecter des changements à l'aide des différences/rapports ;

(ii) l'analyse en composantes principales utilisée lors du cryptage, de la compression, de la détection des changements et de la dimensionnalité multitemporelle des images ;

(iii) les filtres passe-haut qui mettent en valeur la résolution spatiale en combinant l'information spatiale et spectrale des données multispectrales ;

(iv) la substitution des variables par une régression qui est utilisée pour déterminer une combinaison linéaire des images en maximisant la corrélation avec les données individuelles ;

(v) les ondelettes qui créent une image comme combinaison linéaire de fonctions élémentaires à partir d'une des fonctions arbitraires d'un espace fini.

1.3 Caractéristiques du continent Africain

D'une superficie de 30 221 532 km^2, l'Afrique représente 6% du globe terrestre et 20% des terres émergées. Le continent africain traverse l'équateur depuis les zones tempérées du Nord à celles du Sud. La mer Méditerranée au nord induit un climat méditerranéen tandis que d'autres climats sont créés avec la proximité de l'océan Indien au sud-est et celle de l'océan Atlantique à l'ouest. Cela fait de l'Afrique un continent immensément riche sur le plan climatique.

1.3.1 Géomorphologie et pédologie

Le continent africain possède une physiographie unique (figure 1.2) comme le révèlent ses vastes plaines et plateaux ainsi que sa forme en Y intégrée à la grande vallée du Rift (système censé avoir été causé par le mouvement des plaques continentales). La vallée du Rift commence à partir de la mer Rouge et s'étend à travers les hauts plateaux éthiopiens dans la région du lac Victoria, où il se subdivise en des segments Est et Ouest et se poursuit vers le sud par le lac Malawi, au Mozambique. La topographie du continent met en évidence deux zones de haute et de basse altitude qui sont séparées par une ligne reliant le nord de l'Angola au nord-ouest de l'Ethiopie (Stock, 2004). Au nord-ouest de cette ligne, l'altitude moyenne s'élève de moins de 500 m du niveau de la mer tandis qu'au sud-est, les altitudes sont comprises entre 1000 et 2000 m au-dessus du niveau de la mer. Cette apparente régularité se structure autour de trois cratons qui sont des zones d'anciennes formations montagneuses : le craton nord-ouest situé dans la partie occidentale du Sahara ; le craton

Figure I.2 Topographie (d'après UNEP, 2008)

du Congo à l'ouest de l'Afrique Centrale ; et le craton du Kalahari situé en Afrique australe. La plupart des hauts plateaux et montagnes sont le résultat de récentes activités volcaniques. On dénombre par exemple en Afrique de l'Est des montagnes comme le Kilimandjaro (5895 m), le Kenya (5200 m), l'Elgon (4321 m), des hauts plateaux éthiopiens comme le Ras Dashen (4573 m) et en Afrique Centrale des montagnes comme le Cameroun (4070 m).

L'agriculture étant l'activité économique prépondérante pour la majorité des pays, le sol est probablement la ressource la plus importante. En effet, l'agriculture est pratiquée par plus de 60% de la population africaine et elle constitue plus de 30% du produit intérieur brut (World Bank, 2006). À quelques exceptions près, la plupart des sols sont difficiles à cultiver même si des améliorations peuvent être faites pour en augmenter la fertilité naturelle (Eswaran et al., 1996). Toutefois, les meilleurs sols d'un point de vue du potentiel agricole se

trouvent dans les vallées des principaux fleuves, en raison des dépôts d'alluvions qui s'y accumulent, ainsi que dans les zones tropicales humides en raison de la couverture forestière et de la décomposition rapide des matières organiques. Les principaux types de sol se nomment arénosols, leptosols, calcisols, ferralsols, sables mouvants et dunes, cambisols, lixsols (figure 1.3). De par leur minéralogie, ces sols peuvent être argileux et donc incultes à la culture ou alors sableux et donc très perméables (figure 1.4). L'Afrique est principalement composée d'une masse continentale. Mais on trouve un certain nombre d'îles principalement situées dans l'Océan Atlantique qui ne sont pas structurellement différentes de la terre ferme.

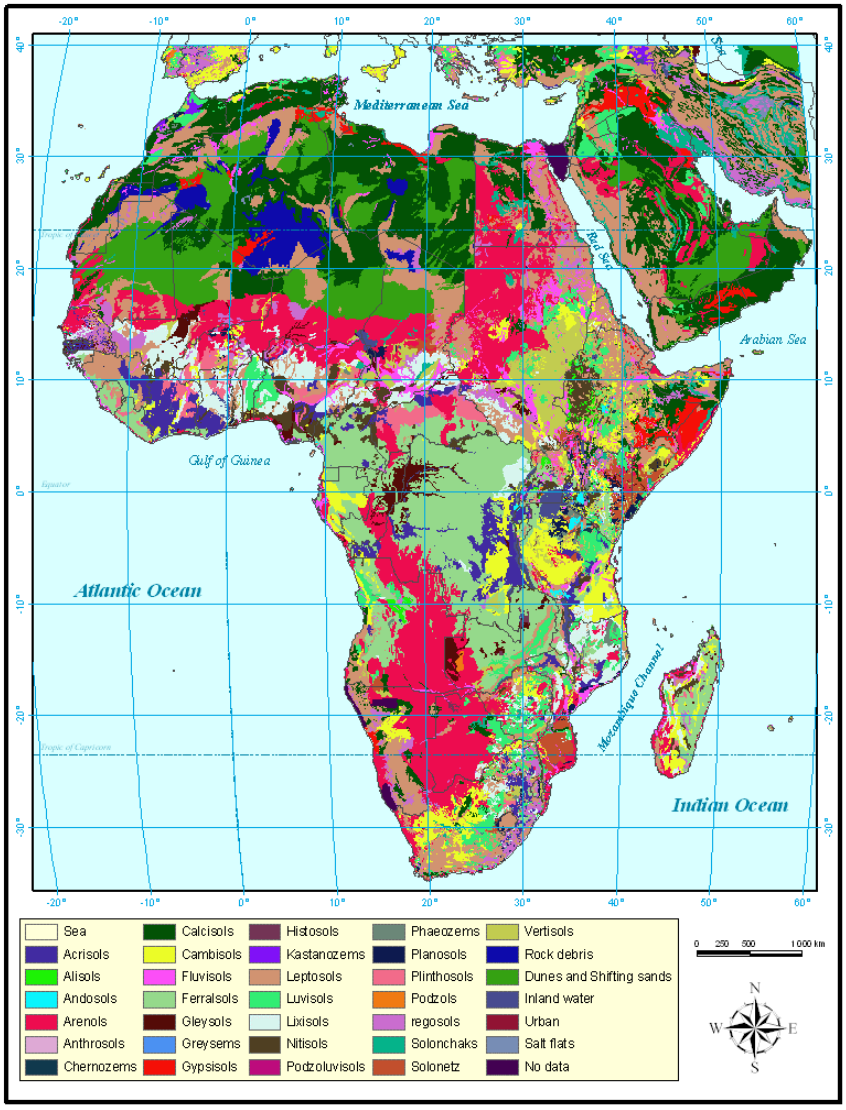

Figure I.3 Principaux types de sols en Afrique (d'après FAO/IIASA, 2009)

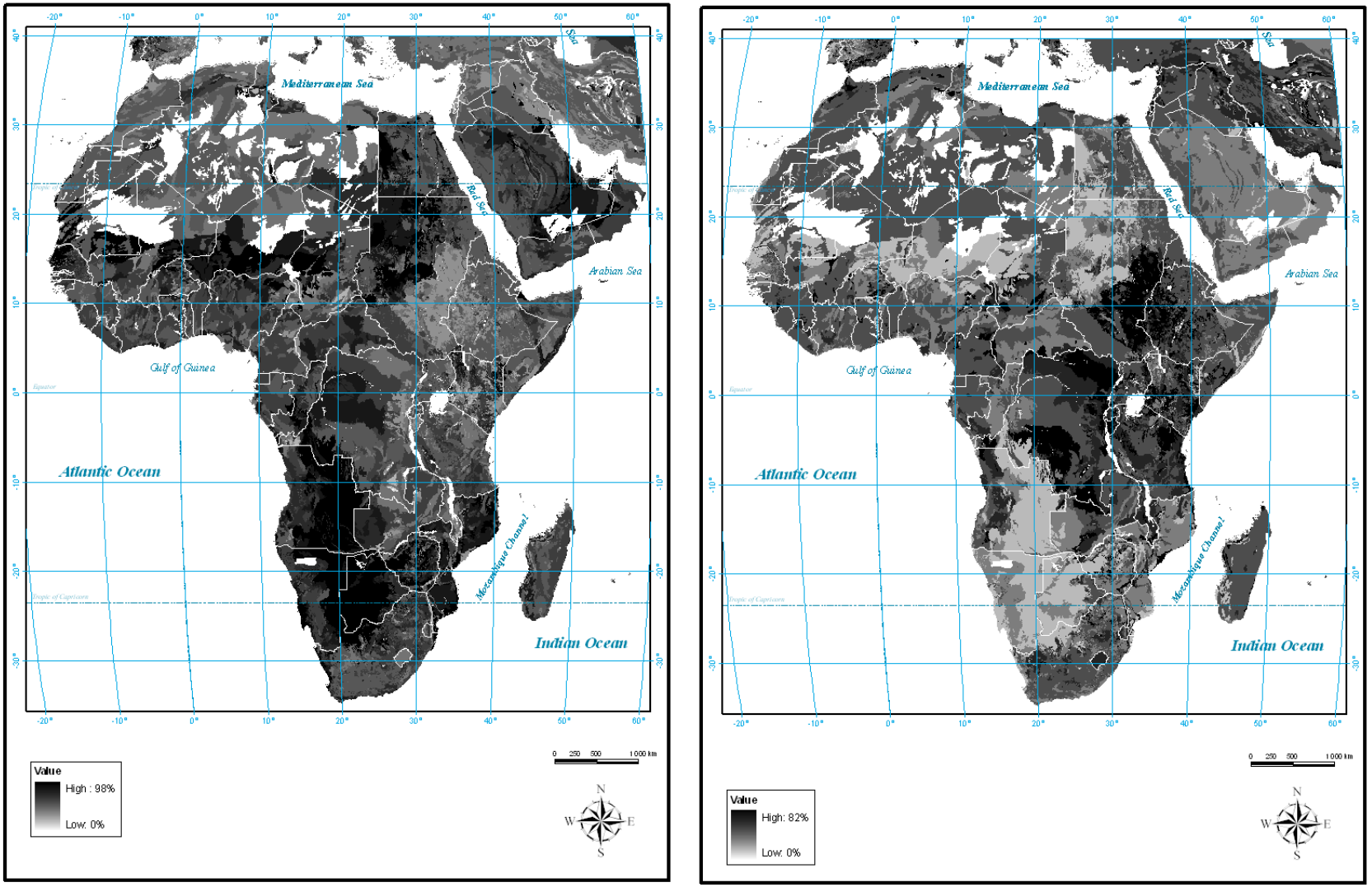

Figure I.4 Pourcentage de sable (gauche) et d'argile (droite) (d'après FAO/IIASA, 2009)

1.3.2 Les ressources en eau

Certains des plus grands fleuves, dont le plus long du monde, se trouvent en Afrique (figure 1.5). C'est le cas du Nil (6670 km), du Congo (4630 km) du Niger (4100 km), du Zambèze (2650 km), de l'Oubangui (2460 km) et de l'Orange (2250 km). En raison du grand nombre de rapides et de cataractes qui les composent, les rivières ne sont pas propres à la navigation sur l'ensemble de leur cours. Cependant, l'Afrique possède le potentiel hydroélectrique le plus élevé du monde (Stock, 2004). Les ressources en eau de l'Afrique sont en permanence affectées par des sécheresses récurrentes, ainsi que par les

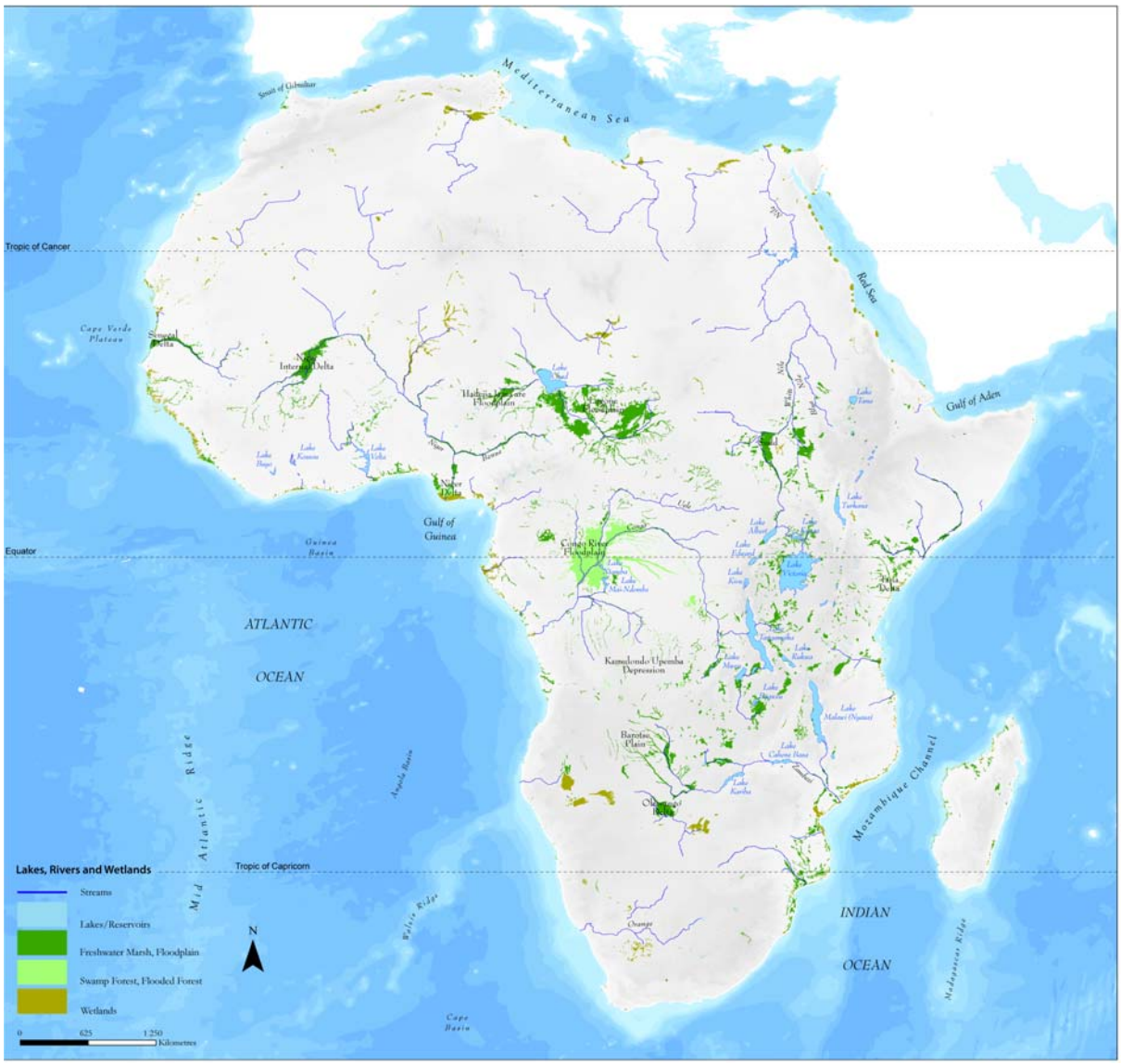

Figure I.5 Ressources hydriques (d'après UNEP, 2008)

changements d'utilisation des sols. C'est par exemple la combinaison de ces deux facteurs qui a conduit au rétrécissement du Lac Tchad de plus de 95% de sa superficie entre 1972 et 2007 alors que c'était la 6^{e} plus grande surface d'eau douce du monde dans les années 1960 (UNEP, 2008). On estime aujourd'hui que plus de 300 millions de personnes doivent faire face à des situations de pénurie en eau en Afrique. Malgré le fait que les eaux souterraines ne représentent que 15% des ressources totales renouvelables en eau du continent, environ 75% de la population africaine tire son eau potable de ces sources souterraines, en particulier en Afrique du Nord et en Afrique Australe.

1.3.3 Climats

A 37° de son point le plus nord et à 35° de son point le plus austral, l'Afrique est le seul continent qui chevauche l'équateur et donc intègre à la fois les tropiques du Cancer et du Capricorne. Cette disposition symétrique en relation avec la latitude caractérise le continent par une série similaire de conditions climatiques et physiques au nord et au sud de l'équateur. Par exemple sur le plan climatique, le désert du Kalahari est le correspondant du Sud au Sahara, le Karoo correspond au Maghreb, alors que les conditions environnementales dans la région du Cap sont presque identiques à celles de la région méditerranéenne. Les principaux climats sont ceux répertoriés comme étant : aride, semi-aride, sub-humide et humide (figure 1.6).

Cette variabilité spatiale extrême constitue une caractéristique fondamentale de l'Afrique où l'on note un gradient prononcé avec des précipitations annuelles totales en déclin au fur et à mesure que l'on s'éloigne de l'équateur. Les précipitations annuelles moyennes varient de moins de 1 mm dans certaines parties du Sahara à environ 10 000 mm dans les montagnes de l'Ouest Cameroun (Debundscha). Elles constituent le facteur climatique dominant pour l'Afrique. Le principal déterminant de la pluviométrie est la zone de convergence intertropicale (ZCIT). Les causes de la variabilité inter-annuelle des précipitations résident dans la structure des vents zonaux AEJ (African Esterly Jet) et TEJ (Tropical Easterly Jet) (Nicholson, 2008).

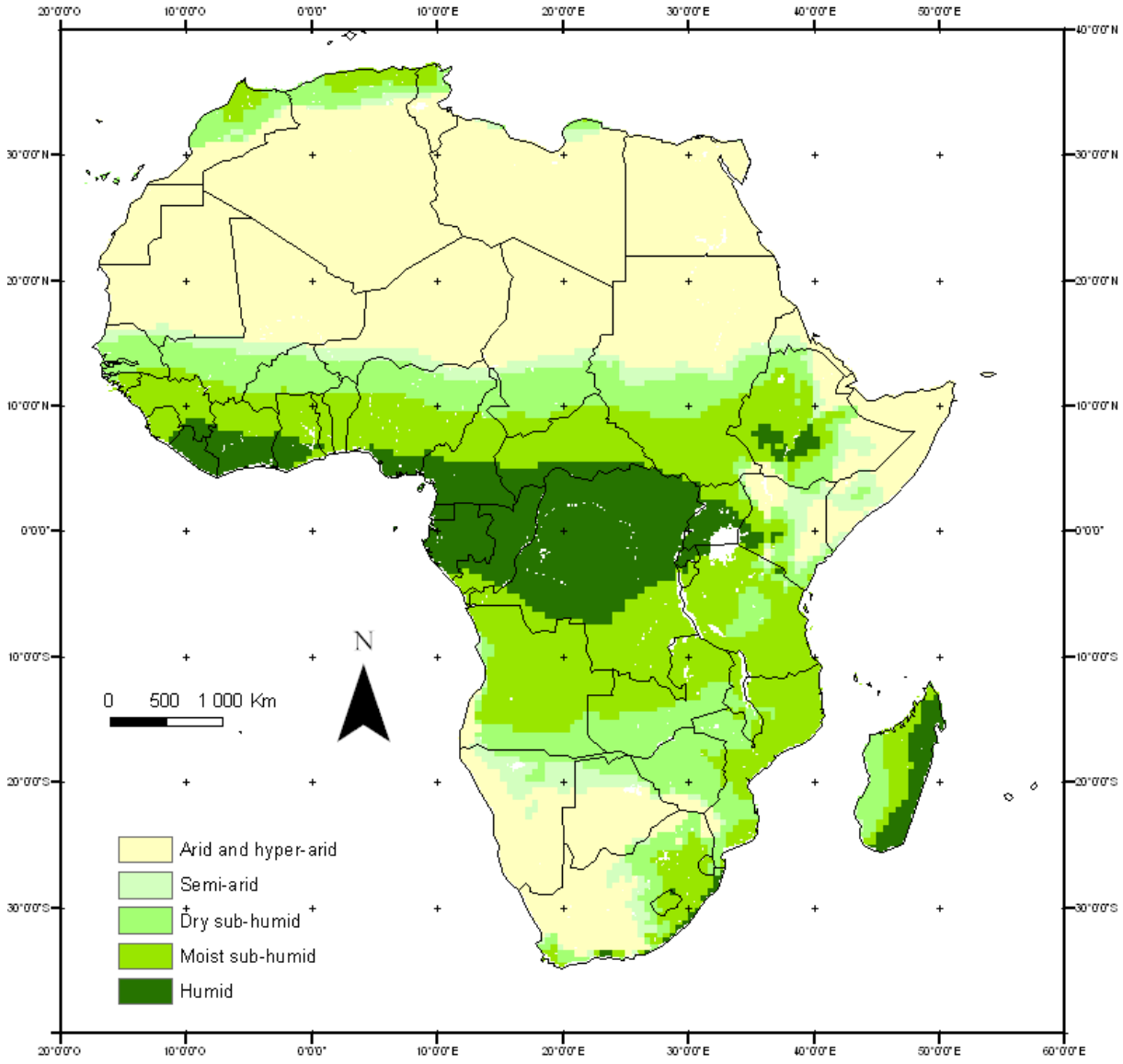

Figure I.6 Climats (d'après Fischer et al., 2002)

Les vents qui soufflent de façon permanente au-dessus de l'océan ont tendance à faire glisser une fine couche de surface de l'eau dans leur direction d'écoulement. Cette couche d'eau océanique, appelée 'courant océanique', dicte les caractéristiques de température et d'humidité des régions côtières sur lesquelles le vent souffle. Les courants océaniques transportent l'eau chaude dans les régions relativement fraîches en passant des basses latitudes proches de l'équateur vers des latitudes légèrement plus élevées. L'Afrique est le continent le plus tropical. Ce qui explique que les températures y sont élevées tout au long de l'année. La température moyenne oscille entre 12 et 32° C avec des plages diurnes et

annuelles qui varient selon la continentalité (faible variation près des côtes et de l'équateur, fortes variations ailleurs). Entre les mois les plus chauds et les mois les plus froids, on observe par exemple une différence de températures moyennes de 24° C au cœur du Sahara contre 1.4° C au Congo (UNEP, 2008).

1.3.4 Ecosystèmes

Un écosystème est un ensemble formé par une association ou communauté d'êtres vivants (ou biocénose) et son environnement géologique, pédologique et atmosphérique (le biotope) (MEA, 2005). Dans tout ce qui suit, le terme 'ecosystème' sera utilisé pour faire référence uniquement à sa composante terrestre végétale. Ainsi le nombre d'écosystèmes présents au sein d'une classe d'occupation des sols estime sa diversité. Par exemple, pour des savanes herbacées présentes au sein d'une même région climatique (en termes d'une unique saison pluvieuse par exemple) peuvent être associés des écosystèmes distincts quand l'une est irriguée et l'autre non. Cette situation peut conduire à 2 périodes distinctes de récolte et la 1ère situation entraînera un profil temporel unimodal du NDVI tandis que le 2ème cas possède un profil bimodal.

L'Afrique compte huit grands biomes décrits au chapitre 3 : la forêt tropicale, la savane humide, la savane sèche et humide, le biome semi-désertique, le biome désertique, les prairies tempérées, le biome méditerranéen, le biome montagneux (figure 1.7) (UNEP, 2008). Le tableau 1.1 illustre les différentes associations entre la végétation, le climat et les types de sol.

1.3.5 L'homme et son environnement

Au sein du système terrestre, la végétation, l'atmosphère, les océans, les glaciers, les processus géophysiques et l'activité humaine interagissent. La théorie de Darwin (1859) sur l'évolution des espèces a exercé une influence décisive sur la compréhension des interactions entre la nature et les activités humaines. Contrairement aux plantes et aux animaux, les êtres humains ont la capacité de s'adapter de manière consciente à de

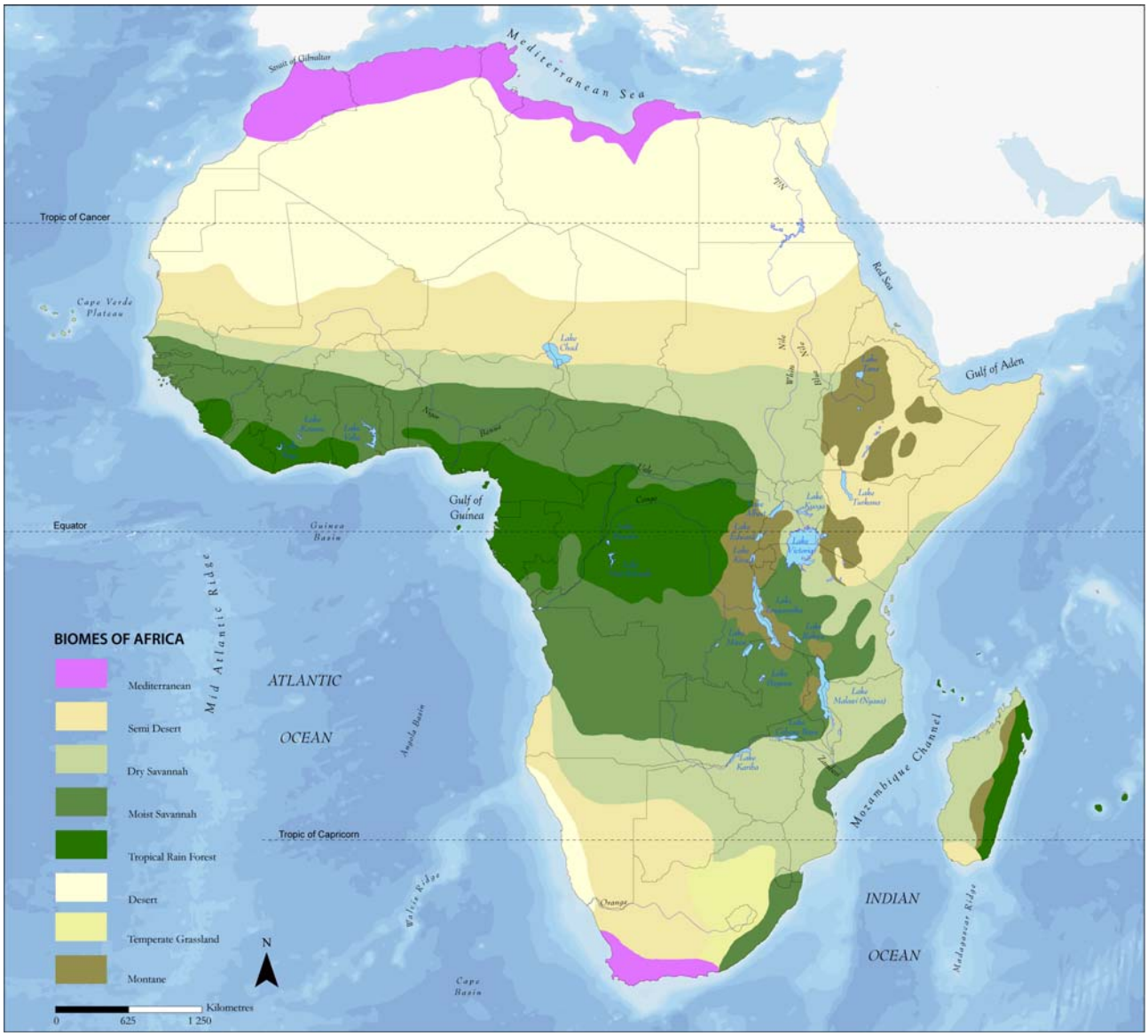

Figure I.7 Biomes (d'après UNEP, 2008)

Biomes	Climat	Type de sol
Forêt tropicale	Humide	Oxisols
Savane arborée humide	Humde sub-humide	Oxisols alfiols, ultisols
Savane sèche	Sec sub humide	Alfisols
Semi-désert	Semi-aride	Alfisols, aridisols
Désert	Aride	Aridisols
Prairie tempérée	Sec subhumide	Alfisols
Méditerranéen	Méditerranéen	Alfisols
Montagne	Hautes altitudes	Variés, peu développés

Tableau I-1 Relations entre les biomes africains, le climat et les types de sols (d'après Stock, 2004)

nouvelles circonstances. Les systèmes naturels et humains co-évoluent, i.e. qu'ils changent ensemble en s'adaptant de manière continue aux changements de l'autre système. Ayant connu une croissance annuelle de 2.32% entre 2000 et 2005 – pratiquement le double du taux mondial moyen de 1.24% par an (UN, 2009), le continent s'urbanise également extrêmement rapidement et devrait connaître le taux d'urbanisation le plus élevé du monde pour de nombreuses décennies (UNFPA, 2007). Cette démographie galopante pose quelques problèmes pour la gestion de l'environnement du continent (vingt des trente pays dont la croissance démographique est la plus rapide du monde se trouvent en Afrique). Avec plus de personnes à nourrir, l'Afrique doit consacrer plus de terres à l'agriculture. Toutefois, la croissance démographique va engendrer l'expansion des aires cultivées pour la production alimentaire, la réduction voire la disparition des ressources naturelles qui sont liées 'élimination d'habitats naturels et le déclin de la population locale (contrainte à l'exode). D'autre part la pauvreté des habitants aggrave et intensifie la pression sur les ressources naturelles.

La richesse biologique du continent est mise en danger par l'une des formes les plus sévères de dégradation des sols à savoir la déforestation. Avec un taux de déforestation deux fois plus élevé que la moyenne mondiale, l'Afrique perd plus de quatre millions d'hectares de forêt chaque année (Kelatwang et Garzuglia, 2006). Les causes directes de la déforestation en Afrique sont :

(i) la conversion des forêts naturelles en des domaines agricoles. La cause primaire de la déforestation est la conversion directe de la forêt en petites parcelles agricoles permanentes (comptant pour 60% des changements de la couverture forestière). La conversion pour de grandes échelles permanentes d'agriculture compte seulement pour 10% et l'intensification de l'agriculture dans des aires d'agriculture itinérante pour moins de 10% ;

(ii) l'exploitation industrielle et la production du bois pour la construction. En effet, les forêts procurent matériaux de construction, et constituent une matière première surexploitée industriellement dans plusieurs pays du bassin du Congo ;

(iii) l'utilisation du bois de chauffage à des fins énergétiques ;

(iv) les feux de forêts en cas de non régénérescence du couvert forestier brûlé ;

(v) la colonisation humaine des espaces pour des fins d'habitation.

La perte des forêts va favoriser l'érosion (éolienne et hydrique) des sols qui est le facteur responsable à 80% de la dégradation des terres en Afrique (Thiombiano et Tourino-Soto, 2007). Une autre forme sévère de dégradation des terres auxquelles l'Afrique doit également faire face est la désertification. Seulement 11% des terres fermes sont situées en zone humide et donc exclues du processus de désertification. La gravité du problème de la désertification en Afrique a conduit à l'adoption d'une convention internationale de lutte contre la désertification (UNCCD, 1994). Le phénomène de désertification est accentué par celui de la déforestation. Les systèmes de culture inappropriés tels que l'agriculture itinérante sur brûlis induisent des feux de savane répétés à grande échelle, contribuant ainsi à reconnaître le continent comme le 'centre de combustion' de la biomasse du globe. En l'an 2000, environ 50% de la biomasse globale consumée provenait des feux de savane de l'Afrique (Tansey et al., 2004) et il faut noter que les 2/3 de la savane mondiale sont situés en Afrique. La combustion de la biomasse contribue ainsi à la pollution transfrontalière de l'air à travers les émissions de particules (aérosols) et de gaz dans l'atmosphère, dont beaucoup ont un impact négatif sur le climat et la santé humaine (Andreae, 1995). Par exemple, les feux rejettent des composés hydrocarbonés dont certains se transforment en ozone troposphérique une fois exposés à la lumière du soleil. Contrairement à l'ozone atmosphérique qui absorbe les rayons ultra-violets, l'ozone troposphérique est un polluant atmosphérique dangereux pouvant provoquer de nombreuses maladies respiratoires et allergies. Les feux de brousse contribuent à au moins 35% de la formation d'ozone troposphérique en Afrique (UNEP, 2008).

En plus de la combustion de la biomasse, une autre source importante de pollution atmosphérique transfrontalière en Afrique est la production des poussières. Véritable danger météorologique lors des phénomènes de tempêtes de sable, ces poussières peuvent être d'origine naturelle -provenant d'éruptions volcaniques ou de l'érosion des sols par les vents-

ou bien elles résultent d'activités humaines telles que les poussières émises par les routes, les usines, les sites de construction ou les feux de ménage. Il faut noter que les poussières d'origine anthropique s'observent moins que les possières d'origine naturelle (Ginoux et al., 2010) et que c'est au Tchad où l'on retrouve la plus grosse région productrice de poussières au monde, à savoir la dépression de Bodélé qui fournit en permanence au Sahara une importante colonne de poussière (Washington et al., 2006). Cette poussière est dispersée à l'ouest au-dessus de l'océan atlantique pour finalement atteindre le bassin de la rivière amazone où il nourrit les sols en nutriments, remplaçant ceux que les fortes pluies tropicales emportent dans leur ruissellement. La dépression de Bodélé ne représente que 0.2% de la surface du Sahara, et pourtant elle contribue à la moitié des 40 millions de tonnes de poussières sahariennes qui atteignent chaque année l'Amazone. Le taux de soulèvement des poussières d'origine minérale est indirectement relié à la pluviosité. Non seulement, l'humidité assure la cohésion entre les particules mais encore la pluie favorise la croissance de la végétation qui sert de facteur d'écran pour le soulèvement de poussières. En résumé, il paraît utile de connaître les types de surface et leur structure saisonnière pour assurer le développement durable dont l'objectif est de définir des schémas viables des trois aspects écologique, social et économique des activités humaines.

1.4 Principales données satellitales utilisées

Au cours des dernières décennies, l'imagerie par télédétection a été acquise par une gamme de systèmes aéroportés et spatiaux qui scrutent la surface terrestre dans différentes longueurs d'onde. Le choix d'une image satellitaire est largement déterminé par quatre facteurs indépendants :

(i) les besoins de l'utilisateur déterminent l'échelle de la zone d'étude qui affecte le choix de la résolution spatiale appropriée ;

(ii) le coût en terme de temps, de labeur nécessaires au télédétecteur pour le traitement des données, mais également pour l'acquisition des images compte tenu de la multitude des capteurs actuellement disponibles ;

(iii) les caractéristiques de la zone d'étude. (topographie, conditions atmosphériques) Le phénomène d'ennuagement permanent observé dans les régions tropicales humides est un obstacle à l'acquisition des données optiques de bonne qualité ;

(iv) les compétences du technicien. L'inaptitude du télédétecteur à corriger le phénomène du chatoiement présent dans les images radar à synthèse d'ouverture (RSO) (en anglais Synthetic Arpeture Radar, SAR) va par exemple le contraindre à choisir des images optiques.

La pertinence de ce choix passe donc inéluctablement par une meilleure connaissance des principales caractéristiques d'un capteur à savoir l'orbite, la plate-forme, le système d'acquisition des images (optique ou radar) et essentiellement de la résolution (voir Annexe A). On parle généralement de résolution radiométrique (capacité de reconnaître de petites différences dans le spectre électromagnétique), résolution spectrale (capacité d'un capteur à utiliser de petites fenêtres de longueurs d'onde), résolution temporelle (durée d'un cycle orbital complet ou répétitivité), résolution spatiale (plus petite distance permettant de discriminer deux objets voisins) (Lefsky et Cohen, 2003). La résolution spatiale peut être scindée en quatre catégories : la très haute résolution (inférieure ou égale à 5 m), la haute résolution (comprise entre 5 et 30 m), la moyenne résolution (comprise entre 30 et 100 m) et la basse résolution (supérieure à 100 m) (Kaptué, 2007).

Dans les régions tropicales, durant la période de croissance où la végétation se développe rapidement, on rencontre très souvent des situations d'obstruction atmosphérique (Vancutsem et al., 2007). D'où la nécessité de disposer des données à une fréquence temporelle élevée. Ces données ne peuvent être acquises que par des observations satellitaires multi-temporelles enregistrées par des capteurs de basse ou moyenne résolution spatiale. D'une part, il n'est pas possible de décrire les différentes phases du cycle phénologique (figure 1.8) à savoir la croissance, la maturité, la sénescence, la dormance et

le retour à croissance, avec une seule image satellitaire (Knight et al., 2006 ; Zhang et al., 2009). D'autre part, l'utilisation de capteurs à haute résolution spatiale (satellitaire ou aéroportée) reste limitée pour la surveillance de grands espaces pendant des intervalles de temps assez courts (Lefsky et Cohen, 2003).

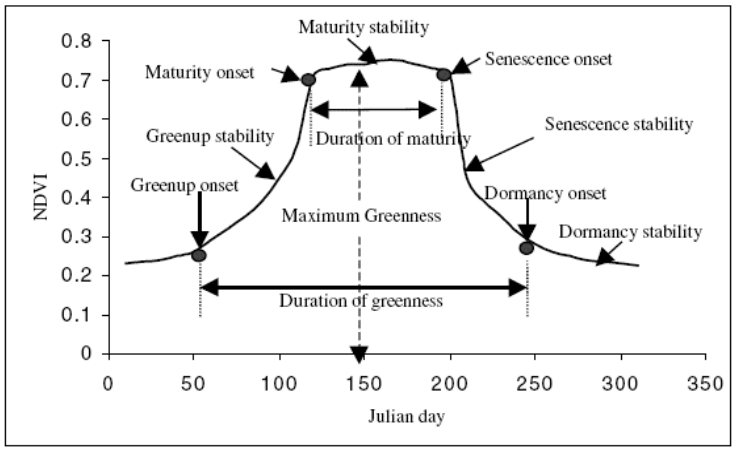

Figure I.8 Transitions entre les phases phénologiques (Knight et al., 2006)

Afin d'obtenir un compromis entre les différentes résolutions d'un capteur et compte tenu du fait que le suivi de la biosphère continentale nécessite une haute répétitivité temporelle, notre choix s'est porté sur les données fournies par le capteur MODIS et celles fournies par le capteur VEGETATION. En effet, la prise en compte des différentes résolutions et des caractéristiques de la zone d'intérêt nous conduit à considérer les données fournies au moins tous les dix jours et cela à une résolution kilométrique. Les données MODIS et VEGETATION (explicitement décrites dans les paragraphes suivants) sont les seules disponibles à une résolution temporelle décadaire et couvrant toute notre période d'intérêt à savoir de 2000 à 2007 ; 2007 étant ici la date de début de ce travail.

1.4.1 L'instrument MODIS

Le spectromètre imageur à moyenne résolution (MODIS) est un instrument de la NASA dédié à la télédétection de l'atmosphère terrestre, des océans et des terres émergées. Il a successivement été embarqué à bord des plates-formes TERRA (le 18 Décembre 1999) et AQUA (le 4 Mai 2002). De tous les capteurs imageurs de couverture globale, il possède le

plus grand nombre de bandes spectrales dans les domaines du visible, du proche infrarouge, court infrarouge, du moyen infrarouge et du long infrarouge ou infrarouge thermique – soit 36 bandes dont 7 sont dédiées à la végétation - (voir tableau 1.2). Les données MODIS couvrant la période 2000 à 2007 ont été obtenues à partir du Distributive Active Archive Center (DAAC), serveur hébergé par la NASA. L'erreur de géolocalisation pour les produits à 1 km de résolution a été évaluée à moins de 50 m (Wolfe et al., 2002), ce qui réduit l'influence des pixels adjacents lors de la discrimination des classes d'occupation du sol. Les produits sélectionnés sont distribués sur une grille sinusoïdale. Nous avons utilisé les produits élaborés à partir d'observations acquises par TERRA du 18 février 2000 au 4 juillet 2002. Pour le reste de la période d'étude, les produits considérés ont été élaborés par fusion des observations acquises par TERRA et AQUA. Les variables d'intérêt pour le présent travail sont le LAI et la réflectance bi-hémisphérique ou 'white sky albedo' (WSA) (Schaaf et al., 2002) de la collection 5 disponibles à une résolution spatiale de 1/120°. Le LAI MODIS se définit comme le ratio de la surface foliaire par unité horizontale surfacique et est déterminée selon l'approche du transfert radiatif de Myneni et al. (1999). L'albédo -rapport du rayonnement solaire réfléchi par la surface de la terre sur l'énergie incidente - contrôle le bilan radiatif à la surface. Par conséquent, l'albédo détermine dans une large mesure la quantité d'énergie disponible qui contrôle l'humidité et les flux turbulents de chaleur (Oleson et al., 2003). Le LAI est obtenu sur une période de 8 jours contrairement aux albédos de surface qui utilisent 16 journées d'observation.

Dans la collection 5, le LAI est basé sur une approche par biome qui utilise la nomenclature à 8 classes de Running et al. (1995). En effet, l'algorithme d'obtention du LAI opère par table de correspondance définie pour chacun des 8 biomes et nécessite également comme données d'entrée les directions d'éclairement et de visée, les coefficients de réflectance bidirectionnelle de chaque bande spectrale. Les résultats sont stockés comme des données de LAI avec des informations sur la qualité des données ou indice de qualité (quality control, QC) utilisées durant la période de composition. Le QC fournit des indications sur l'algorithme utilisé et les conditions environnementales durant la période de composition

(Yang et al., 2006). Les données de LAI avec une bonne précision sont obtenues à partir d'un algorithme principal (main algorithm, MA) fondé sur la physique du transfert radiatif. Un problème de saturation se pose toutefois pour des valeurs de LAI supérieures à 3. En revanche, en cas de défaillance du MA, il est utilisé un algorithme de secours (back-up algorithm, BUA) fondé sur des relations empiriques par biomes entre le LAI et le NDVI. Ces défaillances surviennent lorsque des déficiences sont observées dans la formulation du modèle ou si les incertitudes sur les valeurs de réflectance utilisées en entrée sont supérieures à des seuils prédéfinis

	MODIS	VEGETATION
Champ visuel	110°	101°
Fauchée	2330 km	2200 km
Altitude	705 km	822 km
Inclinaison	98.5° pour TERRA et 98.2° pour AQUA	98.7
Orbite	Polaire héliosynchrone	Polaire héliosynchrone
Durée d'un cycle	2 jours	1 jour
Période de révolution	98.8 minutes	99 minutes pour TERRA et 101 minutes pour AQUA
Heure de passage au nadir	10h 30 pour TERRA et 1h30 pour AQUA	10h30
Nombres de bandes spectrales pour le suivi de la végétation	7 entre 0.405 μm et 14 385 μm	4 entre 0.437μm et 1.692 μm
Résolution spatiale	250m (bandes 1-2), 500 m (bandes 3-7)	1 km
Géolocalisation absolue du pixel	50 m	300 m
Superposition multitemporelle	100 m	325 m

Tableau I-2 Caractéristiques des capteurs VGT-I et II, AQUA et TERRA

L'albédo bi-hémisphérique large-bande MODIS est généré dans trois domaines : visible [0.3-0.7 μm], proche infrarouge [0.7-5.0 μm] et solaire [0.3-5.0 μm]. Cette composante de l'albédo représente la fraction de rayonnement qui est réfléchie sous un ciel parfaitement couvert où l'éclairage peut être considéré comme étant isotrope. La qualité des produits MODIS de LAI et d'albédo a été vérifiée à l'aide de mesures in-situ sur de nombreux sites de la région ouest africaine (Carrer et al., 2010 ; Fensholt et al., 2004 ; Samain et al., 2008).

1.4.2 L'instrument VEGETATION

L'instrument VEGETATION (VGT) a été embarqué en 1998 (sous le nom de VGT-I) à bord de SPOT 4 et en 2002 à bord de SPOT 5 (sous le nom de VGT-II). VGT comprend 4 bandes spectrales dans le bleu, le rouge, le proche infrarouge et le moyen infra-rouge, qui ont été optimisées pour le suivi de la végétation (tableau 1.2). En effet, la végétation active est le seul élément de la surface qui présente une signature spectrale contrastée en raison d'un faible niveau de réflectance dans les longueurs d'onde du rouge et un niveau de réflectance élevé dans les longueurs d'onde du proche infrarouge. La configuration orbitale de SPOT combinée à la fauchée de 2200 km de VGT permettent d'assurer une couverture quotidienne du globe terrestre. Mais cela est en fait un rapport purement théorique et en pratique on aura recours à la composition d'images pour obtenir des scènes sans nuage, en particulier pour les régions proches de l'équateur. Les valeurs de réflectance en haut de l'atmosphère ou Top of Atmosphere (TOA) sont corrigées radiométriquement à l'aide du code SMAC (Simplified Method for the Atmospheric Correction) (Rahman et Dedieu, 1994). Le code SMAC est une version linéarisée du code de transfert radiatif 6S (Vermote et al., 1997) et a pour but la correction des processus d'absorption et de diffusion. Après correction atmosphérique et connaissant les géométries de visée et d'observation, on obtient des produits de synthèse décadaires (S10 et D10), à partir de la composition d'images journalières (Hagolle et al., 2004 ; Maisongrande et al., 2004).

Les données composites S10 sont produites en utilisant l'algorithme de Holben (1986) -le MVC- qui tend à sélectionner les valeurs les plus claires des TOA durant la période de composition en vue d'éliminer les mesures qui sont fortement influencées par la diffusion des aérosols et des conditions atmosphériques plus troublées. Comme la réflectance de surface présente une forte anisotropie, les observations au nadir présentent une forte dépendance de la géométrie soleil-cible-capteur dont les effets sont ignorés dans le produit S10. Le produit composite directionnel, D10, a par la suite été développé à l'aide de l'algorithme de Duchemin et al. (2002), -le BiDirectional Compositing (BDC)- afin de normaliser la réflectance de surface en prenant en considération les effets directionnels dûs aux

différentes géométries de visée et d'éclairement (Roujean et al., 1992). Premièrement, le BDC sélectionne les 10 acquisitions journalières sans nuage les plus récentes pour calibrer un modèle de fonction de distribution de réflectance bidirectionnelle ou Bidirectional Reflectance Distribution Function (BRDF) utilisé pour normaliser les observations non nuageuses. Ensuite la valeur D10 est obtenue en moyennant les observations normalisées obtenues durant la période de composition (Hagolle et al., 2004). Par conséquent, les données D10 lorsqu'elles sont disponibles -i.e. du 1er Juillet 2001 au 31 Décembre 2007- ont été choisies en lieu et place des données S10. Il faudrait noter qu'à l'opposé du MVC, le BDC n'accorde pas de limitation dans la fenêtre temporelle et peut donc sélectionner des données qui ne sont pas comprises dans la période de composition pour calibrer la BRDF durant les dix derniers jours.

La période de composition des images décadaires S10 et D10 a été définie selon le calendrier légal : du 1er au 10 ; du 11 au 20 et du 21 jusqu'au dernier jour du mois. Ces images sont produites par le Centre de Traitement des Images VEGETATION (CTIV) basé en Belgique (Vlaamse Instelling voor Technologisch Onderzoek VITO, www.vgt.vito.be) à une résolution spatiale de 1 km dans le système de projection Universal Transverse Mercator (UTM) avec l'ellipsoïde de référence WGS 84. Chaque image composite consiste en 2 fichiers codés sur 8 bits : un indice de qualité (SM pour Status Map) décrivant la présence des nuages et des problèmes éventuels survenus durant la période de composition, et un code numérique (DN pour Digital Number) converti en NDVI selon VITO comme suit :

$$NDVI = -0.1 + 0.004 * DN \,. \qquad (1.1)$$

1.5 Description des besoins en météorologie

Le climat de la planète est causé principalement par des interactions complexes entre l'atmosphère, les océans, et la surface terrestre, d'où l'exigence d'une bonne paramétrisation

de 4 facteurs dans les modèles météorologiques : la dynamique atmosphérique, la chimie atmosphérique, le cycle de l'eau et les conditions de surface (figure 1.9).

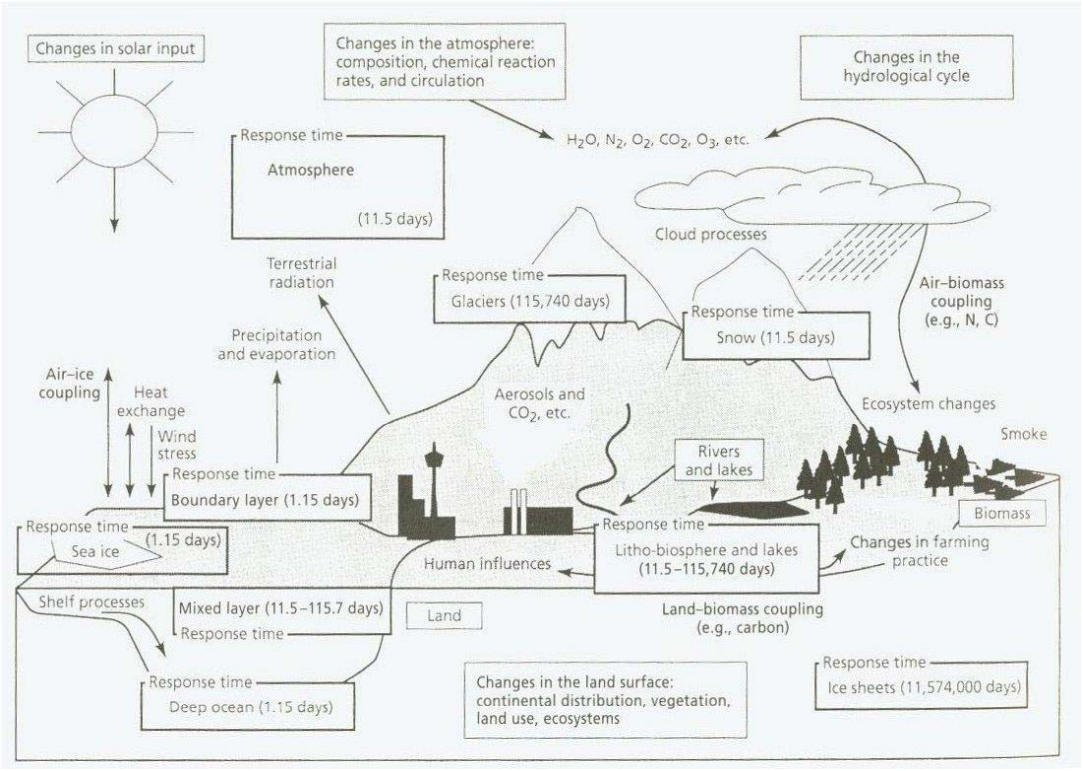

Figure I.9 Différents facteurs météorologiques (d'après Goudie et Cuff, 2002)

La dynamique de l'atmosphère régit la plupart de la variabilité du climat (Shukla, 1981) puisque le mouvement des masses d'air conditionne le climat de diverses régions de la planète. En effet, la circulation atmosphérique est un agent de transport efficace des composés chimiques, des polluants et de la vapeur d'eau. Elle équilibre le contraste entre l'excédent d'énergie (basses latitudes) et les régions déficitaires (moyennes latitudes) et par là même la variabilité du climat à travers le mouvement des masses d'air.

La chimie atmosphérique détermine les effets directs et/ou indirects des constituants chimiques ou particules ou traces de molécules présentes dans l'atmosphère telles que le dioxyde de carbone (CO_2), le méthane (CH_4), l'ozone (O_3) à travers les phénomènes d'absorption et de diffusion du rayonnement solaire dans le spectre du visible et de l'infrarouge (Wuebles et al., 2003). En outre, plusieurs problèmes environnementaux tels que le réchauffement significatif de la surface de la terre et de la troposphère, le refroidissement de la stratosphère et de la mésosphère sont liés aux changements dans la concentration des

principaux constituants atmosphériques que sont l'azote (N_2), l'oxygène (O_2), l'eau (H_2O), l'argon (Ar) et le CO_2.

Le troisième facteur concerne le cycle de l'eau - partie intégrale du système terre-climat – qui se manifeste par de nombreux processus et phénomènes tels que les nuages, les précipitations, le manteau neigeux, les eaux souterraines, les inondations et les sécheresses. Les inondations représentent par exemple un cas typique de débordement d'une rivière qui survient après des précipitations violentes (Stock, 2004).

Enfin, la manière dont les conditions de surface (océanique et continentale) sont représentées dans les modèles climatiques constitue le dernier facteur. Contrairement à l'atmosphère dont la composition change rapidement, l'océan a une forte inertie thermique. Après la période de sécheresse du milieu des années 1960 au milieu des années 1990 avec des périodes moins intenses, on a observé un retour à la normale des précipitations dans la région du Sahel (comme en 2003, 2007) mais aussi une augmentation de la puissance destructrice des cyclones tropicaux comme en 2009 au Sénégal, au Burkina Faso et au Niger.

Bien que les océans soient la principale force qui régit le climat africain (Goudie, 1996), la surface continentale possède un contrôle considérable sur le cycle biogéochimique de la terre à travers par exemple de la dynamique du cycle de l'eau. Les conditions des surface continentales, obtenues à l'aide de l'imagerie satellitaire, sont ainsi utilisées dans les LSM en mode *forçage* ou en mode *réétalonnage* (Lakshmi et al., 2001). Le mode *réétalonnage* consiste à ajuster les variables simulées et les paramètres du modèle afin d'améliorer l'accord entre les valeurs observées et celles qui sont simulées, tandis que le mode *forçage* consiste à utiliser directement les valeurs observées.

« Le seul moyen de faire une méthode instructive et naturelle est de mettre ensemble les choses qui se ressemblent et de séparer celles qui diffèrent les unes des autres », Georges Buffon, Histoire naturelle, 1749.

II. Forçage physiographique sur la zone 'AMMA'[1]

2.1 Introduction

L'Afrique de l'Ouest est une région du monde où les interactions entre la surface et l'atmosphère sont reconnues comme étant fortes. Dans le but d'obtenir une meilleure représentation de ces interactions, les organismes institutionnels ont mis en place régulièrement des programmes scientifiques de visibilité internationale tels que la campagne HAPEX-Sahel (Goutorbe et al., 1997) destinée à fournir les données pour une meilleure compréhension du climat sahélien. Cette campagne s'est poursuivie à travers le programme AMMA (Redelsperger et al., 2006) qui vise à améliorer notre connaissance et notre compréhension de la variabilité intra-saisonnière et inter-annuelle de la mousson africaine. Dans cette partie de l'Afrique, des incertitudes persistent quant aux prévisions pluviométriques (Boko et al., 2007). Certaines études prévoient une humidification progressive (avec une reprise de la couverture végétale) dans le Sahara (Hoerling et al. 2006). D'autres études en revanche prévoient des épisodes plus secs (Jenkins et al., 2005). De telles contradictions sont accentuées par la non prise en compte du changement d'occupation et d'utilisation des sols et des effets rétro-actifs des aérosols dans les modèles (Kamga et al., 2005). En effet, on note sur la région un fort dépôt d'aérosols (produits par les poussières désertiques du Sahara et les fumées résultant de la combustion de la biomasse) (Andreae, 1995) qui influence la formation des nuages et par conséquent le déclenchement de la convection. Les aérosols servent de noyaux de condensation et davantage d'aérosols conduiront à un plus grand nombre de gouttelettes d'eau plus petites et par la suite à un nuage plus réfléchissant. D'autre part, comme ces gouttelettes sont plus petites, elles n'atteindront jamais la taille critique au-delà de laquelle apparaît la précipitation et la persistance du nuage sera augmentée.

[1] Ce chapitre est fondé sur Kaptué, T.A.T, Roujean, J.-L. and Faroux, S., (2010), ECOCLIMAP-II: an ecosystem classification and land surface parameter database of Western Africa at 1 km resolution for the African Monsoon Multidisciplinary Analysis (AMMA) project, *Remote Sensing of Environment*, 114, 961-976.

Avec pour objectif de déterminer l'impact (ou la précision attendue) de l'occupation des sols sur le climat en Afrique de l'Ouest, il est essentiel de développer des forçages physiographiques cohérents à l'échelle régionale. Les progrès récents ont conduit à l'émergence de nouvelles cartes globales d'occupation des sols, produites par analyse de l'imagerie satellitaire variant d'une résolution kilométrique à celles de quelques dizaines de mètres (Arino et al. 2008 ; Bartholomé et Belward, 2005 ; Friedl et al., 2010 ; Hansen et al., 2000 ; Loveland et al., 2000 ; Masson et al., 2003). Les estimations temporelles des variables biophysiques de surface telles que la fraction de végétation (*veg*), le LAI, l'albédo de surface s'appuient sur ces cartes d'occupation des sols. Pour une utilisation dans les modèles de surface ou LSM, ces variables biophysiques de surface sont estimées pour chaque type fonctionnel de surface. L'albédo de surface est exprimé à l'aide de 4 composantes : les albédos direct et diffus pour les bandes spectrales du visible et du proche infrarouge. Pour chaque bande spectrale, l'albédo est décomposé en albédo de la végétation (A_{veg}) et en albédo du sol nu (A_{soil}). Dans ce qui suit, l'albédo du sol nu se réfère à la fraction du rayonnement solaire incident réfléchie par la partie non végétalisée de la surface tandis que l'albédo de la végétation spécifie la réflectance de la partie végétale de la surface continentale.

En parallèle à ces différentes cartes d'occupation des sols et d'estimation des variables biophysiques de surface, nous allons présenter dans ce chapitre un nouveau forçage physiographique sur la portion d'espace [25W-25E]x[5S,20N] (référencée comme AMMA par la suite). Le développement a été justifié par le fait que les cartes d'occupation des sols existantes sont soit obsolètes car les écosystèmes naturels sont valides sur une décennie ou alors ne sont pas suffisamment fiables pour un usage direct dans les modèles bio-géochimiques (Jung et al., 2006). La carte d'occupation des sols de l'Afrique de l'Ouest que l'on cherche à mettre au point se fonde sur les observations fournies par l'instrument MODIS et reprend une partie de l'information fournie par des cartes d'occupation des sols existantes dont celles de GLC2000 et d' ECOCLIMAP-I. Deux types de surface appartenant à la même nomenclature en termes de couvert végétal peuvent cependant représenter des

comportements phénologiques différents (en termes de déclenchement et de fin de la maturité, de la sénescence, de la dormance et de la croissance, voir figure 1.8) en fonction du type de climat auquel ils appartiennent. La carte des climats de Peel et al. (2007) à 0.1° de résolution a été utilisée pour prendre en compte le zonage climatique. On a également utilisé l'information topographique fournie par la carte globale GTOPO30 du service géologique des Etats-Unis (USGS, 1996).

Nos objectifs à travers cette étude consistent à : (i) décrire une méthode de classification non supervisée pour la réalisation d'une nouvelle carte d'occupation des sols, (ii) utiliser cette nouvelle classification de l'occupation des sols pour une estimation des paramètres de surface en sorte que les LSM produisent des résultats plus fiables à l'échelle régionale. La 2ème partie de ce chapitre détaille les données considérées au cours de cette étude ainsi que les pré-traitements effectués. La 3e partie de ce chapitre expose la méthode de classification supervisée tandis que la 4e partie décrit l'estimation des paramètres clés de la végétation et du sol. La 5e partie est dédiée à l'évaluation de la nouvelle classification et des variables biophysiques de surface tandis que dans la dernière partie, nous résumons cette étude et présentons les perspectives de travail à court terme.

2.2 Choix et prétraitements des données

2.2.1 Cartes d'occupation des sols

Durant la production de la nouvelle classification supervisée sur la zone 'AMMA', des cartes d'occupation du sol telles que GLC2000 (VGT) et ECOCLIMAP-I (AVHRR) ont été utilisées en entrée tandis que GLOBCOVER (MERIS) a été utilisée pour la validation de la classification (voir tableau 2.1). Le choix de GLC2000 a été dicté par le fait que c'était le seul produit récent au début de cette étude et qui avait bénéficié d'une importante phase de validation. in-situ. Et GLOBCOVER a été choisi pour produit de validation en raison de sa meilleure résolution spatiale. Toutes ces cartes sont disponibles en libre accès via internet.

Product	ECOCLIMAP	GLC2000	GLOBCOVER
Developer	Masson et al. (2003)	Bartholomé and Belward (2005)	Bicheron et al. (2006)
Sensor	AVHRR	VGT-I	MERIS
Input data	NDVI	4 spectral bands and NDVI	13 spectral bands
Time of data collection	01/04/1992–31/03/1993	01/11/1999–31/12/2000	01/12/2004–31/06/2006
Spatial resolution	1 km	1 km	300 m
Number of classes	27	27	35
Classification name	UMD	LCCS	LCCS
Classification method	Supervised	Flexible classification using local/regional expertise	Flexible classification using local/regional expertise
Internet link	www1	www2	www3

www1 – The EOCLIMAP-I home page http://www.cnrm.meteo.fr/gmme/PROJETS/ECOCLIMAP/page_ecoclimap.htm.
www2 – Land Cover Classification System/GLC 2000 http://ies.jrc.ec.europa.eu/global-land-cover-2000.
www3 – The GLOBCOVER mapping initiative http://postel.mediasfrance.org/en/BIOGEOPHYSICAL-PRODUCTS/Land-Cover/.

Tableau II-1 Cartes utilisées pour cartographier l'occupation des sols sur la zone 'AMMA'

Le Joint Research Center (JRC) a produit une carte d'occupation des sols ayant pour référence l'année 2000. Sur le continent africain, cette carte a été obtenue par une classification non supervisée en utilisant principalement 14 mois de données journalières à une résolution spatiale de 1/112° acquises par le capteur SPOT VGT entre novembre 1999 et décembre 2000 (Mayaux et al., 2004 ; Bartholomé et Belward, 2005). En plus d'un modèle numérique de terrain, trois autres sources de données issues de différents capteurs embarqués sur des plates-formes satellitaires ont été utilisées : les données JERS-I/SAR, ERS-I/SAR et DMSP/OLS. En se basant sur la nomenclature Land Cover Classification System (LCCS) (Di Gregorio and Jensen, 2000), Mayaux et al. (2004) ont identifié 27 classes d'occupation des sols sur la zone 'AMMA'. Au niveau dichotomique, LCCS sépare (i) les surfaces végétalisées des non végétalisées, (ii) les terres émergées des terres immergées et (iii) les domaines cultivés et protégés des espaces naturels et semi-naturels. LCCS décrit le couvert végétal en termes de morphologie, canopée, distribution spatiale, type foliaire, hauteur et phénologie.

ECOCLIMAP-I est une base de données climatologiques produite au Centre National de Recherches Météorologiques (CNRM) afin de répondre aux besoins de la communauté météorologique via la documentation des écosystèmes naturels et fonctionnellement

homogènes dans le cadre de la modélisation de l'étendue des changements climatiques et de leur impact (Champeaux et al., 2005 ; Masson et al., 2003). L'originalité d'ECOCLIMAP-I est de fournir par type d'écosystème et sous une forme tabulaire, un ensemble de paramètres de surface tels que le LAI, le *veg*, le fAPAR, la longueur de rugosité, la résistance stomatique minimale et la profondeur racinaire. Ces paramètres servent à initialiser et contraindre les LSM (Boone et al., 2009). La classification ECOCLIMAP-I sur l'Afrique a été réalisée en combinant l'information issue de la carte d'occupation des sols de Hansen et al. (2000) et la carte des climats de Koeppe et De Lond (1958) par analyse spectrale des données NDVI AVHRR à 1 km de résolution acquises entre avril 1992 et mars 1993. ECOCLIMAP-I comprend 27 classes sur la zone 'AMMA'.

L'ESA a coordonné le projet GLOBCOVER qui est une initiative de cartographie globale de l'occupation des sols afin de parvenir à une nouvelle ère pour la cohérence des estimations globales d'occupation des sols (Arino et al., 2008 ; Bicheron et al., 2008). L'initiative GLOBCOVER avait pour objectifs la détection et le traitement des réglementations environnementales en particulier celles de nature transfrontalière ; ce qui devrait la rendre particulièrement pertinente pour les régions africaines. Les données en entrée sont les réflectances dans 13 fenêtres de longueur d'onde enregistrées par le capteur Medium Resolution Imaging Spectrometer (MERIS) à bord d'ENVISAT à une résolution spatiale de 1/360° (environ 300 m). Tout comme pour GLC2000, les développeurs de GLOBCOVER ont identifié 35 classes d'occupation des sols sur la zone 'AMMA' en se basant sur la nomenclature LCCS.

2.2.2 Données auxiliaires

2.2.2.1 Carte climatique

La carte climatique de Peel et al. (2007) a été utilisée pour prendre en compte le zonage climatique lors de la représentation des écosystèmes. C'est une classification hiérarchique à 3 niveaux basée sur la nomenclature de Köppen (1936) qui résulte d'un large réseau global

de données mensuelles d'observations de température (de 1923 à 1993) et de précipitations (de 1909 à 1991) fourni par le Global Historical Climatology Network (GHCN) (Peterson and Vose, 1997). Comparée à d'autres récentes cartes telles que celles de Mitchell et Jones (2005) ou Kottek et al. (2006), la carte de Peel et al. (2007) est la carte climatique la plus fine en termes de résolution spatiale (0.1° de résolution) et de contenu thématique (nombre de classes).

2.2.2.2 Carte topographique

L'information topographique fournie par la carte GTOPO30 a également été utilisée. GTOPO30 est un modèle numérique de terrain en global à 1 km de résolution qui a été développé par le service géologique américain (USGS, 1996 ; Gesch et Larson, 1996). Cette carte a été produite faisant suite à la fusion de deux types de données topographiques : les données vectorielles de la Digital Chart of the Word au millionième (DCW) (Danko, 1992) et les données rastérisées de la Digital Terrain Elevation Data (DTED) (DMA, 1986).

2.2.3 Prétraitement des données

Les données MODIS d'indice foliare et d'albédo (voir section 1.4.1) ont d'abord été ré-échantillonnées et re-projetées en UTM (Universal Transverse Mercator) (WGS 84) à la résolution spatiale du 1/112° en utilisant la méthode de ré-échantillonnage au plus proche voisin (PPV). Ensuite nous avons effectué une interpolation par enveloppe supérieure pixel par pixel sur les données de LAI et de NDVI en utilisant un polynôme de degré 5 de Savitsky et Golay (1964) afin d'une part de combler le manque de données et d'autre part d'exclure les points isolés. Malgré toutes les méthodes de correction appliquées aux observations satellitaires, des contaminations par les nuages et aérosols persistent et ceci sans compter les éventuelles failles techniques qui peuvent empêcher l'acquisition des données.

Bien que les classifications s'effectuent généralement à partir des IV, notre choix s'est porté sur les données d'indice foliaire MODIS. En effet, les indices de végétation MODIS ont

une résolution temporelle de 16 jours et sont donc inadéquates pour notre étude en raison des contraintes du produit final (cf. section 1.4). En outre, au début de notre étude, une bonne partie des données VGT n'avait pas encoré éte traitée par VITO. La figure 2.1 montre la moyenne des données de LAI avant et après la procédure de lissage pixel par pixel pour la classe d'occupation du sol *forêt dense sempervirente* du produit GLC2000 sur la zone 'AMMA'. Des relations empiriques établies entre le LAI et le NDVI montrent une linéarité pour des faibles valeurs de LAI (Baret and Guyot, 1991 ; Justice, 1986 ; Myneni et al., 1999 ; Pontailler et al, 2003). En excluant des scénarios de LAI MODIS mitigés à cause de la contamination permanente par les nuages et aérosols, la figure 2.2 montre que le LAI MODIS et le NDVI VGT (après son élaboration tout le long de l'année 2007) sont cohérents. Le coefficient de détermination (R^2) entre le LAI et le NDVI est inversement proportionnel à la hauteur de la canopée. R^2 est respectivement égal à 0.29, 0.51 et 0.84 pour les classes : *forêt dense sempervirente*, *savane arborée décidue* et *savane arbustive décidue* de GLC2000. La corrélation entre le LAI MODIS et le NDVI VGT justifie ainsi l'utilisation du LAI MODIS pour la discrimination des types de végétation malgré l'effet de saturation des couverts très denses causé par de fortes valeurs de réflectance dans le proche infrarouge (Liang, 2004). Ainsi le LAI contient des informations majeures sur les propriétés des plantes telles que la physionomie (e.g. longévité foliaire), la phénologie (cycle de vie de la plante) et peut donc être utilisé pour la discrimination des écosystèmes.

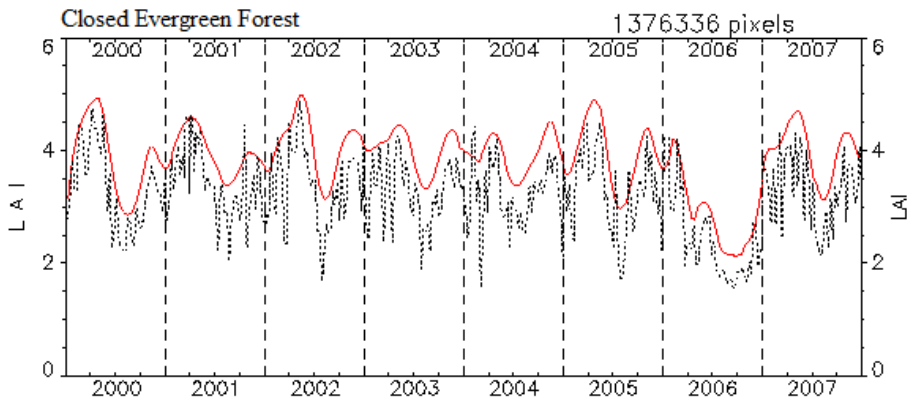

Figure II.1 Séries temporelles du LAI MODIS moyenné en sur la zone 'AMMA' pour la classe foret dense sempervirente de GLC2000 correspondant aux données brutes (noir, pointillé) et aux données lissées (rouge, continu). La courbe rouge est la moyenne du LAI de la classe à partir d'un polynôme de degré 5 appliqué pixel après pixel.

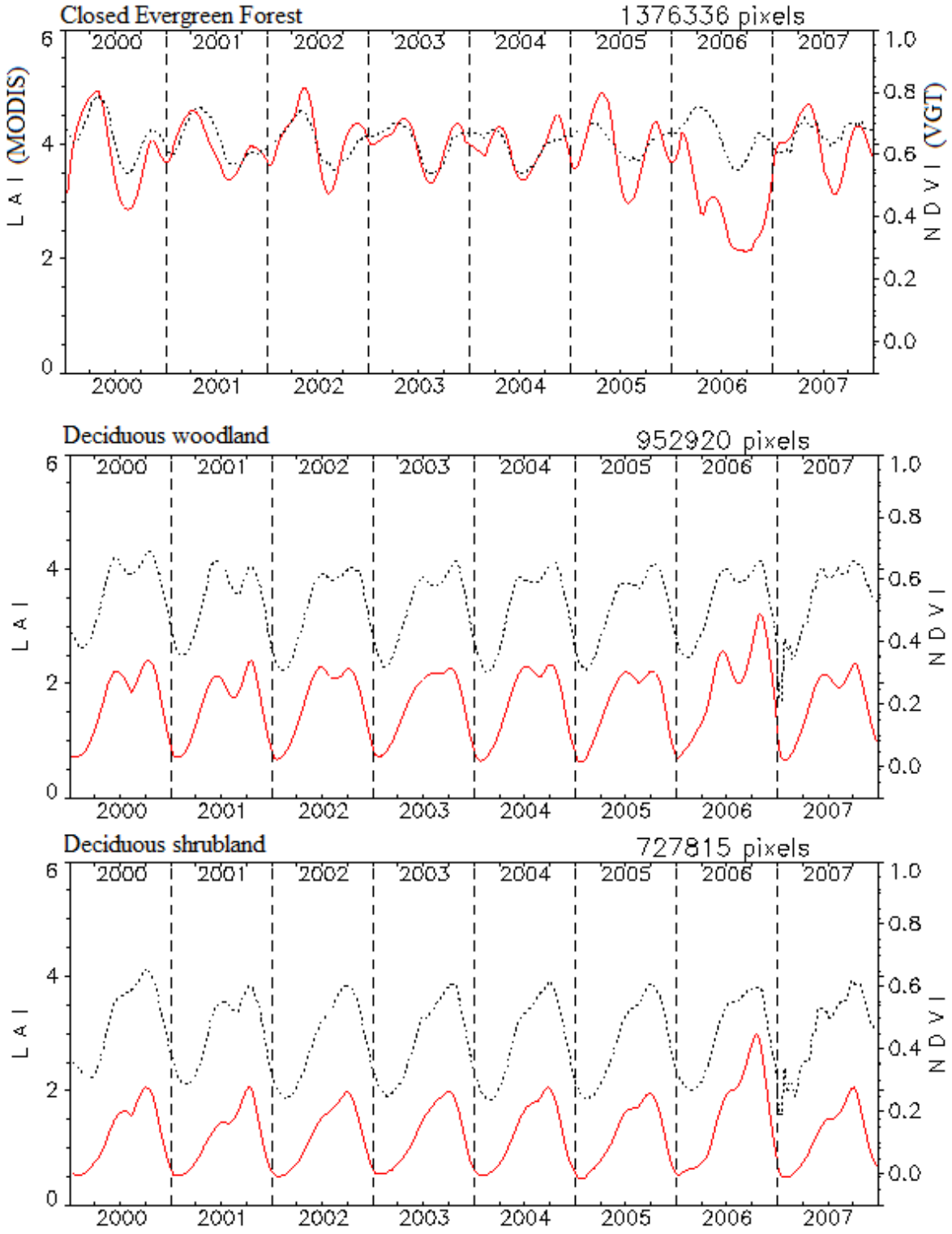

Figure II.2 Séries temporelles du LAI MODIS (rouge, continu) et du NDVI VGT (noir pointillé) de 3 classes GLC2000 (forêt dense sempervirente, savane arborée décidue et savane arbustive décidue) sur la zone 'AMMA' après lissage en utilisant un polynôme de degré 5 appliqué pixel après pixel.

2.3 Méthode de classification supervisée

La présente méthode suit les principales phases de classification qui sont : la sélection des données satellitaires, le prétraitement des données, la technique de classification et l'évaluation de la carte (Lu and Weng, 2007). La technique de classification, effectuée sur la base de l'analyse du profil temporel lissé du LAI MODIS comprend trois étapes (figure 2.3).

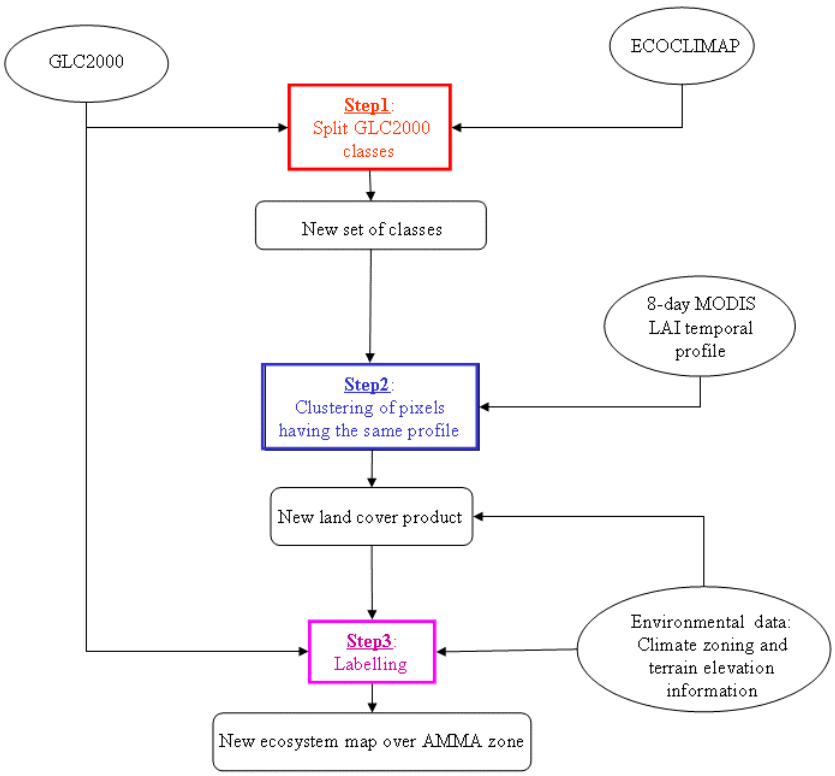

Figure II.3 Méthodologie de classification supervisée

Durant la 1ère étape de la stratégie de classification, les classes GLC2000 sont scindées sur la base d'ECOCLIMAP-I (voir tableau 2.2). Ensuite, pour chaque classe GLC2000, on choisit les clusters les plus représentatifs d'au moins 90% de la superficie, ce qui conduit à sélectionner 149 clusters. Par exemple, la figure 2.4 montre la scission de la classe *forêt dense sempervirente* de GLC2000 en 5 composantes d'ECOCLIMAP-I : *forêt équatoriale* (71.24%), *savane arborée sahélienne* (7.82%), *mosaïque équatoriale de savane arborée et herbacée* (7.71%), *savane herbacée tropicale* (4.13%) et *savane arborée au sud de l'équateur* (2.96%). Il faut noter que : (i) notre méthodologie de classification combine l'information complémentaire des cartes GLC2000 et ECOCLIMAP-I, (ii) le processus de

combinaison des cartes pour la production d'une carte globale d'occupation des sols a déjà été effectuée avec succès par le passé (Masson et al., 2003, Ramankutty et al., 2008), (iii) avec pour objectif de consolider la classification finale, il existe plusieurs raisons de débuter avec le produit GLC2000. Tout d'abord, le produit GLC2000 a été développé à base d'observations acquises durant l'année 2000 à la résolution vraie d'1 km par l'instrument VGT comparée à la taille du pixel AVHRR très variable entre le nadir et en bord de la fauchée. En outre, GLC2000 a une légende thématique qui a été stabilisée par plusieurs mises à jour et grâce à une phase intense de validation (Mayaux et al., 2006).

GLC2000	ECOCLIMAP-I	Classe combinée
1. GLC1	1. ECO1	1. GLC1ECO1
1. GLC1	2. ECO2	2. GLC1ECO2
1. GLC1	3. ECO3	3. GLC1ECO3
…	…	…
27. GLC27	27. ECO27	xxx. GLC27ECO27

Tableau II-2 Exemples de classes d'occupation des sols obtenues par combinaison de GLC2000 et ECOCLIMAP-I

Les profils temporels lissés du LAI MODIS sont analysés un par un de façon supervisée à l'intérieur de chaque classe GLC2000 afin de réduire le nombre de clusters durant la 2ème étape. En guise d'illustration, l'analyse du LAI MODIS en terme de calendrier, de durée et d'intensité de l'activité photosynthétique dans la classe "forêt dense sempervirente" permet la mise en évidence de 3 sous classes (figure 2.5). L'une d'elle regroupe les pixels identifiés comme de la forêt équatoriale dans ECOCLIMAP-I. Une autre est constituée par un mélange composé d'une savane arborée sahélienne et d'une mosaïque de savanes herbacées et arborées. La savane herbacée tropicale et la savane arborée au sud de l'équateur sont combinées pour former une troisième classe à l'intérieur de la classe initiale de GLC2000 (figure 2.5). C'est toujours durant cette étape que l'on regroupe les clusters s'étendant sur de petites surfaces à l'intérieur des écosystèmes mieux représentés. Ceci conduit à une carte de 37 écosystèmes fonctionnellement homogènes. Il est important de noter que d'autres types de données ont été utilisés comme support à l'interprétation et à la fusion des clusters.

Par exemple, l'information orographique GTOPO30 (USGS, 1996) a permis de séparer la classe 'forêt sub-montagnarde' de GLC2000 en deux sous classes. Au cours de la 3ème étape, les différents écosystèmes ainsi identifiés ont été labellisés afin d'améliorer leur discrimination visuelle ainsi que la relation directe aux caractéristiques de la surface physique. Ceci s'effectue par la prise en compte des informations de GLC2000 et des données auxiliaires environnementales (biome, altitude).

Après la production et la labellisation des écosystèmes, la dernière phase du processus de classification consiste à fournir des estimations quantitatives sur l'exactitude de la nouvelle carte d'occupation des sols (voir annexe B). Une pratique largement acceptée consiste à utiliser des informations à haute résolution pour évaluer la qualité des produits à faible résolution (Kloditz et al., 1998). Bien que les produits à haute résolution soient eux-mêmes entachés d'éventuelles erreurs d'interprétation, ce processus est facile à mettre en œuvre lors de la validation d'une carte en comparaison avec l'étape cruciale du choix d'une stratégie d'échantillonnage appropriée telle que décrite par Stehman (2009). Puisque que la validation des produits à l'échelle régionale nécessite des données cohérentes sur le plan temporel, la fiabilité de la nouvelle classification est vérifiée ici par comparaison avec le produit GLOBCOVER réalisé à partir de données acquises en 2005 (voir tableau 2.1). Ce choix se justifie d'une part en raison de sa haute résolution spatiale par rapport aux données d'occupation MODIS et d'autre part, parce que contrairement à GLC2000, le produit GLOBCOVER n'a pas été utilisé comme données d'entrée dans le processus de classification. Compte tenu des résolutions spatiales distinctes de GLOBCOVER et de la nouvelle classification, nous avons d'abord projeté la nouvelle classification dans le même système de projection que celui de GLOBCOVER. La nouvelle carte a d'abord été re-échantillonnée du 112e au 120e de degré en utilisant la règle du PPV. Ensuite, la résolution du 360ème de degré a été atteinte en divisant chaque pixel au 120e de degré en 9 pixels identiques au 1/360e de degré. Les nomenclatures de la nouvelle carte et GLOBCOVER ont été regroupées en 10 grandes catégories : forêt, forêt mixte, savane arborée, savane arbustive, savane herbacée, savane mixte, culture, bâti, sol nu et corps d'eau. La qualité

relative du produit ainsi obtenu est renseignée par une matrice d'erreur comme c'est le cas pour évaluer les classifications d'occupation des sols (cf. annexe B).

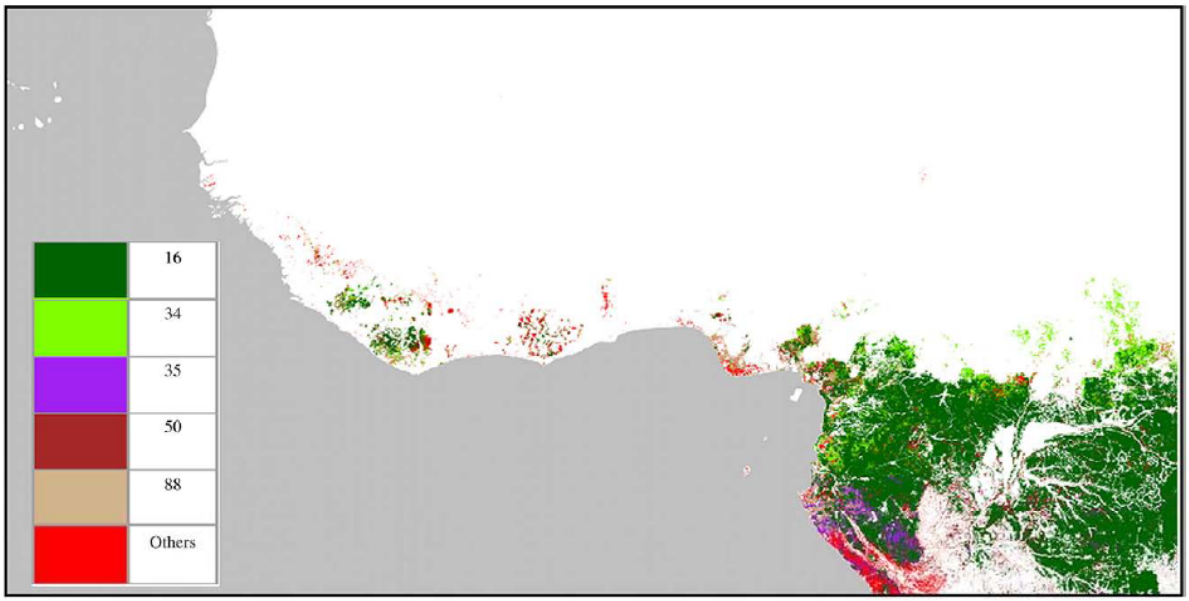

Figure II.4 Scission de la classe forêt dense sempervirente en 5 composantes ECOCLIMAP-I : 'forêt équatoriale' (classe numéro 16), 'savane arborée sahélienne' (classe numéro 34), 'savane arborée au sud de l'équateur' (classe numéro 35), 'mosaïque équatoriale de savane arborée et savane herbacée' (classe numéro 50), et 'savanes herbacées tropicales' (classe numéro 88)

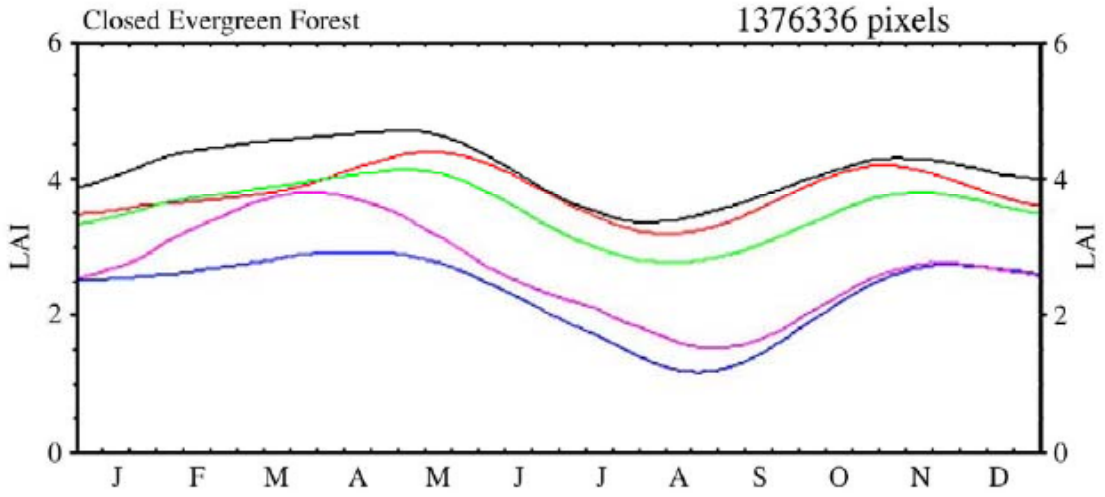

Figure II.5 Moyenne annuelle du profil saisonnier du LAI MODIS sur la zone 'AMMA' pour la classe 'forêt dense sempervirente' de GLC2000 pour les numéros des classes 16 (noir), 34 (rouge), 35 (vert), 50 (bleu), et 88 (violet).

2.4 Estimation des variables biophysiques de surface

La détermination de la fraction de végétation (*veg*) et de l'albédo de surface (A) sont encore perfectibles pour la modélisation des échanges d'énergie et de chaleur entre la surface et l'atmosphère. La grande majorité des LSM considèrent aujourd'hui l'albédo comme une partition entre des valeurs des spectres du visible et du proche infrarouge avec également une séparation entre les composantes directe et diffuse, ce qui répond au besoin de modélisation des processus de surface (Dai et al., 2003 ; Noilhan et Mahfouf, 1996). Dans ces modèles, les albédos du sol nu et de la végétation sont d'abord calculés indépendamment puis l'albédo total est obtenu par addition de ces deux albédos pondérés par la fraction de végétation. Une écriture de l'albédo total de la surface est alors :

$$A = veg * \left(A_{veg} - A_{soil}\right) + A_{soil} \ , \qquad (2.1)$$

Cette fraction de végétation, *veg*, -qui modélise la transmittance- est estimée à partir du LAI comme suit :

$$veg = 1 - \exp\left(- \frac{G(\theta_s)}{< \cos \theta_s >} LAI \right) , \qquad (2.2)$$

où θ_s est l'angle solaire zénithal et $G(\theta_s)$ est la distribution foliaire.

Pour une résolution spatiale modérée, nous avons d'abord émis l'hypothèse selon laquelle le feuillage peut avoir une orientation aléatoire sphérique, ce qui conduit à $G(\theta_s) = 0.5$, indépendamment de la géométrie d'éclairement. Ensuite, au regard de la zone d'intérêt, une valeur moyenne de l'angle solaire zénithal $\theta_s = 30°$ est retenue puisque celui-ci varie entre 0° (à l'équateur) et 60° si on considère qu'il s'agit de l'angle limite à partir duquel, pour une atmosphère standard, le rayonnement direct n'est plus dominant. Soit :

$$veg \approx 1 - \exp\left(-0.6 * LAI\right) , \qquad (2.3)$$

La saisonnalité du *veg* est donc estimée à partir du LAI selon l'équation (2.3), ce qui permet de mettre à jour les albédos. Cependant l'information dérivée du *veg* est limitée durant la période de dépérissement de la végétation. Pour cette raison, nous avons

développé une stratégie particulière pour obtenir des estimations dynamiques des albédos du sol nu et de la végétation en exploitant les caractéristiques de ce paramètre à différentes échelles temporelles. Par exemple, à de hautes fréquences temporelles, allant de quelques heures à quelques jours, l'humidité superficielle influence fortement les changements de l'albédo de surface à partir des variations de l'albédo du sol nu. D'autre part, les changements de l'albédo de la végétation suivent un lent processus généralement observé à des fréquences variant des échelles hebdomadaires à des échelles mensuelles. Néanmoins, ces considérations générales peuvent s'appliquer à l'Afrique de l'Ouest car l'activité de la végétation y est concentrée durant quelques semaines dans la mesure où elle est pilotée par l'alternance des évènements de mousson et des périodes sèches.

En plus des valeurs du *veg* dérivées à partir des données MODIS, l'estimation des albédos de la végétation et du sol nu à partir de l'équation (2.1) est réalisée avec les albédos bi-hémisphériques MODIS du visible et du proche infrarouge. On effectue alors une interpolation par classe sur la base de la nouvelle carte d'occupation des sols pour remplir les trous des cartes de LAI et d'albédo en utilisant la distance de Manathan (voir annexe C). Cela permet d'obtenir des champs spatialement et temporellement continus d'albédo de surface et de fraction de végétation. En outre, *A* et *veg* sont phasés dans le temps en ramenant *veg* à la même échelle temporelle que l'albédo MODIS (annexe C). A ce stade, l'albédo et le *veg* sont associés au même jour julien et se déclinent par huitaines.

Nous appliquons une régression linéaire entre *A* et *veg* sur la base du pixel en considérant tout le jeu de données, i.e. toutes les données de huitaine MODIS durant la période 2000-2007. En utilisant l'équation (2.1), le seuil minimum du nombre d'observations valides a été arbitrairement fixé à 100. Pour des considérations physiques, la quantité $(A_{veg}-A_{soil})$ représentant la pente ne doit pas être positive dans le visible alors qu'elle doit rester positive dans le proche infrarouge. *De facto*, la pente de la relation entre *A* et *veg* est spectralement signée. Si la valeur de la pente qui est obtenue est négative dans le spectre du visible, on considère alors que le terme libre d'albédo correspond à l'albédo du sol nu. Il apparaît pour le visible que 8.73% de la surface continentale de notre zone d'étude

possèdent une pente positive dans le visible. On attribue ce résultat à des inconsistances entre le LAI et l'albédo de MODIS. Cela se produit en raison de l'incertitude sur les champs MODIS, principalement pour des zones dont la densité du couvert végétal est forte et résulte probablement d'un effet de saturation de la réflectance durant le processus de détermination du LAI. Pour ces cas spécifiques, une interpolation par classe des albédos visibles du sol nu a été effectuée à partir des données LAI et d'albédos MODIS cohérentes localisées dans un carré ou voisinage de coté 1 degré (voir annexe C). Les détails de la procédure d'estimation de l'albédo du sol nu dans le visible sont indiqués sur la figure 2.6.

Similairement, mais en considérant cette fois-ci une pente $(A_{veg}-A_{soil})$ positive afin de préserver les valeurs physiques de l'albédo, nous avons également calculé l'albédo du sol nu dans le proche infrarouge. Une fois que ces valeurs d'albédo du sol nu ont été déterminées, nous déduisons pour chacune de ces bandes spectrales (visible et proche infrarouge) A_{veg} directement à partir de l'équation (2.1) par huitaine en faisant varier l'albédo spectral A et le LAI mais en supposant une valeur fixe pour A_{soil}. Il est important de rappeler ici que le but visé est de rendre le LAI et l'albédo MODIS cohérents dans les LSM qui utiliseront cette nouvelle physiographie de la zone 'AMMA'. Cette cohérence pour les spectres du visible et du proche infrarouge est assurée au regard des considérations effectuées dans les équations 2.1 et 2.3. Néanmoins, des difficultés liées à l'usage direct des données MODIS persistent notamment au niveau des pixels pour lesquels des valeurs erronées des pentes ont été obtenues lors du calcul des albédos du sol nu dans les spectres du visible et du proche infrarouge.

L'évaluation de l'estimation des deux composantes d'albédo, A_{soil} et A_{veg}, est accomplie à travers la reconstruction des albédos bi-hémisphériques de surface A_{vis} et A_{nir} de MODIS. Ils sont utilisés pour faire correspondre pixel par pixel, l'albédo MODIS dans le spectre total sur la période des huit ans de 2000 à 2007 par huitaine en utilisant l'équation (2.4), qui différencie le sol de la végétation :

$$A = \alpha + \beta * A_{vis} + \gamma * A_{nir} , \qquad (2.4)$$

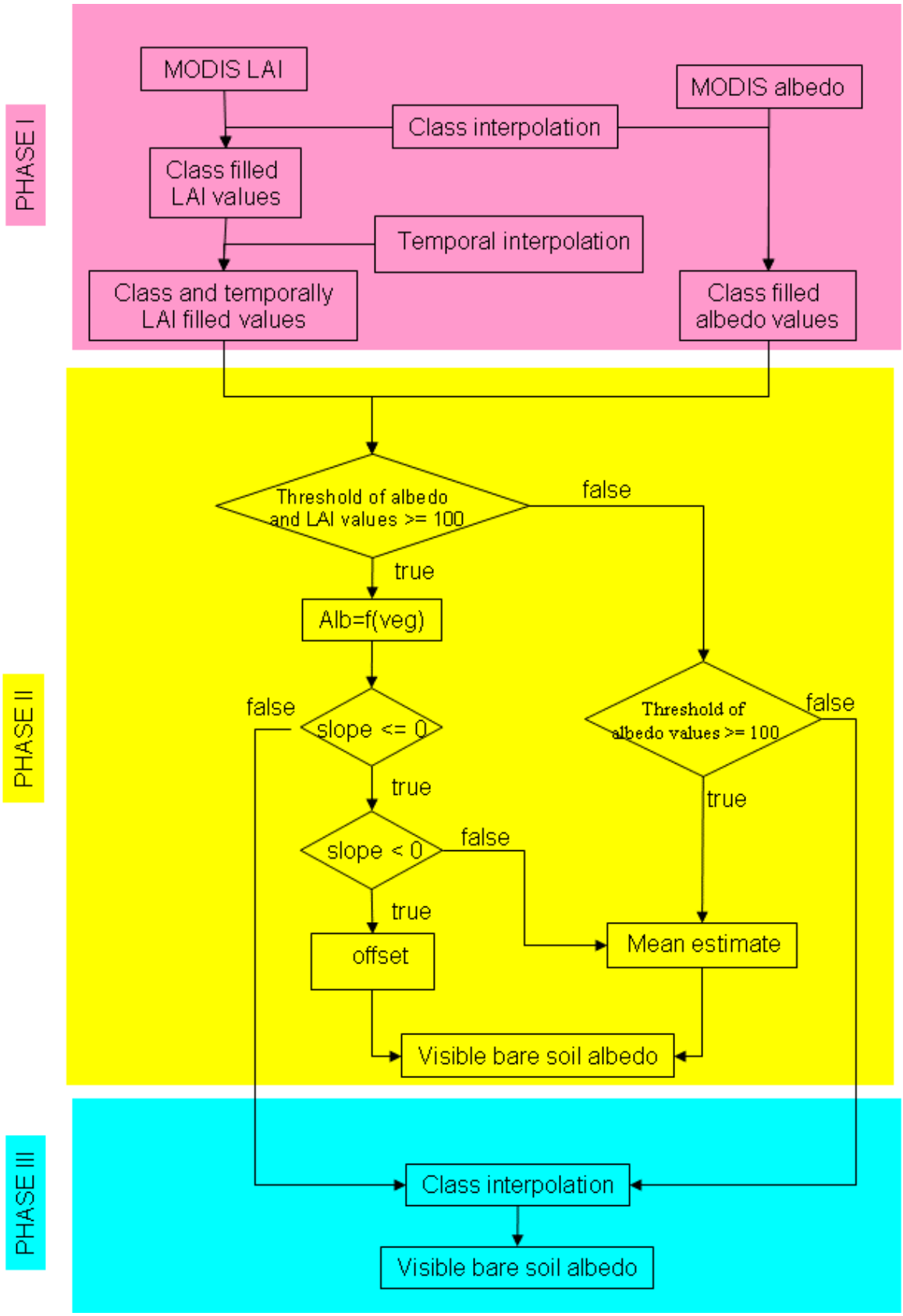

Figure II.6 Méthodologie d'obtention de l'albédo du sol nu

Finalement, des diagrammes de dispersion ont été produits pour comparer ces albédos avec d'autres données telles que celles de MODIS et celles obtenues à partir du capteur Spinning Enhanced Visible and Infrared Imager (SEVIRI) (Geiger et al., 2008) à bord de la plate-forme Meteosat Second Generation (MSG) dont la fréquence temporelle est très élevée (15 min). Afin d'accroître la consistance lors du processus de comparaison, les produits ayant une faible fréquence temporelle tels que MODIS et ceux nouvellement obtenus ont été interpolés à la date pour laquelle les observations à partir de SEVIRI ont été considérées (annexe C).

2.5 Evaluation de la nouvelle physiographie

2.5.1 La carte d'occupation des sols

La figure 2.7 montre la nouvelle carte d'écosystèmes sur la zone 'AMMA' issue de la combinaison des informations de GLC2000 et ECOCLIMAP-I. En règle générale, la distribution spatiale de la végétation en Afrique de l'Ouest correspond étroitement aux zones climatiques. Il existe en effet un gradient latitudinal prononcé sur la zone 'AMMA'. Les zones qui reçoivent un cumul important de précipitations possèdent également un volume important de biomasse ou de productivité primaire. Ainsi, le couvert végétal a ainsi subi d'importants changements à cause de la raréfaction des épisodes pluvieux. La végétation s'adapte ainsi à la pluviométrie car le calendrier des précipitations influence la quantité et le type de végétation pouvant se développer et ceci sur des échelles de temps différentes (Stock, 2004). Par exemple, les savanes clairsemées formées de quelques arbres se trouvent dans des zones de longues saisons sèches, tandis que les forêts tropicales denses apparaissent dans les zones où les précipitations surviennent tout au long de l'année.

Les variations saisonnières du LAI de 2000 à 2007 moyennées par classe sont représentées sur la figure 2.8. Une forte saisonnalité est observée pour la savane. Pour ces surfaces, à l'exception d'évènements tels que les feux de brousse ou des inondations, le paysage du LAI suit bien la distribution de l'occupation des sols avec des valeurs élevées pour la végétation ligneuse et de faibles valeurs pour les herbacées. La variabilité intra-

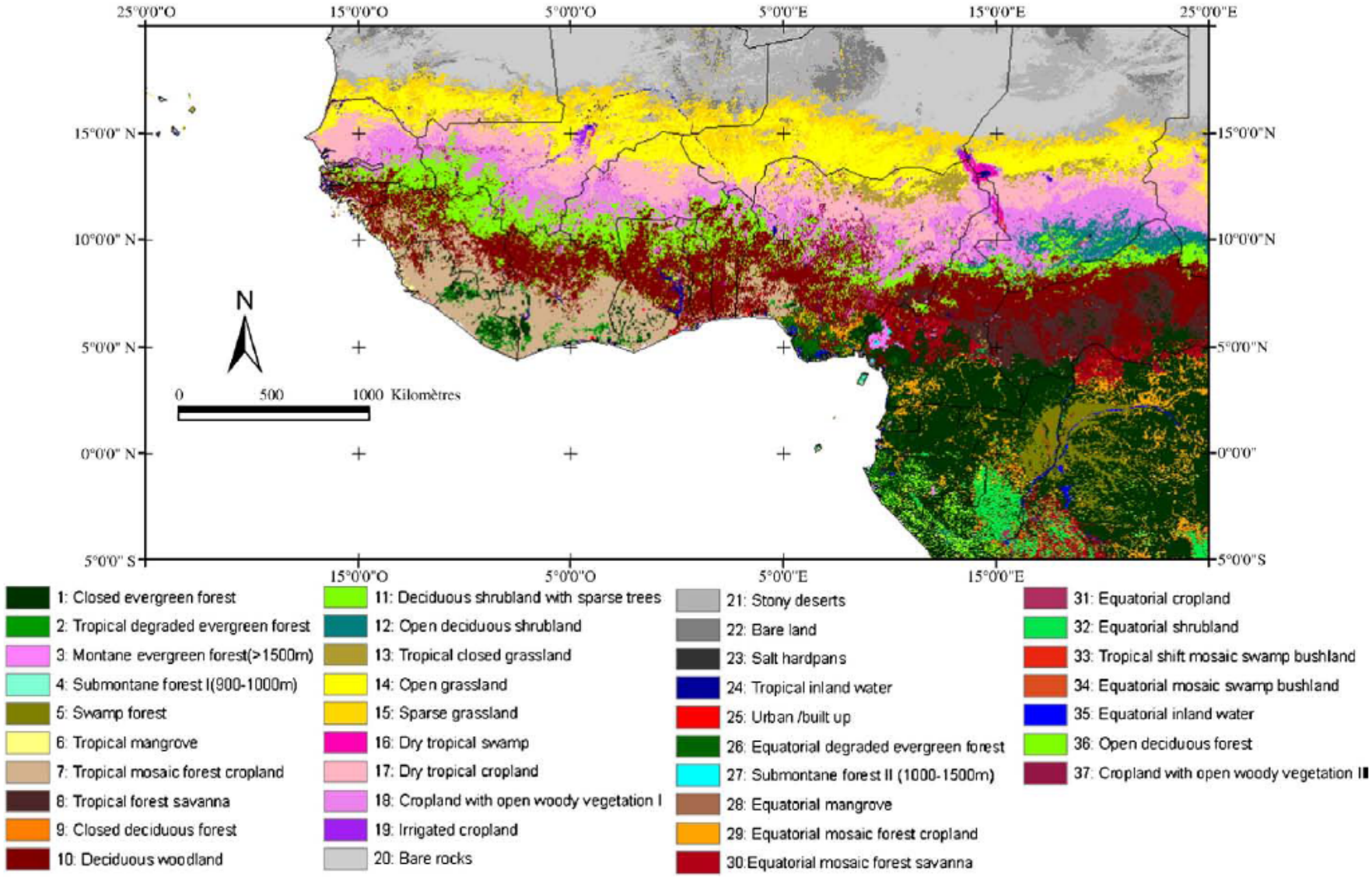

Figure II.7 Classification supervisée à 37 écosystèmes sur la zone 'AMMA'

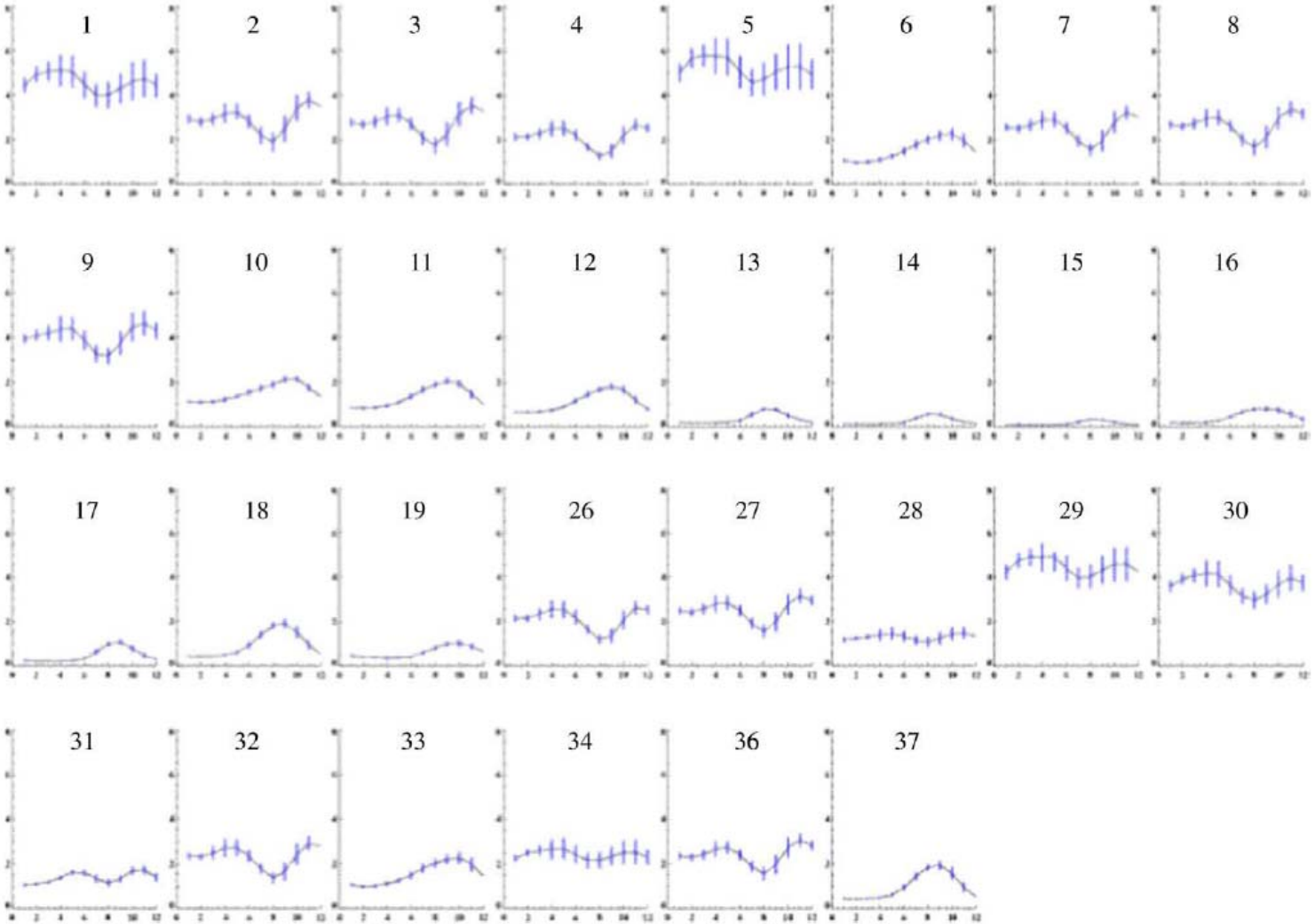

Figure II.8 Moyenne annuelle du LAI saisonnier MODIS sur la zone 'AMMA'. Les barres verticales décrivent la variabilité inter-annuelle

annuelle offre une réponse aux conditions environnementales (température, précipitation, topographie). De fortes valeurs de l'écart type des LAI mensuels sont observées pour les canopées denses avec un pic notable lorsque le LAI atteint un maximum (figure 2.8). Sur la base de cette nouvelle carte, des paramètres tels que la fraction de couverture végétale et la rugosité de surface vont pouvoir être réactualisés car ils sont dépendants du LAI. La variabilité intra-classe et la variabilité inter-annuelle produisent une information utile pour des scénarios d'assimilation des conditions de surface dans les modèles de surface en fournissant des estimations quantitatives sur la pertinence du paramètre LAI.

La projection sur la grille MERIS/GLOBCOVER à 300 m des nouvelles classes a permis d'évaluer l'accord entre les deux produits de classification. À cet égard, il apparaît que la proportion de zones urbaines et les terres cultivées est surestimée dans la nouvelle classification alors que la proportion de terres nues est en sous-estimée. En effet, la faible intensification des techniques agricoles se traduit par le fait que les caractéristiques spectrales et temporelles des cultures restent proches de celles de la végétation environnante naturelle et font partie d'un paysage hétérogène (Ramankutty, 2004). Cela rend particulièrement difficile la caractérisation des terres cultivées en Afrique occidentale.

Un examen statistique des grandes unités du paysage pour la nouvelle carte révèle que le sol nu couvre plus de 22% de la superficie des surfaces continentales, et qu'il se situe pour l'essentiel dans le désert du Sahara. Le couvert forestier représente 17% de la superficie continentale alors que la forêt dense sempervirente est principalement localisée dans le Bassin du Congo. La figure 2.7 permet également d'identifier les clairières dans les zones de déforestation avec par exemple la fragmentation du couvert forestier dans le bassin du Congo. Forêt mixte, savane arborée, savane arbustive et mosaïques couvrent ensemble 32% des terres tandis que les savanes herbacées représentent 13%. Comme le montre le tableau 2.3, l'accord général ou 'overall agreement' (OA) et le coefficient Kappa (κ) sont respectivement de l'ordre de 74,6% et 0,68 pour tous les types agrégés d'occupation des sols, ce qui souligne un bon accord. L'erreur de classification est principalement associée aux catégories "forêt mixte" et "savane mixte". La catégorie 'forêt' est le type de couvert

végétal qui possède la plus forte valeur d'exactitude selon le producteur ou 'producer accuracy' (PA) de 73,05%, et la valeur la plus élevée de l'exactitude selon l'utilisateur ou 'user accuracy' (UA), de 69,42% (voir annexe B pour la définition de toutes ces quantités).

GLOBCOVER													
		Sea	Inland water	Forest	Mixed forest	Woodland	Shrubland	Grassland	Mixed savanna	Agriculture	Urban	Bare land	Total	UA (%)
ECOCLIMAP II	Sea	65,349,670	7133	18,508	13,326	1954	2876	544	6304	1100	401	6512	65,408,328	99.91
	Inland water	29,628	270,888	58,413	78,135	9215	16,612	15,671	33,460	19,575	457	28,088	560,142	48.36
	Forest	20,106	75,263	11,180,153	2,671,681	275,441	338,990	33,444	1,499,499	8461	542	858	16,104,438	69.42
	Mixed forest	8647	30,841	3,295,241	522,042	2,067,718	708,674	11,895	4,334,841	73,897	1079	3182	11,058,057	4.72
	Woodland	1542	25,407	524,262	596,269	5,216,525	3,409,177	11,555	1,323,891	342,274	933	11,006	11,462,841	45.51
	Shrubland	2593	22,703	197,660	867,675	1,544,031	4,147,901	141,970	506,474	916,982	455	12,286	8,360,730	49.61
	Grassland	3713	12,567	4680	816,457	8064	8754	6,326,750	1,797,146	261,529	1436	3,112,745	12,353,841	51.21
	Mixed savanna	1507	45,734	8346	34,673	18,993	48,679	2430	91,014	31,059	116	9715	292,266	31.14
	Agriculture	6710	39,601	16,890	1,039,872	174,258	2,205,582	192,358	3,154,252	7,738,616	5838	457,382	15,031,359	51.48
	Urban	1290	2125	1332	2720	143	1475	1354	3204	2946	40,496	929	58,014	69.8
	Bare land	4196	3671	0	91,320	4	123	1,235,171	6289	3443	275	19,965,492	21,309,984	93.69
	Total	65,429,602	535,933	15,305,485	6,734,170	9,316,346	10,888,843	7,973,142	12,756,374	9,399,882	52,028	23,608,195	162,000,000	
	PA (%)	99.87	50.55	73.05	7.75	55.99	38.09	79.35	0.71	82.33	7784	84.5		

Kappa coefficient = 0.68. Overall accuracy = 74.6%.

Tableau II-3 Matrice de confusion entre GLOBCOVER et la classification supervisée sur la zone 'AMMA'

2.5.2 Les variables biophysiques

La grandeur moyenne des valeurs de LAI présente une forte variabilité spatiale avec des valeurs de LAI avoisinant 6 (sur les zones de forêt dense) à des valeurs quasiment nulles en période de sécheresse (sur les zones de savanes herbacées). Cette variabilité spatiale s'accompagne d'une variabilité inter-annuelle assez prononcée sur les couverts forestiers (figure 2.8). Ceci est lié aux trois principales sources d'incertitudes pouvant affecter la précision des données LAI MODIS (Baret et Buis, 2008 ; Yang et al., 2006) : i) le modèle utilisé pour construire les tables de correspondance ; ii) la classe du biome utilisée en entrée ; iii) la réflectance de la surface en entrée. Par exemple, des couverts denses en bonne santé présentent un phénomène de saturation de la réflectance de surface aux valeurs élevées de LAI. Pour de pareils cas, la précision sur la réflectance est comparable au bruit résiduel (étalonnage, nuages, aérosols) et *de facto* le LAI ne peut être correctement obtenu.

L'incertitude sur la détermination du LAI en fonction du modèle utilisé et d'autres facteurs tels que les nuages, les aérosols, et la mauvaise classification du biome est analysée sur la base des informations du paramètre QC (Quality Control). Par exemple une mauvaise

classification du biome peut propager des erreurs à deux niveaux : (i) sur l'utilisation erronée d'une table de correspondance durant le processus d'estimation et (ii) au sein de la procédure d'apprentissage de l'algorithme lors du développement des tables de correspondance. Cette analyse a été effectuée pour les deux classes 'forêt dense sempervirente' et 'savane arborée décidue' (figure 2.9). Ces classes qui appartiennent à différentes zones écoclimatiques (Leibovici, et al., 2007). Elles couvrent de vastes zones, et des valeurs extrêmes de la variance inter-annuelle du LAI sont observées (figure 2.8). La classe 'forêt dense sempervirente' montre un pic mineur (pour 35% de pixels) par utilisation du MA avec les meilleurs résultats obtenus aux alentours du mois de juillet à cause de l'absence prolongée d'ennuagement pendant la courte période sèche. Mais le principal pic apparaît vers la mi-janvier (55% de pixels) au cours de la longue saison sèche (figure 2.9). La contamination par les nuages est présente pour les deux classes. Plus de 60% des valeurs de LAI MODIS sont obtenues pour les deux classes avec une bonne classification des biomes utilisés en entrée. La figure 2.9 montre également que la contamination moyenne mensuelle par les aérosols varie dans un sens opposé à la saisonnalité de l'ennuagement. Nous pensons que la contamination nuageuse empêche la détection de la contamination des aérosols et également que l'augmentation des aérosols pendant la saison sèche est due aux activités de combustion de la biomasse dans le bassin du Congo. D'autre part, les incertitudes sur la réflectance de la surface spectrale peut entraîner des variations importantes des valeurs de LAI à cause du phénomène de saturation et du caractère instable de la procédure d'estimation du LAI pour la classe 'forêt dense sempervirente'.

Afin de séparer l'effet de la végétation de celui de la surface sous-jacente représentée soit par la litière, soit par le sol, nous avons considéré la fraction de couvert végétal *veg* provenant du LAI en utilisant l'équation (2.3). La figure 2.10 montre la moyenne trimestrielle du *veg* s'étendant sur 8 années (2000-2007). L'analyse des variations saisonnières confirme le pic attendu du *veg* en été (juin, juillet et août) et les valeurs faibles en hiver (décembre, janvier et février). Comme prévu, les valeurs les plus élevées de *veg* sont occurrentes dans les domaines couverts par la forêt et la végétation ligneuse durant la saison des pluies.

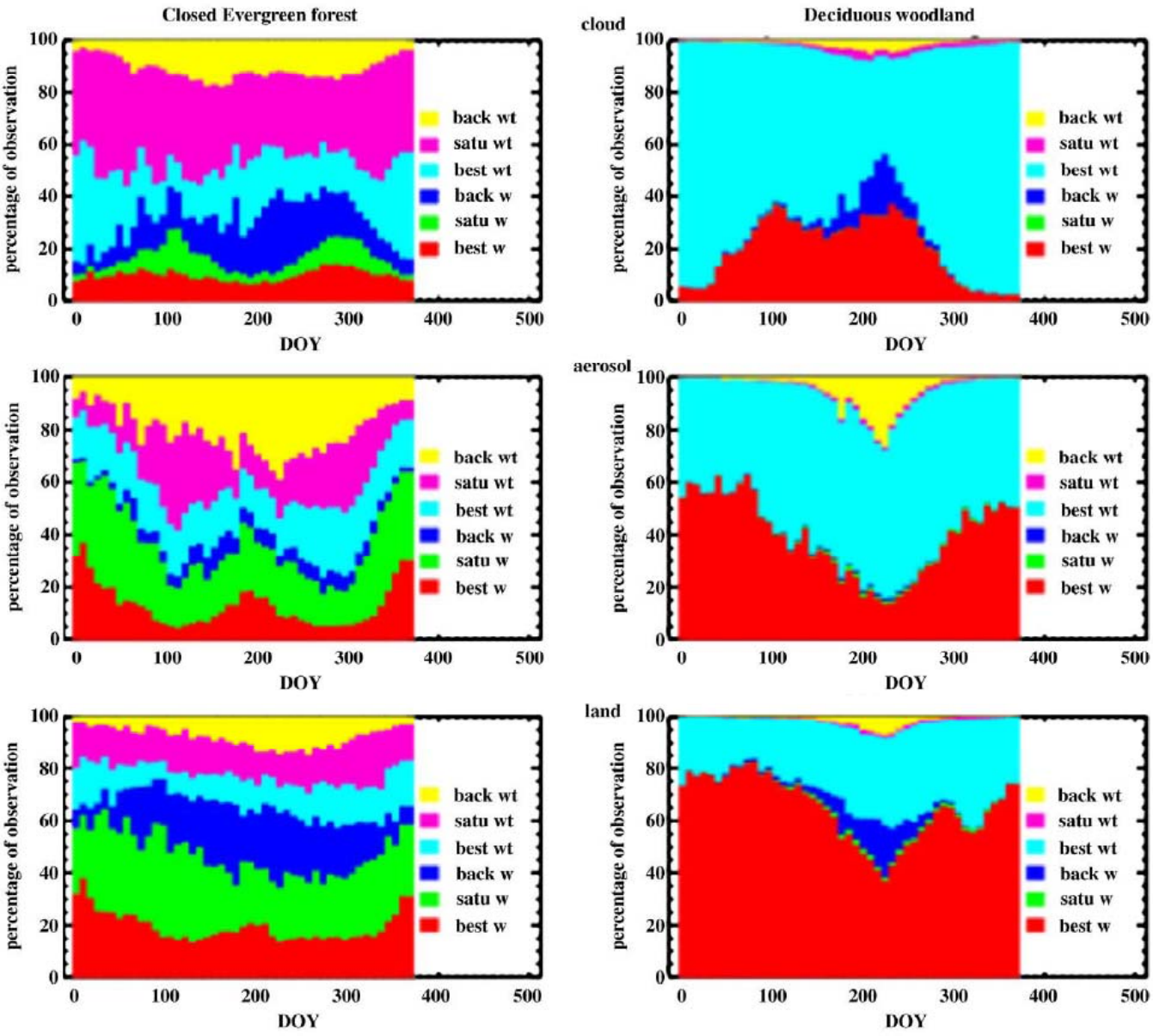

Figure II.9 Impact des nuages, aérosols et mauvaise classification des biomes sur la zone 'AMMA' lors de l'estimation du LAI MODIS pour les classes 'forêt dense sempervirente' (droite) et 'savane arborée décidue' (gauche). *wt* et *w* signifient sans (without) et avec (with), respectivement, *best* (*satu*) signifie que le MA a été utilisé en produisant un résulat meilleur (saturé) et *back* signifie que le BUA a été utilisé. En haut à gauche, le rouge représente le pourcentage d'observation nuageuses quand le MA est utilisé avec les meilleurs résultats, tandis que le bleu foncé représente le pourcentage d'observations nuageuses pour lesquelles le BUA a été utilisé durant la période 2000 à 2007.

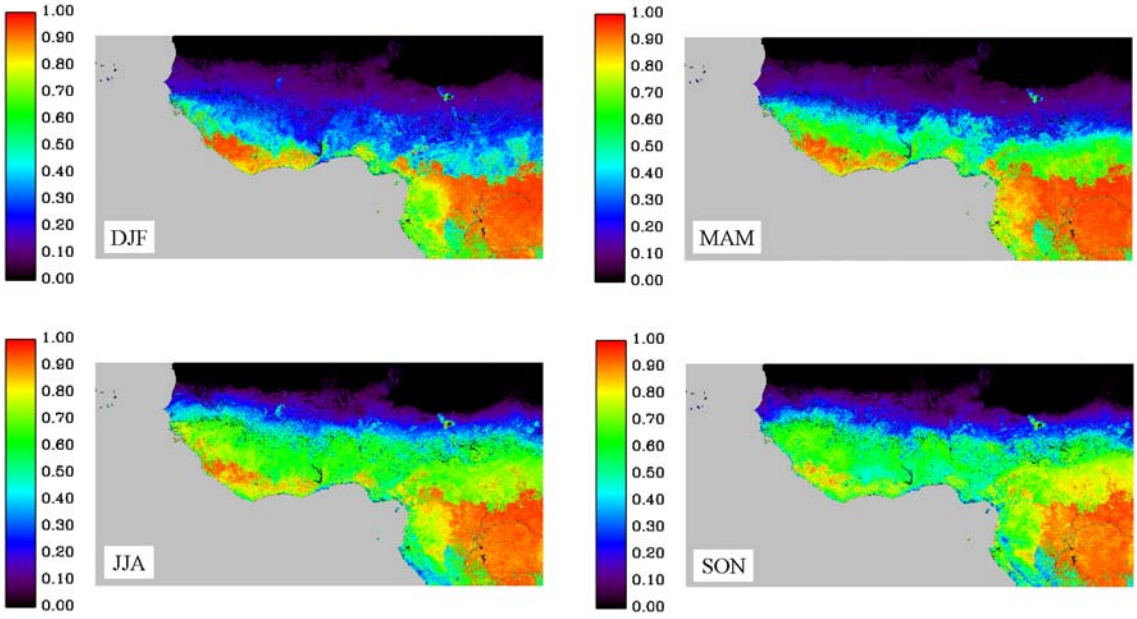

Figure II.10 Variations de la fraction de végétation

L'équation (2.1) a permis de déterminer les composantes du sol et de la canopée des albédos de surface MODIS. L'albédo du sol nu révèle des détails de la structure du paysage (figure 2.11), offrant ainsi une vision claire des domaines de sol nu dans la partie nord-est de la zone 'AMMA'. Les roches nues et les forêts présentent respectivement les plus faibles et les plus fortes valeurs de l'albédo bi-hémisphérique du sol nu et sont également facilement identifiables. Le ratio de l'albédo du proche infrarouge et celui du visible pour le sol nu varie de 1,62 à 1,97 en accord avec des études précédentes (Zhou et al., 2005). L'albédo total large bande de la nouvelle physiographie est déterminé en utilisant les coefficients de conversion obtenus par un ajustement des albédos MODIS visible et proche infrarouge à large bande (Eq. (2.4)). Pour les zones de végétation, la valeur moyenne des coefficients de l'équation (2.4) sont α=0.002, β=0.461, et γ=0.455. Ils deviennent 0.003, 0.271 et 0.606 pour le sol nu. Ces coefficients sont sensiblement égaux à ceux obtenus avec l'albédo total MODIS larges bandes (voir tableau 2.4).

Ainsi, les coefficients de l'équation (2.4) sont supposés fiables pour l'estimation de l'albédo sur la zone 'AMMA' parce que la conversion des petites bandes en larges bandes MODIS indique que le visible et le proche infrarouge sont à peu près pondérés

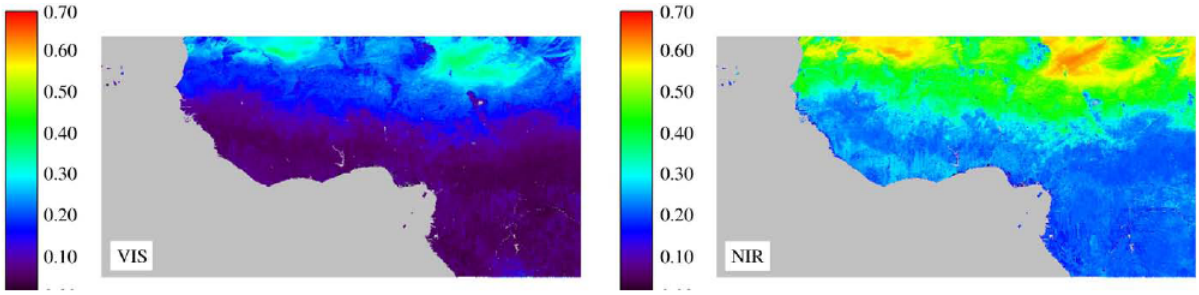

Figure II.11 Albédo du sol nu dans le visible et le proche infrarouge sur la zone 'AMMA'

		MODIS		New physiography	
		Vegetated	Other	Vegetated	Other
MODIS	α	0.006	0.005	0.002	0.003
	β	0.477	0.289	0.461	0.271
	γ	0.487	0.594	0.455	0.606
	rms	0.003	0.004	0.021	0.009
	r^2	0.998	0.996	0.893	0.979

Tableau II-4 Paramètres et statistiques de conversion entre divers albedo MODIS larges bandes sur la zone 'AMMA'

équitablement dans le calcul de l'albédo solaire (Liang et al., 1999). Bien qu'il y ait une légère dispersion entre les albédos reconstitués dans les spectres du visible et du proche infrarouge par rapport à ceux de MODIS (figure 2.12), le tableau 2.4 montre que les nouveaux albédos reconstitués donnent de bonnes valeurs. En effet, le tableau 2.4 indique des valeurs proches de l'albédo MODIS dans le visible, le proche infrarouge et le spectre total des ondes courtes. Le résultat pour les surfaces continentales, avec et sans végétation, donne une variance expliquée de 89% avec une erreur quadratique moyenne (RMS) d'environ 2%. La figure 2.12 montre que l'albédo total large bande dans tout le spectre des ondes courtes possède des valeurs avoisinant 0.2 sur la forêt et 0.25 sur les surfaces herbacées. Cette figure révèle également que le coefficient de détermination entre les albédos ondes courtes reconstitués avec d'autres jeux de données tels que MODIS et SEVIRI diminue pendant les périodes humides. Les albédos MODIS et SEVIRI sont généralement supérieurs aux albédos reconstitués en période humide tandis que l'albédo journalier SEVIRI est supérieur à l'albédo reconstitué durant les périodes sèches.

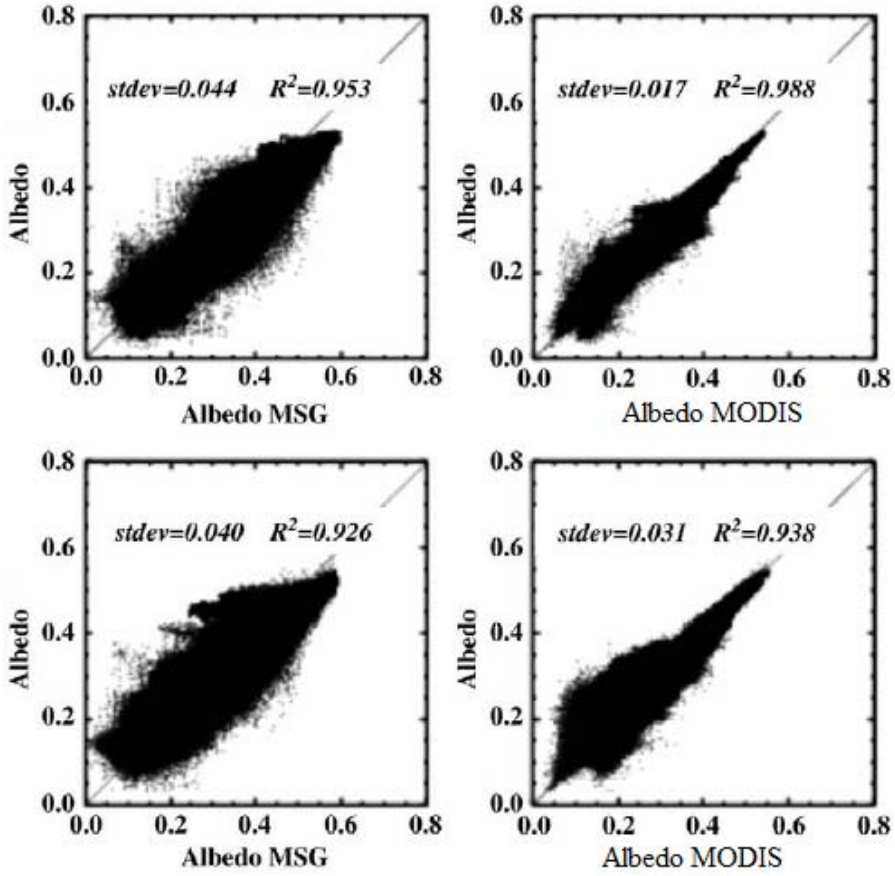

Figure II.12 Diagrammes de dispersion sur la zone 'AMMA' entre les albédos larges bandes SEVIRI (gauche) et MODIS (droite) et la nouvelle physiographie. Graphes du 6 mars 2006 (en haut) et du 6 septembre 2006 (en bas)

2.6 Résumé et conclusion

Dans ce chapitre, nous avons présenté une nouvelle classification composée de 37 écosystèmes pour la zone 'AMMA'. La méthode présentée ici diffère des classifications conventionnelles car elle se fonde sur la compilation de cartes d'occupation des sols GLC2000 et ECOCLIMAP-I, et sur une analyse de la dynamique du LAI MODIS et de l'information topographique et climatique. L'idée était de détecter les écosystèmes présents au sein des classes de GLC2000. À la résolution modérée offerte par MODIS qui est de 1 km, la mixité des classes déjà présente pour la majorité des pixels demeure à l'exception toutefois de la végétation et des déserts qui ressortent bien séparés. La nouvelle classification proposée, exhaustive au vu de la grande surface échantillonnée, a le mérite d'être compatible avec des produits de meilleure résolution tels que GLOBCOVER. Chaque

pixel étant traité de manière indépendante, la cohérence spatiale atteste indirectement de la robustesse de cette méthode. Un atout de la nouvelle physiographie réside dans la cohérence des variables biophysiques principalement utilisées en météorologie que sont l'albédo de surface et le LAI. Contrairement à ECOCLIMAP-I, le nouveau forçage physiographique intègre la variabilité inter-annuelle des paramètres biophysiques de surface et offre par conséquent de nouvelles perspectives intéressantes en termes de suivi pour des études climatiques. Nous avons présenté une technique permettant d'estimer l'albédo du sol nu et celui de la végétation à partir des données de télédétection afin de produire de meilleures estimations de l'albédo total de surface dans les LSM. L'albédo du sol nu est ainsi dérivé d'une analyse de l'albédo de surface totale et de la fraction de végétation. Parce que la procédure intègre les conditions sèches et humides, l'albédo du sol nu résultant peut être plus faible que celui d'autres bases de données qui se sont focalisées uniquement sur les conditions sèches (voir section 3.6).

Par la suite, une étude d'impact de la nouvelle physiographie sur des flux de surface à partir du LSM ISBA sera réalisée (voir chapitre 4). L'incertitude sur les paramètres de surface de la nouvelle physiographie, en fonction de la variabilité intra-classe et inter-annuelle, favorise leur utilisation lors de l'assimilation de données de carbone et lors de la modélisation du bilan hydrique. À cet égard, les écarts d'un albédo de sol nu statique, tels qu'observés par un capteur géostationnaire semblent pouvoir être utilisés pour déduire par exemple l'humidité superficielle. Il est important de noter que l'amélioration de la cohérence entre l'albédo de surface et le LAI par référence à MODIS permettra d'améliorer la situation, d'où l'importance de concevoir une carte de changements d'utilisation du sol à 1 km qui reflètera les transformations progressives des surfaces continentales en Afrique de l'Ouest, i.e. les modifications et les conversions en terme de type et d'intensité. Ce travail ayant été essentiellement effectué sur l'Afrique de l'Ouest, il est nécessaire de l'étendre sur l'ensemble du continent africain, ce qui est réalisé au chapitre suivant.

« Qui voit de haut, voit bien ; qui voit de loin voit juste », Victor Hugo, l'exilé satisfait dans L'art d'être grand père, 1877.

III. Le produit ECOCLIMAP-II sur l'Afrique[2]

3.1 Introduction

Dans le chapitre précédent, nous avons présenté une carte des composantes végétalisées des écosystèmes de l'Afrique de l'Ouest à 1 km de résolution, réalisée dans le cadre du programme AMMA, et obtenue par classification supervisée du LAI satellitaire. Au cours de cette étude, on a pu mettre en évidence d'importantes différences d'occupation du sol sur la zone dite 'AMMA', cela plaide en faveur d'une réactualisation de l'occupation des sols et par conséquent des variables biophysiques de surface à l'échelle du continent. La corrélation entre le LAI MODIS et le NDVI VGT ayant également été établie, le choix de l'un ou de l'autre produit comme donnée d'entrée pour une cartographie des écosystèmes devrait normalement conduire à des résultats identiques ou alors presque similaires.

Une cartographie précise et efficace des écosystèmes par télédétection exige de choisir l'approche de classification la mieux appropriée. Celle-ci doit notamment tenir compte de l'échelle de la zone d'étude, des résolutions spatiale, temporelle et spectrale des données, et des besoins de l'utilisateur (Cihlar, 2000). En raison de leurs capacités spécifiques, des classificateurs différents produiront des résultats différents même s'ils utilisent le même jeu de données. Par exemple, en utilisant des données mensuelles composites de NDVI AVHRR acquises entre avril 1992 et Mars 1993, Loveland et al. (2000) ont identifié, avec pour classificateur les "k-plus proches voisins" (k-nearest neighbour, k-NN) qui est une stratégie de classification non supervisée, 17 classes en écosystèmes globaux. En utilisant le même jeu de données AVHRR, avec cette fois-ci un arbre décisionnel (Decision Tree, DT) pour classificateur qui est une stratégie de classification supervisée, Hansen et al. (2000) ont

[2] Ce chapitre est fondé sur :
(i) Kaptué, T. A. T., De Jong, S. M., Roujean, J.-L, Favier, C. and Mering, C. (2010), Ecosystem mapping at the African continent scale using a hybrid clustering approach based on 1-km resolution multi-annual data from SPOT/VEGETATION, *Remote Sensing of Environment*, **doi:10.1016/j.rse.2010.09.015**.
(ii) Kaptué, T. A. T, Roujean, J.-L and De Jong, S. M. (2010), Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale, *International Journal of Applied Earth Observation and Geoinformation*, **.doi: 10.1016/j.jag.2010.11.005**.
(iii) Kaptué, T. A. T., Bégué, A., Roujean, J.-L, Garrigues, S., Los, S. O., Grey, W. M. F., Boone, A., Mahfouf, J.-F. and Badiane, D. (2010), ECOCLIMAP-II: a new characterisation of the land surface heterogeneity at the African continental scale for use in land surface models, *Global Change Biology*, **submitted**.

identifié 14 classes d'écosystèmes globaux. Un autre aspect d'importance primordiale pour la cartographie des écosystèmes par télédétection sur de larges échelles est la prise en compte du contexte spatial puisque l'hétérogénéité du paysage peut induire des variations importantes de la réflectance et par conséquent du NDVI au sein du même écosystème (voir section 1.3.4). Pour pallier cet inconvénient, la zone d'intérêt est généralement stratifiée en écorégions, i.e. des régions identiques d'un point de vue écologique et similaires d'un point de vue comportement spectral. A partir de l'information latitudinale par exemple, Han et al. (2004) ont divisé la France en deux parties tandis que Bicheron et al. (2008) ont stratifié le monde à partir de considérations climatiques. Dans les deux cas, le processus de stratification a été déterminé de façon plus ou moins arbitraire, plutôt que par le développement d'un formalisme, ce qui constituerait un procédé reproductible et applicable à n'importe quel domaine.

Dans ce chapitre, nous proposons une nouvelle approche hybride de classification (que nous appellerons *Stacked Segment Classification*) basée sur l'information spectrale et temporelle des données décadaires de NDVI VGT et qui intègre les principes de partition et de hiérarchie. Cette méthode a été spécialement développée pour la cartographie des écosystèmes aux échelles continentales afin de mettre à jour la base de données ECOCLIMAP (Masson et al., 2003 ; Champeaux et al., 2005) sur l'Afrique. Pour rappel, nous avons décrit au chapitre 2 ECOCLIMAP comme étant une base de données développée à l'initiative du Centre National de Recherches Météorologiques (CNRM) et ayant pour but premier de répondre aux besoins de la communauté météorologique. Cette base nourrit la description des écosystèmes naturels/protégés fonctionnellement homogènes dans le cadre de la modélisation de l'étendue des changements climatiques et de leur impact. A travers cette nouvelle carte des écosystèmes, on vise une meilleure discrimination des types de végétation comparés à leur représentation sur les cartes pancontinentales existantes telles qu'ECOCLIMAP-I, GLC2000, GLOBCOVER et MODIS (voir tableau 3.1). Contrairement à ces cartes pancontinentales produites en utilisant seulement une année de données satellites, l'utilisation de séries temporelles pluriannuelles va nous permettre de générer une

carte couvrant un certain nombre d'années et non pas contrainte par des événements annuels. En effet, le profil temporel de la réflectance d'un pixel au cours d'une année peut être atypique en raison de la sensibilité aux conditions climatiques. Un exemple typique est un couvert herbacé stressé tout au long d'une année et qui recouvre son activité photosynthétique normale l'année suivante.

Notre méthode hybride de classification diffère des méthodes de classification existantes parce qu'elle combine les principes de partition au sein d'une stratification hiérarchique. Celle-ci s'effectue à partir de la Transformée de Fourier Discrète (TFD) appliquée sur 8 années de données de NDVI VGT (2000-2007). Ensuite, chaque écoregion est affinée à l'aide du classificateur k-NN. La classification résultante est comparée avec celle obtenue par utilisation directe du classificateur k-NN sur le même jeu de données. Ce processus a conduit à l'obtention de 3 cartes originales à 1 km de résolution : une carte représentant les écorégions bioclimatiques et deux autres représentant la distribution spatiale des écosystèmes.

Ainsi ce chapitre a pour objectifs de proposer et de démontrer une méthode hybride innovante de classification et son application pour la cartographie des écosystèmes sur de larges échelles, basée sur l'information spectrale et temporelle offerte par les données décadaires VGT. Plus précisément, les objectifs spécifiques de ce chapitre sont :

(i) la description de la *stacked segment classification* ;

(ii) l'illustration de sa valeur ajoutée pour l'amélioration de l'exactitude de la classification à travers son application sur des séries temporelles de NDVI couvrant l'intégralité du continent africain et des comparaisons par pixel entre la classification résultante et les classifications continentales existantes et récentes ;

(iii) l'estimation des variables biophysiques par type d'écosystème ainsi cartographiés et l'illustration de leur variabilité interannuelle.

Product	Reference	Sensor	Input Data	Time of data collection	Spatial resolution	Projection	Number of classes	Nomenclature	Classification method	Supplemental data	Validation scheme
IGBP-Dis	Loveland et al. (2000)	AVHRR	NDVI	01/04/1992-31/03/1993	1/120°	Plate Carrée	14	IGBP	Unsupervised classification using k-NN	Digital chart of the world of Danko et al. (1992)	Confusion matrix
UMD	Hansen et al. (2000)	AVHRR	5 spectral bands and BNDVI	01/04/1992-31/03/1993	1/120°	Plate Carrée	11	UMD	Supervised classification using decision tree	None	Confusion matrix
ECOCLIMAP-I	Masson et al. (2003)	AVHRR	NDVI	01/04/1992-31/03/1993	1/120°	Plate Carrée	51	UMD	Supervised classification	UMD product Climate map of Koeppe and De Long (1958)	None
GLC2000	Mayaux et al. (2004)	VGT-I	4 spectral bands and NDVI	01/11/1999-31/12/2000	1/112°	Plate Carrée	27	LCCS	Flexible classification using local/regional expertise	JERS/SAR ERS-I/SAR DMSP/OLS GTOPO30	Intercomparison Statistical sampling High resolution data
GLOBCOVER	Arino et al. (2008)	MERIS	13 spectral bands	01/12/2004-31/06/2006	1/360°	Plate Carrée	42	LCCS	Flexible classification using local/regional expertise	None	Statistical sampling
MODIS LC-I	Friedl et al. (2010)	MODIS	7 spectral bands, LST and NDVI	01/01/2005-31/12/2005	1/240°	Integrid Sinusoidal Grid	14	IGBP	Supervised classification using decision tree	None	High resolution data Confidence values Confusion matrix
ECOCLIMAP-II	This study	VGT-I and VGT-II	NDVI	01/01/2000-31/12/2007	1/112°	Plate Carrée	73	LCCS	Hybrid unsupervised classification using the Fourier Transform	None	Intercomparison at national and continental scale High resolution data

Tableau III-1 Principales caractéristiques des cartes pancontinentales d'occupation des sols en Afrique d'une résolution kilométrique à une centaine de mètres

3.2 Prétraitements des données d'observations de la terre

L'information qualitative fournie par le fichier SM des données VEGETATION (cf. section 1.4.2), et de façon plus précise celle les bits de numéro 6, 5, 1 et 0, a été utilisée pour qualifier les données de NDVI comme *acceptables*, *bonnes* et *parfaites*. En effet, d'après VITO (http://www.vgt.vito.be/faq/FAQS/faq15.html), les bits de numéro 6 et 5 sont utilisés pour décrire la qualité radiométrique des bandes du rouge et du proche infrarouge respectivement tandis que les deux autres bits de numéro 1 et 0 sont utilisés pour décrire les conditions du ciel durant la période de composition. Toutes les valeurs de NDVI obtenues sont considérées comme *acceptables*, les données *acceptables* obtenues avec de bonnes valeurs de la réflectance dans les domaines spectraux du rouge et du proche infrarouge sont *bonnes* et les données *parfaites* sont de *bonnes* données traitées dans des conditions de ciel clair. On observe un nombre inférieur de données *acceptables* par comparaison avec le nombre de données *parfaites*. Pour les forêts denses, le pourcentage des images *acceptables* sur les 288 images composites disponibles (images composites décadaires, soit 36 images par an sur les 8 années considérées) est d'environ 80% (figure 3.1-a), le ratio du nombre des données *bonnes* sur celui du nombre des données *acceptables* avoisine les 50% (figure 3.1-b) tandis que le ratio du nombre des données *parfaites* sur celui du nombre des données *bonnes* est d'environ 40% (figure 3.1-c). La sélection des *bonnes* valeurs de NDVI est donc vue comme le meilleur compromis.

La suppression des pixels contaminés en raison de la couverture nuageuse est une étape importante lors du pré-traitement des images et indispensable pour effectuer une classification. Afin de combler les lacunes et les trouées (obtenues soit par défaillance du capteur ou à la suppression du pixel en cas de contamination), nous avons effectué un lissage pixel par pixel à l'aide d'une fonction polynomiale de degré 5 (cf. section 2.2.2). Les

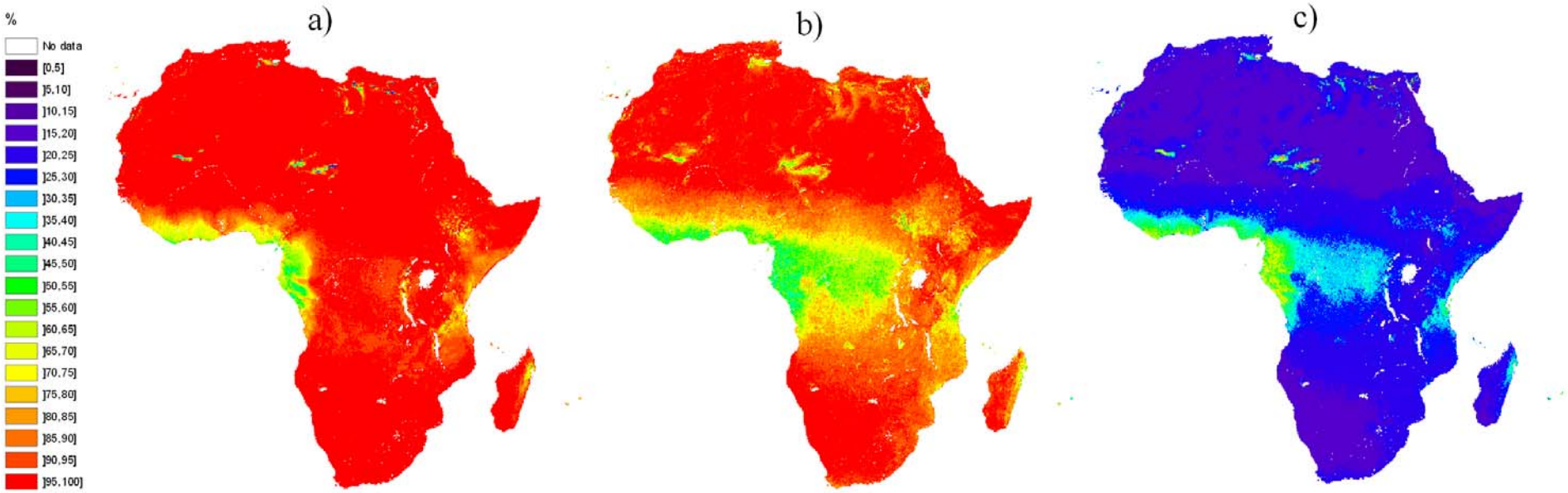

Figure III.1 Qualité des données NDVI VGT de 2000 à 2007

a) Rapport des données *acceptables* sur le nombre de décades.

b) Rapport des données *bonnes* (i.e. des données *acceptables* obtenues avec de bonnes valeurs du NDVI dans les spectres du visible et du proche infrarouge) sur les données *acceptables*.

c) Rapport des données *parfaites* (i.e. des données *bonnes* obtenues en condition de ciel clair) sur les données *bonnes*

surfaces artificielles telles que l'urbain ou le bâti ont été obtenues à partir de la carte GLC2000 (Mayaux et al., 2004). Ceci parce qu'il n'est pas possible de détecter les aires urbaines par la seule utilisation des données d'indice de végétation (Schneider et al., 2010). Les zones d'eau sont les pixels pour lesquels une valeur constante de NDVI égale à –0.1 est observée durant les 8 années de données (i.e DN=0 en vertu de l'équation 3.1). D'autres zones d'eau ont été identifiées par référence à la classe 'dépôts salins' de GLC2000. Un masque a été appliqué sur les écosystèmes 'surfaces artificielles' et 'plans d'eau' ainsi obtenus durant le processus de classification.

3.3 Classification non supervisée

L'analyse des images pré-traitées s'est effectuée en deux étapes : (i) la classification et (ii) l'évaluation (figure 3.2 (a)). Le processus de classification comprend deux techniques de classification non supervisée : la première est fondée sur l'application d'un regroupement dynamique tandis que la seconde est une approche hybride qui intègre à la fois le regroupement hiérarchique et le regroupement dynamique.

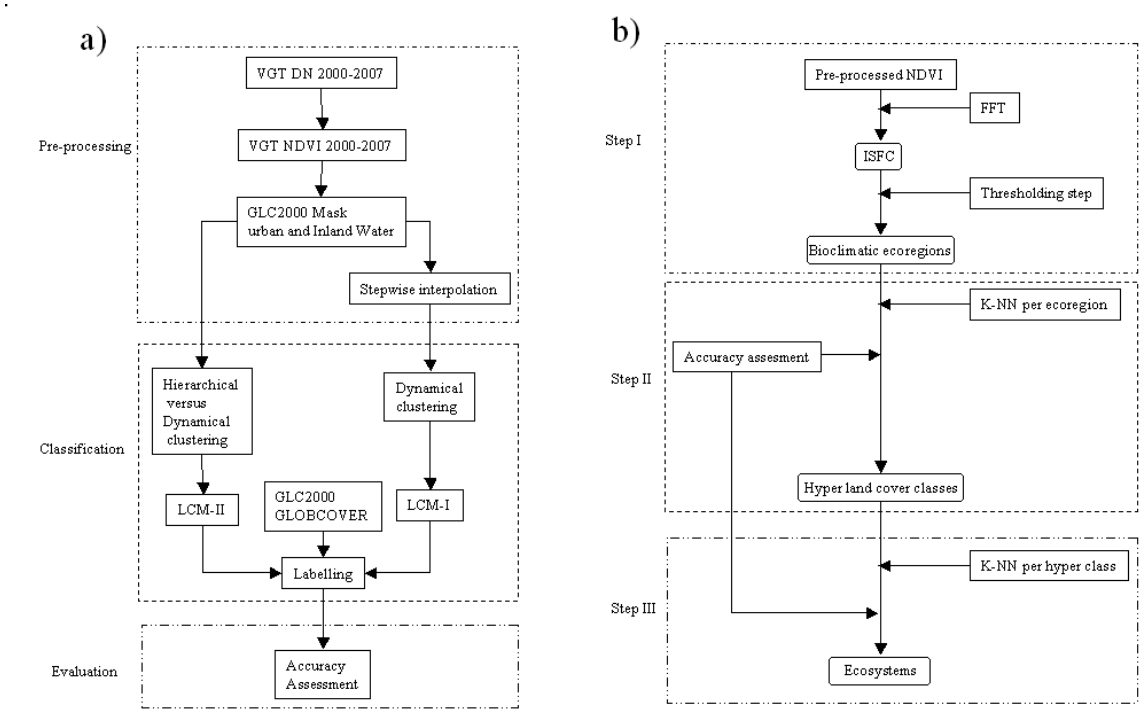

Figure III.2 Schéma méthodologique : a) Diagramme du projet ; b) ***Stacked segment classification***

3.3.1 Classificateur dynamique

Le classificateur k-NN (k-Nearest Neighbour) est une méthode de regroupement dynamique qui subdivise des données en k clusters représentés par leur centroïdes à travers un processus itératif dans le but d'optimiser une fonction de choix (Jain et al., 1999). C'est l'une des techniques de classification automatique couramment utilisée pour de grands ensembles de données en raison de : (i) la simplicité de sa mise en œuvre, (ii) son temps de calcul qui est du même ordre que la taille des données ($0(lkn)$ où *n* désigne le nombre de pixels à classifier au sein d'une image, *k* le nombre de clusters et *l* le nombre d'itérations), et (iii) son indépendance au niveau de l'ordre de présentation des centroïdes. Dans ce qui suit, on notera $NDVI_d^y$ la valeur du NDVI de la décade *d* de l'année *y* et par *Y* le nombre d'années (i.e. *Y*=8), *D* le nombre de décades (i.e. *D*=36) et ζ la fonction de choix définie par :

$$\zeta = \sum_{r=1}^{k}\sum_{u=1}^{n_r}\sum_{y=1}^{Y}\sum_{d=1}^{D}\left|NDVI_d^y\left(u^{(r)}\right) - g_d^y\left(C^{(r)}\right)\right|^2, \qquad (3.1)$$

où $C^{(r)}$, $u^{(r)}$, n_r et g_d^y sont respectivement le $r^{ième}$ cluster , le $u^{ième}$ pixel appartenant à $C^{(r)}$, la cardinalité de $C^{(r)}$ et le centroïde de $C^{(r)}$ de la $d^{ième}$ décade de l'année *y*.

Le k-NN procède comme suit pour minimiser la fonction de choix ζ :

(i) Détermination du nombre de clusters, *k*

(ii) Choix de *k* centroïdes (durant les itérations) ou de *k* points au hasard (comme sélection initiale) à partir du jeu de données

(iii) Création de *k* clusters par affectation de chaque pixel du jeu de données au centroïde le plus proche en accord avec le critère de la distance minimale

(iv) Calcul du nouveau centroïde de chaque cluster ainsi formé

(v) Répétition des étapes (iii) et (iv) jusqu'à ce que la convergence de ζ soit vérifiée. Ceci se produit lorsqu'il n'y a plus de ré-affectation de pixels aux centroïdes des nouveaux clusters ou bien lorsque le nombre d'itérations fixé par l'utilisateur est atteint.

Comme toutes les classifications par partition, l'implémentation du k-NN présente quelques défauts. L'un deux se situe au niveau de la détermination du nombre optimal *k* de clusters (Boudraa, 1999 ; Dubes, 1987 ; Everitt, 1979 ; McCullagh et Yang, 2008 ; Thorndike, 1953). Le nombre de clusters affecte l'aptitude du classificateur k-NN à saisir la part importante de la variabilité des écosystèmes dans le sens où un petit nombre induirait une mauvaise classification alors qu'un nombre élevé exigerait une fusion des clusters résultants. Un autre problème concerne la sensibilité des classes à la sélection initiale. Ceci oriente la convergence de l'algorithme vers une solution localement optimale qui n'est pas toujours la solution globalement optimale comme le montre la figure 3.3. Le choix des points A, B et C pour centroïdes initiaux conduit à la partition $\{\{A\}, \{B, C\}, \{D, E, F, G\}\}$ (illustrée en rouge) en lieu et place de la partition optimale $\{\{A, B, C\}, \{D, E\}, \{F, G\}\}$ (illustrée en bleu) obtenue en choisissant par exemple pour sélection initiale, les points A, D et F. Pour surmonter ces difficultés, on a choisi : (i) d'opter pour un nombre élevé de clusters parce que cela a pour action de réduire l'impact sur les paramètres de contrôle (Cihlar, 2000) et (ii) pour sélection initiale les centroïdes spatialement distribués des 214 clusters représentatifs (ayant une cardinalité supérieure à 5000) qui résultent du croisement entre GLC2000 et GLOBCOVER (cf. tableau 2.3).

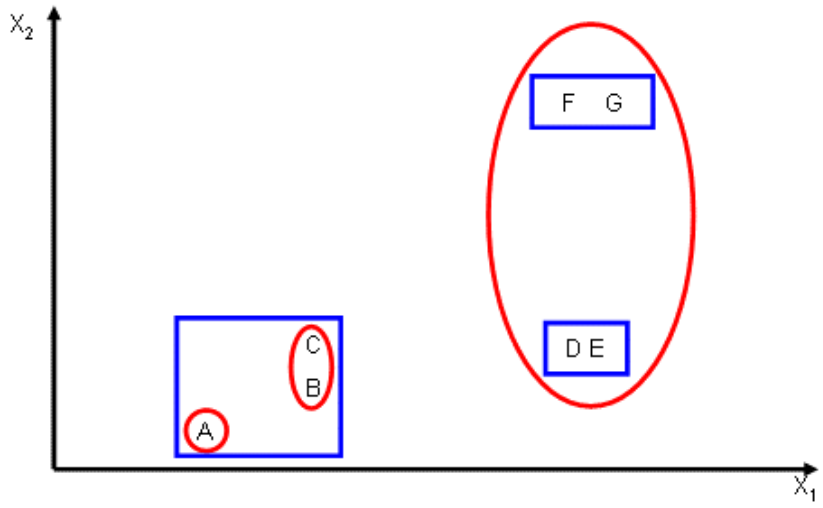

Figure III.3 Sensibilité du k-NN à la sélection initiale d'après Jain et al. (1999)

Level I: hyper-class	Level II: class
1:Forest	1:Closed lowland forest
	2:Closed to open lowland deciduous forest
	3: Montane Forest
2: Mixed forest	4:Flooded forest
	5:Mosaic forest/cropland
	6:Mosaic forest/savannah
3:Woodland/Shrubland	7:Closed deciduous woodland
	8:Closed deciduous shrubland
	9:Open deciduous shrubland
4:Cropland	10:Rainfed cropland
	11:Mosaic cropland/vegetation
	12:Irrigated or post flooding cropland
5:Grassland	13:Closed grassland
	14:Closed to open grassland
	15:Open grassland
	16:Sparse grassland
6:Barren areas	17:Bare land and Dunes
	18: Stony
	19:Rock
7: Inland water	20: Inland water
8: Urban and Built up	21: Urban and Built up

Tableau III-2 Légende hiérarchique des cartes LCM-I et LCM-II (8 hyper classes au niveau I et 21 classes au niveau II)

Après l'application du k-NN aux 214 profils du NDVI, une analyse interactive a été effectuée visuellement afin de regrouper certains clusters et ceci a conduit à l'obtention d'une carte avec 90 classes. Ces classes représentent des groupements homogènes du comportement de la saisonnalité et des différentes phases de l'activité photosynthétique qui sont : la maturité, la sénescence, la dormance et la croissance (voir figure 1.8). Les 90 classes résultantes ont été labellisées selon la nomenclature LCCS de la FAO (Di Gregorio et Jansen, 2000). On a finalement rajouté les deux écosystèmes représentant les surfaces artificielles et les zones d'eau, ce qui a ainsi abouti à 92 écosystèmes que nous appellerons dans la suite LCM-II au niveau III. Comme décrit dans le tableau 3.2, les écosystèmes de LCM-I ont été fusionnés aux niveaux supérieurs de la légende hiérarchique sur la base de leurs caractéristiques fonctionnelles. La carte d'occupation des sols consiste aussi en 8 hyper-classes au niveau I et en 21 classes au niveau II.

3.3.2 Classification hiérarchique versus classification par partition

L'Afrique s'étendant des 2 cotés de l'équateur, l'inversion du cycle saisonnier est observée. Tandis que la période de croissance d'une savane arborée dans le Sahel est juillet et août, le même comportement phénologique sera observé en novembre et décembre pour une savane arborée localisée sur l'île de Madagascar. Et on peut ainsi observer d'importantes variations du NDVI au sein d'une même classe d'occupation des sols et obtenir des résultats incertains en utilisant une approche de classification par pixel telle que celle du k-NN. Par conséquent, une étape de pré-classification en ensembles connexes (ou patchs) homogènes peut être utilisée comme canaux additionnels afin d'améliorer la classification (voir par exemple, Andrieu et Mering, 2008 ; Vintrou et al., 2009).

3.3.2.1 La transformée de Fourier Discrète (TFD)

La première stratification du continent en zones homogènes est basée sur la nature physique et structurelle des composantes du NDVI en utilisant la transformée de Fourier

discrète (TFD). Des études précédentes ont démontré l'utilité de la TFD pour l'étude des cycles phénologiques de la végétation (Andres et al., 1994 ; Azzali et Menenti, 1999 ; Immerzeel et al., 2005 ; Moody et Johnson, 2001 ; Scharlermann et al., 2008). D'après la théorie de Fourier, toute fonction *f* *p*-périodique peut s'écrire sous la forme d'une combinaison linéaire de fonctions trigonométriques élémentaires:

$$f(t) = \sum_{r=0}^{\infty} \tilde{f}_r e^{(-2\pi rtj/p)}, \qquad (3.2)$$

Les constantes $\tilde{f}_r$ sont déterminées à partir de la forme sinusoïdale de *f(t)*. Pour *r>0*, le terme de la série définie par l'équation (3.2), a une période égale à *p/r*. Pour construire une série de Fourier qui représentera les valeurs décadaires de NDVI, nous commençons par définir une nouvelle fonction, *NDVI*(t)*, ayant une périodicité annuelle :

$$NDVI^*(d + yD) = \frac{1}{Y}\sum_{y=1}^{Y} NDVI_d^y . \qquad (3.3)$$

Puis, à l'aide de l'équation (3.2) et de la transformation des fonctions sinusoïdales, la fonction théorique $NDVI^*(t)$ peut s'écrire sous la forme :

$$NDVI^*(d) = \sum_{r=0}^{D-1} f_{*,r} \cos(2\pi dr/(\Delta D) + \varphi_r), \qquad (3.4)$$

où Δ est l'intervalle d'échantillonnage en unité journalière (i.e. Δ=10), $f_{*,r}$ fait référence à l'amplitude et φ_r à la phase.

L'amplitude mesure la variabilité maximale des valeurs de NDVI tandis que la phase est le temps de réponse entre le NDVI maximum et le début de la 1ère décade.

3.3.2.2 L'indice de segmentation des composantes de Fourier (ISFC)

Pour synthétiser les réponses spectrales des composantes de Fourier, nous avons introduit un nouvel indice à savoir l'indice de Segmentation des Composantes de Fourier (Index of Segmentation of Fourier Component, ISFC), défini comme suit :

$$ISFC = \sum_{r=0}^{D-1} w_r \, f_{*,r} \,, \tag{3.5}$$

où w_r est un coefficient de pondération qui se définit comme étant la contribution relative de la $r^{ième}$ composante de Fourier à la puissance totale (Priestley, 1981), avec :

$$w_r = \frac{f_{*,r}^2}{\sum_{r=0}^{D-1} f_{*,r}^2} \,. \tag{3.6}$$

3.3.2.3 Application de l'ISFC

Cette nouvelle approche, à savoir la *stacked segment classification*, combine classification hiérarchique et regroupement dynamique. Une stratification du continent en écorégions bioclimatiques est effectuée en utilisant la TFD appliquée aux 8 ans (2000-2007) de données brutes décadaires de NDVI de SPOT/VGT. La segmentation des valeurs de l'ISFC en 5 intervalles numériques a été effectuée de telle sorte que l'on retrouve les principaux climats rencontrés sur le continent. A cet effet, plusieurs tests ont été effectués avant de retenir les intervalles numériques qui nous ont permis de définir 5 écorégions bioclimatiques (régions similaires d'un point de vue comportement spectral). Les écorégions ont été identifiées comme étant : aride (comprenant également les zones hyperarides pour des valeurs de l'ISFC inférieures à 0,02), semi-aride (de 0,02 à 0,11), subhumide sèche (de 0,11 à 0,37), subhumide humide (de 0,37 à 0,56), et humide (pour des valeurs de l'ISFC supérieures à 0,56). L'objectif de cette première stratification du continent en écorégions est double : (i) la réduction de la variabilité du NDVI pour améliorer la précision de classification et, (ii) l'amélioration des tendances des caractéristiques régionales surfaciques telles que la saisonnalité de la végétation, les conditions atmosphériques, la couverture nuageuse. Une telle approche quantifie l'hétérogénéité spatiale qui existe à divers niveaux d'échelle à l'intérieur du jeu de données considéré afin de minimiser l'hétérogénéité locale qui masque souvent le comportement des unités spatiales plus étendues. Ceci comprend la complexité,

la variabilité et la fragmentation spatiale qui peuvent influencer le caractère, l'amplitude et l'intérêt du fonctionnement des écosystèmes.

Ecoregion	Hyper classe
Aride et hyper-aride	1, 2, 3, 4, 5, 6
Semi-aride	2, 3, 4, 5
Sec	1, 2, 3, 4, 5
Sub-humide	1, 2, 3, 4, 5
Humide	1, 2

Tableau III-3 Sous-ensembles obtenus après l'application du classificateur par écorégions. Se référer au tableau 3.2 pour la légende des numéros des hyper classes.

L'application du classificateur k-NN effectuée sur chacune des 5 écorégions précédemment délimitées a conduit à 22 sous-ensembles (voir tableau 3.3). En analysant la classification obtenue, nous avons observé que les limites entre les hyper classes *forêts* et *forêts mixtes* étaient localement aberrantes et avons décidé de les fusionner en une seule hyper classe *forêt*. Cela a conduit à 18 sous-ensembles (les deux hyper classes étaient présentes dans les domaines aride et hyper-aride, sec et humide). En accord avec les hyper classes définies au tableau 3.2, chaque sous-ensemble a été ensuite labellisé. À ce niveau d'information sur les écosystèmes, chaque hyper-classe est une partition de sous-ensembles des écorégions bioclimatiques. Par exemple, les pixels de l'hyper classe *forêt* sont disséminés dans chacune des zones bioclimatiques (tableau 3.3). Enfin, un k-NN effectué séparément sur chacun des 18 sous-ensembles, afin de les affiner, a conduit à l'obtention de 192 clusters. Les types de végétation rencontrés dans une même écorégion sont fusionnés en 71 classes et facilement labellisés. Enfin, nous avons ajouté les deux écosystèmes « urbain » et « plans d'eau » pour obtenir la seconde carte des écosystèmes (appelée LCM-II dans la suite) à 73 classes au niveau III. Les différentes étapes de cette approche hybride de classification sont résumées dans la figure 3.2-b. Tout comme avec LCM-I, nous avons également construit LCM-II (ou classification ECOCLIMAP-II dans tout ce qui va suivre) aux mêmes niveaux d'information d'occupation du sol que ceux définis dans le tableau 3.2.

3.3.3 Evaluation de la classification

3.3.3.1 La matrice de contingence

L'évaluation d'une carte d'occupation des sols consiste à fournir aux utilisateurs le degré de correspondance entre l'information contenue dans la carte à la véritable valeur de la quantité mesurée (Congalton et Green, 2009). Pour ce faire, une stratégie consiste à élaborer la matrice de contingence résultant d'une analyse par pixel afin de comparer les champs classifiés en utilisant comme données de référence, l'information provenant d'une source qui est supposée juste. Les éléments dérivés de la matrice de contingence comprennent le coefficient kappa (**K**), les précisions selon l'utilisateur, le producteur et totale (voir Annexe B). La carte LCM-II a été validée par inter-comparaison sur une analyse pixel par pixel avec 3 cartes du continent : la carte à 1 km GLC2000 pour l'année 2000, la carte à 300 m GLOBCOVER pour l'année 2005 et la première couche de la carte à 500 m de la collection du produit MODIS (MODIS LC-I ci-après) pour l'année 2005. Les différentes combinaisons de 2 produits choisis parmi ces 4 classifications (les 3 classifications ci-dessus énumérées et ECOCLIMAP-II) ont ainsi conduit à l'élaboration de 6 matrices de contingence. Il est important de mentionner que : (i) tous ces produits sont basés sur des images satellites acquises durant les années 2000 (cf. tableau 3.1), (ii) MODIS LC-I, aussi connu sous le nom de MODIS IGBP, contient le plus grand nombre de classes d'occupation du sol et (iii) le choix de l'année 2005 est dicté par le fait que cette année est représentative des résultats de la période pluriannuelle 2000 à 2009 (Friedl et al., 2010).

3.3.3.2 Harmonisation des cartes d'occupation des sols

Bien que les cartes GLC2000, GLOBCOVER, MODIS LC-I et ECOCLIMAP-II ont pour but de décrire la distribution spatiale des types de végétation rencontrés sur le continent, des différences remarquables au niveau de leur légende sont observées (voir tableau 3.4). Afin de faciliter la comparaison directe entre ces cartes, les diverses classes ont été combinées en 7 catégories sur la base des standards de LCCS (Di Grégorio et Jansen, 2000) :

GLC2000		GLOBCOVER		MODIS LC-I		ECOCLIMAP-II at level II	
0	Sea	11	Post-flooding or irrigated croplands	2	Evergreen broadleaf forest	0	Sea
1	Closed Evergreen Forest	12	Post-flooding or irrigated shrubs or tree crops	4	Deciduous broadleaf Forest	1	Closed lowland forest
2	Degraded Evergreen forest	13	Post-flooding or irrigated herbaceous crops	5	Mixed forest	2	Closed to open lowland deciduous forest
3	Submontane forest(900-1500m)	14	Rainfed croplands	6	Closed shrubland	3	Montane Forest
4	Montane forest(>1500 m)	15	Rainfed herbaceous crops	7	Open shrubland	4	Flooded forest
5	Swamp forest	16	Rainfed shrub or tree crops	8	Woody savanna	5	Mosaic forest/cropland
6	Mangrove	20	Mosaic cropland / vegetation (grassland/shrubland/forest)	9	Savanna	6	Mosaic forest/savanna
7	Mosaic forest/croplands	21	Mosaic cropland / grassland or shrubland	10	Grasslands	7	Closed deciduous woodland
8	Mosaic forest/savanna	30	Mosaic vegetation (grassland/shrubland/forest) / cropland	11	Permanent wetland	8	Closed deciduous shrubland
9	Closed deciduous woodland	31	Mosaic grassland or shrubland / cropland	12	Cropland	9	Open deciduous shrubland
10	Open deciduous woodland	32	Mosaic forest / cropland	13	Urban and built up	10	Rainfed cropland
11	Deciduous shrubland with sparse trees	40	Closed to open broadleaved evergreen or semi-deciduous forest	14	Mosaic Cropland/natural vegetation	11	Mosaic cropland/vegetation
12	Open deciduous shrubland	41	Closed broadleaved evergreen and/or semi-deciduous forest	16	Barren or sparsely vegetated	12	Irrigated or post flooding cropland
13	Closed grassland	42	Open broadleaved semi-deciduous and/or evergreen forest with emergents	17	Water	13	Closed grassland
14	Open grassland with sparse shrubs	50	Closed broadleaved deciduous forest			14	Closed to open grassland
15	Open grassland	60	Open broadleaved deciduous forest/woodland			15	Open grassland
16	Sparse grassland	70	Open broadleaved deciduous forest/woodland			16	Sparse grassland
17	Swamp bushland and grassland	90	Open needleleaved deciduous or evergreen forest			17	Bare land and Dunes
18	Croplands(> 50%)	100	Closed to open mixed broadleaved and needleleaved forest			18	Stony
19	Cropland with open woody vegetation	101	Closed mixed broadleaved and needleleaved forest			19	Rock
20	Irrigated croplands	110	Mosaic forest or shrubland / grassland			20	Inland water
21	Tree Crops	120	Mosaic grassland / forest or shrubland			21	Urban and Built up
22	Sandy desert and dunes	130	Closed to open shrubland				
23	Stony desert	131	Closed to open broadleaved or needle-leaved evergreen shrubland				
24	Bare rock	134	Closed to open broadleaved deciduous shrubland				
25	Salt hardpans	140	Closed to open herbaceous vegetation				
26	Water bodies	141	Closed grassland				
27	Cities	143	Closed to open grassland				
		150	Sparse vegetation				
		151	Sparse grassland				
		160	Closed to open broadleaved forest regularly flooded				
		161	Closed to open broadleaved forest on (semi-)permanently flooded land				
		162	Closed to open broadleaved forest on temporarily flooded land				
		170	Closed broadleaved forest or shrubland permanently flooded				
		180	Closed to open grassland or woody vegetation on regularly flooded soil				
		181	Closed to open woody vegetation on regularly flooded or waterlogged soil				
		185	Closed to open grassland on regularly flooded or waterlogged soil				
		190	Artificial surfaces and associated areas				
		200	Bare areas				
		201	Consolidated bare areas (hardpans, gravels, bare rock, stones, boulders)				
		202	Non-consolidated bare areas (sandy desert)				
		203	Salt hardpans				
		210	Water bodies				

Tableau III-4 Légende des cartes d'occupation des sols du tableau 3.1 produites à partir des images satellites acquises durant les années 2000

(i) la catégorie forêt/forêt mixte qui contient les forêts sempervirentes ou décidues qu'elles soient denses ou pas, les forêts inondées, la mosaïque forêt/savane et la mosaïque forêt/culture (ii) la catégorie savane arborée/arbustive qui comprends les sols recouverts d'une savane arborée ou d'une savane arbustive, (iii) la catégorie savane herbacée constituée des différents types d'herbacés, (iv) la catégorie culture qui représente les cultures irriguées ou marécageuses et les cultures sous pluie, (v) la catégorie du bâti qui correspond aux surfaces artificielles, (vi) la catégorie du sol nu qui comprend les sols nus ou couverts de végétation clairsemée ou recouverts par le sable, la roche et les dunes, et (vii) la catégorie des plans d'eau qui comprend les eaux naturelles et artificielles.

La conversion des classes d'occupation du sol en l'une de ces catégories (voir tableau 3.5) a été effectuée sur la base des définitions des légendes de chaque classe initiale du tableau 3.4. Il est important de rappeler ici que la définition de la végétation arborée dans les 4 produits n'est pas la même. Par exemple, la classe 'savanna' de MODIS, correspondant aux couverts herbacés ayant une canopée forestière, est contenue dans la catégorie 'savane herbacée' tandis que sa classe correspondante du produit GLOBCOVER, intitulée 'closed to open broadleaf deciduous shrubland' est contenue dans la catégorie 'savane arbustive/savane arborée'. Tout en reconnaissant que le regroupement de nombreuses classes d'occupation du sol en quelques catégories peut conduire à une mauvaise représentation du jeu de données initiales à cause du problème de transférabilité des classes d'une légende à une autre (Latifovic et Olthof, 2004), on considère que ce compromis arbitraire préserve la répartition géographique du paysage africain.

	Forest/Mixed forest (1)	Woodland/Shrub land (2)	Cropland (3)	Grassland (4)	Bare land (5)	Inland water (6)	Urban and built-up (7)
GLC2000	1-8	9-12	18-21	13-17	22-24	25-26	27
GLOBCOVER	32, 40-42, 50, 60, 70, 90, 100, 101, 110, 160, 161, 162, 170	130, 131, 134, 181	11-16, 20, 21	30-31, 120, 140-141, 143, 150-151, 180, 185	200-202	203, 210	190
MODIS LC-I	2, 4, 5	6-9	12, 14	10	16	0	13
ECOCLIMAP-II	1-6	7-9	10-12	13-16	17-19	20	21

Tableau III-5 Classes agrégées. (Cf. le tableau 3.4 pour la définition du numéro de chaque classe)

3.3.3.3 Comparaison spatiale

On a choisi la résolution spatiale du 112ème de degré et le système géodésique WGS 84 de la projection plate carrée comme grille spatiale de référence pour l'intercomparaison car ECOCLIMAP-II et GLC2000 étaient déjà disponibles dans ce système spatial (cf. tableau 3.1). On a ensuite projeté les produits MODIS LC-I et GLOBCOVER sur cette grille de la façon suivante :

(i) rééchantillonage au 1/120° en utilisant la règle de la classe majoritaire basée au sein du voisinage considéré. Le centre du pixel au 120^{e} de degré est le centre d'un voisinage de 2x2 et 3x3 pour MODIS LC-I et GLOBCOVER respectivement (figure 3.4). En cas de non-unicité d'une classe majoritaire, on va augmenter la taille du voisinage en se limitant à l'intérieur d'une boîte de 5x5 km² comme le montre la figure 3.5. Cet accroissement est justifié par le fait que le capteur ne repasse pas exactement sur sa précédente trace après un cycle orbital complet,

(ii) projection des pixels au 1/120^{e} sur la grille de référence (i.e. au 112^{e}) en utilisant la règle du PPV qui préserve le mieux les valeurs originales des cartes au 1/120^{e} de degré.

Afin de percevoir la manière dont les classes sont décrites à un endroit donné, on a généré des cartes d'accord en comptant le nombre d'occurrences pour lequel les 4 produits de classification GLC2000, GLOBCOBER, MODIS LC-I et ECOCLIMAP-II s'accordent. L'information produite par toutes ces cartes a été synthétisée de la façon suivante :

- Pas d'accord pour les pixels identifiés comme tels par une seule classification,
- Faible accord pour les pixels identifiés comme tels par seulement deux classifications,
- Accord moyen pour des pixels ayant été identifiés comme tels par les deux premières classifications tandis que les deux autres sont également en accord mais pour une autre catégorie,
- Fort accord pour des pixels identifiés comme tels par 3 des 4 classifications,
- Accord total pour un consensus entre les 4 produits

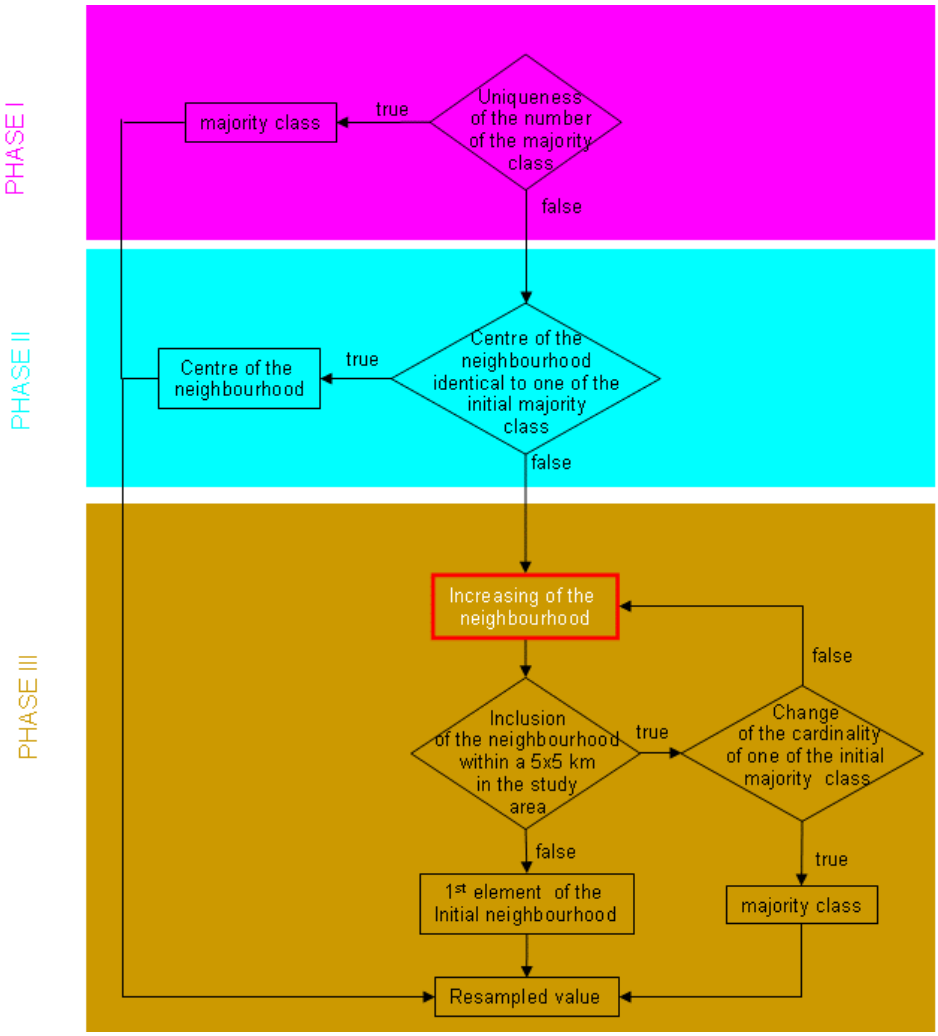

Figure III.4 Diagramme du rééchantillonage par classe majoritaire mis sur pied. Noter que la phase II n'a pas été utilisée lors du rééchantillonage de MODIS LC-I

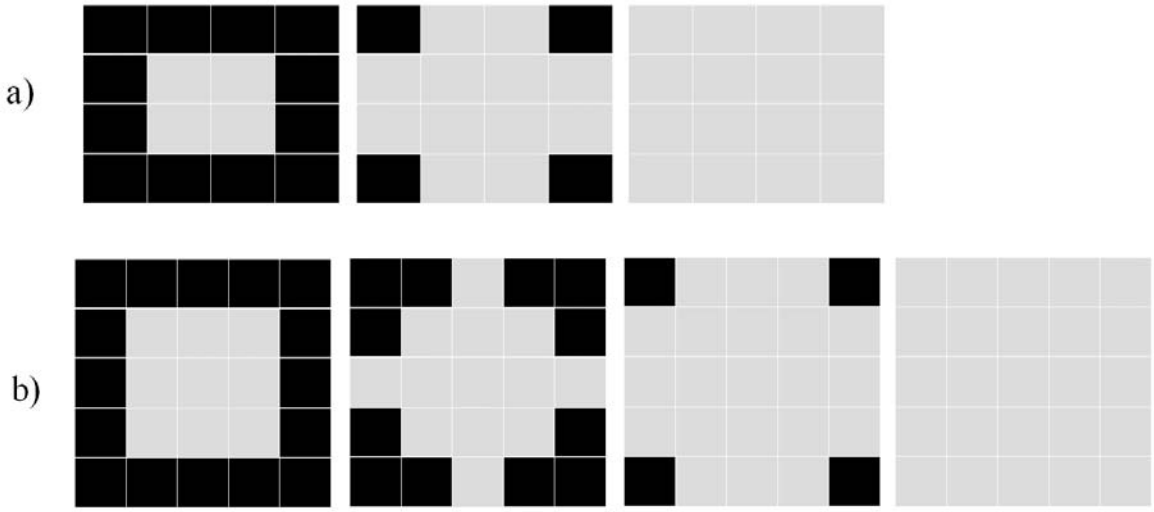

Figure III.5 Accroissement du voisinage lors du ré-échantillonnage par classe majoritaire (figure 3.4) par mutation des cellules grisées en cellules noires pour MODIS LC-I (en haut) et GLOBCOVER (en bas)

3.4 Validation des produits

3.4.1 Ecorégions bioclimatiques

La figure 3.6-a montre les valeurs de l'ISFC obtenues par transformation des composantes de Fourier (voir section 3.3.2.2). Il y a un zonage latitudinal assez marqué avec les valeurs les plus basses comprises entre 20° et 30° N et les valeurs les plus élevées situées dans la région équatoriale entre 10° N et 10° S. Au nord de 10° S, les valeurs de l'ISFC montrent une baisse générale direction ouest-est, tandis qu'au sud de 10° S, la situation est inversée et donc les valeurs de l'ISFC tendent à se réduire en direction est-ouest (figure 3.6-b). Des valeurs élevées de l'ISFC sont obtenues dans les zones où l'on note une forte variabilité du NDVI. C'est ainsi que l'on a facilement pu identifier la zone bioclimatique humide et les limites entre la zone sahélienne de la zone saharienne. La variabilité bioclimatique régionale dépend de la localisation géographique à proximité de l'équateur et du littoral. La principale raison de la délimitation des écorégions est de définir de façon naturelle plutôt qu'administrativement ou politiquement les limites de régions pour résoudre les problèmes environnementaux (Bonan et al., 2002). En effet, le découpage politique (cités, pays) pour lequel des décisions socio-économiques sont souvent prises ne correspond pas à des unités cartographiques de l'environnement naturel telles que le type de sol et de climat.

Une comparaison visuelle avec la carte globale des zones agro-écologiques (figure 3.6-c) produite par Fischer et al. (2002) donne une appréciation du contenu thématique des écorégions bioclimatiques. Cette comparaison entre la figure 3.6-c et la carte pluviométrique du continent Africain (Goudie, 1996) montre que la distribution pluviométrique est un facteur clé de la longueur de la période de croissance de la végétation (Length of Growing Period, LGP) et c'est ce qui justifie les différences importantes entre les écorégions délimitées. La LGP est inférieure à 60 jours pour les sols hyperarides et arides tandis qu'elle est comprise entre 60 et 165 jours au niveau des sols semi-arides et subhumides secs. Dans les zones humides et subhumides froides, la LGP est supérieure à 165 jours. Contrairement aux

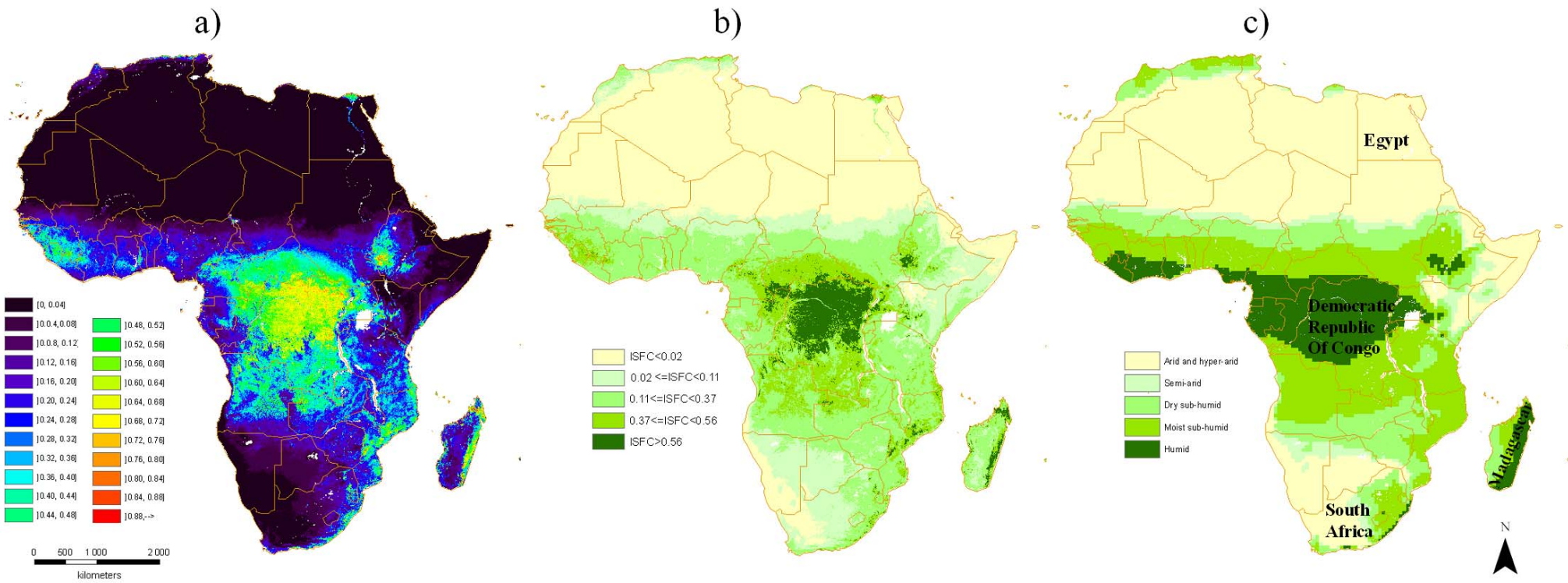

Figure III.6 : a) l'Index de Segmentation des Composantes de Fourier calculé à partir de l'équation (3.6) ; b) les écorégions bioclimatiques ; c) les majeures zones agro-écologiques zones basées sur les données de l'IIASA et de la FAO (Fischer et al., 2002).

cartes agro-écologiques existantes, cette nouvelle carte reflète la distribution spatiale de la végétation d'un point de vue bioclimatique et écologique. Cependant on observe une lacune de l'information fournie par l'ISFC sur les zones pour lesquelles le pourcentage des données *acceptables* est inférieur à 50% à cause de l'ennuagement permanent (cf. figure 3.6-a), ce qui justifie les différences observées sur les zones humides des deux cartes des écorégions.

3.4.2 Unités d'écosystèmes

Dans certaines régions en Afrique australe et orientale, on trouve des forêts tropicales sèches tandis qu'ailleurs dans certaines zones en Afrique occidentale et centrale, on rencontre plus souvent des forêts tropicales humides. Chacune de ces formations forestières est caractérisée par des conditions biotiques telles que la structure de la végétation (la hauteur du peuplement, la stratification verticale) et une composition floristique spécifique (Kelatwang et Garzuglia, 2006). Par exemple, dans les forêts denses sempervirentes, la strate supérieure possède une densité de 70% pour une hauteur comprise entre 35 et 45 m, ce qui empêche le développement de la strate arbustive et herbacée. Une transition du couvert forestier dans les domaines arides et semi-arides est marquée par une distribution éparse de savane arborée et arbustive avec une possible association de savane herbacée. Une telle strate herbacée caractérise les régions qui sont saisonnièrement arrosées. Elle renferme principalement les savanes et les jachères qui peuvent toutefois être mélangées aux cultures.

Les figures 3.7 et 3.8 illustrent la localisation et l'étendue des types d'écosystèmes en Afrique pour les deux cartes LCM-I et LCM-II produites selon la méthodologie illustrée sur la figure 3.2. Le niveau III est le niveau d'information des écosystèmes le plus détaillé (cf. légendes figures 3.7 et 3.8) par rapport aux niveaux I et II du tableau 3.3. Ce niveau d'information permet d'estimer la diversité des écosystèmes au sein d'une même classe. La classe *forêt inondée* comprend, au niveau III, 2 types d'écosystèmes de LCM-I et 5 types d'écosystèmes de LCM-II.

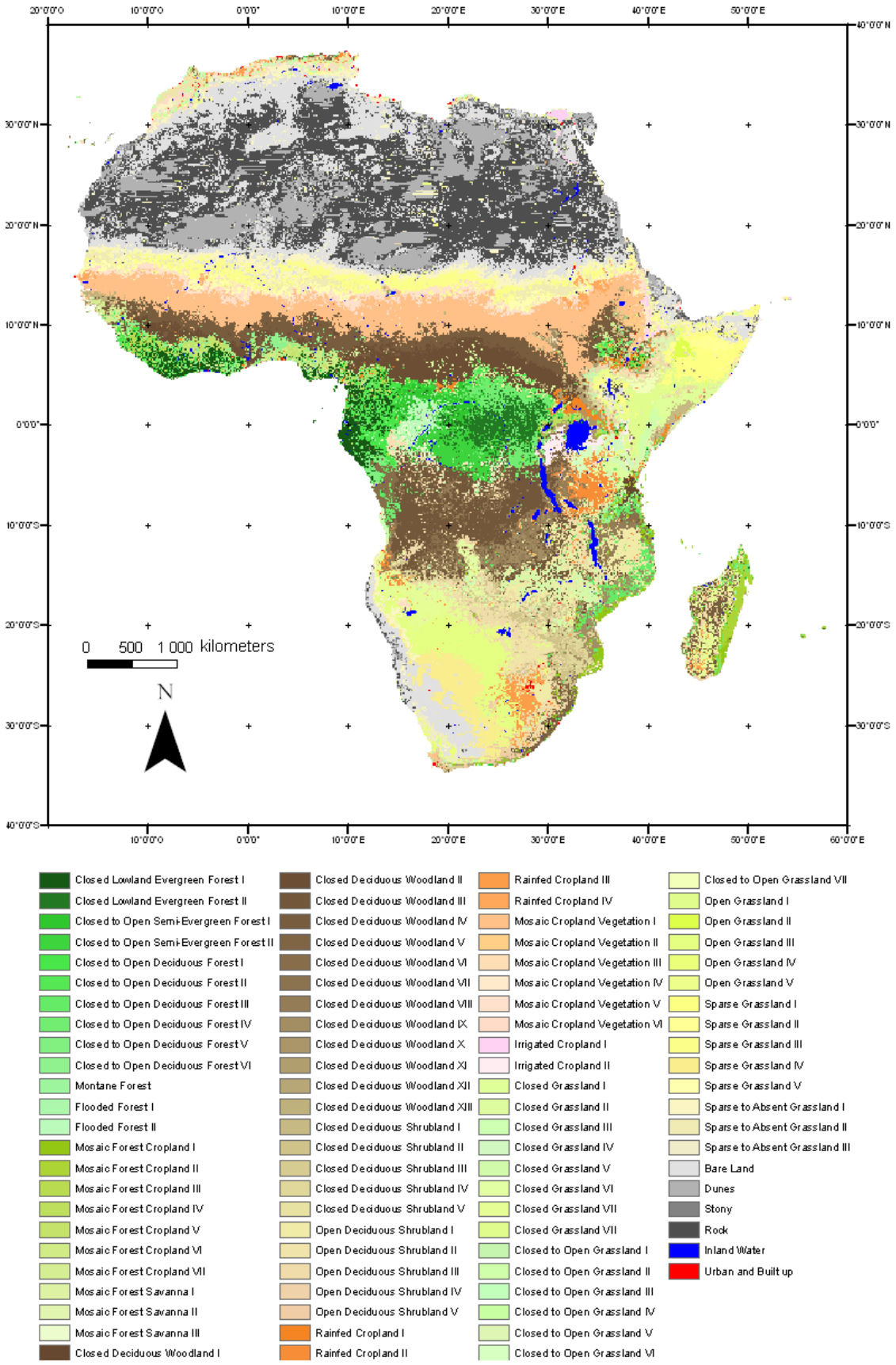

Figure III.7 LCM-I au niveau III

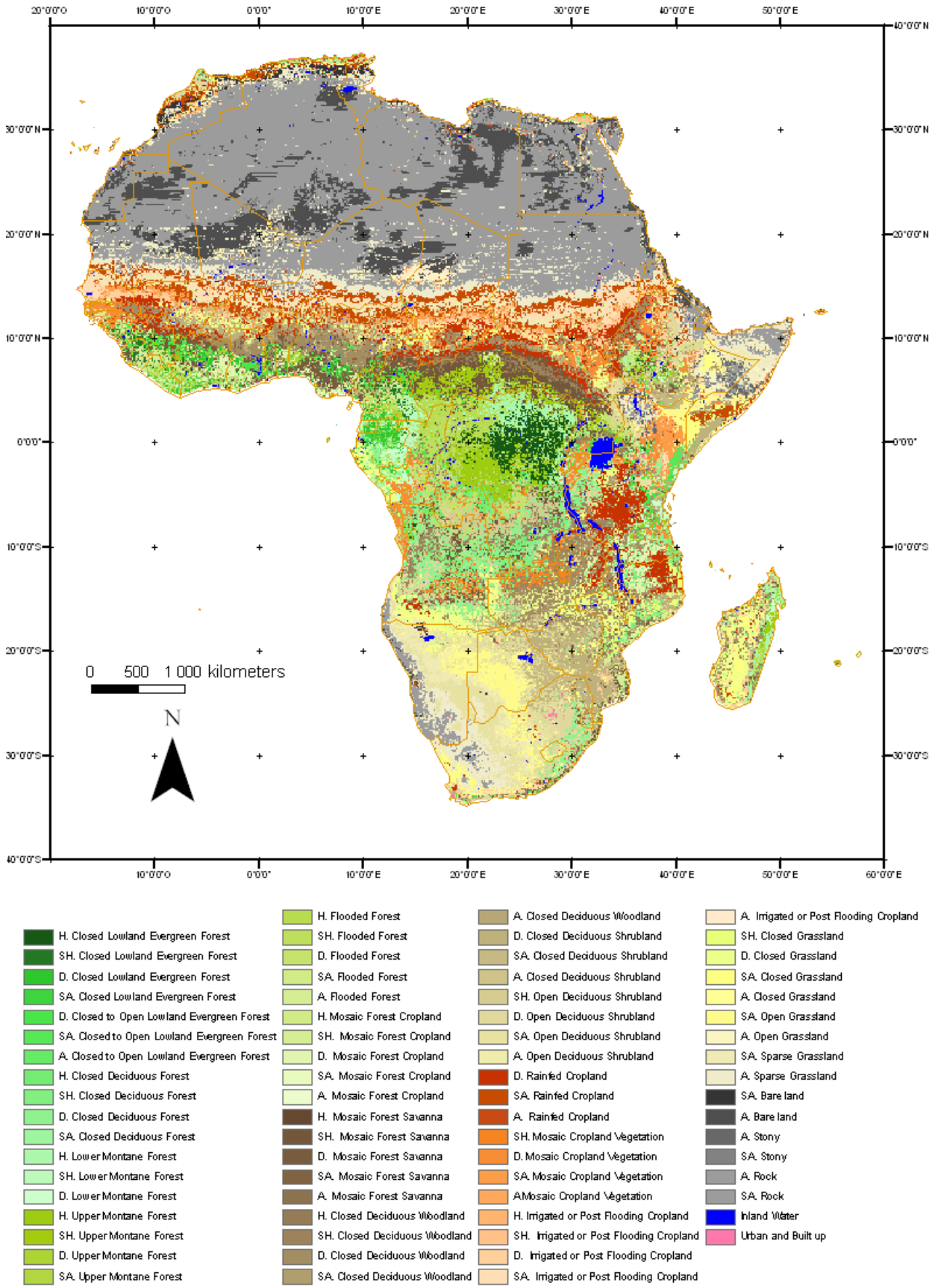

Figure III.8 LCM-II au niveau III.

Les *cultures irriguées* se trouvent principalement le long de la vallée du Nil. On note un type de savane herbacée ouverte (appelée type III pour LCM-I et semi-aride pour LCM-II) presque exclusivement situé en Afrique du Sud. La *forêt de feuillus décidus* est principalement localisée sur certains versants humides et les ravins du Chaillu ou Mayombe au nord du Congo. On note quatre faciès de la *forêt dense sempervirente* dans LCM-I : les faciès central et littoral sont sempervirents tandis que les faciès du nord et du sud-ouest sont semi-sempervirents. Tous ces faciès correspondent à une formation qui s'étend de la côte Atlantique à la savane soudano-zambézienne située au sud-ouest de la République Centrafricaine ou au sud-est du Cameroun. Ceci a permis l'isolement de la forêt du Gabon, contenue dans le faciès littoral, de la forêt transfrontalière du Mayombe entre l'Angola, la République Démocratique du Congo (RDC) et la République du Congo, et des autres forêts internes en RDC. Le paysage Batéké est entouré par une forêt dense de basse altitude et par une savane herbacée le long du fleuve Congo.

Au niveau II, de nombreuses classes ont des représentations similaires dans les deux cartes (cf. tableau 3.6), notamment la *forêt dense sempervirente de basse altitude*, la *mosaïque forêt/culture* et la *savane herbacée dense*. Toujours à ce niveau, les classes les plus représentatives sont la *savane dense arborée décidue* pour la partie végétalisée et la roche pour le sol nu. Certaines hyper classes présentent des différences assez nettes au niveau de leur étendue dans les deux cartes. L'une d'elles est l'hyper classe *culture* qui comprend des terres de différentes intensités agricoles et de différentes tailles. Ceci explique sa fragmentation et le fait qu'elle soit potentiellement intégrée à d'autres types d'écosystèmes. En effet, à cause du faible niveau d'intensification des pratiques agricoles (Hannerz et Lotsh, 2008), les champs sont petits et mélangés à d'autres types d'écosystèmes et ceci explique les inconsistances entre plusieurs inventaires statistiques agricoles. Une autre est l'hyper classe *sol nu* pour laquelle des proportions différentes des classes d'occupation des sols ont été obtenues au niveau II avec les deux méthodes : la *stacked segment classification* et le k-NN (comme le montrent les figures 3.6 et 3.7).

	LCM-I			LCM-II		
	N.P.	%	N.E.	N.P.	%	N.E.
1: Closed lowland evergreen forest	1 338 208	4.17	4	1 130 630	3.52	6
2: Closed to open lowland deciduous forest	1 289 343	4.02	6	1 044 183	3.25	5
3: Montane Forest	43 857	0.14	1	1 463 861	4.56	7
4: Flooded forest	184 618	0.58	2	920 784	2.87	5
5: Mosaic forest/cropland	775 575	2.41	7	648 559	2.02	5
6: Mosaic forest/savannah	231 131	0.72	3	1 201 946	3.75	5
7: Closed deciduous woodland	5 108 520	15.92	13	2 031 766	6.33	5
8: Closed deciduous shrubland	641 389	2.00	5	1 450 098	4.52	3
9: Open deciduous shrubland	1 855 656	5.78	5	1 958 790	6.11	4
10: Rainfed cropland	682 250	2.13	4	1 989 651	6.20	3
11: Mosaic cropland/vegetation	2 900 627	9.04	6	1 282 751	4.00	4
12: Irrigated or post-flooding cropland	168 883	0.53	2	1 102 749	3.44	5
13: Closed grassland	964 699	3.01	6	1 187 776	3.70	2
14: Closed to open grassland	435 910	1.36	9	454 801	1.42	2
15: Open grassland	1 087 443	3.39	6	691 510	2.16	2
16: Sparse grassland	3 175 249	9.90	7	2 439 137	7.60	2
17: Bare land and Dunes	5 811 899	18.11	2	2 110 326	6.58	2
18: Stony	208 606	0.65	1	251 412	0.78	2
19: Rock	4 831 795	15.05	1	8 374 928	26.10	2
20: Inland water	316 108	0.99	1	316 108	0.99	1
21: Urban	32 511	0.1	1	32 511	0.1	1
Total		100	92	32 084 277	100	73

Tableau III-6 Nombre des pixels (N.P) et des écosystèmes (N.E) au niveau II

Le tableau 3.6 montre la surface occupée par chaque écosystème et leur pourcentage sur la surface continentale des deux cartes. Ces statistiques montrent qu'environ 33% de la surface continentale présente très peu ou presque pas de végétation, la savane herbacée représente 15%, la savane arbustive ou arborée constitue 16% et 10% des terres sont cultivées. Enfin, un peu plus de 20% de la surface est couvert de forêts (mixtes ou non) dans les deux cartes. Les forêts du bassin du Congo représentent environ 50% de la superficie

forestière totale du continent et ces forêts constituent la deuxième plus grande zone de forêt dense humide tropicale du monde après la forêt amazonienne.

3.4.3 LCM-I versus LCM-II

Dans LCM-I, certaines classes peuvent être constituées d'une seule unité végétalisée à partir de pixels contigus et disséminés ; une autre classe peut représenter des unités distinctes de végétation potentiellement différentes au plan structurel alors que des types de végétation adjacents ou voisins peuvent se distinguer au sein de cette même classe. La 3^{e} catégorie de la classe *mosaïque culture végétation* de LCM-I (qui agrège la savane au Congo, les cultures au Kenya et la savane en Ethiopie) et la 4e catégorie de la classe *Forêt semi-dense décidue* de LCM-I (qui agrège les forêts tropicales semi-sempervirentes dans l'est de la RDC, au nord de l'Angola et quelques forêts de montagne en Ethiopie et Madagascar) en sont quelques exemples. En outre, des forêts contiguës peuvent inclure différentes classes. Par exemple la 4^{e} catégorie de la classe *Forêt semi-dense décidue* de LCM-I se distingue aisément du faciès central de la *forêt dense sempervirente*. Ceci souligne l'attention particulière que génère la difficulté de labelliser chaque classe tout en respectant la subtile différence entre les classes aussi bien d'un point de vue régional que d'un point de vue continental.

LCM-II (fournit un comportement fin des forêts tropicales en Afrique Centrale. Le comportement général coïncide bien avec les larges différences phytogéographiques connues. Les forêts sempervirentes continentales et littorales sont bien distinguées, notamment la forêt inondée dans le bassin du Congo. Les forêts restantes correspondent aux forêts semi-sempervirentes. Entre elles, il existe un gradient pluviométrique saisonnier nord-sud et un autre d'ouest en est. Les régions pour lesquelles apparaît une intense conversion de la forêt en agriculture apparaissent comme des mosaïques de forêts mixtes au sud du Cameroun et à l'uest de La RDC. En RDC, dans la partie Est où la conversion des forêts atteint des proportions remarquables, les forêts ne se distinguent plus des aires

fragiles. Dans les savanes localisées au nord et au sud de la foret tropicale en Afrique Centrale, la succession des classes des zones humides à arides correspond de façon générale à celles identifiées par White (1983), avec un gradient structurel associé au gradient climatique. Il est important de noter que la transition forêt-savane se retrouve très bien, avec des pixels écotones individualisés dans une classe particulière.

3.4.4 Intercomparaison des produits d'occupation des sols

3.4.4.1 Les surfaces forestières

La figure 3.9-a montre qu'il existe un consensus au niveau de la présence des aires forestières dans le bassin du Congo, la chaîne montagneuse de l'île de Madagascar et le long du littoral ouest-africain. Par contre, au Soudan, en Angola, Zambie, Tanzanie, et sur le littoral Mozambicain, il n'existe aucun consensus au niveau de la localisation des forêts. Il y apparaît plutôt un fort accord pour l'hyper classe savane arborée/savane arbustive. Ce désaccord entre les cartes d'occupation des sols peut ne pas être dû à des erreurs de classifications mais plutôt à cause du changement d'occupation des sols entre les 6 années de l'acquisition des données GLC2000 et les autres données (tableau 3.1). En effet, l'Afrique perd plus de 4 millions d'hectares par an (Kelatwang et Garzuglia, 2006). Comme la Zambie, la Tanzanie et le Mozambique font partie des 10 pays ayant la plus forte couverture forestière en Afrique (FAO, 2006), ceci peut illustrer la forte décroissance des aires forestières en Afrique orientale. Comme mentionné à la section 1.3.5, les principales causes de la déforestation comprennent la sécheresse, les guerres civiles, les feux de brousse et les systèmes agricoles inappropriés tels que l'agriculture itinérante sur brûlis (voir section 1.3.5).

3.4.4.2 Les surfaces cultivées

La figure 3.9-d montre la comparaison des surfaces cultivées présentes sur le continent entre les différents produits. Une ceinture latitudinale d'un accord très fort localisé dans la

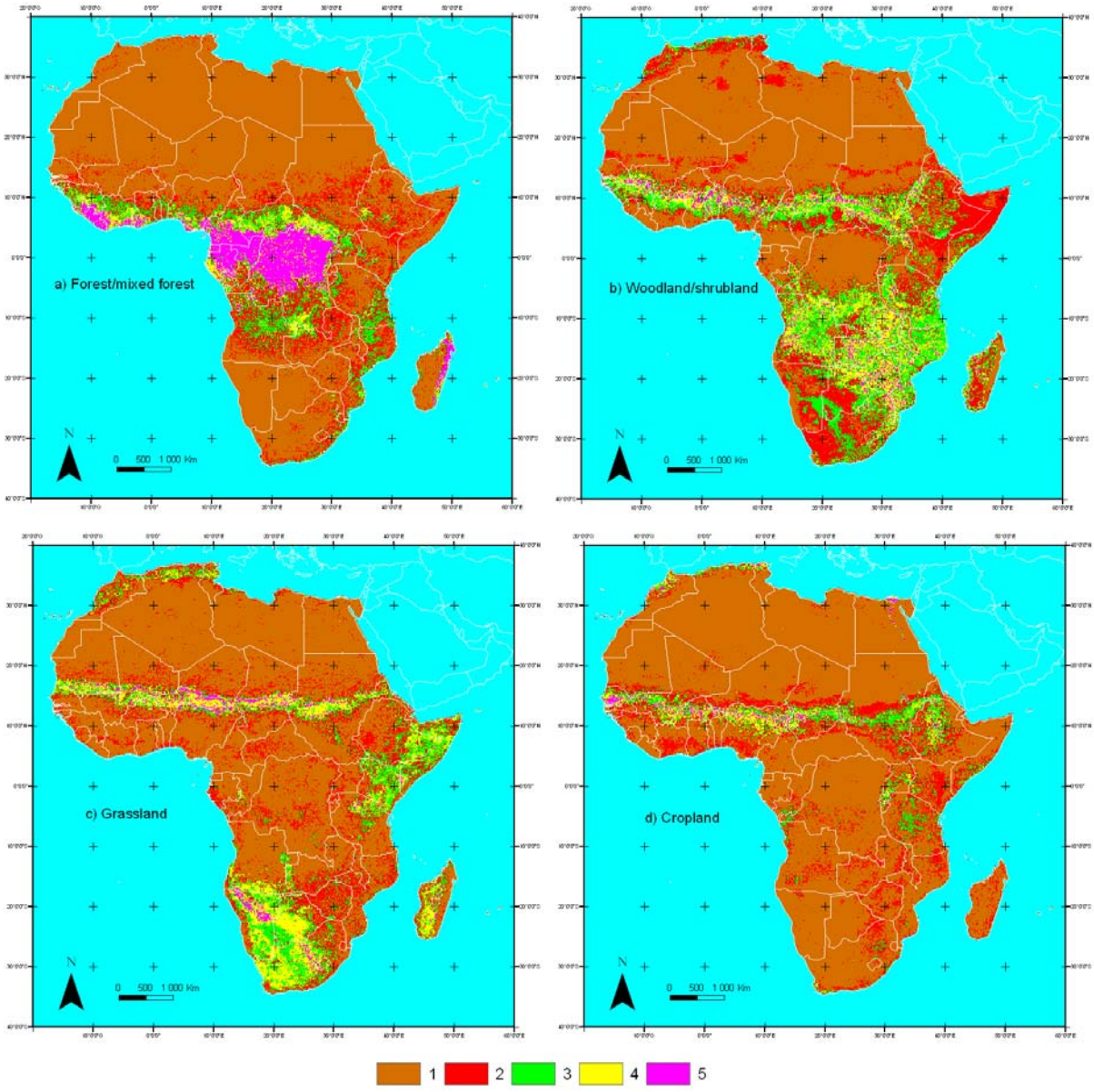

Figure III.9 Comparaison spatiale entre GLC2000, GLOBCOVER, MODIS LC-I et ECOCLIMAP-II pour les classes agrégées : a) Forêts et forêts mixtes, b) Savanes arborées et savanes arbustives, c) Savanes herbacées et d) Cultures. Le nombre de la légende indique le nombre de cartes d'occupation des sols ayant identifié la classe agrégée concernée à une position donnée. Par exemple, les régions en marron qui ont été identifiées comme la classe agrégée en question par un seul des quatre produits d'occupation des sols.

zone de transition entre les domaines soudanien et désertiques -qui s'étend du Sénégal à l'Ouest à l'Ethiopie à l'Est- est facilement distinguable. En outre, des aires d'un accord faible vers un accord modéré (i.e ; les aires de faible intensité agricole) apparaissent de façon récurrente le long de cette ceinture. Une autre ceinture de culture d'un accord faible vers un

accord modéré s'étend du sud (portion orientale de l'Afrique du Sud) à l'est (Ethiopie). Cette ceinture comprend de fortes concentrations de cultures dans la région du Cap au sud-ouest de l'Afrique et dans des régions avoisinant le nord du rivage du Lac Victoria. D'autres fortes concentrations de cultures sont visibles le long du Nil et de la côte méditerranéenne

3.4.4.3 (Dés)Accord spatial

La figure 3.10 montre qu'un schéma différent lorsqu'on synthétise les différentes comparaisons spatiales par hyperclasse et ce pour toutes les hyperclasses. Les régions qui montrent un consensus entre les 4 produits de classifications sont les régions arides ou de végétation éparse –comprenant de larges déserts tels que le Sahara, le Namib, et le Kalahari (Kgalagadi)- et également les forêts équatoriales. Le désaccord entre les classifications apparaît principalement entre les zones de transition comme le montre la figure 3.10. Les aires de consensus sont adjacentes aux aires de fort accord et les aires d'un accord modéré sont adjacentes aux aires de faible accord. Seulement 29.7% du continent possèdent un accord fort (accord de 3 sur 4) tandis que 40.9% des pixels montrent un consensus entre les 4 cartes. Il est important de mentionner que 85.6% des pixels pour lesquels on a un consensus sont localisés dans des zones arides ou des zones forestières. Quoique 1.4% du continent ne possède aucun accord, 10.4% possède un accord moyen (i.e. 2 cartes sont en accord pour une classe tandis que les 2 autres sont en accord pour une autre classe).

Le tableau 3.7 résume l'évaluation entre les combinaisons par 2 des 4 cartes du tableau 3.1. Les combinaisons MODIS LC-I/GLOBCOVER et MODIS LC-I/GLC2000 sont respectivement les plus proches (OA=56.3%, κ=0.45) et les plus éloignées (OA=68.8%, κ=0.56). Ceci est en accord avec les résultats de la figure 3.9. Le tableau 3.7 montre également que la classe agrégée "agriculture" possède la plus faible valeur de l'exactitude selon le producteur (16.2%) avec la combinaison ECOCLIMAP-II/MODIS LC-I tandis que la plus faible valeur de l'exactitude selon l'utilisateur (22.6%) est atteinte avec la combinaison

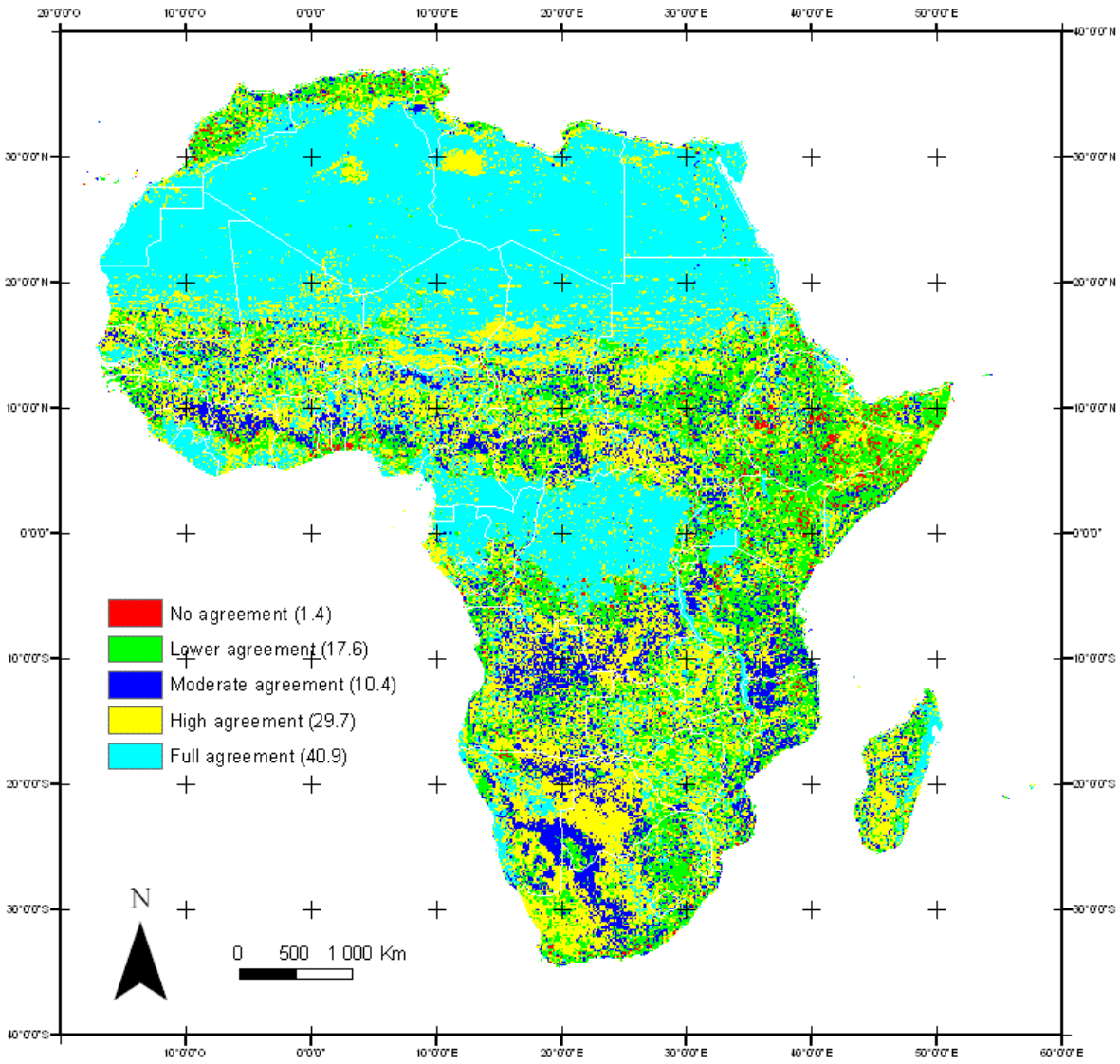

Figure III.10 Accord entre les classifications GLC2000, GLOBCOVER, MODIS LC-I et ECOCLIMAP-II. La valeur entre parenthèses indique le pourcentage de couverture de la classe correspondante.

MODIS LC-I/GLOBCOVER pour la classe agrégée savane herbacée. Il est utile de préciser que la parfaite correspondance entre ECOCLIMAP-II et GLC2000 pour la catégorie 'plans d'eau' et 'surfaces artificielles' découle du procédé de classification d'ECOCLIMAP-II.

3.4.5 Intercomparaison des variables biophysiques

Les variables biophysiques sont estimées selon la méthodologie décrite à la section 2.4.

Predicted map	ECOCLIMAP-II						MODIS LC-I				GLOBCOVER	
Reference map	MODIS LC-I		GLOBCOVER		GLC2000		GLOBCOVER		GLC2000		GLC2000	
Parameters	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA
Forest/Mixed Forest	**88.08**	40.37	55.65	71.08	89.69	53.39	*31.56*	88.00	66.08	85.75	81.02	37.16
Woodland/Shrubland	31.24	76.75	39.42	21.94	40.00	54.61	**82.61**	*18.73*	86.20	47.90	27.19	66.69
Cropland	32.71	*16.15*	43.61	16.42	43.95	35.34	51.41	39.24	32.62	53.16	31.30	**66.84**
Grassland	33.51	*19.83*	34.76	58.00	45.69	60.11	22.63	63.88	27.77	61.78	**69.09**	54.47
Bare land	91.68	*89.97*	89.71	91.07	91.48	91.12	92.43	95.62	93.73	**95.14**	94.48	92.71
Inland water	94.29	*65.14*	77.13	78.56	**100**	**100**	69.21	98.56	*65.14*	94.29	78.56	77.13
Urban nad built up	37.18	47.11	74.20	79.66	**100**	**100**	44.51	37.99	47.11	37.18	79.66	74.20
	OA=57.03 *κ =0.45*		OA=60.12 κ =0.49		OA=65.27 κ =0.56		*OA=56.25* *κ =0.45*		**OA=68.76 κ =0.59**		OA=65.23 κ =0.56	

Tableau III-7 Comparaison entre les cartes d'occupation des sols en prenant comme référence les cartes de la 2nde ligne. Les valeurs minimales sont en italiques et soulignées et les valeurs maximales sont en gras

3.4.5.1 Albédo du sol nu

Il n'a pas été possible de déterminer l'abédo du sol nu dans le visible pour 1.66% des surfaces continentales (figure 3.11). Pour certains de ces pixels, le nombre seuil de données disponibles du LAI n'était pas atteint à cause d'une valeur erronée du biome en entrée de la chaîne de traitement de MODIS. D'autres pixels concentrés dans certaines régions du Nord et du Sud de l'Afrique, illustrent l'inconhérence entre les données MODIS de LAI et d'albédo. Les cartes d'albédo du sol nu possèdent une variabilité spatiale prononcée dans les spectres du visible et du proche infrarouge comme le montre clairement la figure 3.12. Les valeurs élevées d'albédo du sol nu se rencontrent dans les régions arides –en particulier les déserts du Sahara, du Kalahari et du Karoo.

Nous avons par la suite évalué les nouvelles cartes d'albédo du sol nu en les comparant avec d'autres produits d'albédo du sol nu notamment ECOCLIMAP-I (Masson et al., 2003), ISLSCP-II (Hall et al., 2006), JULES (Houldcroft et al., 2009). A cette fin, les grilles à 1° et 0.05° respectivement d'ISLSCP-II et Jules ont été reprojetées au 1/120° et tous les pixels considérés comme données manquantes dans les 4 bases de données ont été écartées.

D'après le tableau 3.8 qui résume les statistiques ainsi obtenues par type dominant de sol (figure 1.3) ainsi que leurs composants tels que les proportions d'argile et de sable (figure 1.4) fournies par la FAO, les sols les plus réflectifs sont les champs de dunes suivis par les calcisols, les débris de roche et les gypsisols tandis que les sols les moins réflectifs sont les alisols, les ferralsols et les greyzems. Plusieurs sols ayant de fortes valeurs d'albédo possèdent aussi en général un coefficient de variation élevé. ECOCLIMAP-II est plus proche de JULES que des autres produits. Nous pensons que cela doit être dû au fait que les deux produits soient des dérivés des données MODIS. Ceci est également mis en évidence par la figure 3.13 sur laquelle on observe également que l'albédo du sol nu d'ECOCLIMAP-II est nettement supérieur à celui d'ECOCLIMAP-I. Cependant des trouées sont observées sur le produit JULES, précisément dans les hyper classes forêt et forêt mixte et on note une différence de l'ordre de 0.2 dans le PIR avec les pixels contigus à ces données manquantes. Certains types de sols (e.g. Gypsisols) possèdent clairement un histogramme bimodal des

valeurs d'albédo (figure 3.14). On observe clairement que les différences entres les modes des bandes du visible et du proche infrarouge d'albédo du sol nu d'ECOCLIMAP-II varient entre 0.15 et 0.25. Le tableau 3.8 illustre également le besoin de mettre à jour l'information texturale du continent africain car il est surprenant d'y observer que les arenosols el podzols aient un pourcentage de sable supérieur à celui des dunes.

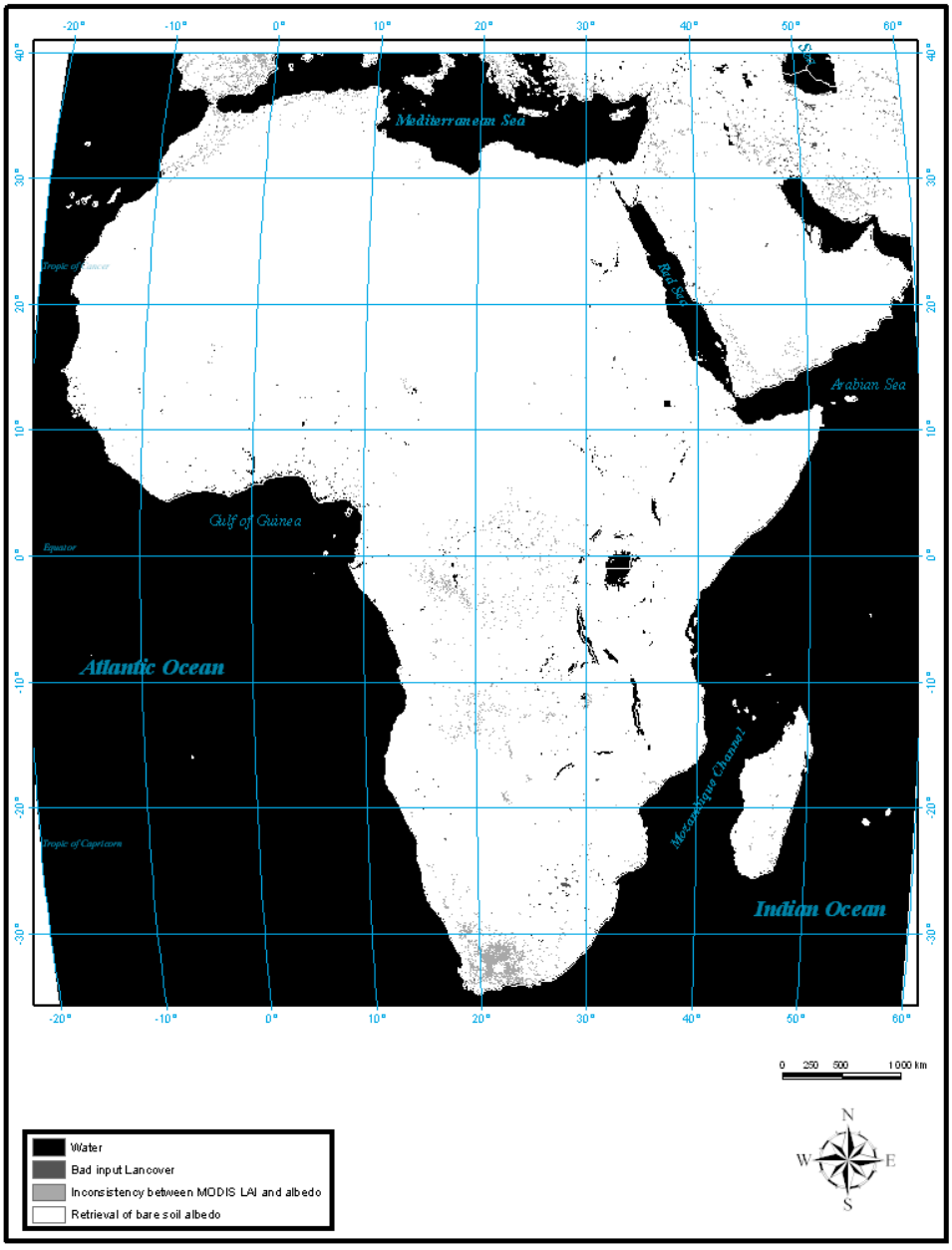

Figure III.11 Localisations des valeurs aberrantes obtenues lors de la détermination de l'albédo du sol nu dans le visible (i.e. en utilisant l'équation 2.1).

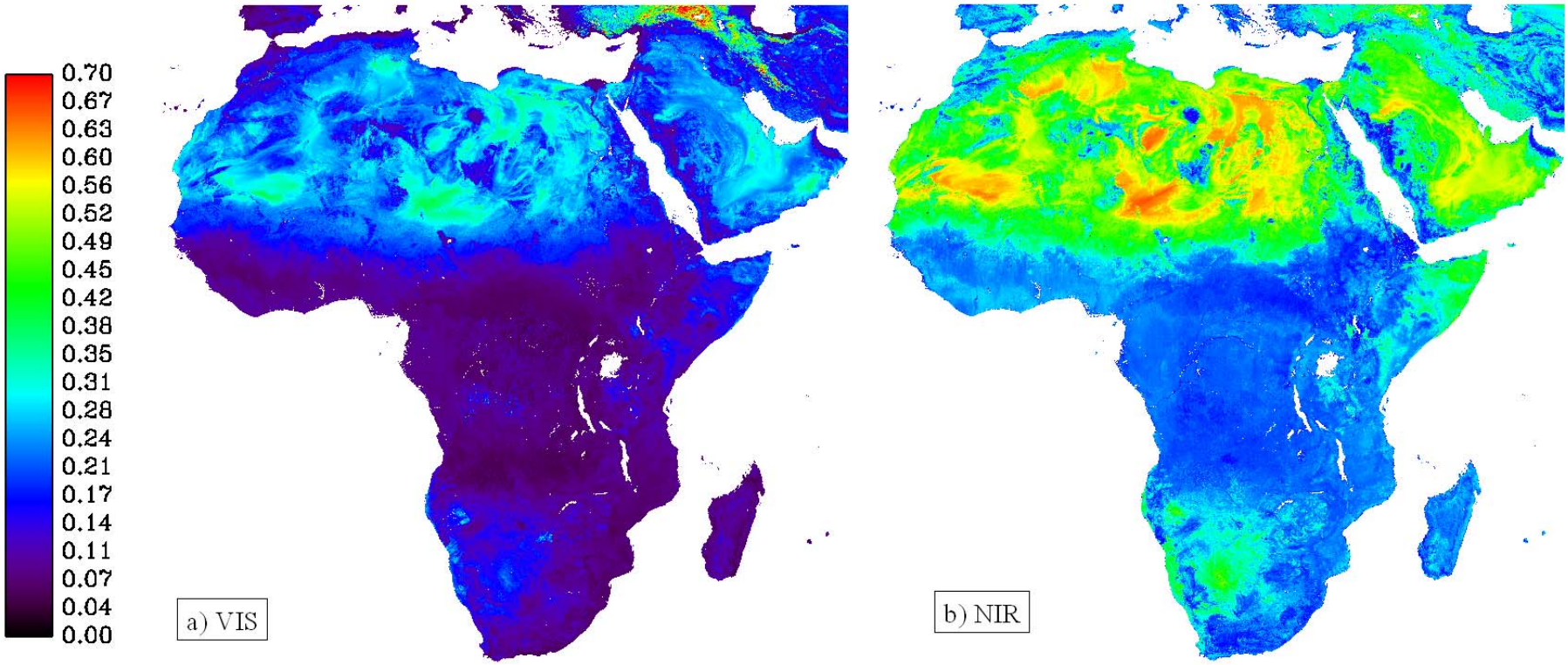

Figure III.12 Albédo du sol nu dans le visible (gauche) et proche infrarouge (droite)

Soil	Pixels	Texture (%)		visible (%)				near infrared (%)			
		Sand	Clay	ISLSCP-II	ECOCLIMAP-I	JULES	ECOCLIMAP-II	ISLSCP-II	ECOCLIMAP-I	JULES	ECOCLIMAP-II
Acrisols	493341	57.5 (10.5)	22.8 (6.2)	11.5 (3.8)	9.6 (3.7)	14.2 (2.8)	8.6 (8.6)	28.7 (3.7)	19.5 (3.7)	14.2 (9.2)	23.5 (3.2)
Alisols	45440	54.2 (18.0)	26.3 (8.7)	10.3 (2.0)	8.3 (2.7)	12.8 (2.2)	7.9 (7.9)	27.1 (1.3)	18.3 (2.7)	8.9 (7.1)	21.3 (2.5)
Andosols	38884	44.1 (8.7)	23.1 (9.2)	11.3 (2.2)	8.7 (2.7)	14.6 (2.3)	9.8 (9.8)	28.9 (3.1)	18.7 (2.7)	16.8 (7.7)	24.3 (2.3)
Arenosols	5333249	86.2 (9.2)	7.6 (3.8)	24.9 (9.7)	22.9 (9.9)	21.2 (5.8)	18.2 (17.8)	37.4 (8.6)	32.2 (9.9)	38.1 (15.3)	38.9 (12.1)
Chernozems	8488	28.3 (4.7)	37.4 (8.2)	14.5 (3.8)	9.0 (3.1)	14.5 (2.7)	10.7 (10.7)	32.2 (3.2)	19.0 (3.1)	19.2 (8.6)	24.6 (2.9)
Calcisols	4397305	45.2 (11.5)	19.7 (4.0)	28.0 (7.1)	23.6 (5.3)	23.7 (5.4)	22.5 (22.4)	39.1 (7.1)	33.7 (5.3)	45.0 (10.7)	44.5 (10.3)
Cambisols	1417327	49.0 (12.8)	27.9 (8.1)	15.6 (6.1)	14.2 (7.5)	14.7 (3.5)	11.3 (11.3)	30.6 (5.2)	24.2 (7.5)	22.1 (10.6)	26.1 (6.1)
Fluvisols	802773	41.7 (10.1)	24.7 (7.5)	19.7 (7.7)	16.7 (7.6)	18.7 (4.9)	15.4 (15.4)	33.4 (6.4)	26.7 (7.6)	29.9 (14.4)	32.2 (10.3)
Ferralsols	1213949	55.1 (11.6)	28.9 (8.5)	10.4 (2.0)	8.5 (2.8)	13.6 (2.3)	7.5 (7.6)	27.5 (2.3)	18.6 (2.8)	11.5 (7.6)	22.5 (2.0)
Gleysols	254688	46.4 (17.6)	24.5 (12.4)	18.4 (8.7)	14.9 (6.4)	17.9 (4.9)	12.9 (13.0)	34.6 (6.6)	24.9 (6.4)	27.9 (13.9)	29.4 (8.7)
Greyzems	601	36.5 (7.2)	33.4 (5.8)	9.8 (0.0)	6.6 (0.1)	15.9 (0.7)	8.1 (8.1)	26.7 (0.2)	16.6 (0.1)	6.8 (3.7)	25.0 (1.4)
Gypsisols	419758	48.9 (5.1)	20.3 (6.6)	26.9 (5.2)	24.2 (6.3)	21.8 (6.1)	20.9 (20.9)	38.5 (5.6)	34.1 (6.3)	41.6 (10.2)	41.5 (9.1)
Histosols	24011	26.0 (6.2)	25.5 (9.6)	12.5 (3.6)	8.1(2.5)	14.7 (3.9)	7.6 (7.6)	28.3 (2.2)	18.1 (2.5)	19.5 (18.9)	21.8 (2.5)
Kastanozems	31413	30.5 (7.7)	34.6 (9.1)	12.6 (2.3)	10.2 (4.2)	17.5 (1.9)	11.8 (11.8)	32.1 (2.6)	20.2 (4.2)	26.5 (4.4)	27.6 (3.2)
Leptosols	4959778	47.1 (7.2)	23.3 (4.2)	22.0 (8.6)	18.6 (7.3)	18.6 (6.1)	16.4 (15.9)	33.6 (8.1)	28.6 (7.3)	31.7 (14.8)	33.7 (12.3)
Luvisols	1162786	54.0 (14.4)	25.9 (9.6)	15.0 (5.6)	12.3 (6.4)	14.7 (3.0)	10.7 (10.7)	31.0 (4.3)	22.4 (6.4)	21.5 (8.8)	25.7 (5.2)
Lixisols	1461832	59.7 (11.0)	20.4 (6.3)	15.5 (6.7)	11.8 (4.5)	15.9 (3.7)	11.2 (11.2)	32.7 (5.2)	21.8 (4.5)	21.8 (10.7)	26.4 (5.3)
Nitisols	399596	39.5 (14.0)	34.2 (13.2)	12.2 (3.9)	9.9 (3.6)	14.5 (2.8)	9.4 (9.4)	29.4 (3.6)	19.9 (3.6)	15.5 (8.6)	23.5 (3.5)
Phaeozems	119644	43.0 (14.0)	35.5 (10.9)	11.4 (2.7)	9.3 (3.4)	13.4(2.6)	9.1 (9.1)	29.0 (3.4)	19.3 (3.4)	14.6 (8.5)	24.0 (3.2)
Planosols	150113	49.7 (12.4)	24.6 (7.0)	17.3 (7.5)	11.8 (5.5)	17.3 (4.4)	13.1 (13.1)	33.3 (5.7)	21.8 (5.5)	26.4 (12.2)	29.2 (6.9)
Plinthosols	390381	53.7 (6.9)	22.6 (3.7)	12.6 (4.8)	10.2 (3.8)	14.1 (3.5)	9.5 (9.5)	30.1 (4.2)	20.2 (3.8)	15.3 (10.2)	24.2 (4.4)
Podzols	15590	77.9 (1.1)	9.0 (0.4)	10.1 (0.2)	13.5 (4.7)	15.8 (1.6)	8.3 (8.3)	28.6 (1.3)	23.5 (4.7)	22.0 (6.5)	24.5 (1.8)
Regosols	914157	55.4 (10.9)	18.5 (4.9)	19.9 (7.4)	17.4 (7.8)	18.0 (5.1)	15.2 (15.2)	33.4 (6.3)	27.4 (7.8)	29.7 (13.3)	31.3 (10.1)
Solonchaks	296051	44.7 (14.0)	27.6 (10.6)	24.9 (8.0)	23.0 (8.0)	22.6 (6.5)	20.9 (20.9)	37.2 (8.0)	32.9 (8.0)	39.3 (12.6)	38.4 (11.6)
Solonetz	381663	58.7 (16.9)	23.8 (10.9)	17.1 (6.9)	12.4 (5.5)	16.2 (3.8)	11.6 (11.7)	33.2 (5.1)	22.4 (5.5)	23.3 (11.0)	27.6 (6.6)
Vertisols	1128509	25.5 (7.3)	49.1 (6.3)	18.2 (6.0)	12.1 (4.3)	14.3 (3.6)	11.0 (11.0)	33.5 (4.2)	22.2 (4.4)	20.0 (9.2)	24.1 (5.1)
Rock debris	473341	44.1 (1.9)	21.2 (0.5)	27.1 (4.6)	20.2 (1.2)	23.3 (4.4)	22.9 (22.9)	35.7 (6.3)	30.0 (1.2)	42.7 (8.8)	42.5 (8.9)
Dunes	826508	55.5 (16.5)	16.6 (5.4)	33.7 (4.4)	32.1 (3.8)	27.3 (4.9)	27.1 (27.2)	44.8 (5.8)	42.4 (3.8)	50.9 (7.8)	50.5 (7.7)

Tableau III-8 Valeurs moyennes et deviations standards entre parentheses (%) à l'échelle continnetale africaine pour l'albédo du sol nu et par type de sol.

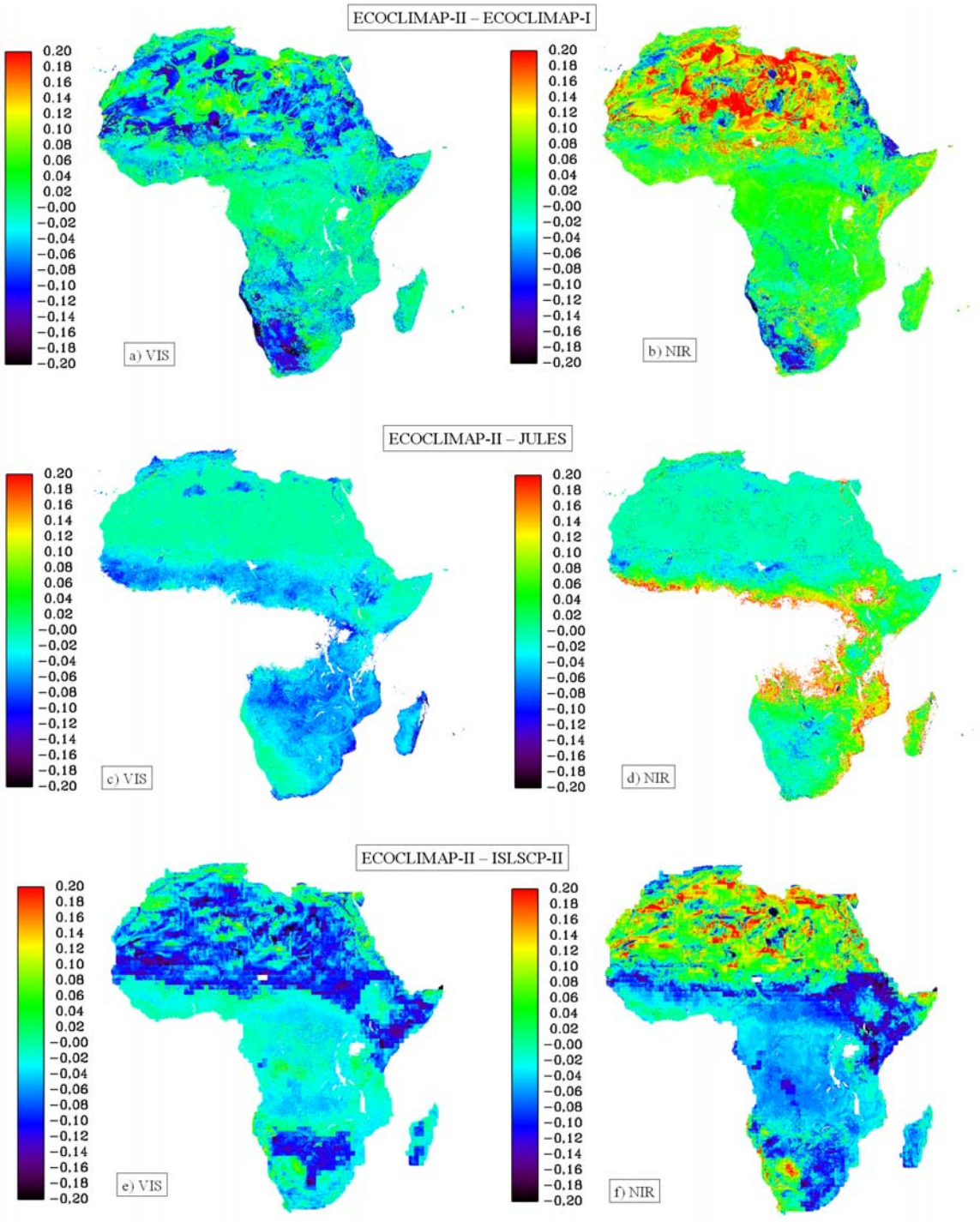

Figure III.13 Différence d'albédo du sol nu d'ECOCLIMAP-II et ECOCLIMAP-I (haut), JULES (milieu) et ISLSCP-II (bas) dans les spectres du visible (gauche) et du proche infrarouge (droit)

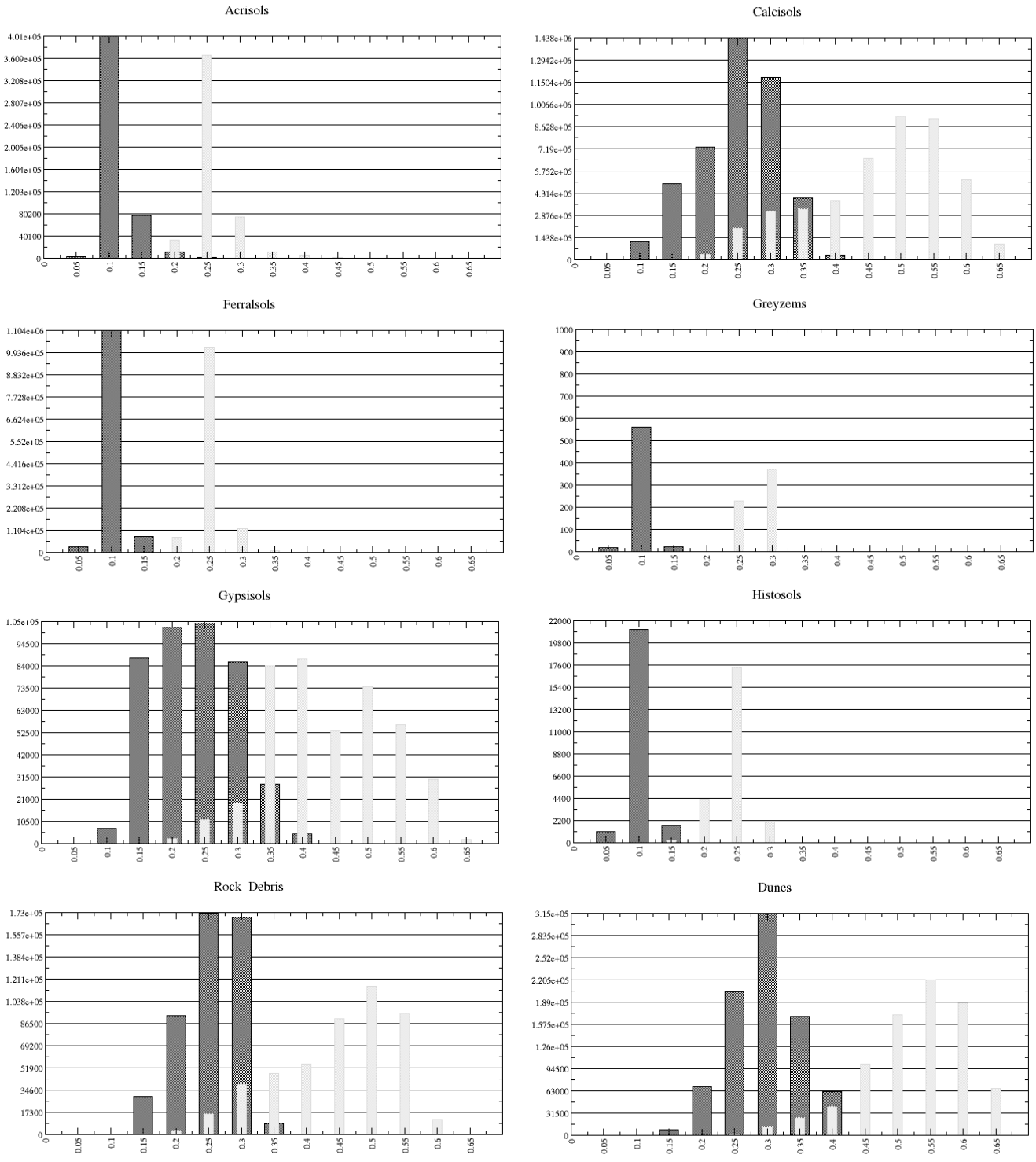

Figure III.14 Histogrammes d'albédo du sol nu pour certains types de sol : visible noir) et proche infrarouge (gris)

3.4.5.2 Albédo de la végétation

Les valeurs d'albédo du sol nu obtenues permettent de produire, à partir de l'Eq. 2.1, les valeurs de l'albédo de la végétation pour chaque huitaine. Ces valeurs ont par la suite été moyennées pour produire une climatologie de valeurs fixes (figure 3.15 a-b). Plus la végétation est haute, moins les valeurs d'albédo de la végétation sont élevées. Les différences d'albédos de la végétation entre les 2 versions d'ECOCLIMAP (illustrées à la figure 3.15 c-d) permettent de conclure que les albédos de la végétation d'ECOCLIMAP-II sont en général supérieurs à ceux d'ECOCLIMAP-I. Cependant on observe des valeurs aberrantes ou incohérentes dans certaines localités situées en Turquie et en Iran, en particulier dans le spectre du visible (voir également figures 3.12 a et 3.15 a). Ces incohérences sont dûes à un mauvais filtrage des nuages dans la chaîne de production des albédo MODIS. En effet, ces incohérences s'observent principalement sur la période novembre à mars (avec un pic en janvier) qui correspond bien à la saison pluvieuse dans ces zones et atteste ainsi de la présence des nuages.

3.4.5.3 Albédo total

On a évalué les différences d'albédo de surface entre MODIS et les 2 versions d'ECOCLIMAP en moyennant mensuellement pour chaque produit toutes les valeurs obtenues durant les 8 années. En janvier et juillet, dans les spectres du visible et du PIR, les différences entre les albédos MODIS et ECOCLIMAP-I sont importantes sur les déserts et considérablement réduites dans les cartes de différences en utilisant les albédos ECOCLIMAP-II comme le montre la figure 3.16. Cette mauvaise représentation de l'albédo de surface pour ces régions dans ECOCLIMAP-I, s'explique par les valeurs erronées d'albédo du sol nu (voir aussi figure 3.13 a-b). Le tableau 3.9 montre que l'usage des valeurs améliorées d'albédo à la place de celles existantes réduit l'erreur moyenne quadratique de 0.072 à 0.010 et de 0.069 à 0.009 respectivement sur les sols nus en janvier et juillet dans les spectres du visible et du PIR.

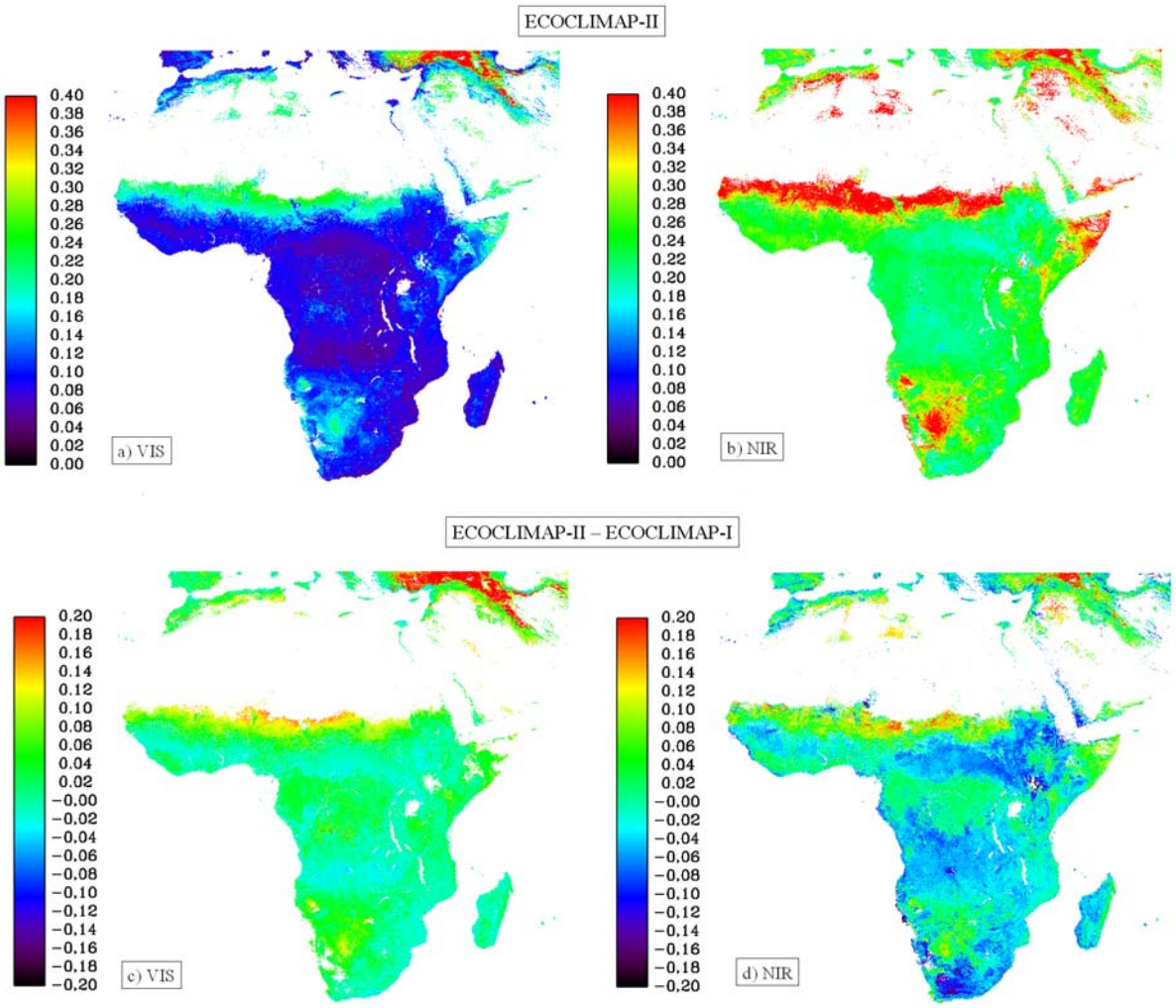

Figure III.15 Comportement spatial de l'albédo de la végétation d'ECOCLIMAP-II (haut) et différence d'albédo de la végétation entre ECOCLIMAP-II et ECOCLIMAP-I (bas) dans le visible (gauche) et dans le proche infrarouge (droit)

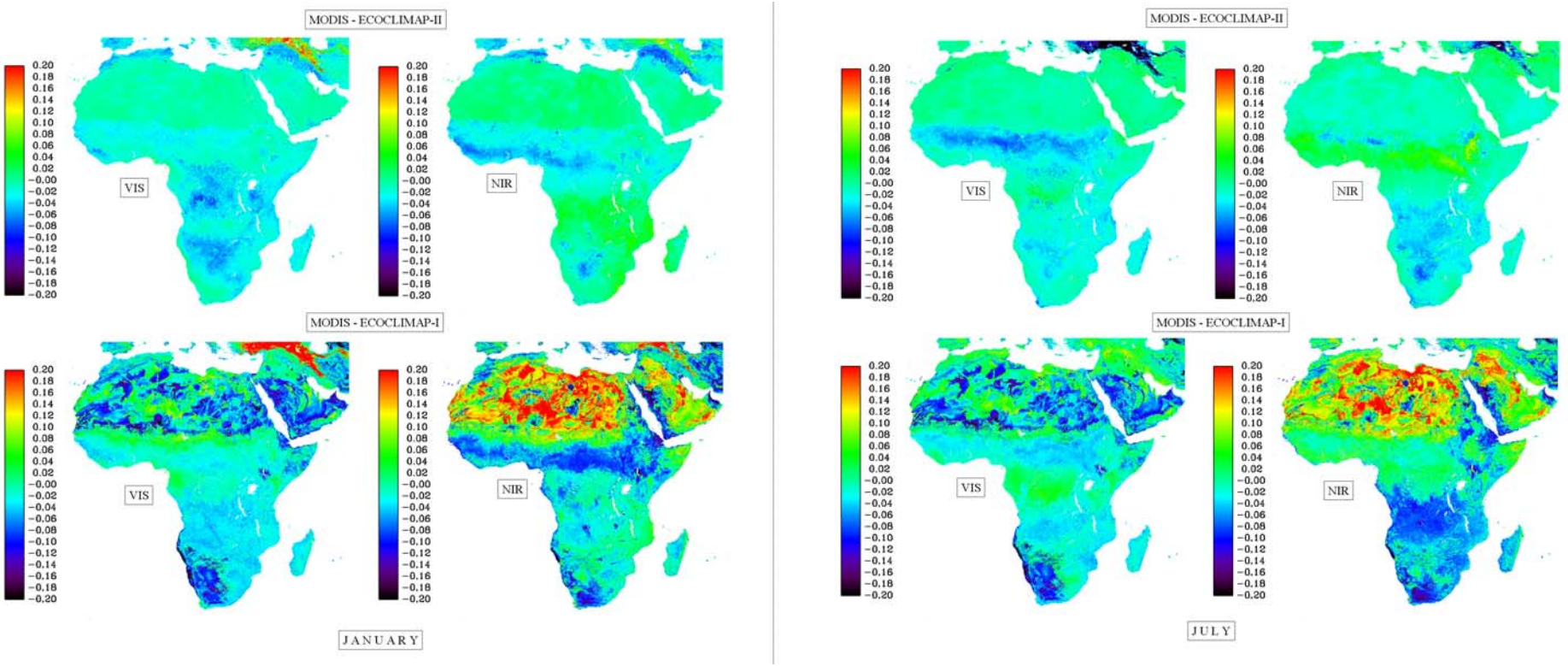

Figure III.16 Comportement spatial des différences MODIS-ECOCLIMAP-II (haut) et MODIS-ECOCLIMAP-I (bas) en janvier (gauche) et juillet (droite) dans les spectres du visible et du proche infrarouge

Compte tenu du fait que le niveau de précision des albédos MODIS est égal à 0.002 en valeur absolue (Jin et al., 2003 ; Salomon et al., 2006) et en prenant ces albédos MODIS comme valeurs de référence, les différences d'albédo entre les 2 versions d'ECOCLIMAP inférieures à ± 0.002 sont statistiquement insignifiantes. Le tableau 3.10 montre qu'en janvier 60% des surfaces continentales vérifient cette condition dans ECOCLIMAP-II contre seulement 23% dans ECOCLIMAP-I. Puisque l'un des objectifs de cette étude est de mettre à jour les variables biophysiques en utilisant les données satellitales, la différence MODIS-ECOCLIMAP-I illustre le biais alors que la différence MODIS-ECOCLIMAP-II illustre l'amélioration de la nouvelle base de données.

	ECOCLILMAP-I				ECOCLOMAP-II			
	visible		near infrared		visible		near infrared	
month	vegetated	bare land	vegetated	bare land	vegetated	bare land	vegetated	bare land
1	0.04164	0.07164	0.05837	0.13851	0.0294	0.00971	0.02707	0.01466
2	0.04185	0.07033	0.05699	0.13695	0.02825	0.0074	0.02201	0.01151
3	0.04186	0.06948	0.05394	0.1355	0.02873	0.00651	0.0195	0.0095
4	0.04505	0.06895	0.05291	0.12265	0.03308	0.00754	0.02174	0.01209
5	0.04398	0.06878	0.05724	0.12602	0.03082	0.00736	0.02255	0.00901
6	0.04096	0.06882	0.06021	0.12488	0.02725	0.00713	0.02569	0.00939
7	0.04056	0.06899	0.06125	0.12315	0.02811	0.00862	0.02931	0.01142
8	0.03967	0.0698	0.05519	0.12512	0.02934	0.00808	0.02476	0.00911
9	0.04032	0.0696	0.05115	0.12935	0.03218	0.0075	0.02417	0.00768
10	0.03944	0.07004	0.05139	0.13654	0.03048	0.00668	0.02243	0.01056
11	0.04061	0.07124	0.05252	0.11772	0.02919	0.00875	0.02448	0.01923
12	0.04103	0.07223	0.05842	0.13821	0.0295	0.011	0.02764	0.01632

Tableau III-9 Erreur moyenne quadratique entre les moyennes climatologiques d'albédo de surface de MODIS comparées à celles d'ECOCLIMAP en prenant la classification ECOCLIMAP-II pour référence. Il faut noter que les pixels urbains ont été écartés lors de cette comparaison

	ECOCLIMAP-I		ECOCLIMAP-II	
month	vis	nir	vis	nir
1	31.51823	23.30084	60.14239	63.75103
2	32.22372	22.90928	65.23942	76.71568
3	32.63275	22.22867	65.90112	82.52068
4	27.16052	22.36796	59.27165	77.01801
5	29.07064	21.21774	58.5331	78.68882
6	31.26039	21.60544	62.15512	71.91695
7	32.36913	22.29374	65.31283	65.57117
8	35.71674	23.13333	67.92836	73.3475
9	35.50812	22.92936	63.96376	77.64245
10	34.45034	20.38952	63.71481	79.33225
11	31.46769	21.36093	56.78001	64.21023
12	30.86068	23.98891	55.23712	62.48656

Tableau III-10 Pourcentage des différences non significatives des moyennes mensuelles climatologiques d'albédo de surface d'ECOCLIMAP (I et II) en prenant MODIS comme référence

3.4.5.4 Fraction de végétation

La figure 3.17 montre la fraction de végétation estimée selon l'approximation 2.3 ainsi que celle issue des produits ECOCLIMAP-I et CYCLOPES pour les mois de février et août. Les valeurs CYCLOPES sont sous-estimées dans le bassin du Congo (inférieures à 0.7) et on observe également la délimitation des régions constituées de roche nues (figure 3.17.a), ce qui illustre des défaillances techniques lors du processus d'estimation du veg dans la chaine CYCLOPES. Dans le bassin du Congo, le type fonctionnel de plante (TFP) majoritaire est « la forêt de feuillus sempervirent » et dans ECOCLIMAP-I, la valeur du veg de ce TFP est constante et égale à 0.95 (figure 3.17 c), ce qui masque les differentes stages phénologiques des plantes dans cette région (figure 1.8). Par contre, les valeurs d'ECOCLIMAP-II (figure 3.17 b) semblent plus proches de celles simulées par des modèles tel que l'aDGVM spécialement développé sur l'Afrique (Scheiter et Higgins, 2006) que celles des 2 autres produits ECOCLIMAP-I et CYCLOPES.

3.5 Résumé et conclusion

Ce chapitre a eu pour objectif de présenter les différents changements apportés à la base de données ECOCLIMAP sur l'Afrique. Une méthode hybride de classification a été développée afin de permettre la cartographie des classes d'occupation des sols. Des méthodes spécifiques décrites dans le chapitre précédent ont par la suite été utilisées pour estimer les variables biophysiques de premier ordre telles que l'albédo du sol nu et l'albédo de la végétation. Tout en réduisant ainsi le biais existant entre les variables biophysqiues des données ECOCLIMAP-I et MODIS, nous avons amélioré la cohérence entre les données de LAI et d'albédo. La procédure de '*dévégétalisation*' de l'albédo présentée au chapitre précédent a été appliquée sur l'intégralité du continent africain. Les nouvelles estimations des albédos ont par la suite été comparées aux albédos ECOCLIMAP-I et MODIS, ce qui a permis la mise en évidence d'un bon accord entre les produits MODIS et ECOCLIMAP-II.

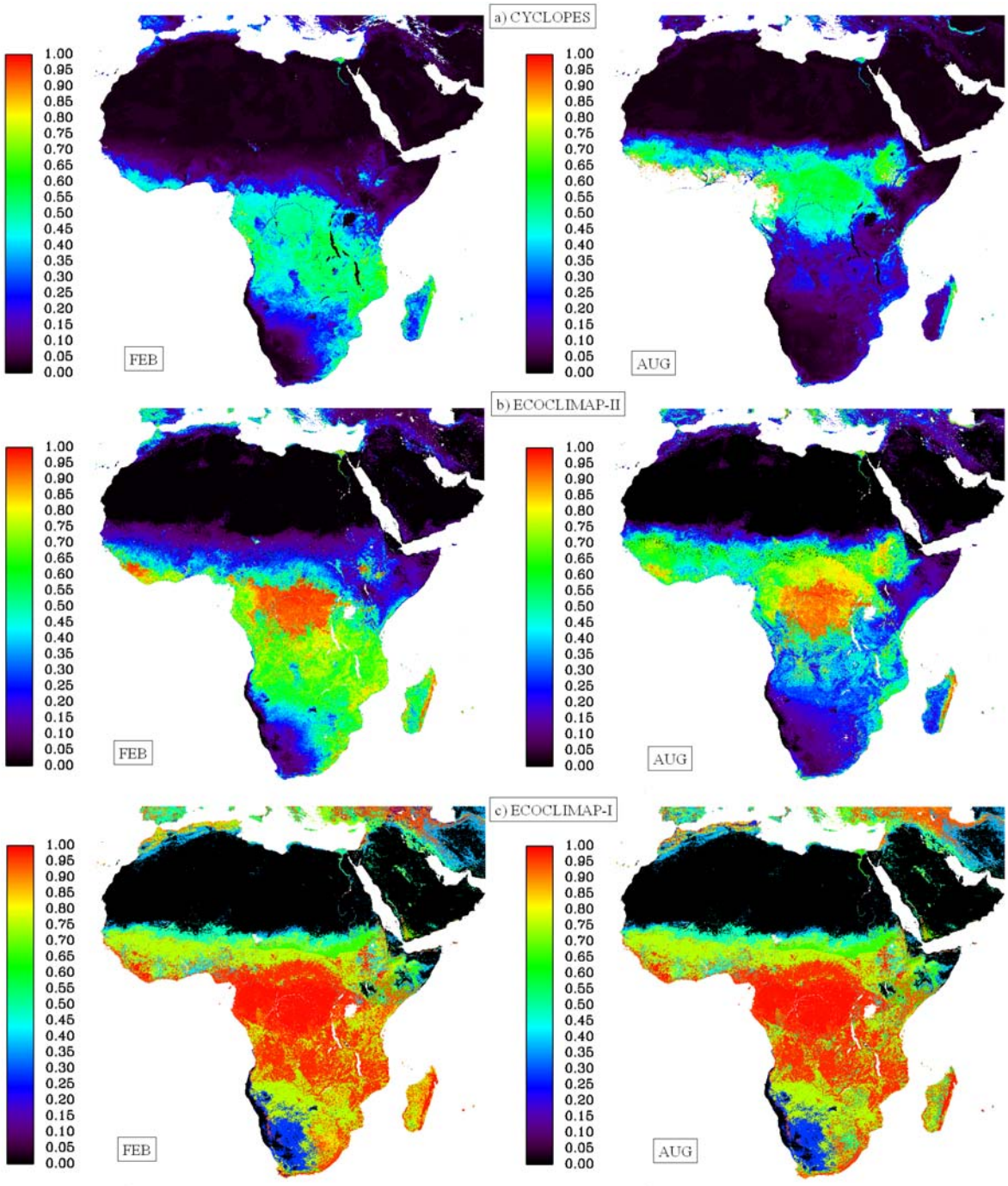

Figure III.17 Comparaison spatiale de la fraction de végétation issue des produits CYCLOPES (a) , ECOCLIMAP-II (b) et ECOCLIMAP-I (c)

Les variables biophysiques ainsi obtenues nous permettent de modéliser la surface, ce qui fait l'objet de notre prochain chapitre. Cependant les algorithmes de la chaîne de traitement MODIS doivent être améliorés de sorte à ce que cette procédure n'induise plus de valeurs aberrantes des albédos du sol nu et de la végétation pour certaines localités.

Une carte délimitant la distribution spatiale de 5 écorégions bioclimatiques à l'échelle du continent africain a été produite par utilisation de l'iISFC, un nouvel indice de segmentation basé sur l'utilisation de la TFD. Cet indice peut être utilisé pour effectuer des analyses d'images satellitales à n'importe quelle échelle (locale, nationale, régionale, continentale ou globale) indépendamment de la résolution (basse, moyenne ou haute) et du domaine de la télédétection (optique ou radar).

Une nouvelle méthodologie de classification, la *stacked segement classification*, basée sur la prise en compte de l'information bioclimatique et des classifications successives par l'algorithme du k-NN, a été développée afin de produire une carte d'occupation du sol à 1 km à partir d'une série pluri-annuelle d'images satellitales. L'utilisation de cette méthode permet une identification et une labellisation des classes d'occupation des sols plus flexible et plus correcte que celle effectuée par la seule utilisation du k-NN, ce qui lui confère le mérite d'être facilement reproductible sur d'autres domaines. La nouvelle carte a été comparée à d'autres produits de classification -notamment GLC2000, GLOBCOVER et MODIS- réalisés à partir d'images satellites acquises dans les années 2000, ce qui a permis de mettre en évidence d'importantes disparités entre ces produits précisément au niveau des paysages hétérogènes. Les séries d'images à haute résolution doivent donc être mises à la disposition de la communauté scientifique pour une cartographie effective des zones de disparité. Ceci pourrait être utilisé pour étendre le projet AFRICOVER (http://www.africover.org/) à l'ensemble du continent africain avec des observations LANDSAT récemment disponibles en libre accès (Roy et al., 2010, Wulder et al., 2008).

« Triste époque que celle où il est plus difficile de briser un préjugé qu'un atome », Albert Einstein, Comment je vois le monde, 1934

IV. Sensibilité de la modélisation des flux hydriques dans ISBA au forçage physiographique[3]

4.1 Introduction

ISBA (Noilhan et Planton, 1989 ; Noilhan and Mahfouf, 1996) est le module de la plateforme de modélisation de surface SURFEX (Le Moigne et al., 2009) qui permet de simuler les échanges d'eau et d'énergie entre le sol, la végétation et les basses couches de l'atmosphère. ISBA est un schéma de surface en pleine évolution développé par le CNRM et utilisé dans plusieurs modèles atmosphériques à Météo France pour la prévision numérique du temps (ARPEGE) ou les recherches sur la méso-échelle et le climat (MESO-NH) via l'interfaçage de SURFEX (figure 4.1) (e.g. Mahfouf et al., 1995). Comme tout LSM, ISBA a besoin de la spécification d'un certain nombre de variables décrivant les caractéristiques de la végétation telles que l'albédo, l'indice foliaire (LAI) ou la fraction de végétation (*veg*). En effet, l'hétérogénéité spatiale et les caractéristiques des classes d'occupation des sols modulent les flux d'énergie, d'eau et de carbone entre la surface et la couche limite (Baret al., 2007 ; Roujean et Lacaze, 2002). Le LAI contrôle l'évapotranspiration et le rayonnement solaire absorbé par la photosynthèse, le facteur *veg* partitionne l'évaporation de surface entre la contribution des plantes et celle du sol nu, tandis que l'albédo module le bilan d'énergie en contrôlant la quantité d'énergie solaire absorbée à la surface de la terre.

L'impact des modifications des caractéristiques de surface telles que l'occupation des sols, l'indice foliaire, la fraction de végétation, et les albédos a été examiné dans de nombreuses études antérieures (e.g. Barlage and Zeng, 2004; Buermann et al., 2001; Houldcroft et al., 2009). Certains biais des modèles rapportés dans ces études résultent de l'inexactitude ou de l'incohérence de certains paramètres biophysiques qui peuvent induire un mauvaise partitionnement : (i) de l'énergie radiative disponible entre flux de chaleur

[3] Ce chapitre est fondé sur Kaptué, T. A. T., Boone, A., Kergoat, L., Faroux, S., Le Moigne, P., Alkama, M. R., Mahfouf, J.-F. and Roujean, J.-L. (2010), Impact of a new MODIS-based physiography forcing on Land Surface Model simulations over the Western African Region, *Journal of Geophysical Research*, **submitted**.

latente et sensible et (ii) des précipitations entre l'évapotranspiration et le ruissellement (Zhao et al., 2001). Dans le Common Land Model (CLM) utilisé par Barlage and Zeng (2004), il existe des incohérences entre les valeurs de LAI et de *veg* pour certains pixels car les 17 classes du produit IGBP-DIS ont été utilisées lors de la dérivation du *veg* contre seulement 6 classes pour la dérivation du LAI. Le LAI et le veg dérivent des données de l'Advanced Very High Resolution Radiometer (AVHRR), dont la qualité est dégradée par le mauvais étalonnage, les larges variations de la géométrie de visée et la dérive du capteur (Gutman, 1999).

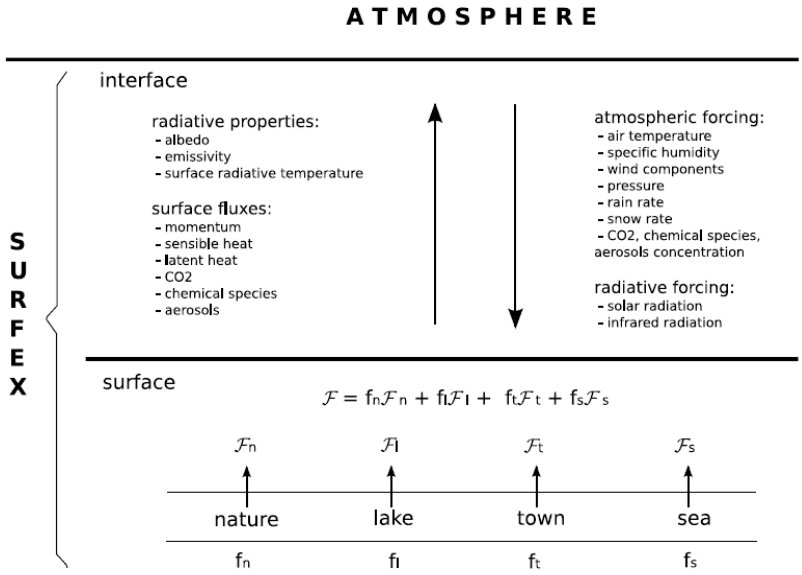

Figure IV.1 Description des échanges entre un modèle atmosphérique envoyant des flux radiatifs et un forçage météorologique à la surface et le module SURFEX décrivant un ensemble de 4 modèles physiques avec une interface où les variables moyennées sont renvoyées à l'atmosphère (d'après Le Moigne et al., 2009)

Le chapitre 2 a montré le développement d'une nouvelle classification de la surface comprenant une carte d'écosystèmes associée à des séries temporelles de LAI, de *veg* et d'albédo. L'albédo se compose d'un albedo du sol nu constant et d'un albedo de la végétation temporellement fixe pondéré par la fraction de végétation variable dans le temps. Au regard de la classification ECOCLIMAP-I (Champeaux et al., 2005) existante dans ISBA, la nouvelle cartographie induit 3 améliorations résultant de l'utilisation des données MODIS :

(i) une meilleure répartition des écosystèmes par classe, (ii) des valeurs récentes d'albédo et de LAI incluant la variabilité inter-annuelle et (iii) une cohérence entre les albédos et le LAI tous les deux dérivés des mêmes données satellitaires.

Le but de ce chapitre est de remplacer la base de données ECOCLIMAP-I prescrite dans ISBA par la nouvelle classification de la surface et ensuite d'évaluer dans ISBA son impact sur les bilans énergétiques et hydriques. Les simulations décrites dans les sections suivantes font partie intégrante du projet ALMIP d'intercomparaison des LSM sur la région Ouest Africaine. La 2ème partie décrit brièvement le projet ALMIP tandis que la 3^{e} partie décrit les observations et la méthodologie utilisée dans cette étude. Les 4^{e} et 5^{e} parties illustrent comment les variables climatiques ont changé à l'échelle régionale après l'usage du nouveau forçage physiographique et la 6^{e} partie s'achève par des discussions et quelques conclusions.

4.2 Le modèle de surface ISBA standard

Il existe plusieurs versions du module de surface ISBA, toutes basées sur l'approche « *force-restore* » de Deardorff (1977). L'approche « *force-restore* » est une simplification des principes physiques décrivant les transferts d'eau et d'énergie dans le sol en supposant que la température et la quantité d'eau évoluent sous l'action d'une contrainte (ou forçage), et sont rappelées vers une situation d'équilibre (ou rappel). Nous utiliserons tout au long de notre étude la version d'ISBA à 3 couches (ou ISBA 3-L) proposée par Boone et al. (1999), selon laquelle l'hydrologie du sol contient : une fine couche de surface (1 cm), la zone racinaire et une troisième couche (appelée couche sous-racinaire) qui permet de distinguer la profondeur racinaire de la profondeur totale du sol. La figure 4.2 synthétise les processus hydrologiques et énergétiques simulés par le modèle ISBA 3-L. Le réservoir de surface (w_1), utilisé pour le calcul de l'évaporation du sol nu (E_{soil}), est inclus dans le réservoir racinaire (w_2) qui pilote en partie la transpiration des plantes (E_{transp}). Les différents réservoirs échangent de l'eau par diffusion (D_i) et drainage gravitationnel (K_i). Le réservoir de pluie interceptée par la végétation (w_r) se vide par ruissellement (d_r) et évaporation potentielle

(E_{canop}). Les pluies non interceptées, s'infiltrent dans le sol (I_r) ou provoquent un ruissellement de surface (Q_s). Pour une meilleure compréhension de la modélisation des flux hydriques et énergétiques dans ISBA, le lecteur pourra se référer à l'annexe D.

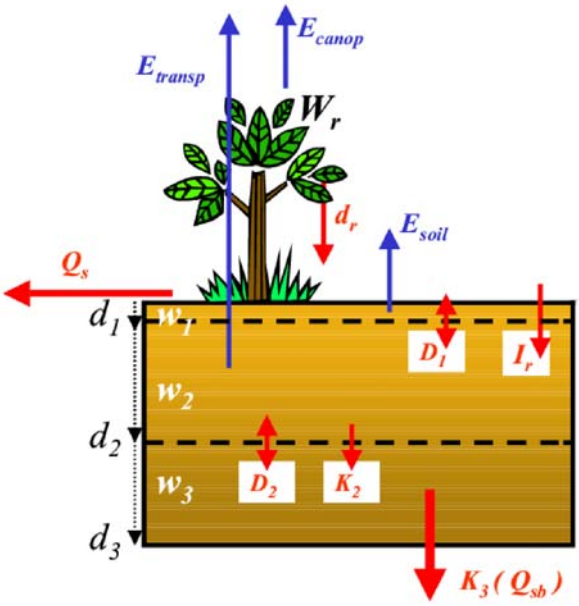

Figure IV.2 Schéma de surface ISBA-3L (adapté d'après Douville, 2006)

4.2.1 Les paramètres d'ISBA

Les paramètres d'entrée ont été choisis de façon à caractériser les principaux processus physiques observés à la surface tout en essayant de réduire le nombre de variables indépendantes Ces paramètres peuvent être regroupés en deux grandes catégories :

(i) Les **paramètres primaires** qui décrivent la nature de la surface continentale et la texture du sol ont besoin d'être spécifiés à chaque point de grille du modèle à l'aide des combinaisons linéaires des types fonctionnels de surface, les pourcentages de sable et d'argile et de sable et le masque terre-mer ;

(ii) Les **paramètres secondaires** sont associés au type de sol et estimés à partir des paramètres primaires notamment :

- Les *paramètres hydrologiques liés au sol* tels que le contenu en eau à la saturation (w_{sat}), le contenu en eau à la capacité du champ (w_{fc}), le contenu en eau au point de flétrissement (w_{wilt}), la conductivité hydraulique (k_{sat}), les

constantes associées aux vitesses de drainage (C_i), les propriétés thermiques, etc.

- Les *paramètres liés à la végétation* tels que la fraction de végétation (*veg*), l'indice foliaire (LAI), l'albédo (*A*), l'émissivité (ε) et les longueurs de rugosité dynamique (z_{0m}) et thermique (z_{0h}) qui évoluent avec le temps, la résistance stomatique minimale (R_{smin}) et les horizons de sols (d_i) qui restent constants.

4.2.2 Le projet ALMIP

Le projet ALMIP a vu le jour afin de coordonner les activités de modélisation des processus de surface au sein du programme AMMA à travers une intercompraison des simulations de 11 LSM parmi lesquels ISBA (Boone et al., 2009). En effet, une meilleure estimation des flux de surface peut être obtenue en combinant des simulations multi-modèles sur de grandes échelles régionales (Dirmeyer et al., 2006). Les objectifs d'ALMIP sont :

(i) l'étude de la sensibilité des modèles aux paramétrisations et aux forçages d'entrée,

(ii) la détermination des processus modélisés de façon inadéquate dans les LSM

(iii) l'examen des réponses des LSM aux changements d'échelles spatiale : locale à régionale en passant par celle du paysage (telles que décrites dans l'introduction)

(iv) le développement d'une climatologie 'réaliste' à haute résolution spatiale de l'humidité des sols, des flux de surface et des variables diagnostiques des bilans hydriques et énergétiques de surface,

(v) l'évaluation de la réponse de la végétation dans les LSM au forçage atmosphérique sur des échelles de temps saisonnières et interannuelles.

ALMIP est un projet en cours dont la 1ère phase (ALMIP-I), récemment achevée, s'est attelée à évaluer les thèmes (i) et (iv) à partir de 2004 à 2007 en analysant les résultats des simulations de 11 LSM (parmi lesquels ISBA) à l'échelle régionale sur l'Afrique de l'Ouest. La

phase ALMIP-I a ainsi permis de montrer que l'utilisation des forçages basés sur des données satellites permet d'améliorer de façon significative les résultats des simulations des LSM par exemple pour l'évapotranspiration et aussi les autres flux hydriques. La 2ème phase du projet (ALMIP-II) va débuter cet automne et s'attellera à évaluer les thèmes restants (à savoir (ii), (iii) et (v)). Le travail qui suit fait donc partie intégrante de la phase ALMIP-II et en particulier du thème (iii).

4.3 Données et méthodologie

Afin de déterminer les biais de la base de données ECOCLIMAP-I, dérivée des données AVHRR acquises entre 1992 et 1993 et actuellement utilisée dans ISBA, nous avons généré une nouvelle base de paramètres biophysiques à partir de la récente version 5 des données MODIS de LAI et d'albédo (voir chapitre 2).

Nous avons effectué deux simulations de 10 ans du LSM ISBA en mode forcé (c.-à-d. découplé d'un modèle atmosphérique) à partir de janvier 2002 à décembre 2007 sur la zone s'étendant approximativement de 20°O à 25°E et de 5°S à 20°N (figure 4.3) à une résolution spatiale de 0.5° et à une résolution temporelle de 3h ; ceci en utilisant la base de données ECOCLIMAP-I et la nouvelle base de données. Les simulations avec la classification ECOCLIMAP-I seront référencées "simulations de contrôle" (CTL) et celles avec la nouvelle physiographie "simulations expérimentales" (EXP). La comparaison de ces deux simulations illustre les modifications majeures du nouveau forçage physiographique. 5 itérations complètes de l'année 2002 ont été utilisées pour l'initialisation des humidités et températures dans le sol (i.e. comme spin-up) et seules les variables des 3 dernières années (2005-2007) sont utilisées pour la comparaison. On a utilisé le même forçage atmosphérique pour les deux simulations et ainsi, les différences dans les résultats des simulations seront uniquement dues aux propriétés des surfaces..

Les conditions initiales et aux limites telles que la température de l'air à 2 m, la pression atmosphérique à la surface, l'humidité spécifique et les composantes (vitesse et direction) du

vent à 10 m ont été fournies par le Centre Européen de Prévisions Météorologiques à Moyen Terme (CEPMMT) (http://www.ecmwf.int/). La pluviométrie est celle estimée par les données Tropical Rainfall Measurement Mission (TRMM) 3B-42 (Huffman *et al.*, 2007). Les rayonnements solaire et atmosphérique descendants ont été obtenus du CEPMMT pour les années 2002 et 2003, du SAF OSI (Oceans and Ice Satellite Applications Facility, http://www.osi-saf.org) pour 2004 et du SAF LAND (Satellite Applications Facility Land, http://landsaf.meteo.pt/) (Geiger *et al.*, 2008) pour les autres années. A noter que les produits du SAF OSI sont de moindre résolution sur terre mais ont permis de démarrer l'étude avant l'arrivée des données du SAF LAND. Toutes ces variables ont été reprojetées dans une projection Plate carrée (ellipsoïde de référence, WGS84) à une résolution spatiale de 0.5 degré et ensuite interpolées à une résolution temporelle de 3 h.

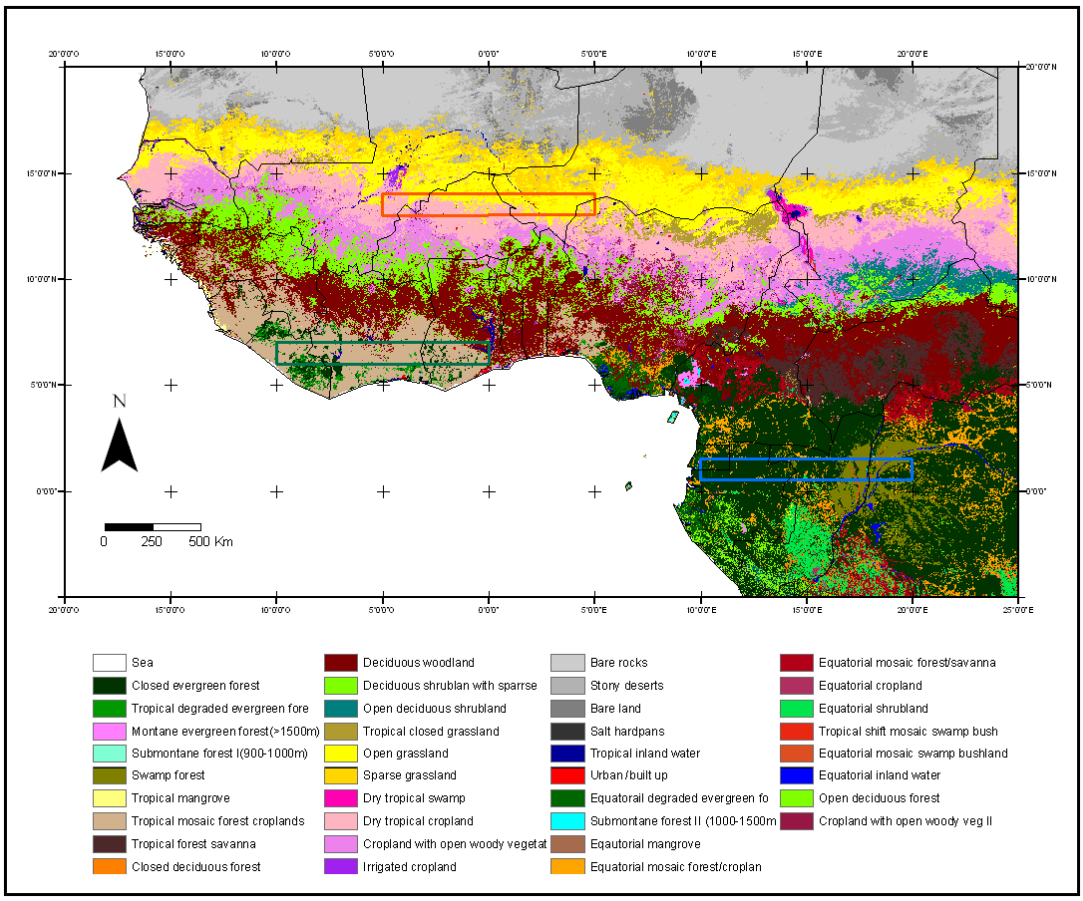

Figure IV.3 Occupation des sols de la nouvelle classification et localisation des 3 transects latitudinaux considérés pour des analyses approfondies : sahélo-soudanien (0.5N-1.5N, 10E-20E) en rouge, guinéen (6N-7N, 10W-0W) en vert et équatorial (13N-14N, 5W-5E) en bleu.

4.4 Distribution spatiale des différences à l'échelle régionale

4.4.1 Différences des classes d'occupation des sols à 0.5°

La figure 4.4 montre que des différences significatives d'occupation des sols existent entre la classification ECOCLIMAP-I et celle utilisée dans cette étude. Puisque la nouvelle classification combine l'information issue des cartes ECOCLIMAP-I et GLC2000 (voir chapitre 2), ces différences sont principalement dûes aux procédés de classification de ces deux produits. En considérant des cellules de 0.5° de résolution spatiale, 19.61% des surfaces continentales de la zone d'étude présentent un pourcentage de différence supérieur à 80% notamment le long d'une bande latitudinale au nord du Sahel s'étendant du Sénégal au Soudan, le long du littoral ouest africain et au Centrafrique.

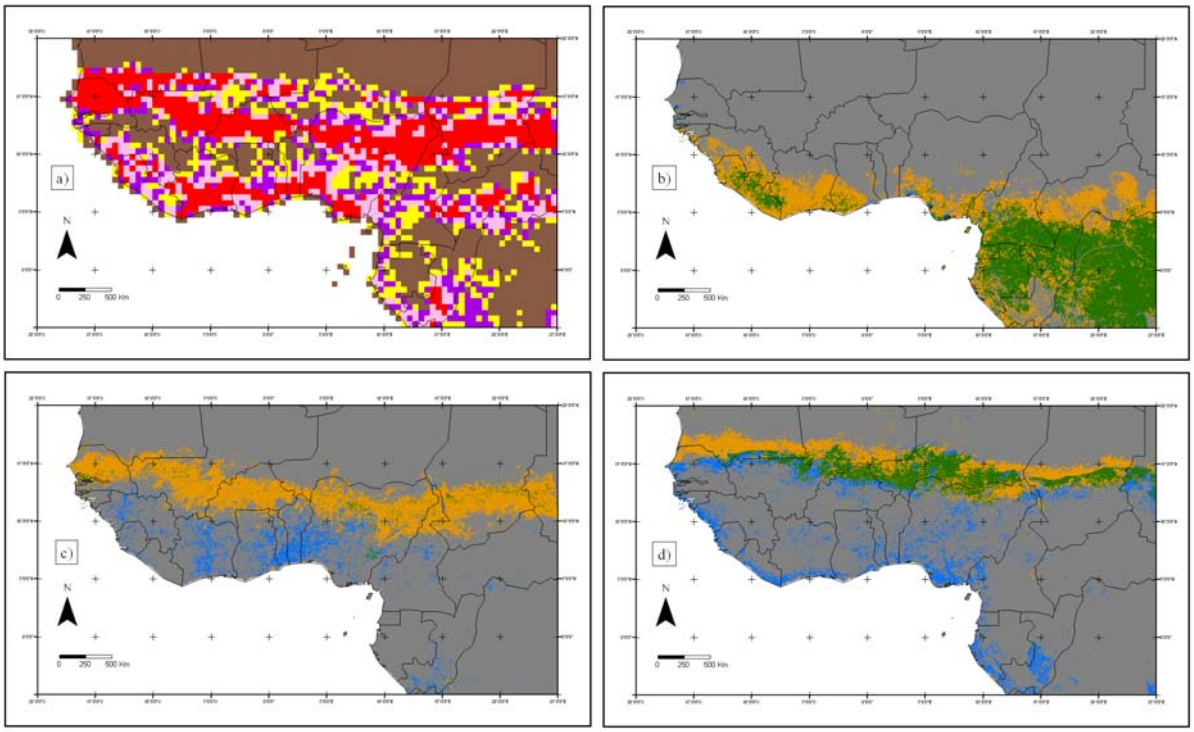

Figure IV.4 Différence d'occupation des sols entre la classification ECOCLIMAP-I et la nouvelle classification pour les principaux écosystèmes : forêt, savane, cultures : a) Pourcentage de changement dans des cellules de 0.5°x0.5° (marron=0-20%, jaune=20-40%, violet=40-80%, rose=60-80%, rouge=80-100%), b) Forêt, c) Cultures, d) savane arbustive/arborée. Légende des couleurs Orange=seulement la nouvelle classification, bleu=seulement ECOCLIMAP-I, vert=identique dans ECOCLIMAP-I et la nouvelle classification

4.4.2 Différences de la fraction de végétation à 0.5°

Les différences de l'occupation des sols dans les deux produits de classification considérés s'accompagnent évidemment des changements des valeurs des paramètres biophysiques. La figure 4.5 a illustre la moyenne du paramètre *veg* pour la période de juin à septembre de 2005 à 2007 issu des deux bases de données utilisées ainsi que la différence entre ces dernières. On y observe clairement 2 gradients latitudinaux : le *veg* EXP est supérieur au CTL d'au moins 0.4 sur le gradient situé au nord du Sahel (de la Mauritanie au Tchad en passant par la frontière entre le Niger et le Mali). Les valeurs EXP sont inférieures de 0.4 sur le gradient situé au centre du Sahel (du Sénégal au Soudan). En plus de ces transects, des biais importants sont observés dans le bassin du Congo et le long de la côte ouest africaine.

4.4.3 Différences de la profondeur racinaire

La figure 4.5 b montre que la profondeur racinaire des simulations expérimentales est en général inférieure à celle des simulations de contrôle de l'ordre de 0.4 m. Cependant, ce biais peut être plus important au Centrafrique, au Congo et au Gabon. Par contre cette tendance est inversée en Sierra Léone, au Libéria et dans le bassin du Congo. Les changements de la profondeur racinaire vont directement influencer l'humidité du sol (comme le montre l'eq. D.8) et le bilan hydrique (voir la section 5 pour plus de discussions).

4.5 Résultats

L'influence du forçage physiographique dans ISBA été étudiée par comparaison entre les simulations expérimentales (EXP) et les simulations dites de contrôle (CTL). Afin d'évaluer de manière générale la qualité des simulations, la figure 4.6 illustre la distribution spatiale de la pluviométrie et celle de la température à 2m durant la période juin à septembre 2005-2007. On y observe un déplacement en direction du nord de la mousson caractérisée par une augmentation de la pluviométrie vers le Nord (avec un pic en juillet et août) et la

diminution des valeurs le long du littoral proche du golfe de Guinée. Les valeurs maximales de la pluviométrie sont atteintes au Sud-Ouest du Cameroun et spatialement corrélées avec les valeurs minimales de la température de l'air.

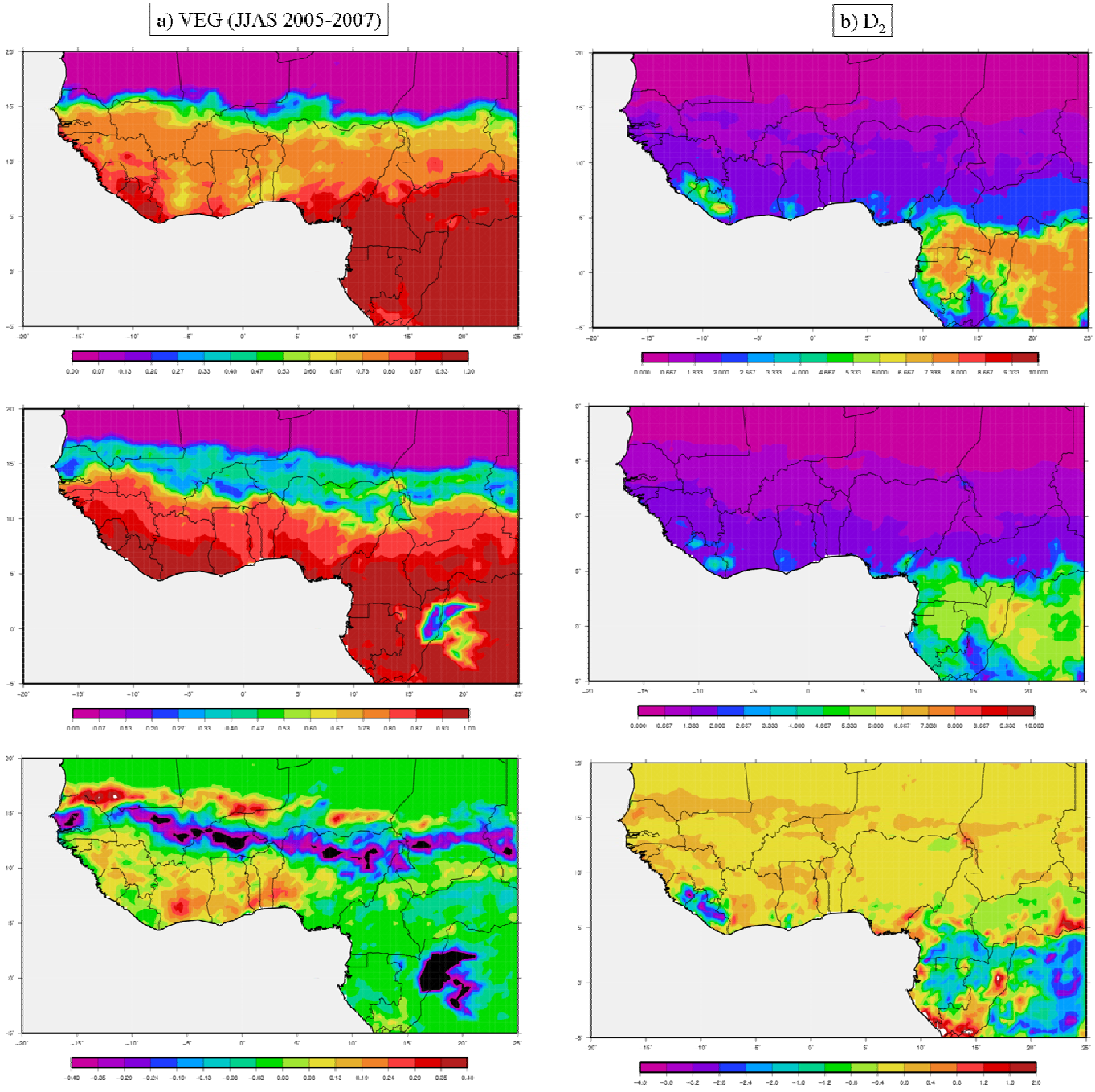

Figure IV.5 Modification de la distribution spatiale du *veg* (moyenné de juin, juillet août, septembre 2005-2007 1ère colonne) et de la profondeur racinaire (2ème colonne) : pour CTL (1ère ligne), EXP (2nde ligne) et EXP – CTL (3e ligne)

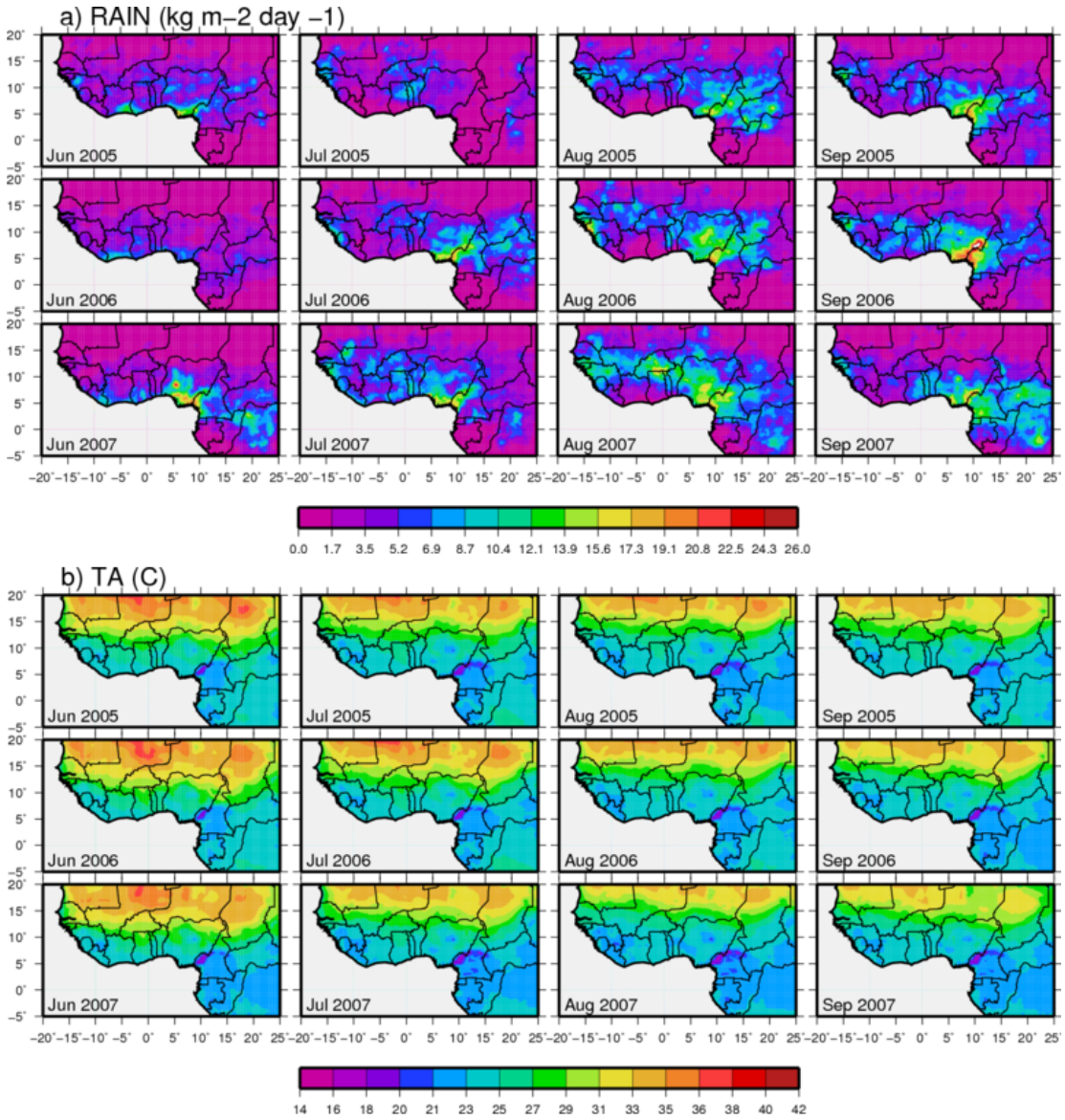

Figure IV.6 Distribution de la pluviométrie (haut) et de la température à 2 m (bas) pour les années 2005 (1ère ligne), 2006 (2nde ligne) et 2007 (3e ligne)

4.5.1 Variations saisonnières à l'échelle régionale

4.5.1.1 Composantes du bilan hydrique

L'une des composantes les plus importantes du bilan hydrique est l'évaporation obtenue selon l' Eq. 4.14. En effet, cette variable constitue un lien critique entre les bilans hydriques et énergétiques. La figure 4.7 montre que l'évaporation EXP est presque identique à l'évaporation CTL. Bien que l'évaporation soit pilotée par la quantité de précipitations reçues, on observe une très forte corrélation spatiale entre la moyenne (de juin, juillet, août septembre 2005-2007) de la distribution spatiale du biais obtenu entre le *veg* et celui de l'évaporation (voir figure 4.5). En effet, sur les régions où l'on note un biais de l'ordre de 1mm/day en valeur absolue pour l'évaporation, un biais de l'ordre de 0.4 en valeur absolue est également observé pour le *veg*, résultant de la conversion de : la savane arbustive/arborée en cultures ou de la forêt dense en savane arborée (figure 4.4). L'humidité du sol (cf. Eq. 4.9) EXP est presque identique à CTL (figure 4.8 a-b). Cependant, on note des zones biaisées au Nord du Mali, dans le bassin du Congo, au Libéria et en Sierra Léone. Le biais observé est spatialement corrélé à celui obtenu avec la profondeur sous racinaire. La décroissance de la profondeur racinaire EXP en Sierra Léone (figure 4.5 b) induit une décroissance de l'humidité du sol (figure 4.8 c)

4.5.1.2 Bilan énergétique

La figure 4.9 représente les valeurs moyennes du flux de chaleur sensible pour les mois de juin, juillet, août et septembre obtenus au cours des années 2005, 2006 et 2007. Le biais du flux de chaleur sensible est en général positif et de l'ordre de 6 à 10 W m^{-2}. Malgré cette différence, on observe de façon nette une corrélation spatiale entre le flux de chaleur latente et le *veg*. On observe également que le paramètre *veg* agit sur la partition de l'énergie solaire disponible en flux de chaleurs latente et sensible.

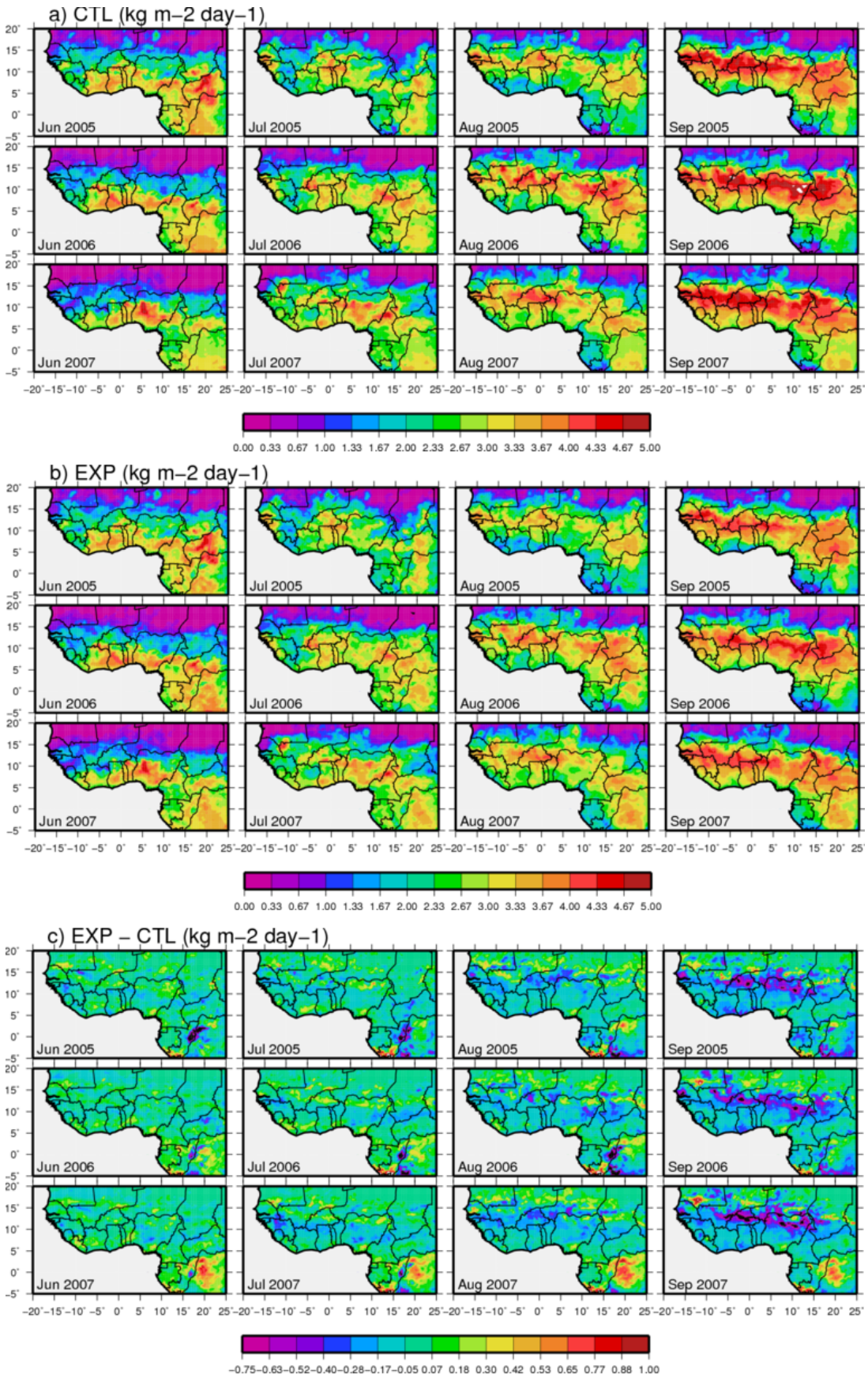

Figure IV.7 Distribution de l'évaporation durant juin à septembre 2005-2007

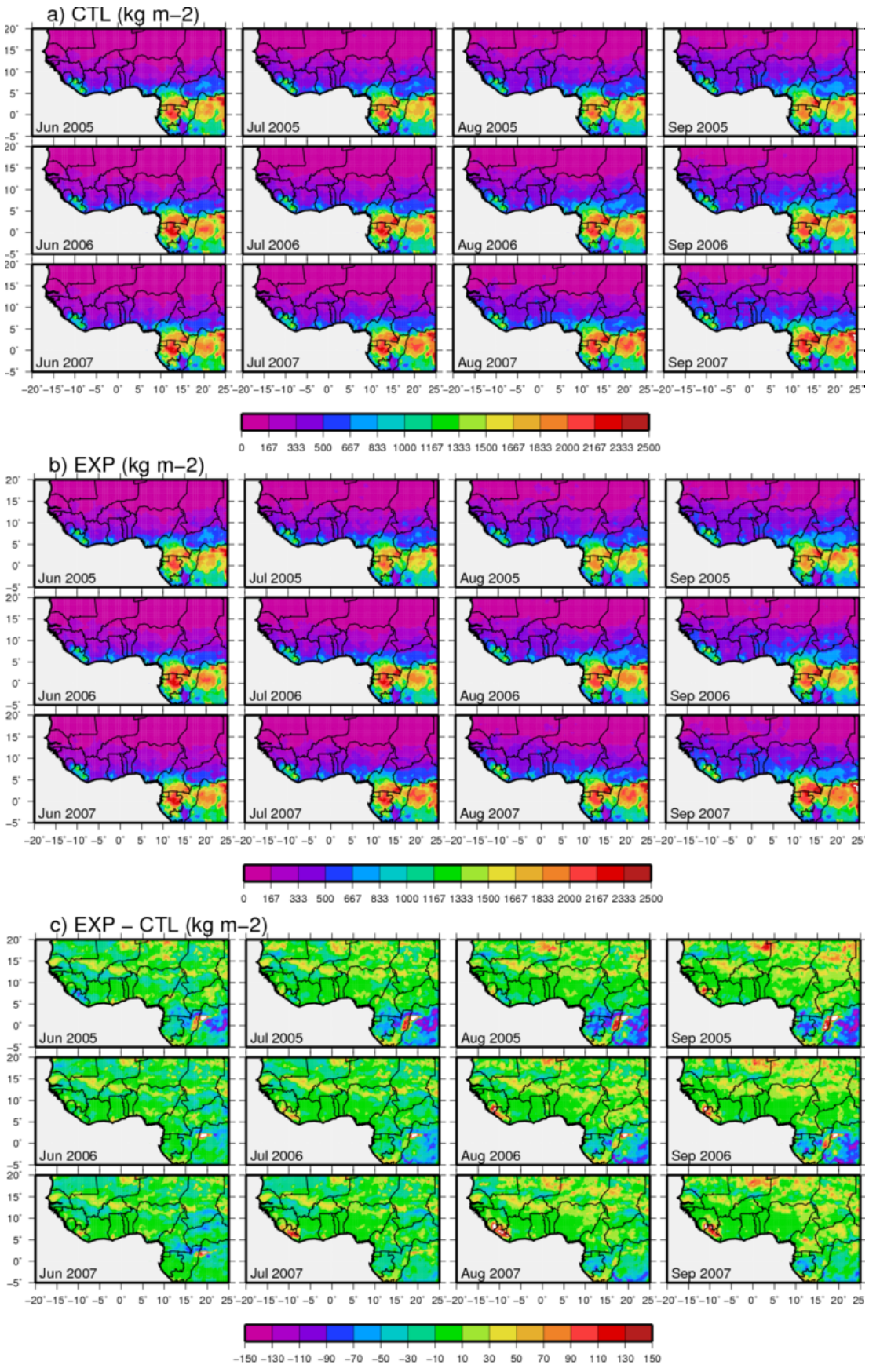

Figure IV.8 Identique à la figure 4.5 pour l'humidité du sol

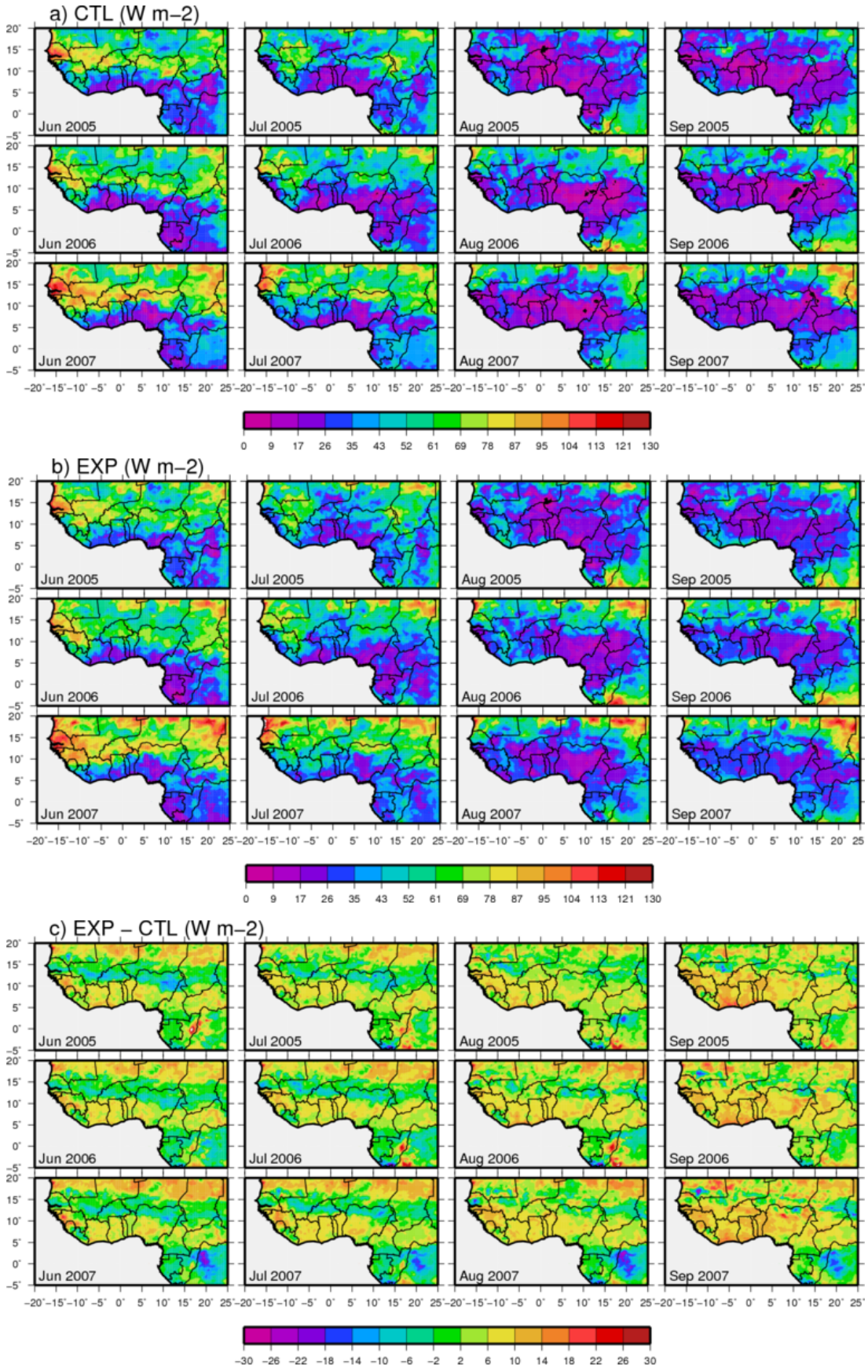

Figure IV.9 Identique à la figure 4.5 pour le flux de chaleur sensible

4.5.2 Variations saisonnières à l'échelle du paysage

Il est important de mentionner que la résolution spatiale de 0.5° ne permet pas de bien illustrer la variabilité interannuelle des paramètres biophysiques de la nouvelle classification à cause de la mixité des classes d'occupation des sols au sein d'une cellule de 0.5° ; d'où la délimitation des 3 transects latitudinaux (figure 4.3) contenus dans des zones écoclimatiques distinctes (Leibovici et al., 2007) et majoritairement constitués d'une seule classe d'occupation des sols (voir figure 4.11).

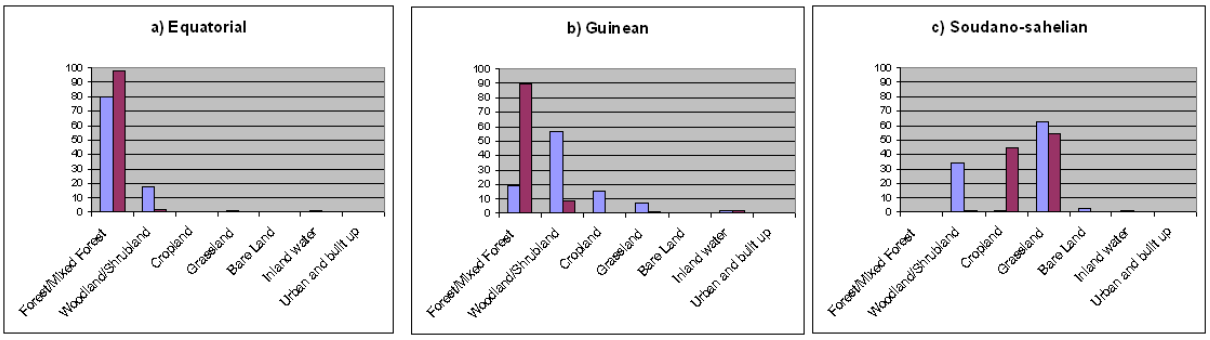

Figure IV.10 Pourcentage des classes agrégées contenues dans les 3 transects délimités à la figure 4.3 de la nouvelle classification (violet) et d' ECOCLIMAP-I (bleu). Se référer à la section 3.3.3.2 pourla définition des classes agrégées

Dans le transect équatorial (0.5° - 2.5°N, 9.5° - 19.5°E), l'ancien LAI est inférieur au nouveau LAI d'environ 1 durant la première saison pluvieuse de l'année (figure 4.12 a). Le cycle saisonnier du LAI d'ECOCLIMAP-I n'est donc pas en phase avec le cycle phénologique attendu de la région, probablement à cause des problèmes de qualité associés aux données AVHRR. En effet, le LAI de ce transect, principalement constitué d'une seule classe (70% de la classe "Closed Evergreen Forest" dans la nouvelle classification contre 79% de la classe "Equatorial African Forest" d'ECOCLIMAP-I), doit être en phase avec le NDVI issu de SPOT/VGT; ce qui est le cas pour le nouveau LAI dérivé de la collection 5 de MODIS comme le montre la figure 2.2. Le flux de chaleur sensible croît d'environ 10 W/m² durant la première saison pluvieuse de l'année. Des résultats similaires s'observent dans le

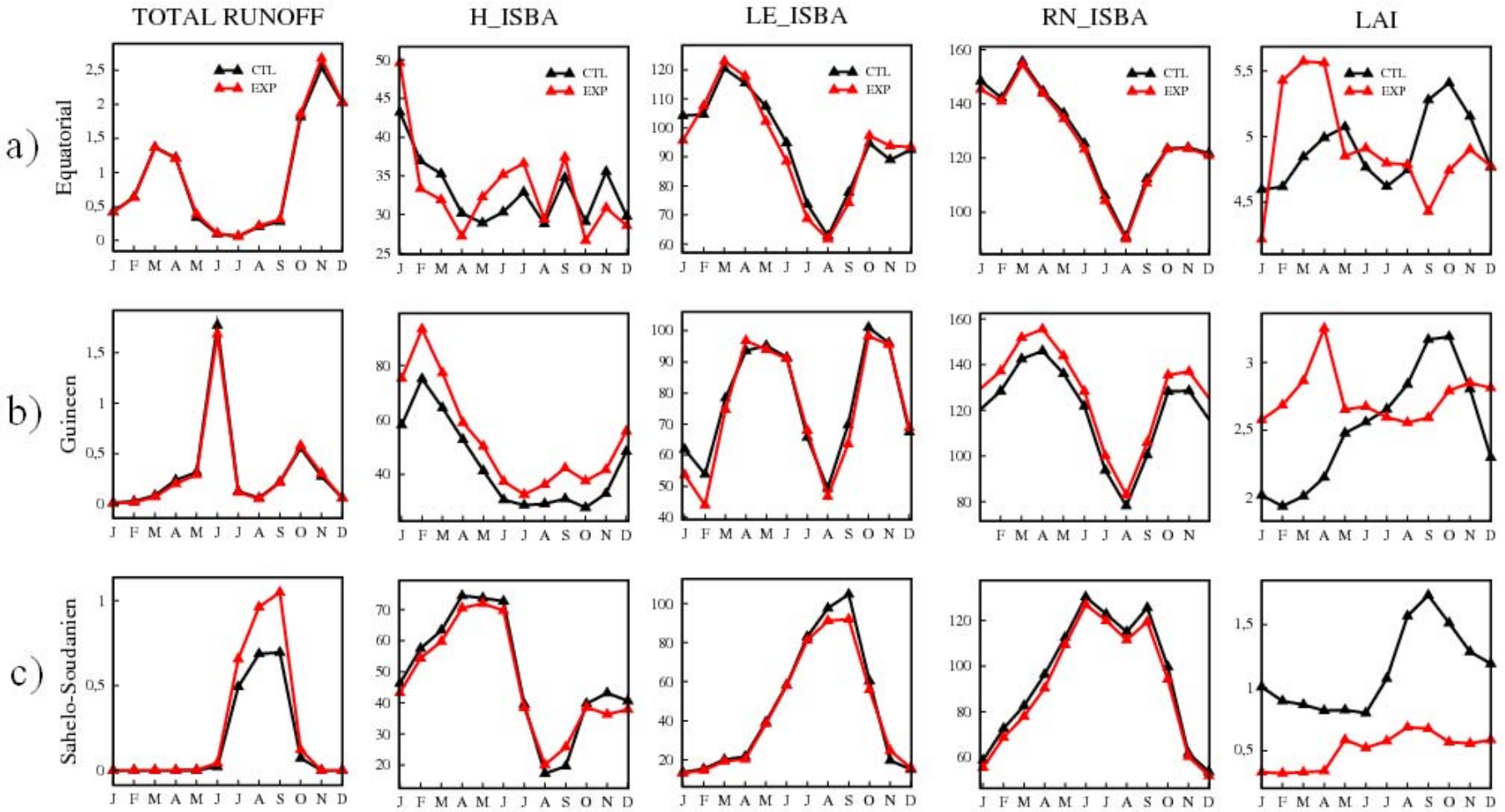

Figure IV.11 Moyennes mensuelles des variables climatiques pour les simulations de contrôle (ECOCLIMAP-I) et expérimentale (nouvelle physiographie) dans les 3 transects de la figure 4.3 où le ruissellement total est exprimé en kg m^{-2} day^{-1}, le rayonement net et les flux de chaleur latente et sensible en W m^{-2}

transect guinéen (15°-17°N, 7.5°W - 2.5°E) équatorial, où le LAI était également sous-estimé le début de l'année dans ECOCLIMAP-I (figure 4.12 b).

Des paramètres de la végétation autres que le LAI peuvent avoir également contribu en partie aux changements observés dans la région sahélo-soudanienne (13° – 15°N, 5°W – 5°E) (figure 4.12 c). Par exemple, quand la classe d'occupation passe d'une mosaïque de savane arborée/arbustive à des cultures (voir figure 4.11.c), les racines des arbres peuvent avoir accès à l'eau contenu dans la zone racinaire inaccessible à celle des herbacés.

4.6 Discussion et conclusion

Ce chapitre a examiné l'amélioration des variables climatiques dans ISBA en utilisant un nouveau forçage physiographique (d'indice foliaire, d'albédo, de fraction de végétation et des types fonctionnel de plantes) dérivé des données issues du capteur de dernière génération MODIS. Les anciennes valeurs d'ECOCLIMAP surestiment le LAI et donc surestiment aussi la fraction de végétation *veg* sur une large bande sahélienne. Les résultats montrent que ces changements modifient de façon substantielle les flux de surface (avec par exemple une diminution de l'ordre de 20 W m^{-2} pour le flux de chaleur latente). Une suite de ce travail serait de vérifier l'amélioration de l'humidité du sol dans ISBA avec les données gravimétriques GRACE (Tapley et al., 2004) ou encore l'amélioration de l'humidité de la couche de surface avec des données AMSR (Pellarin et al., 2009) . Il est important de rappeler qu'au Sahel, l'évolution temporelle de la pluviométrie et des ressources en eau n'est pas toujours correlée (e.g. Mahé et al., 2005 ; Gerbaux et al., 2009) Ce processus connu sous le nom de 'paradoxe sahélien' s'explique aujourd'hui par les changements liés à l'utilisation et à l'occupation des sols.

Le nouveau forçage physiographique réduit considérablement l'évaporation et augmente l'humidité du sol précisément au Nord du Sahel. L'amélioration de la qualité des données de LAI et d'albédo dérivées de MODIS comparée à celle de l'AVHRR permettent de mettre à jour ces deux variables de sorte qu'elles se rapprochent d'avantage de la réalité. On pourrait

également vérifier à l'aide des données d'albédo MSG la façon avec laquelle l'usage du nouveau forçage physiographique améliore l'absorption et la partition de l'énergie entre la canopée et le sol et, en procédant ainsi, des humidités superficielles plus réalistes dans ISBA devraient être produites. Il est mis en évidence le besoin de disposer d'une estimation la plus précise possible des variables biophysiques, en particulier le LAI, l'albédo, le *veg*. Ayant effectué des tests de sensibilité avec ISBA en mode forcé, une suite de ce travail serait d'effectuer les simulations en mode couplé en utilisant la technique dite de « *nudging* » (). En effet, le fait d'imposer un forçage atmosphérique à 2 m statique sans possible rétroaction de la surface sur l'atmosphère peut potentiellement atténuer la portée de l'impact d'une nouvelle physiographie.

« Tant que les lions n’auront pas leurs propres historiens, les histoires de chasse continueront de glorifier le chasseur », dit le proverbe africain

Conclusion

Cette étude s'est focalisée sur la classification de l'occupation des sols et sur la restitution des variables biophysiques afin de pouvoir caractériser l'hétérogénéité du continent africain par traitement des séries d'images pluriannuelles de données satellitales. fournies par des capteurs à basse résolution. Le sujet de cette thèse, les données traitées, les modèles utilisés et les thématiques abordées sont au cœur des réflexions développées actuellement dans la région afin d'améliorer nos capacités de modélisation du cycle hydrologique. Une originalité essentielle de ce travail réside dans le caractère fortement fédérateur des travaux entre deux disciplines, la télédétection et la météorologie, et présente un intérêt réel pour ces deux communautés. Nous avons contribué à l'utilisation des systèmes d'observation de la terre à des fins scientifiques touchant le fonctionnment des surfaces continentaales, et aussi pour des fins thématiques de surveillance et de gestion de l'environnement, dans un contexte de changement climatique et anthropique fort.

Deux techniques de classification des écosystèmes ont été explorées. Ayant noté l'existence d'importantes disparités entre la classification ECOCLIMAP-I et la carte GLC2000 (pouvant être considéres comme des informations complémentaires), la première approche de classification essaye de détecter les écosytèmes présents au sein des classes de GLC2000 en les fusionnant aux classes d'ECOCLIMAP-I par une analyse interactive du LAI MODIS sur la région ouest africaine. A contrario, des procédés existants de fusion des cartes d'occupation des sols, la méthode ainsi présentée ne nécessite pas l'usage de données supplémentaires spécifiques à certaines régions comme l'incorporation des données d'inventaire agricoles lors de la réalisation de la carte de distribution des cultures SAGE (Ramankutty et al., 2008) ou alors l'incorporation des classifications nationales d'occupation des sols lors de la réalisation d'ECOCLIMAP-I (Masson et al., 2003).

Dans la seconde approche, nous avons developpé un processus automatique et hybride de classification d'occupation des sols par utilisation de la transformée de Fourier discète et du classificateur dynamique k-NN. La méthode permet de s'affranchir de : (i) l'usage des cartes climatiques comme lors des classifications existantes (Arino et al., 2008) par la définition d'un formalisme pour la stratification du continent en écorégions, (ii) des difficultés liées à la labellisation des classes comme ce fut le cas dans la précédente approche. En outre, le caractère non supervisée de cette approche de classification lui confère le mérite d'être généralisé à l'échelle globale sans grandes difficultés en comparaison alors que l'utilisation de la précédente approche de classication reste limitée au regard du nombre exponentiel de classes que l'on pourrait être amené à analyser (éventalité facilement réalisable lorsque l'information fournie par les cartes en entrée est très disparate). La méthode a ensuite été appliquée aux données NDVI de SPOT/VEGETATION pour produire la classification ECOCLIMAP-II à l'échelle continentale africaine. Par utilisation d'une longue série pluri-annuelle d'observation, les résultats de la méthode hybride ne sont pas contraints par des évènements annuels comme cela peut être le cas dans les cartes continentales existantes telles que GLC2000, GLOBCOVER ou MODIS LC-I.

La classification ECOCLIMAP-II a été comparée avec les cartes continentales d'occupation des sols existantes produites à partir des données acquises durant les années 2000. Après avoir surligné, les différences et similarités des quatre produis, une carte d'accord spatial a été réalisée par analyse spatiale des informations issues de GLC2000, GLOBCOVER, MODIS LC-I, et ECOCLIMAP-II. Les fortes disparités rencontrées dans les modes de caractérisation et de réprésentation des paysages confirment le besoin d'améliorations nécessaires. Au cours de cette comparaison, nous avons également proposé une solution au problème du choix de la classe majoritaire en cas de non-unicité lors du sous-échantillonnage des classifications. L'implémentation de cet algorithme dans les systèmes d'informations géographiques (SIG) devrait combler les lacunes observées lors du rééchantillonnage des variables discrètes. Cette comparaison n'a été possible qu'après

uniformisation des nomenclatures des différentes classifications en sept classes agrégées de niveau I de la nomenclature LCCS (Di Gregorio et Jansen, 2000). La non-prise en compte du dégré d'appartenance d'une classse originale au sein d'une classe agrégée (Fritz et See, 2008) peut induire un faible accord à cause des erreurs de classification. En outre, des erreurs de location (mauvais géoréférencement) vont induire des erreurs de classification lors de la superposition des différentes classifications.

Des techniques statistiques de restitution des paramètres biophysiques tels que l'albédo de surface, l'indice foliaire et la fraction de végétation, utiles dans les modèles de surface ont été implémentées. Le processus de « *dévégétalisation* » de l'albédo de surface a permis de discriminer l'albédo du sol nu de celui de l'albédo de la végétation tel que requis dans les modèles de surface en utilisant les données mesurées par le capteur MODIS. La méthode s'avère assez performante et moins consommatrice en temps de calcul que les techniques de transfert radiatifs (e.g. Alton, 2009). La technique de dévégétalisation de l'albédo a d'abord été testée sur la région ouest-africaine et sa robustesse a été démontrée lors de son applicabilité sur l'intégralité du continent africain. La base de données ECOCLIMAP-II qui en dérive bénéficie alors : (i) d'une meilleure identification de l'occupation des sols dérivée de l'approche hybride de classification basée sur l'utilsation de la transformée de Fourier, (ii) des valeurs récentes d'albédo et de LAI incluant la variabilité inter-annuelle et (iii) d'une consistance entre les paramètres biophysique de surface dérivés des données récentes de télédétection qui possèdent une précision améliorée tant sur le plan temporel, spectral que spatial en comparaison avec les données AVHRR de la base de données ECOCLIMAP-I.

Le forcage physiographique, dévéloppé sur la région ouest africaine, a été implémenté dans ISBA pour simuler les flux hydriques et énergétiques. Le forçage atmosphérique utilisé lors de ces simulations est le même que celui de la 3^{e} expérience du projet projet ALMIP d'intercomparaison des modèles de surface sur la région ouest-africaine. Ce qui a permis de mener une étude de sensibilité sur l'impact du changement d'occupation des sols. Les

résultats montrent que la fraction de végétation (et par conséquent l'indice folaire à partir duquel elle est calculée) pilote les flux de chaleur latente et sensible. Le modèle ISBA peut donc être utilisé pour prédire l'impact du changement de l'occupation des sols et par conséquent des actions anthropiques sur le bilan hydrique. Il serait intéressant d'examiner la température à 2 m, l'humidité spécifique de l'air ainsi que la quantité d'eau précipitable en effectuant des simulations sur l'ensemble du continent africain en couplant le modèle de surface ISBA à un modèle régional climatique à l'instar d'ALADIN. Ces simulations pourraient être effectuées dans le cadre du projet CORDEX d'intercomparaison des modèles régionaux climatiques (http://www.meteo.unican.es/en/projects/CORDEX). Ceci pourrait permettre de mieux comprendre la sensibilité de ces variables à l'hétérogénéité de surface, d'ajuster de façon plus précise certaines de ces variables et par conséquent accroître la précision des simulations.

Au-delà de la caractérisation de l'hétérogénéité de surface du continent africain, principale motivation de cette étude, différentes perspectives apparaissent pour enrichir la réflexion méthodologique menée jusqu'ici. L'une des principales limitations de notre étude découle de la rareté des meures in situ et donc l'impossibilité de vérifier si l'utilisation des varaibles biophysques dérivées du capteur MODIS de dernière génération améliorait la restitution des flux hydriques et énergétiques. De ce fait, il ne nous a pas également été possible de valider de façon directe la nouvelle caractérisation de l'hétérogénéité de surface de la base de données ECOCLIMAP-II (occupation des sols, variables biophysiques). Une perspective de ce travail serait d'utiliser les données LANDSAT récemment mises à la disposition de la communauté scientifique en libre accès (Roy et al., 2010) afin de mettre à jour la base de données AFRICOVER et ensuite de l'étendre sur l'intégralité du continent africain en utilisant un algorithme de classification plus approprié pour lequel il n'existerait plus une inconsistance à cause de l'hétérogénéité d'interprétation d'une scène LANDSAT à l'autre. La nouvelle carte d'occupation des sols à 30 m de résolution ainsi produite servirait de référence pour l'estimation de l'exactitude spatiale et la détermination des sources

d'erreur géométriques des quatre cartes d'occupation des sols GLC2000, GLOBCOVER, MODIS LC-I et ECOCLIMAP-II et également à l'indication des cartes appropriées pour des applications spécifiques.

Un nouvel indice, (l'indice de segmentation des composantes de Fourier, ISFC) de segmentation des surfaces continentales en écorégions à n'importe quelle échelle (locale, nationale, régionale, continentale ou globale) indépendamment de la résolution (basse, moyenne ou haute) et du domaine de la télédétection (optique ou radar) a été présenté. Cependant le phénomène d'ennuagement permanent, notamment au niveau du littoral ouest-africain, induit un mauvais découpage des écorégions. Afin de contourner ce problème, il serait intéressant de procéder à une fusion au niveau du pixel (e.g. l'analyse en composantes principales) des données optiques SPOT/VEGETATION à des données micro-ondes puisqu'elles sont quasi-transparentes pour l'atmosphère afin d'obtenir une meilleure délimitation des écorégions déduites de l'ISFC. Ceci pourrait s'effectuer en utilisant les données issues du capteur ENVISAR/ASAR qui semble le plus approprié car il possède plusieurs modes de polarisation et une résolution kilométrique en mode global.

Nous pensons que plusieurs pistes relatives à l'utilisation de l'ISFC restent encore à explorer, notamment pour caractériser l'évolution spatio-temporelle de la biodiversité compte tenu du fait que ses quatres composantes majeures à savoir la topographie, la productivité de la végégation, la fragmentation de l'occupation des sols et la pertubation (Duro et al., 2007) sont des caractéristiques de l'hétérogénéité de surface. L'application de l'ISFC à des séries temporelles en écartant une année spécifique peut être utilisée pour analyser l'année en question par comparaison avec une tendance climatologique construite en utilisant toutes les années. Parce que l'ISFC se fonde sur l'hypothèse seolon laquelle les series temporelles présentent une certaine périodicité et donc ont une durée infinie, il ne peut pas révéler la variabilité inetrannuelle nécessaire pour diagnostiquer une altération ou une pertubance de l'hétrogénéité de surface. A cet effet, l'SFC pourrait être combiné à l'usage des ondelettes

qui semble mieux permettre une périodicité irrégulière afin d'estimer la pluviométrie (Quiroz et al. 2010), la phénologie (Martinez et Gilabert, 2009) ou le changement d'occupation des sols (). Il serait également intéressant de combiner l'ISFC à d'autres indices tels que le dynamic habitat Index (DHI) (Coops et al., 2009) afin d'étudier le rôle conditions du sol sur l'habitat à diverses échelles ou alors de construire un démonstrateur illustrant l'impact de la fragmentation des forêts sur attributs africain de la biodiversité. Il est important de appeler que le DHI est une mesure tripartite de la productivité de la végétation pour surveiller les conditions d'habitat de façon répétitive sur de grands domaines.

Bien que la reproductibilité de la *dévégétalisation* sur n'importe quel domaine ait été établie, il serait intéressant d'étudier son usage avec d'autres données de télédétection pour lesquelles des produits composites d'albédo et de LAI sont disponibles, notamment les données du SAFLAND. L'usage de la dévégétalisation reste limité en cas d'inconsistance entre les données d'albédo et d'indice foliaire utilisées en entrée. Il est donc recommandé aux développeurs des produits d'incorporer cet aspect dans les prochaines versions de leurs données. A ce propos, la mise à jour des produits albédos des composantes du sol nu et de la végétation établis séparamment sera réalisée dans le cadre du SAF LAND à l'aide des données MSG/SEVIRI et à partir d'une méthode qui restera consistente avec la restitution des paramètres de la végétation.

Bibliographie

Alton, P. (2009), A simple retrieval of ground albedo and vegetation absortance from MODIS satellite data for parameterisation of global Land Surface Models, *Agricultural and Forest Meteorology*, **149**, 1769-1775.

Anderson, J. R., Hardy, E. E., Roach, J. T., Winner, R. E., 1976. *A land use and land cover classification system for use with remote sensor data*. US Geological Survey Professional Paper, N° 964. USGS, Washington, D. C.

Andreae, M. O. (1995), Climatic effects of changing atmospheric aerosols levels, in *World survey of climatology,* **16***, Futures climates of the world*, Henderson-Sellers A., Ed., Elsevier, Amsterdam, pp 341-392.

Andres, L., Salas, W. A. and Skole, D. (1994), Fourier analysis of multitemporal AVHRR data applied to a land cover classification, *International Journal of Remote Sensing*, **15**, 1115-1121.

Andrieu, J. and Mering, C. (2008), Cartographie par télédétection des changements de la couverture végétale sur la bande littorale ouest-africaine: exemple des rivières du sud et du delta du Saloum (Sénégal) au Rio Geba (Guinée-Bissau), *Télédétection*, **8**, 93-118.

Anyamba, A., Tucker, C. J. and Mahoney, R. (2002), From El Niño to La Niña: vegetation responses patterns over east and southern Africa during the 1997-2000 period, *Journal of Climate*, **15**, 3096-3103.

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R. and Weber, J.-L. (2008), GLOBCOVER: The most detailed portrait of the earth, *E.S.A bulletin,* **136**, 24-31.

Avissar, R. and Pielke, R. A. (1989), A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology, *Monthly Weather Review*, **117**, 2113–2136.

Azzali, S. and Menenti, M. (1999), Mapping isogrowth zones on continental scale using temporal fourier analysis of AVHRR-NDVI data, *International Journal of Applied Earth and Geoinformation*, **1**, 9-19.

Bacour, C, Baret, F., Beal, D., Weiss, M. And Pavageau, K. (2006), Neural network estimation of LAI, fAPAR, fCover and LAI×C_{ab}, from top of canopy MERIS reflectance data: Principles and validation, *Remote Sensing of Environment*, **105**, 313-325.

Baret, F. and Buis, S. (2008), Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems, in *Advances in land remote Sensing*, S. Liang, Ed., Springer, pp. 173-202.

Baret, F. and Guyot, G. (1991), Potential limits of vegetation indices for LAI and APAR assessment, *Remote Sensing of Environment*, **35**, 161−173.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M. et al. (2007), LAI, fPAR and fCover CYCLOPES global products derived from VEGETATION Part1: Principles of the algorithm, *Remote Sensing of Environment*, **110**, 275-286.

Barlage, M. and Zeng, X. (2004), The effects of observed fractional vegetation cover on the land surface climatology of the community land model, *Journal of Hydrometeorology*, **5**, 823-830.

Bartholomé, E. and Belward, S. (2005), GLC2000: a new approach to global land cover mapping from Earth observation data, *International Journal of Remote Sensing*, **26**, 1959-1977.

Benjaminsen, T. A. (2009), Does supply-induced scarcity drive violent conflicts in the African Sahel? The case of the Tuareg rebellion in Northen Mali, *Journal of Peace Research*, **45**, 819-836.

Berberoglu, S., Llyod, C. D., Atkinson, P. M. and Curran, P. J. (2000), The integration of spectral and textural information using neural network for land cover mapping in the Mediterranean, *Computers & Geosciences*, **26**, 385-396.

Bicheron, P., Defourny, P., Brockman, C., Schouten, L., Vancutsem, C., Huc, M. et al. (2008), *GLOBCOVER: product description and validation report*, See http://postel.mediasfrance.org/en/DOWNLOAD/Documents/#globcover (Last accessed May 2009).

Boko, M., Niang, I., Nyong, A., Vogel, C., Githeko, A., Medany, M., Osman-Elasha, B., Tabo, R. and Yanda, P. (2007), Africa in *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, M. L. Parry et al., Eds., Cambridge University Press, Cambridge, pp. 433-467.

Bonan, G. (2002), *Ecological climatology: concepts and application*, University Press, Cambridge, 678 pp.

Bonan, G. B., Levis, S., Kergoat, L. and Oleson, K. W. (2002), Landscape as patches of plant functional types: An integrating concept for climate and ecosystems models, *Global Biogeochemical Cycles*, **16**, 1021.

Bonan, G. B., Levis, S., Sitch, S. Vertenstein, M. and Oleson, K. W. (2003), A dynamic global vegetation model for use in climate models: concepts and description of simulated vegetation dynamics, *Global Change Biology*, **9**, 1543-1566.

Boone, A., Calvet, J.-C. and Noilhan, J. (1999), Inclusion of a third soil layer in a land-surface scheme using the force-restore method, *Journal of Applied Meteorology*, **38**, 1611-1630.

Boone, A., De Rosnay, P., Balsamo, G., Beljaars, A., Chopin, F., Decharme, B. et al. (2009) The AMMA land surface intercomparison project (ALMIP), *Bulletin of the American Meteorological Society*, **90**, 1865−1880.

Boudraa, A.-O. (1999), Dynamic estimation of number of clusters in data sets, *Electronic Letters*, **35**, 1606-1608.

Buermann, W., Dong, J., Zeng, X., Myneni, R. B. and Dickinson, R. E. (2001), Evaluation of the utility of satellite-based vegetation leaf area index data for climate simulations, *Journal of Climate*, **14**, 3536–3551.

Burke, M. B., Miguel, E., Satyanath, S., Dykema, J. A. and Lobell, D. B. (2009), Warming increases the risk of civil war in Africa, *Proceedings of the National Academy of Sciences*, **106**, 20670-20674.

Callendar, G. S. (1938), The artificial production of carbon dioxide and its influence on temperature, *Quarterly Journal of Royal Meteorology*, **64**, 223-240.

Carrão, H., Gonçalves, P., and Caetano M. (2007), Contribution of multispectral and multitemporal information from MODIS images to land cover classification, *Remote Sensing of Environment*, **112**, 986-997.

Carrer, D., Roujean, J.-L and Meurey, C. (2010), Comparing Operational MSG/SEVIRI Land Surface Albedo Products From Land SAF With Ground Measurements and MODIS, *IEEE Transactions on Geoscience and Remote Sensing*, **48**, 1714-1728.

Cihlar, J. (2000), Land cover mapping of large areas from satellites: status and research priorities, *International Journal of Remote Sensing*, **21**, 1093-1114.

Champeaux, J.-L., Masson, V. and Chauvin, F. (2005), ECOCLIMAP: a global database of land surface parameters at 1 km resolution, *Meteorological Applications*, **12**, 29-32.

Charney, J. G. (1975), Dynamics of deserts and drought in the Sahel, *Quarterly Journal of the Royal Meteoroligal Society*, **101,** 193–202.

Cohen, J. (1960), A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, **20**, 37–46.

Congalton, R. G. and Green, K. (2009), *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd edition, Boca Raton, FL: Lewis, 208 pp.

Coops, N. C., Wulder M. and Iwanicka, D. (2009), Demonstration of a satellite-based index to monitor habitat at continental-scales, *Ecological indicators*, **9**, 948-958.

Cox, P., Betts, R., Bunton, C., Essery, R., Rowntree, P. and Smith, J. (1999), The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, *Climate Dynamics*, **15**, 183-203.

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., et al. (2003), The Common Land Model, *Bulletin of the American Meteorological Society*, **84**, 1013-1023.

Danko, D. M. (1992), The Digital Chart of the World, *GeoInfo Systems*, **2**, 29–36.

Darwin, C. R. (1859), *On the origin of species by means of natural selection*, *or the preservation of favoured races in the struggle for life*, John Murray, London, 502 pp.

Dawson, T. P., North, P. R. J., Plummer S. E. and Curran, P. J. (2003), Forest ecosystem chlorophyll content: implications for remotely sensed estimates of net primary productivity, *International Journal of Remote Sensing*, **24**, 611-617.

De Jong, S. M., Hornstra, T. and Maas, H.-G. (2001), An integrated spatial and spectral approach to the classification of Mediterranean land cover types: the SSC method, *International Journal of Applied earth Observation and Geoinformation*, **3**, 176-183.

De Jong, S. M., Van Der Meer, F. D. and Clevers, J. G. P. W. (2004), Basic of remote sensing, in *Remote sensing images analysis: including the spatial domain*, S. M. De Jong and F. D. Van Der Meer, Eds.,Springer, Dordrecht, 1-11.

De Wit, M. and Stankiewicz, J. (2006), Changes in water supply across Africa with predicted climate change, *Science*, **311**, 1917-1921.

Deardorff, J. W. (1977), A parameterization of ground surface moisture content for use in atmospheric prediction models, *Journal of Applied Meteorology*, *16*, 1182-1185.

Deckers, J., Nachtergaele, F. and Spaargaren, O. (2003), Tropical soils in the classification systems of USDA. *Evolution of tropical soil science: Past and future*, G. Stoop, Ed., Food and Agricultural Organisation, pp. 79-94.

Defries, R. S., Hansen, M. and Townshend, J. R. G. (1995), Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data, *Remote Sensing of Environment*, **54**, 209-222.

Di Gregorio, A. and Jansen, L. (2000), *Land cover classifications system (LCCS), classification concepts and user manual*, Food and Agriculture Organisation of the United Nations, Rome, Italy, 179 pp.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z. Oki, T. and Hanasaki, N. (2006), GSWP-2: Multimodel analysis and implications for our perception of the land surface, *Bulletin of the American Meteorological Society*, **87,** 1381–1397.

DMA (1986), *Defense Mapping Agency product specifications for digital terrain elevation data (DTED)*, 2nd edition, Defense Mapping Agency Aerospace Center, St. Louis, Missouri, 26 pp.

Douville, H. (2006) Influence des surfaces continentales sur la variabilité hydrologique des échelles inter-annuelles à multi-décennale. Thèse de HDR, Institut National Polytechnique de Toulouse, France, 95 pp.

Dubes, R. C. (1987), How many clusters are best?-An experiment, *Pattern Recognition*, **20**, 645-643.

Duchemin, D., Berthelot, B., Dedieu, G., Leroy, M. and Maisongrande, P. (2002), Normalisation of directional effects in 10-day global syntheses derived from VEGETATION/SPOT: II. Validation of an operational method on actual data sets, *Remote Sensing of Environment***, 81**, 101-113.

Duro, D. C., Coops, N. C., Wulder, M. A. and Han, T. (2007), Development of a large area biodiversity monitoring system driven by remote sensing, *Progress in Physical Geography*, **31**,1-26.

EC (1993), *CORINE land cover: Technical guide. Environment, nuclear safety and civil protection series*, Commission of the European Communities, Office for Official Publication of the European Communities, Luxembourg, 144 pp.

EC (1995), *Regionalization and stratification of European forest ecosystems*, Internal special publication of the Joint Research Centre of the European Commission, S.I.P.95.44, EC, SAI, EMAP, Italy, 69 pp.

Eswaran, H., Almaraz, R., Van Den Berg, E. and Reich, P. (1996), *An Assessment of the soil Resources of Africa in Relation to Productivity*. See http://soils.usda.gov/use/worldsoils/papers/africa1.html (Last accessed March 2010).

Everitt, B. S. (1979), Unresolved problems in clusters analysis, *Biometrics*, **35**, 169-181.

FAO (2004), *The Africover Initiative, Food and Agricultural Organisation*, See www.africover.org/africover_initiative.htm (Last accessed December 2009).

FAO (2006), *Global forest resources assessment 2005. Progress towards sustainable forest management*. FAO forestry paper 147. Rome.

FAO (2008), *Climate change and food security: a framework document*, Food and Agricultural Organisation of the United Nations, Rome, Italy.

FAO/IIASA (2009) *Harmonized World Soil Database (version 1.1)*. FAO, Rome, Italy and IIASA, Laxenburg, Austria

Fernandes, R. A., Butson, C., Leblanc, S. and Latifovic, R. (2003), A Landsat-5 TM and 12 Landsat-7 ETM+ based accuracy assessment of leaf area index products for Canada 13 derived from SPOT4/VGT data, *Canadian Journal of Remote Sensing,* **29**, 241-258.

Fensholt, R., Sandholt, I. and Rasmussen, M. S. (2004), Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ-measurements, *Remote Sensing of Environment*, **91**, 490-507.

Fischer, G., Van Velthuizen, H., Sha M. and Nachtergaele F. (2002), *Global agro-ecological assessment for agriculture in the 21st century: Methodology and results*, International Institute for Applied Systems Analysis, Laxenburg, Austria, Food and Agricultural Organization, Rome, Italy, CD-ROM.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X.Y., Muchoney, D., Strahler, A. H., Nickeson, J. E. et al. (2002), Global land cover mapping from MODIS: algorithms and early results, *Remote Sensing of Environment*, **83,** 287-302.

Friedl, M. A., Sulla-Menashe, D, Tan, B., Schneider, A., Ramankutty, N., Sibley, A. and Huang, X. (2010), MODIS Collection 5 global land cover: algorithms refinements and characterization of new datasets, *Remote Sensing of Environment*, **114,** 168-182.

Fritz, S. and See, L. (2008), Identifying and quantifying uncertainty and spatial disagreement in the comparions of global land cover for dofferent applications, *Global Change Biology*, **14**, 1057-1075.

Gabban, A., San-Miguel-Ayanz, J., Barbosa, P. and Liberta, G. (2006), Analysis of NOAA-AVHRR NDVI inter-annual variability for forest fire risk estimation, *International Journal of Remote Sensing*, **27**, 1725–1732.

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M. et al., (2008), Validation and intercomparison of global Leaf Area Index products derived from remote sensing data, *Journal of Geophysical Research*, **113**, G02028.

Gason, F., Gastellu-Etchegorry, J.-P., Lefevre-Fonollosa, M.-J. and Dufrene, E. (2004), Retrieval of forest biophysical variables by inverting a 3-D radiative transfer model and using high and very high resolution imagery, *International Journal of Remote Sensing*, **25**, 5601-5616.

Geiger, B., Meurey, C., Lajas, D., Franchistéguy, L. Carrer, D. and Roujean, J.-L. (2008), Near real-time provision of downwelling shortwave radiation estimates derived from satellite observations, *Meteorological Applications*, **15**, 411-420.

Gerbaux, M., Hall, N., Dessay, N. and Zin, I. (2009), The sensitivity of Sahelian runoff to climate change, *Hydrological Sciences Journal,* **54**, 5-16.

Gesch, D. B. and Larson, K. S. (1996), *Techniques for development of global 1-kilometer digital elevation models*, Pecora Thirteen, Human Interactions with the environment-Perspectives from space, Sioux falls, South Dakota, August 20-22, 1996.

Ginoux, P., Garbuzov, D. and Hsu, N. C. (2010), Identification of anthropogenic and natural dust sources using Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue level 2 data, Journal of Geophysical Research, **115**, D05204.

Githeko, A. K. and Ndegwa, W. (2001), Predicting malaria in the east African highlands: a tool for decision maker, *Global Change & Human Health*, **2**, 54-63.

Gómez, J. M., Valladares, F and Puerta-Piñero, C. (2004), Differences between structural and functional environmental heterogeneity caused by seed dispersal, *Functional Ecology*, **18**, 787–792.

Goudie, A. S. (1996), Climate: Past and Present, in *The physical geography of Africa*, W. Adams et al., Eds., Oxford University Press, Oxford, pp. 34-59.

Goudie, A. S. and Cuff, D. J. (2002), *Encyclopedia of Global Change: Environmental change and human society*, Oxford University Press.

Goutorbe, J. P., Lebel, T., Dolman, A. J., Gash, J. H. C., Kabat, M., Kerr, Y. H. et al. (1997), An overview of Hapex-Sahel: a study in climate and desertification, *Journal of Hydrology*, **188-189**, 4-17.

Gutman, G. (1999), On the use of long-term global data of land reflectances and vegetation indices derived from the advanced very high resolution radiometer, *Journal of Geophysical Research*, **104**, 6241-6255.

Hagolle, O., Lobo, A., Maisongrande, P., Cabot, F., Duchemin, B. and De Pereyra A. (2004), Quality assessment and improvements of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images, *Remote Sensing of Environment*, **94**, 172-186.

Hall, F. G. De Colstum, E. B., Collatz, G., J., Landis, D., Dirmeyer, P., Betts, A., Huffman, G. J., Bounoua, L. and Meeson, B. (2006), The ISLSCP Initiative II global datasets: surface boundary conditions and atmospheric forcings for land atmosphere studies, *Journal of Geophysical Research*, **111**, D22S01

Han, K.-S., Champeaux, J.-L and Roujean, J.-L. (2004) Land cover mapping over France in using SPOT-4/VEGETATION data, *Remote Sensing of Environment*, **92**, 52-66.

Hannerz,, F. and Lotsch, A. (2008), Assessment of remotely sensed and statistical inventories of african agricultural fields, *International Journal of Remote Sensing*,. **29**, 3787–3804.

Hansen, M. C., Defries, R. S., Townshend, J. R. G. and Sohlberg, R. A. (2000), Global land cover classification at 1 km spatial resolution using a classification tree approach, *International Journal of Remote Sensing*, **21**, 1331-1364.

Hansen, M., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C. and Sohlberg, R. A. (2003), Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interactions*, **7**, 1-15.

Hill, M. J. and Donald, G. E. (2003), Estimating spatio temporal pattern of agricultural productivity in fragmented landscapes using AVHRR NDVI time series, *Remote Sensing of Environment*, **84**, 367-384.

Holben, B. N. (1986), Characteristics of maximum-value composite images from temporal AVHRR data, *International Journal of Remote Sensing*, **7**, 1417–1434.

Hoerling, M., Hurrel, J., Eischeid, J. and Phillips, A. (2006), Detection and attribution of twentieth-century northern and southern African rainfall change, *Journal of Climate*, 19, 3989-4008.

Houldcroft, C. J., Grey, W. M. F., Barnsley, M., Taylor, C. M., Los, S. O. and North, P. R. J. (2009), New Vegetation Albedo Parameters and Global Fields of Soil Background Albedo Derived from MODIS for Use in a Climate Model, *Journal of Hydrometeorology*, **10**, 183–198.

Huang, Z. and Lees, B. G. (2004), Combining non-parametric models for multisource predictive forest mapping, *Photogrammetric Engineering and Remote Sensing*, **70**, 415–425.

Huete, A. K. (1988), A soil-adjusted vegetation index (SAVI), *Remote Sensing of Environment*, **25**, 295-309.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P. , Hong, Y., Stocker, E. F. and Wolff, D. B. (2007), The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale, *Journal of Hydrometeorology*, **8**, 38-55.

Immerzeel, W. W., Quiroz, R. A. and. De Jong, S. M. (2005), Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series, *International Journal of Remote Sensing*, **26**, 2281-2296.

Jain, A. K., Murty, M. N. and Flynn, P. J. (1999), Data clustering: a review, *ACM Computing Surveys*, **31**, 264-323.

Jenkins, G. S., Gaye, A. T. and Sylla, B. (2005) Late 20th century attribution of drying in the Sahel from the regional climate model (RegCM3), *Geophysical Research Letters*, **32**, L22705.

Jin, M. and Liang. S. (2006), An improved land surface emissivity parameter for land surface models using global remote sensing observations, *Journal of Climate*, **19**, 2867-2881.

Jin, Y., Schaaf, C., Woodcock, C., Gao, F., Li, X., Strahler, A., Lucht, W. and Liang, S. (2003), Consistency of MODIS surface bidirectional reflectance distribution function and albedo retrievals: 2. Validation, *Journal of Geophysical Research*, **108**, 4159.

Jung, M., Henkel, K., Herold, M. and Churkina, G. (2006), Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sensing of Environment*, **101**, 534–553.

Justice, C. O. (1986), Monitoring east African vegetation using AVHRR data, *International Journal of Remote Sensing*, **6**, 1335-1372.

Kahl, C. H. (2006), *States, scarcity, and civil strife in the developing world*, Princeton University Press, Princeton, 333 pp.

Kamga, A. F., Jenkins, G. S., Gaye, A. T., Garba, A., Sarr, A. and Adedoyin, A. (2005), Evaluating the National Center for Atmospheric Research climate system model over West Africa: Present day and the 21st century A1 scenario, *Journal of Geophysical Research*, 110, D03106.

Kaptué, A. (2007), *Analysis of the spatial mutation of the Bangr-Weogo urban park in the city of Ouagadougou*, Master thesis, International Institute for Water and Environmental Engineering of Ouagadougou, Burkina Faso, 27 pp.

Kaufman, Y. J. and Tanré, D. (1992), Atmospherically resistant vegetation index (ARVI) for EOS-MODIS, *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 261-270.

Kelatwang, S. and Garzuglia, M. (2006) Changes in forest area in Africa 1990-2005, *International Forestry Review*, **8**, 21-30.

Kloditz, C., Van Boxtel, A., Carfagna, E., and Van Deursen, W. (1998), Estimating the accuracy of coarse scale classification using high scale information, *Photogrammetric Engineering and Remote Sensing*, **64**, 127−133.

Knight, J. F., Lunetta, R. L., Ediriwickrema, J. and Khorram, S. (2006), Regional scale land-cover characterization using MODIS-NDVI 250 m multitemporal imagery: A phenology based approach, *GIScience and Remote Sensing*, **43**, 1−23.

Knipling, E. B. (1970), Physical and physiological bases for the reflectance of visible and near infrared radiation from vegetation, *Remote Sensing of Environment*, **1**, 155-159.

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Diner, D. J. and Running, S. W. (1998), Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data, *Journal of Geophysical Research,* **103**, 32257-32275.

Koeppe, C. E. and De Long, G. C. (1958), *Weather and Climate*, New-York, McGraw-Hill, 341 pp.

Kolasa, J. and Rollo, C. D. (1991), Introduction: the heterogeneity of heterogeneity: a glossary, in *Ecological heterogeneity*, J. Kolasa and S. T. A. Pickett, Eds.,. Springer- Verlag, New-York, pp. 1-23.

Köppen, W. (1936) Das geographisca System der Klimate, in *Handbuch der Klimatologie*, W. Köppen and G. Geiger, Eds, Vol.1, Part 1, Part C, Borntraeger, pp. 1-44.

Kottek, M., Rieser, J., Beck, C., Rudolf, B. And Rubel, F. (2006), World Map of the Köppen-Geiger climate classification updated, *Meteorologische Zeitschrift*, **15**, 259-263.

Krinner, G., Viovy, N., De Noblet-Ducoudré N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C. (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochemical Cycles*, 19, GB1015.

Lacaze, R., Donier, S,. Lacarrère, P. and Roujean, J.-L. (2003), AVHRR-Derived Land Surface Conditions for Flux Simulations with a Mesoscale Model over the HAPEX-Sahel Study Area, *Journal of Applied Meteorology*, **42**, 686-700.

Lakshmi, V., Albertson J. and Schaake J. (2001), *Land Surface Hydrology, Meteorology, and Climate Observations and Modelling*, Amercican Geophysical Union, Washington, 246 pp.

Landgrebe, D. A. (2003), *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley & Sons, Hoboken, 508 pp.

Latifovic, R. and Olthof, I. (2004), Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. *Remote Sensing of Environment*, **90**, 153-165.

Le Moigne, P., Bonne, A., Calvet, J.-C, Decharme, B., Faroux, S., Gibelin, A.-L. et al. (2009), *SURFEX scientific documentation*, Technical note n° 87, GMME/CNRM, Météo France, 211 pp.

Lefsky, M. A. and Cohen, W. B. (2003), Selection of remotely sensed data, in *Remote sensing of Forest Environments: Concepts and case studies*, M. A. Wulder and S. E. Franklin, Eds., Kluwer Academic Publishers, pp 13-46.

Leibovici, D., Quillevere, G. and Desconnets, J.-C. (2007), A method to classify ecoclimatic arid and semiarid zones in circum-Saharan Africa using monthly dynamics of multiple indicators, *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 4000-4006.

Liang, S. (2004), *Quantitative Remote Sensing of Land surface*, John Wiley and Sons, Hoboken, 534 pp.

Liang, S., Strahler, A. H. and Walthal, C. (1999), Retrieval of land surface albedo from satellites observations: A simulation study, *Journal of Applied Meteorology*, **38**, 712-715.

Linderman, M., Rowhani, P., Benz, D., Serneels, S. and Lambin, E. F. (2005), Land-cover change and vegetation dynamics across Africa, *Journal Geophysical Research*, **110**, D12104.

Lo, C. P. and Choi, J., 2004, A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM +) images, *International Journal of Remote Sensing*, **25**, 2687–2700.

Los, S. O., Collatz, G. J., Sellers, P. J., Malmström, C. M., Pollack, N. H., Defries, R. S., Bounoua, L., Parris, M. T., Tucker, C. J. and Dazlich, D. A. (2000), A Global 9-yr Biophysical Land Surface Dataset from NOAA AVHRR Data, *Journal of Hydrometeorology*, **1**, 183-199.

Loveland, T. R. and Belward, A. S. (1997), The IGBP-DIS global 1 km land cover data set, DISCover first results, *International Journal of Remote Sensing*, **18**, 3289-3295.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. and Merchant, J. W. (2000), Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *International Journal of Remote Sensing*, **21**, 1303-1330.

Lovett, J. C., Midgely, G. F. and Barnard, P. B. (2005), Climate change and ecology in Africa, *African Journal of Ecology*, **43**, 167-169.

Lu, D. and Weng, Q. (2007), A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, **28**, 823−870.

Maignan, F., Bréon, F.-M., Bacour, C., Demarty, J. and Poirson, A. (2008), Interannual vegetation phenology estimates from global AVHRR measurements - Comparison with in situ data and applications, *Remote Sensing of Environment*, **112**, 496–505.

Maisongrande, P., Duchemin, B. and Dedieu, G. (2004), VEGETATION/SPOT: an operational mission for the earth monitoring; presentation of new standard products, *International Journal of Remote Sensing*, **25**, 9-14.

Mahé, G., Paturel, E. J., Servat, E., Conway, D. and Dezetter, A. (2005), The impact of land use change on soil water holdingcapacity and river flow modelling in the Nakambe River, Burkina-Faso. *Journal of Hydrology*, **300**, 33–43.

Mahfouf, J.-F., Manzi, A. O., Noilhan, J., Giordani, H., DéQué, M. (1995), The land surface scheme ISBA within the Météo-France climate model ARPEGE. Part I. implementation and preliminary results, *Journal of Climate*, **8**, 2039-2057.

Martinez, B. and Gilabert, M. A. (2009), Vegetation dynamics from NDVI time series analysing using the waveltt transform, *Remote Sensing of Environment*, **113**, 1823-1842.

Masson, V., Champeaux, J.-L., Chauvin, F., Meriguer, C. and Lacaze, R. (2003), A global database of land surface parameters at 1-km resolution in meteorological and climate models, *Journal of Climate*, **16**, 1261-1282.

Matthews, E. (1983), Global vegetation and land use: new high resolution data bases for limited studies, *Journal of Climatology and Applied Meteorology*, **22**, 474–487.

Mayaux, P., Bartholomé, E., Fritz, S. and Belward, A. (2004), A new land cover map of Africa for the year 2000, *Journal of Biogeography*, **31**, 861-877.

McCallum, I., Obersteiner, M., Nilson, S. and Shvidenko, A. (2006), A spatial comparison of four satellite derived 1 km global land cover data sets, *International Journal of Applied Earth Observation and Geoinformation*, **8**, 246-255.

McCullagh, P. and Yang, J. (2008), How many clusters, *Bayesian Analysis*, **3**, 101-120.

McLean, C. J., Lovett, J. C., Küper, W., Hanah, L., Sommer, J. H., Barthlott, W. et al. (2005), African plan diversity and climate change, *Annals of the Missouri Botanical Garden*, **92**, 139-152.

McMichael, A., Woodruff, R. and Hales, S. (2006), Climate change and human health: present and future risks, *The Lancet*, **367**, 859-869.

MEA (2005), *Ecosystems and Human Well-being: Our human planet*, Millennium Ecosystem Assessment, Island Press, Washington, DC, 109 pp.

Mitchell, T. D. and Jones, P. D. (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *International Journal Of Climatology*, *International Journal of Climatology,* **25,** 693–712.

Molesworth, A. M., Cuevas, L. E., Connor, S. J., Morse, A. P. and Thomson, M. C. (2003), Environmental risk and meningitis epidemics in Africa, *Emerging Infectious Disease*, **9**, 1287-1293.

Moody, A. and Johnson, D. M. (2001), Land-Surface Phenologies from AVHRR Using the Discrete Fourier Transform, *Remote Sensing of Environment*, 75, 305-323.

Myneni, R. B., Knyazikhin, Y., Zhang, Y., Tian, Y., Wang, Y., Lotsch, A. et al. (1999), Modis Leaf Area Index (LAI) and Fraction of Phosynthetically Active Radiation Absorbed by Vegetation (fPAR) Product, *Algorithm theoretical basis document,* Boston University, Boston, USA, 126 pp.

Nakache, J.-P. and Confais, J. (2004), *Approche pragmatique de la classification*, EditionsTechnip, 262 pp.

Nelleman, C. and Corcoran, E. (2006), *Our precious coasts: marine pollution, climate change and the resilience of coastal ecosystems*, United Nations Environment Programme, Norway, 40 pp.

Nelson, G. C., Rosegrant, M. W., Koo, J., Robertson, R., Sulser, T., Zhu, T. et al (2009), *Climate Change Impact on Agriculture and Costs of Adaptation*, International Food Policy Research Institute, Washington, USA, 19 pp.

Nicholson, S. E. (2001), Climatic and environmental change in Africa during the last two centuries, *Climate Research*, **17**,123–144.

Nicholson, S. E. (2008), The intensity, location and structure of the tropical rainbelt over west africa as factors in interannual variability,. *International Journal of Climatology*, **28**, 1775-1785.

Noilhan, J. and Mahfouf, J.-F. (1996), The ISBA land surface parametrisation scheme, *Global Planetary Change*, **13**, 145-159.

Noilhan, J. and Planton, S. (1989), A simple parameterization of land surface processes for meteorological models, *Monthly Weather Review*, **117**, 536-549.

Nordas, R. and Gleditsch, N. P. (2007), Climate change and conflict, *Political Geography*, **26**, 627-638.

Oleson, K. W., Bonan, G. B., Schaaf, C., Gao, F., Jin, Y., and Strahler, H. (2003), Assessment of global climate model land surface albedo using MODIS data, *Geophysical Research Letters*, **30**, 1443.

Olson, J. S., Watts, J. A. and Allison L. J. (1983), *Carbon in Live Vegetation of Major World Ecosystems*, Oak Ridge National Laboratory -5862. Oak Ridge Tennessee, 164 pp.

Patz, J. A., Campbell-Lendrum, D., Holloway, T. and Foley, J. A. (2005), Impact of regional climate change on human health, *Nature*, **438**, 310-317.

Peel, M. C., Finlayson, B. L. and McMahon, T. A. (2007), Updated world map of the Köppen-Geiger climate classification, *Hydrology and earth Sciences*, **11**, 1633-1644.

Pellarin, T., Tran, T., Cohard, J.-M, Galle, S., Laurent, J.-P., De Rosnay, P. and Vischel, T. (2009), Soil moisture mapping over West Africa with a 30-min temporal resolution using AMSR-E observations and a satellite-based rainfall product, *Hydrology and Earth System Science.*, **13**, 1887-1896.

Peterson, T. C. and Vose, R. S. (1997) An overview of the Global Historical Climatology Network temperature database, *Bulletin of American Meteorology and Society*, **78**, 2837–2849.

Philippon, N., Jarlan, L., Martiny, N., Camberlin, P. and Mougin, E. (2007), Characterization of the interannual and intraseasonal variability of West African vegetation between 1982 and 2002 by means of NOAA AVHRR NDVI data, *Journal of Climate*, **20**, 1202-1218.

Pielke, R. A., Lee, T. J., Copeland, J. H., Eastman, J. L., Ziegler, C. L. and Finley, C. A. (1997), Use of USGS-provided data to improve weather and climate simulations, *Ecological Applications*, **7**, 3–21.

Pinty, B. and Verstraete, M. M. (1991), Extracting information on surface properties from bidirectional reflectance measurements, *Journal of Geophysical Research*, **96**, 2865-2874.

Pinty, B. and Verstraete, M. M. (1992), GEMI: a non linear index to monitor global vegetation from satellite, *Vegetatio*, **10**, 15-20.

Pitman, A. J. (2003), The evolution of, and revolution in, land surface schemes designed for climate models, *International Journal of Climatology*, **23**, 479-510.

Price, M. F. (1989), Global change*:* defining the ill-defined, *Environment*, **31,** 18-20, 42-44.

Priestley, M. B. (1981) *Spectral Analysis and Time Series: Univariate Series*, vol 1, Academic Press, London, 653 pp.

Pohl, C. and Van Genderen, J. L. (1998), Multisensor image fusion in remote sensing: concepts, methods, and applications, *International Journal of Remote Sensing*, **19**, 823–854.

Quiroz, R., Yarlequé, C., Posadas, A., Mares, V. And Immerzel, W. W. (2010), Improving daily rainfall estimation from NDVI using a wavelett trasform, *Environmental Modelling & Software*, doi:10.1016/j;envsoft.2010.07.06

Rahman, H. and Dedieu, G. (1994), SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum, *International Journal of Remote Sensing*, **15**, 123–143.

Ramankutty, N. (2004), Croplands in West Africa: a geographically explicit dataset for use in models, *Earth Interactions*, **8**, 1-21.

Ramankutty, N., Evan, A. T., Monfreda, C. and Foley, J. A. (2008), Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, *Global Biogeochemical Cycles*, **22**, GB1003.

Redelsperger, J. L., Thorncroft, C., Diedhiou, A., Lebel, T., Parker, D. and Polcher, J. (2006), African Monsoon Multidisciplinary Analysis (AMMA): An International Research Project and Field Campaign, *Bulletin of the American Meteorological Society*, **87**, 1739-1746.

Roujean, J.-L and Lacaze, R. (2002), Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation, *Journal of Geophysical Research*, **107**, D12.

Roujean, J.-L., Leroy, M. and Deschamps, P.-Y. (1992), A bidirectional reflectance model of the earth's surface for the correction of remote sensing data, *Journal of Geophysical Research*, **97**, 20455-20468.

Rouse, J. W., Haas, R. H., Schell, J. A. and Deering, D. W. (1974), Monitoring vegetation systems in the Great Plains with ERTS, *Proceedings of the 3rd Earth Resources Technology Satellite-1 Symposium, Greenbelt*, NASA SP-351 I, 301-317.

Roy, D. P., Ju, J., Mbow, C., Frost, P. and Loveland, T. (2010) Accessing free Landsat data via the Internet: Africa's challenge, *Remote Sensing Letters*, **1**, 111-117.

Running, S. W., Loveland, T. R., Pierce, L. L., Nemani, R. and Hunt, E. R. (1995). A remote sensing based vegetation classification logic for global land-cover analysis, *Remote Sensing of Environment*, **51**, 39−48.

Salmun, H. and Molod, A. (2006), Progress in modelling the impact of land cover change on the global climate, *Progress in Physical Geography*, **30**, 737-749.

Salomon, J. G., Schaaf, C. B., Strahler, A. H., Gao, F. and Jin, Y. F. (2006), Validation of the MODIS bidirectional reflectance distribution function and albedo retrievals using combined observations from the aqua and terra platforms, *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 1555–1565.

Samain, O., Kergoat, L., Hiernaux, P., Guichard, F., Mougin, E., Timouk, F. and Lavenu, F. (2008), Analysis of the in situ and MODIS albedo variability at multiple timescales in the Sahel, *Journal of Geophysical Research*, **113**, D14119.

Saux-Picart, S., Ottlé, C., Perrier, A., Decharme, B., Coudert, B., Zribi, M., Boulain, N., Cappelaere, B and Ramier, D. (2009), SetHyS_Savannah: A multiple source land surface model applied to Sahelian landscapes, *Agricultural and Forest Meteorology*, **149**, 1421-1432.

Savitzky, A. and Golay, M. J. E. (1964), Smoothing and Differentiation of Data by Simplified Least Squares Procedures, *Analytical Chemistry*, **36**, 1627–1639.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X. W., Tsang, T. et al. (2002), First operational BRDF, albedo nadir reflectance products from MODIS, *Remote Sensing of Environment*, **83**, 135-148.

Schneider, A, Friedl, M. A. and Potere, D. (2010), Mapping global urban areas using MODIS 500-m data: New methods and datasets based on `urban ecoregions', *Remote Sensing of Environment*, **114**, 1733-1746.

Scarlermann, J. P. W., Benz, D., Hay, S. I., Purse, B. V., Tatem, A. J., Wint, G. R. W. and Rogers, D. J. (2008), Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data, *Plosone*, **1**, e1408.

Scheiter, S. and Higgins, S. (2009), Impacts of climate change on the vegetation of Africa: an adaptative dynamic vegetation modelling approach, *Global Change Biology*, **15**, 2224-2246.

Sellers, P. J., Los, S. O., Tucker, C. J., Justice, C. O. , Dazlich, D. A., Collatz, G. J. and Randall, D. A. (1996), A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data, *Journal of Climate,* **9,** 706–737.

Shalan, M. A., Arora, M. K. and Ghosh, S. K. (2003), An evaluation of fuzzy classifications from IRS 1 C LISS III imagery: a case study, *International Journal of Remote Sensing*, **64**, 273-280.

Scharleman J. P. W., Benz D, Hay S. I., Purse BV, Tatem A. J., Wint, W. R. G. and Rogers, D. J. (2008), Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data, *PLoS One*, doi:10.1371/journal.pone.0001408

Shepard, M. K., Campbell, B. A., Bulmer, M. H., Farr, T. G., Gaddis, L. R. and Plaut, J. J. (2001), The roughness of natural terrain: A planetary and remote sensing perspective, *Journal of Geophysical Research*, **106**, 32777-32795.

Shukla, J. (1981), Dynamical predictability of monthly means, *Journal of Atmospheric Science*, **38**, 2547-2572.

Simmons, R., Barnard, P., Dean, W. R. J., Midgley, G. F., Thuiller, W. and Hughes, G. (2005), Climate change and birds: perspectives and prospect s from Southern Africa, *Ostrich*, **75**, 295-308.

Small, J., Goetz, S. J. and Hay, S. I. (2003), Climatic suitability for malaria transmission in Africa, *Proceedings of the National Academy of Sciences*,**100**, 15341-15345.

Spalding, M. D. (2001), *World Atlas of Coral Reefs*, University of California Press, Berkeley, 424 pp.

Stefanov, W. L., Ramsey, M. S. and Christensen, P. R. (2001), Monitoring urban land cover change: an expert system approach to land cover classification of semiarid to arid urban centers, *Remote Sensing of Environment*, **77**, 173–185.

Stehman, S. V. (2009), Sampling designs for accuracy assessment of land cover, *International Journal of Remote Sensing*, **30**, 5243-5272.

Stenberg, P. (2006), A note on the G-function for needle leaf canopies, *Agricultural and Forest meteorology*, **136**, 76-79.

Stern, N. (2007), *The economics of climate change: the Stern review*, Cambridge University Press, Cambridge, 692 pp.

Stock, R. (2004), *Africa South of the Sahara: A Geographical Interpretation*, Guilford Press, New York, 477 pp.

Sun, W., Liang, S., Xu, G., Fang, H. and Dickinson, R. (2008), Mapping plant functional types from MODIS data using multisource evidential reasoning, *Remote SEnsing of Environment*, **112**, 1010-1024.

Sutton, A. E., Dohn, J., Loyd, K., Tredennick, A., Bucini, G. and Solorzano, A. (2010) Does Warming increase the risk of civil war in Africa?, *Proceedings of the National Academy of Sciences*, **107**, E 102.

Tansey, K., Grégoire, J.-M, Stroppiana, D., Sousa, A., Silva, J., Pereira, J. M. C. et al. (2004), Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data, *Journal of Geophysical Research*, **109**, D14S03.

Tapley B. D., Bettadpur, S., Ries, J. C., Thompson P. F. and Watkins, M. M. (2004), GRACE measurements of mass variability in the Earth system, *Science*, **305**, 503–505.

Thiombiano, L. and Tourino-Soto, I. (2007), Status and trends in land degradation in Africa, in *Climate and Land Degradation*, M. V. K. Sivakumar and N. Ndiang'ui, Eds., Springer-Verlag, pp. 39-53.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J, Collingham, Y. C. et al. (2004), Extinction risk from climate change, *Nature*, **427**, 145-148.

Thorndike, R. L. (1953), Who belongs in a family?, *Psychometrika*, **18**, 267-276.

Tourre, Y. M., Paz, S. Kushnir, Y. and White, B. (2010), Low-frequency climate variability in the atlantic basin during the 20th century, *Atmospheric Science Letters*, **11**, 180-185.

Tso, B. and Mather, P. M. (2001), *Classification Methods for Remotely Sensed Data*, Taylor and Francis Inc., New York, 332 pp.

UN (2009), *World population prospects: The 2008 Revision Population Database*, United Nations Population Division. See http://esa.un.org/unpp/ (Last accessed August, 2010)

UNCCD (1994), *Convention 47/188 of the United Nations to Combat Desertification*, United Nations, Paris, 58 pp.

UNEP (2006), *Africa environment outlook 2: Our environment, our health*. Division of Early Warning and Assessment, United Nations Environment Programme, Kenya, 542 pp.

UNEP (2008), *Africa: Atlas of our changing environment*. Division of Early Warning and Assessment, United Nations Environment Programme, Kenya, 374 pp.

UNFPA (2007), *State of World Population 2007: Unleashing the Potential of Urban Growth*, United Nations Population Fund. See http://www.unfpa.org/swp/swpmain.htm (Last accessed August, 2010).

USGS (1996) *GTOPO 30: Global 30 arc second elevation data set*. See http://edcdaac.usgs.gov/gtopo30/gtopo30.html (Last accessed February 2010).

Vancutsem, C., Pekel, J.-F, Bogaert, P. and Defourny, P. (2007), Mean Compositing, an alternative strategy for producing temporal syntheses. Concepts and performance assessment for SPOT VEGETATION time series, *International Journal of Remote Sensing*, **28**, 5123-5141.

Vermote, E. F., Tanré, D., Deuze, J. L , Herman, M. and Morcette, J.-J. (1997), Second simulation of the satellite signal in the solar spectrum, 6S: an overview, *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 675–686.

Vintrou, E., Houles, M., Seen, D. L., Feau, C., Laine, G. and Bégué, A. (2009), Mapping cultivated area in West Africa using MODIS imagery and agroecological stratification, *International Geoscience And Remote Sensing Symposium*, **5**, 393-396, Cape Town, 12-17 July 2009.

Vörösmarty, C. J., Douglas, E. M., Green P. A. and Revenga, C. (2005), Geospatial indicators of emerging water stress: an application to Africa. *Ambio*, **34**, 230-236.

Washington, R., Todd, M. C., Lizcano, G., Tegen, I., Flamant, C. et al. (2006). Links between topography, wind, deflation, lakes and dust: the case of the Bodélé depression, Chad, *Geophysical Research Letters*, **33**, L09401.

WHO (2002), *The World Health Report 2002*, World Health Organization, Geneva, Switzerland.

Wilson, M. F. and Henderson-Sellers, A. (1985), A global archive of land cover and soils data for use in general circulation models, *Journal of Climatology*, **5**, 119–143.

Wolfe, R., Nishihama, M., Fleig, A., Kuyper, J., Roy, D., Storey, J. and Patt, F. (2002), Achieving subpixel geolocation accuracy in support of MODIS land science, *Remote Sensing of Environment*, **83**, 31-49.

World Bank (2006), *World Development Indicators of the year 2006*, The World Bank, Washington, DC, USA.

Wuebles, D. J., Brasseur, G., P., Rhode, H., Barrie, L. A., Crutzen, P. J. et al. (2003), Changes in the chemical composition of the atmosphere and potential impacts, in *Atmospheric chemistry in a changing world*, G., Brasseur, et al., Eds., Springer-Verlag, 1-17.

Wulder, M. A., White, J. C., Goward, S. N., Masek, J. G., Irons, J. R., Herold, M. et al. (2008), Landsat continuity: Issues and opportunities for land cover monitoring, *Remote Sensing of Environment*, **112**, 955–969.

WWF (2000), *The Africa Water Vision for 2025. Equitable and sustainable use of water for socioeconomic development*, Word Water Forum, The Hague, The Netherlands, 30 pp.

Xie, Y., Sha, Z. and Yu, M. (2008), Remote sensing imagery in vegetation mapping: a review, *Journal of Plant Ecology*, **1**, 9-23.

Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N. V., Wang, Y. et al. (2006), MODIS leaf area index products: From validation to algorithm improvement, *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 1885– 1898.

Zeng, X. (2001), Global Vegetation Root Distribution for Land Modeling, *Journal of Hydrometeorology*, **2**, 525-530

Zeng, X., Dickinson, R. E., Walker, A.,Shaiki, M., Defries, R. S. and Qi, J. (2000), Derivation and evaluation of global 1-km fractional vegetation cover data for land modelling, *Journal of Applied Meteorology*, **39**, 826-839.

Zhang, X., Friedl, M. A. and Schaaf, C. B. (2009), Sensitivity of vegetation phenology detection to the temporal resolution of satellite data, *International Journal of Remote Sensing of Environment*, **30**, 2061-2074.

Zhao, M., Pitman, J. A. and Chase, T. (2001), The impact of land cover change on the atmospheric circulation, *Climate Dynamics*, **17**, 467-477.

Liste alphabétique des sigles

AEJ	African Esterly Jet
AMMA	Analyses Multidisciplinaires de la Mousson Africaine
AMESD	African Monitoring of Environment for Sustainable Development
ALMIP	AMMA Land Surface Model Intercomparison Project
ARVI	Atmospherically Resistant Vegetation Index
AVHRR	Advanced Very High Resolution Radiometer
BRDF	Bidirectional Reflectance Distribution Function
BSA	Black Sky Albedo
BUA	Back-Up Algorithm
CCM	Community Climate Model
CLM	Common Land Model
CNRM	Centre National de Recherches Météorologiques
DAAC	Distributive Active Archive Center
DCW	Digital Chart of the World
DMA	Defense Mapping Agency
DMSP	Defense Meteorological Satellite Program
DN	Digital Number
DT	Decision Tree
DTED	Digital Terrain Elevation Data
EC	European Community
ERS	European Remote Sensing
ESA	European Space Agency
FAO	Food and Agricultural Organization
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation by the canopy
GCOS	Global Climate Observing System
GEMI	Global Environment-Monitoring Index

GES	Gaz à Effet de Serre
GEWEX	Global Energy and Water Cycle Experiment
GHCN	Global Historical Climatology Network
GMES	Global Monitoring for Environment and Security
IGBP	International Geosphere Biosphere Program
IIASA	International Institute for Applied Systems Analysis
ISBA	Interactions entre le Sol, la Biosphère et l'Atmosphère
ISFC	Index of Segmentation of Fourier Component
IV	Indice de Végétation
JERS	Japanese Earth Remote Sensing
JRC	Joint Research Centre
JULES	Joint UK Land Environment Simulator
k-NN	k-Nearest Neighbour
LAI	Leaf Area Index
LGP	Length of Growing Period
LSM	Land Surface Model
MA	Main Algorithm
MODIS	Moderate Imaging Spectroradiometer
MSG	Meteosat Seconde Generation
MERIS	Medium Resolution Imaging Spectrometer
NASA	National Space Agency
NDVI	Normalised Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
OLS	Operational Linescan System
OMD	Objectifs du Millénaire pour le Développement
ONU	Organisation des Nations Unies
PPV	Plus proche voisin
QC	Quality Control

RVI	Ratio Vegetation Index
SAR	Synthetic Aperture Radar
SAVI	Soil Adjusted Vegetation Index
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SM	Status Map
SMAC	Simplified Method for the Atmospheric Correction
SPOT	Satellite Probatoire pour l'Observation de la Terre
TEJ	Tropical Easterly Jet
TOA	Top Of Atmosphere
TFD	Transformée de Fourier Discrète
TFP	Type fonctionnel de plante
UMD	University of Maryland
UNCCD	United Nations Convention To Combat Desertification
UNEP	United Nations Environment Protection
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
VGT	VEGETATION
WGS	World Geodetic System
WHO	World Health Organisation
WRCP	World Climate Research Programme
WSA	White Sky Albedo
ZCIT	Zone de Convergence InterTropicale

Liste alphabétique des symboles

A	Albédo diffus de la surface
C_H	Coefficient de transfert turbulente
C_p	Chaleur spécifique de l'air
C_T	Coefficient d'inertie thermique de la surface
$(C_i)_{1\leq i\leq 4}$	Constante associée à la vitesse de drainage
E_{canop}	Evaporation de l'eau interceptée par la végétation
E_{soil}	Evaporation à la surface
E_{transp}	Transpiration
E_v	Evapotranspiration par la végétation
$(F)_{1\leq i\leq 4}$	Coefficients associés au stress hydrique, rayonnement solaire, température et humidité atmosphérique
G	Flux de chaleur par conduction
H	Flux de chaleur sensible
I_r	Infiltration
L	Chaleur latente de vaporisation
LE	Flux de chaleur latente
P	Précipitation
Q_s	Ruissellement de surface
R_A	Rayonnement infrarouge atmosphérique incident
R_a	Résistance aérodynamique
R_g	Rayonnement solaire incident
R_n	Rayonnement net

R_s	Résistance stomatique
R_t	Ruissellement total
$R_{s\,\min}$	Résistance minimale de surface
T_a	Température de l'air au voisinage de la surface
T_s	Température de la surface
T_2	Température en profondeur du sol
V_a	Vitesse du vent
w_r	Quantité d'eau interceptée par la végétation
$w_{r\,\max}$	Quantité d'eau maximale interceptée par la végétation

---

d_r	Drainage
$(d_i)_{1\le i\le 3}$	Profondeur des réservoirs de surface, racinaire et sous-racinaire
h_{sat}	Conductivité hydraulique
h_u	Humidité relative de la surface
h_v	Coefficient de Halstead
q_a	Humidité spécifique de l'air
q_{sat}	Humidité spécifique à la saturation de l'air
veg	Fraction de végétation
w_{fc}	Contenu eau à la capacité au champ
w_r	Réservoir d'eau intercepté par la végétation
w_{sat}	Contenu en eau à la saturation
w_{seq}	Contenu volumique en eau à l'équilibre de la surface
w_{wilt}	Contenu en eau au point de flétrissement

$(w_i)_{1\le i\le 3}$	Réservoirs de surface, racinaire et sous-racinaire
z	Niveau du sol
z_{0m}	Rugosité dynamique
z_{0h}	Rugosité thermique

δ	Fraction du feuillage recouverte par l'eau interceptée
ε_s	Emissivité de la surface
λ	Conductibilité thermique du sol
ρ_a	Densité de l'air
ρ_w	Densité de l'eau liquide
σ	Constante de Stefan-Bolltzmann
τ	Constante temporelle d'un jour

Liste des tableaux

Tableau I-1 Relations entre les biomes africains, le climat et les types de sols (d'après Stock, 2004) 30
Tableau I-2 Caractéristiques des capteurs VGT-I et II, AQUA et TERRA..... 37
Tableau II-1 Cartes utilisées pour cartographier l'occupation des sols sur la zone AMMA.... 46
Tableau II-2 Exemples de classes d'occupation des sols obtenues par combinaison de GLC2000 et ECOCLIMAP-I..... 52
Tableau II-3 Matrice de confusion entre GLOBCOVER et la classification supervisée sur la zone AMMA..... 63
Tableau II-4 Paramètres et statistiques de conversion entre divers albedo MODIS du spectre total sur la zone AMMA 67
Tableau III-1 Principales caractéristiques des cartes pancontinentales d'occupation des sols en Afrique d'une résolution kilométrique à une centaine de mètres..... 74
Tableau III-2 Légende hiérarchique des cartes LCM-I et LCM-II (8 hyper classes au niveau I et 21 classes au niveau II)..... 80
Tableau III-3 Sous-ensembles obtenus après l'application du classificateur par écorégions. Se référer au tableau 3.2 pour la légende des numéros des hyper classes. 84
Tableau III-4 Légende des cartes d'occupation des sols du tableau 3.1 produites à partir des images satellites acquises durant les années 2000 86
Tableau III-5 Classes agrégées. (Cf. le tableau 3.4 pour la définition du numéro de chaque classe)..... 87
Tableau III-6 Nombre des pixels (N.P) et des écosystèmes (N.E) au niveau II 96
Tableau III-7 Comparaison entre les cartes d'occupation des sols en prenant comme référence les cartes de la 2nde ligne. Les valeurs minimales sont en italiques et soulignées et les valeurs maximales sont en gras 102
Tableau III-8 Valeurs moyennes et deviations standards entre parentheses (%) à l'échelle continnetale africaine pour l'albédo du sol nu et par type de sol. 106
Tableau III-9 Erreur moyenne quadratique entre les moyennes climatologiques d'albédo de surface de MODIS comparées à celles d' ECOCLIMAP en prenant la classification ECOCLIMAP-II pour référence. Il faut noter que les pixels urbains ont été écartés lors de cette comparaison 112
Tableau III-10 Pourcentage des différences non significatives des moyennes mensuelles climatologiques d'albédo de surface d'ECOCLIMAP (I et II) en prenant MODIS comme référence 112

Liste des figures

Figure I.1 Approche par tuiles 13
Figure I.2 Topographie (d'après UNEP, 2008) 23
Figure I.3 Principaux types de sols en Afrique (d'après FAO/IIASA, 2009) 24
Figure I.4 Pourcentage de sable (gauche) et d'argile (droite) (d'après FAO/IIASA, 2009) 25
Figure I.5 Ressources hydriques (d'après UNEP, 2008) 26
Figure I.6 Climats (d'après Fischer et al., 2002) 28
Figure I.7 Biomes (d'après UNEP, 2008) 30
Figure I.8 Transitions entre les phases phénologiques 35
Figure I.9 Différents facteurs météorologiques (d'après Goudie et Cuff, 2002) 40
Figure II.1 Séries temporelles du LAI MODIS moyenné en sur la zone AMMA pour la classe foret dense sempervirente de GLC2000 correspondant aux données brutes (noir, pointillé) et aux données lissées (rouge, continu). La courbe rouge est la moyenne du LAI de la classe à partir d' un polynôme de degré 5 appliqué pixel après pixel 49
Figure II.2 Séries temporelles du LAI MODIS (rouge, continu) et du NDVI VGT (noir pointillé) de 3 classes GLC2000 (forêt dense sempervirente, savane arborée décidue et savane arbustive décidue) sur la zone AMMA après lissage en utilisant un polynôme de degré 5 appliqué pixel après pixel. 50
Figure II.3 Méthodologie de classification supervisée 51
Figure II.4 Scission de la classe forêt dense sempervirente en 5 composantes ECOCLIMAP-I : 'forêt équatoriale' (classe numéro 16), 'savane arborée sahélienne' (classe numéro 34), 'savane arborée au sud de l'équateur' (classe numéro 35), 'mosaïque équatoriale de savane arborée et savane herbacée' (classe numéro 50), et 'savanes herbacées tropicales' (classe numéro 88) 54
Figure II.5 Moyenne annuelle du profil saisonnier du LAI MODIS sur la zone AMMA pour la classe 'forêt dense sempervirente' de GLC2000 pour les numéros des classes 16 (noir), 34 (rouge), 35 (vert), 50 (bleu), and 88 (violet). 54
Figure II.6 Méthodologie d'obtention de l'albédo du sol nu 58
Figure II.7 Classification supervisée à 37 écosystèmes sur la zone AMMA 60
Figure II.8 Moyenne annuelle du LAI saisonnier MODIS sur la zone AMMA. Les barres verticales décrivent la variabilité inter-annuelle 61
Figure II.9 Impact des nuages, aérosols et mauvaise classification des biomes sur la zone AMMA lors de l'estimation du LAI MODIS pour les classes 'forêt dense sempervirente' (droite) et 'savane arborée décidue' (gauche). *wt* et *w* signifient sans (without) et avec (with), respectivement, *best* (*satu*) signifie que le MA a été utilisé en produisant un résulat meilleur (saturé) et *back* signifie que le BUA a été utilisé. En haut à gauche, le

rouge représente le pourcentage d'observation nuageuses quand le MA est utilisé avec les meilleurs résultats, tandis que le bleu foncé représente le pourcentage d'observations nuageuses pour lesquelles le BUA a été utilisé durant la période 2000 à 2007. 65
Figure II.10 Variations de la fraction de végétation 66
Figure II.11 Albédo du sol nu dans le visible et le proche infrarouge sur la zone AMMA 67
Figure II.12 Diagrammes de dispersion sur la zone AMMA entre les albédos larges bandes SEVIRI (gauche) et MODIS (droite) et la nouvelle physiographie. Graphes du 6 mars 2006 (en haut) et du 6 septembre 2006 (en bas) 68
Figure III.1 Qualité des données NDVI VGT de 2000 à 2007 76
Figure III.2 Schéma méthodologique : a) Diagramme du projet ; b) *Stacked segment classification* 77
Figure III.3 Sensibilité du k-NN à la sélection initiale d'après Jain et al. (1999).....79
Figure III.4 Diagramme du rééchantillonage par classe majoritaire mis sur pied. Noter que la phase II n'a pas été utilisée lors du rééchantillonage de MODIS LC-I 89
Figure III.5 Accroissement du voisinage lors du ré-échantillonnage par classe majoritaire (figure 3.4) par mutation des cellules grisées en cellules noires pour MODIS LC-I (en haut) et GLOBCOVER (en bas) 89
Figure III.6 : a) l'Index de Segmentation des Composantes de Fourier calculé à partir de l'équation (3.6) ; b) les écorégions bioclimatiques ; c) les majeures zones agro-écologiques zones basées sur les données de l'IIASA et de la FAO (Fischer et al., 2002). 91
Figure III.7 LCM-I au niveau III.....93
Figure III.8 LCM-II au niveau III. 94
Figure III.9 Comparaison spatiale des classes agrégées entre GLC2000, GLOBCOVER, MODIS LC-I et ECOCLIMAP-II (a-Forêts et forêts mixtes ; b-Savanes arborées et savanes arbustives ; c-Savanes herbacées ; d- Cultures). Le nombre de la légende indique le nombre de cartes d'occupation des sols ayant identifié les classes agrégées à une position donnée. Par exemple, 0 signifie que le pixel n'a pas été identifié comme la classe agrégée en question par aucun des produits d'occupation des sols. 99
Figure III.10 Agrément entre les classifications GLC2000, GLOBCOVER, MODIS LC-I et ECOCLIMAP-II. La valeur entre parenthèses indique le pourcentage de couverture de la classe correspondante. 101
Figure III.11 Localisations des valeurs aberrantes obtenues lors de la détermination de l'albédo du sol nu dans le visible (i.e. en utilisant l'équation 2.1). 104
Figure III.12 Albédo du sol nu dans le visible (gauche) et proche infrarouge (droite) 105

Figure III.13 Différence d'albédo du sol nu d'ECOCLIMAP-II et ECOCLIMAP-I (haut), JULES (milieu) et ISLSCP-II (bas) dans les spectres du visible (gauche) et du proche infrarouge (droit) 107
Figure III.14 Histogrammes d'albédo du sol nu pour certains types de sol : visible (blanc) et proche infrarouge (gris) 108
Figure III.15 Comportement spatial de l'albédo de la végétation d'ECOCLIMAP-II (haut) et différence d'albédo de la végétation entre ECOCLIMAP-II et ECOCLIMAP-I (bas) dans le visible (gauche) et dans le proche infrarouge (droit) 110
Figure III.16 Comportement spatial des différences MODIS-ECOCLIMAP-II (haut) et MODIS-ECOCLIMAP-I (bas) en janvier (gauche) et juillet (droite) dans les spectres du visible et du proche infrarouge 111
Figure III.17 Comparaison spatiale de la fraction de végétation issue des produits CYCLOPES (a) , ECOCLIMAP-II (b) et ECOCLIMAP-I (c) 114
Figure IV.1 Description des échanges entre un modèle atmosphérique envoyant des flux radiatifs et un forçage météorologique à la surface et le module SURFEX décrivant un ensemble de 4 modèles physiques avec une interface où les variables moyennées sont renvoyées à l'atmosphère (d'après Le Moigne et al., 2009)..... 118
Figure IV.2 Schéma de surface ISBA-3L (adapté d'après Douville, 2006) 120
Figure IV.3 Occupation des sols de la nouvelle classification et localisation des 3 transects latitudinaux considérés pour des analyses approfondies : sahélo-soudanien (0.5N-1.5N, 10E-20E) en rouge, guinéen (6N-7N, 10W-0W) en vert et équatorial (13N-14N, 5W-5E) en bleu. 123
Figure IV.4 Changement d'occupation des sols entre la classification ECOCLIMAP-I et la nouvelle classification pour les principaux écosystèmes : forêt, savane arborée/arbustive, savane herbacée, cultures et sol nu : a) Pourcentage de changement dans des cellules de 0.5°x0.5° (marron=0-20%, jaune=20-40%, violet=40-80%, rose=60-80%, rouge=80-100%), b) Forêt, c) Cultures, d) savane arbustive/arborée. Légende des couleurs Orange=seulement la nouvelle classification, bleu=seulement ECOCLIMAP-I, vert=identique dans ECOCLIMAP-I et la nouvelle classification..... 124
Figure IV.5 Modification de la distribution spatiale du *veg* (moyenné de juin, juillet août, septembre 2005-2007 1ère colonne) et de la profondeur racinaire (2nde colonne) : pour CTL (1ère ligne), EXP (2nde ligne) et EXP – CTL (3e ligne) 126
Figure IV.6 Distribution de la pluviométrie (haut) et de la température à 2 m (bas) pour les années 2005 (1ère ligne), 2006 (2nde ligne) et 2007 (3e ligne) 127
Figure IV.7 Distribution de l'évaporation durant juin à septembre 2005-2007 129
Figure IV.8 Identique à la figure 4.5 pour l'humidité du sol ..... **Erreur ! Signet non défini.**
Figure IV.9 Identique à la figure 4.5 pour le flux de chaleur sensible..... 131

Figure IV.10 Pourcentage des classes agrégées contenues dans les 3 transects délimités à la figure 4.3 de la nouvelle classification (violet) et d' ECOCLIMAP-I (bleu). Se référer à la section 3.3.3.2 pourla définition des classes agrégées 132
Figure IV.11 Moyenne mensuelles des variables climatiques pour les simulations de contrôle et expérimentale dans les 3 transects de la figure 4.3 où le ruissellement total est exprimé en kg m^{-2} day^{-1}, le rayonement net net et les fflux de chaleur latente et sensible en W m^{-2} 133

Annexe A : Systèmes spatiaux d'acquisition d'images

A.1 Bref historique

La télédétection se définit ici comme l'ensemble des techniques, méthodes et instruments utilisés pour observer ou surveiller un objet ou phénomène sans un contact direct avec celui-ci. Bien que le terme de télédétection ait été inventé par Evelyne Pruit du Bureau de la Recherche Navale des Etats-Unis dans les années 1960, l'histoire de la télédétection remonte bien plus loin (voir tableau A.1).

1800	Découverte de l'infrarouge par F. W. Herschel
1839	Début de la photographie par l'invention du DAGUERREOTYPE
1850	Théorie ondulatoire de la lumière par J. B. L. Foucault
1858	1ère photographie à partir d'un ballon par G.P. Tournachon
1909	1ère photographie à partir d'un avion par W. Wright
1873	Théorie du spectre électromagnétique par J. C. Maxwell
1900	Quantification des ondes électromagnétiques par M. Planck
1909	Photographie à partir des avions
1914	1ère guerre mondiale : reconnaissance aérienne de la position et des forces ennemies
1935	Développement du radar en Allemagne
1940	2nde guerre mondiale :Utilisation de l'infrarouge pour distinguer les filets de camouflage de la vegetation
1950-	Recherches militaires et développements
1959	1ère photographie spatiale de la terre (EXPLORER-6)
1960	Lancement du 1er satelitte météorologique TIROS
1972	Lancement du 1er de la série des satellites LANDSAT Développements rapides dans le traitement d'images satellites
1973	Mise en orbite de la 1ère station spatiale SKYLAB
1978	Lancement du 1er satellite du système de positionnement GPS qui ne fut opérationnel qu'à partir de 1995 Lancement du 1er satellite civil radar SEASAT
1990-	Commercialisation des données satellitales
1999	Lancement d' IKONOS : 1er satellite civil ayant une très haute résolution spatiale (i.e <=5 m) Lancement du satellite TERRA avec à son bord le 1er capteur multispectral de dernière génération (36 bandes)
2000	Lancement du 4e satellite de la série SPOT avec à son bord le capteur de dernière génération VEGETATION
2001	Lancement de QUICKBIRD 2nd satellite civil ayant une très haute résolution spatiale
2002	Lancement d'ENVISAT avec à son bord 10 capteurs multispectraux de dernières génération dont MERIS
2009	Lancement de SMOS pour la fourniture des cartes globales d'humidité du sol

Tableau A.1 Repères dans l'histoire du développement de la télédétection (modifié de De Jong et al., 2004)

En télédétection, le rayonnement électromagnétique (REM) provenant de l'objet étudié (qui est la surface de la Terre dans notre cas) est mesuré et traduit en une information relative à l'objet ou en un processus relatif à l'objet. Durant la 1ère phase de télédétection dite phase de mesurage, quatre composantes sont d'importance priomordiale :

(i) la source du REM ;

(ii) le chemin du REM à travers l'atmosphère ;

(iii) l'interaction du REM avec l'objet ;

(iv) l'information enregistrée par le capteur.

La 2nde phase de télédétection comprend les composantes suivantes :

(i) la transmission, la réception et le (pré)traitement du rayonnement enregistré ;

(ii) l'interprétation et l'analyse des données ;

(iii) la création d'un produit final (ou carte dans notre cas).

A.2 Termes et concepts de base

A.2.1 Rayonnement électromagnétique

Le REM est une énergie qui se propage dans l'espace et interagit avec la matière. Il est composé d'un champ électrique (E) et d'un champ magnétique (M) (figure A1). Le champ électrique varie en grandeur et est orienté de façon perpendiculaire à la direction de propagation du rayonnement. Le champ magnétique est orienté de façon perpendiculaire au champ électrique. Les deux champs se déplacent à la vitesse de la lumière.

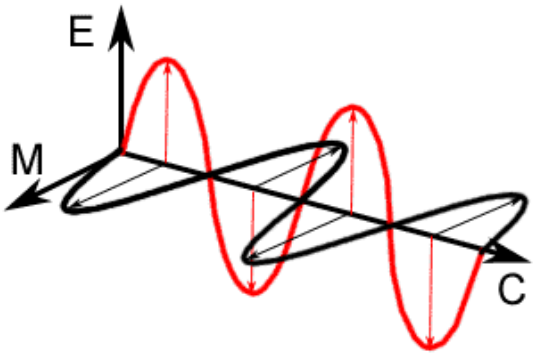

Figure A 1 :Décomposition d'une onde électromagnétique en un champ électrique (E) et un chaamp magnétique (M) disposés dans des plans perpendiculaires l'un par rapport à l'autre

Cinq grandeurs caractérisent le rayonnement électromagnétique:

(i) l'intensité qui exprime le pouvoir de réflexion (d'émission) ;

(ii) la fréquence ν ou la longueur d'onde λ qui détermine l'information spectrale et offrant ainsi un moyen efficace de différentiation des classes d'objets ;

(iii) la polarisation P qui fournit de bonnes indications sur l'état de rugosité du sol ;

(iv) la phase $\varphi(r,t)$ principalement exploitée dans l'analyse du signal radar ;

(v) la direction de propagation.

Les longeurs d'ondes les plus utilisées en télédétection sont :

(i) le visible $(0.4 < \lambda < 0.7\ \mu m)$ comprend la plage des longueurs d'onde correspondant à la *vue* humaine. Ici on a une lumière polychromatique ou blanche que l'on peut décomposer en une succession de longueurs d'onde que l'œil perçoit comme des couleurs (expérience d'Einstein de la décomposition de la lumière à travers un prisme) Les couleurs qui la composent vont du rouge (qui correspond au minimum d'énergie), jusqu'au bleu et au violet (énergie maximale), en passant par le jaune et le vert. La principale source de production de ce rayonnement est le soleil ;

(ii) l'infrarouge (IR) $(0.75\mu m < \lambda < 1\,mm)$ est une catégorie d'onde invisible de faible énergie. Il est émis par des corps modérément chauds. L'IR se divise en 2 grandes catégories : l'IR réfléchi et l'IR émis. L'IR réfléchi comprend le proche IR (PIR) de $0.75\ à\ 1.5\ \mu m$ très utilisé pour l'étude de la détéection de la végétation et de l'eau et le moyen IR (MIR) de $1.5\ à\ 3\ \mu m$ utilisé pour l'évaluation de l'humidité dans le sol, la détection de la neige et de la glace. L'IR émis comprend l'IR thermique (IRth) qui s'étend de $3\ à\ 15\ \mu m$: c'est la chaleur émise par la surface de la Terre et l'IR lointain de $15\ \mu m\ à\ 1\ mm$. L'IR émis fournit les températures de surface et permet donc d'obtenir des informations sur l'état hydrique et énergétique des surfaces ;

(iii) les hyperfréquences (HF) ou micro-ondes $(1\,mm < \lambda < 10\,cm)$ dont la principale caractéristique est la quasi-transparence de l'atmosphère.

Ces formes de REM, qui peuvent initialement apparaître comme des phénomènes séparés, font, en fait partie d'un spectre continu encore appelé spectre électromagnétique (figure A2).

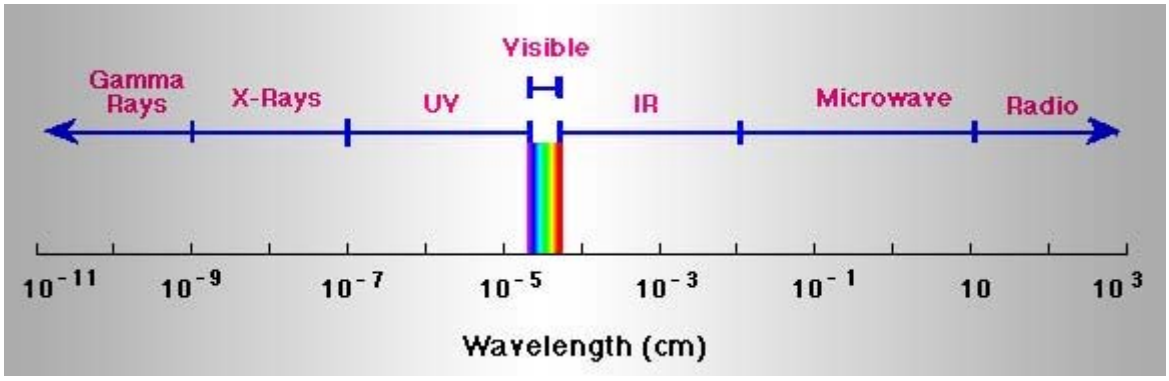

Figure A 2 : Spectre électriomagnétique

Les ondes électromagnétiques sont modifiées par l'atmosphère à travers les phénomènes de diffusion ou d'absortion et par la surface de Terre à travers les phénomènes de réflexion, d'absortion ou de transmission (voir figure A.3).

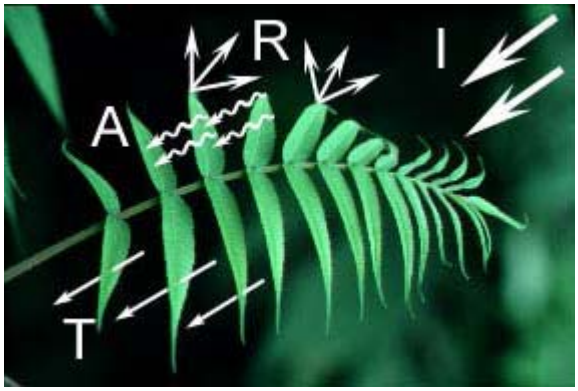

Figure A.3 Interaction du rayonnement avec la surface (Noter que selon le premier principe de la thermodynamique, la somme des énergies réfléchies, absorbées et transmises est égale à l'énergie incidente.)

Plusieurs satellites de télédétection ont ainsi été conçus pour mesurer ces interactions du REM et en particulier, le REM réfléchi par les cibles. Il est important de noter qu'un satellite est constitué de deux sous-ensembles :

(i) la charge utile, support sur lequel sont installés l'ensemble des capteurs et de leurs accessoires ;

(i) la plate-forme qui maintient le satellite sur son orbite selon l'orientation demandée et assure la liaison avec les stations à la surface de la terre, qui la supporte la charge utile et lui fournit les ressources dont elle a besoin pour son fonctionnement. Ce qui implique des contraintes de masse, de puissance et de trajectoire.

A.2.2 Orbitographie

L'altitude, la position des satellites sont dictées par un certain nombre de lois dont certaines sont bien connues (loi gravitationnelle de Newton, les lois de Kepler). L'orbite (ou trajectoire d'un satellite) est choisie en fonction de la capacité des capteurs qu'il transporte et des objectifs de sa mission. Le choix d'une orbite est déterminé par l'altitude (la hauteur du satellite au-dessus de la surface de la terre), l'orientation et la rotation du satellite par rapport à la terre.

Certains satellites ont une altitude très élevée et regardent toujours la même région de la surface de la terre (figure A.4.a), ils ont une orbite géostationnaire (leur période de révolution est environ celle de la rotation de la terre i.e. 23 h 56 min 4.1 s ou 0.986°/j et l'inclinaison du satellite est nulle). Ces satellites géostationnaires ont une altitude d'environ 36 000 kilomètres et se déplacent à une vitesse qui correspond à celle de la Terre, donnant ainsi l'impression qu'ils sont stationnaires (e.g. METEOSAT, GOES). Cette configuration orbitale permet au satellite d'observer et d'amasser continuellement de l'information sur une région spécifique. Les satellites de communication et d'observation des conditions météorologiques sont situés sur de telles orbites. Trois satellites géostationnaires suffisent, s'ils sont bien

placés les uns par rapport aux autres, à « couvrir » l'ensemble des régions habitées de la terre, par exemple pour émettre des signaux de télévision.

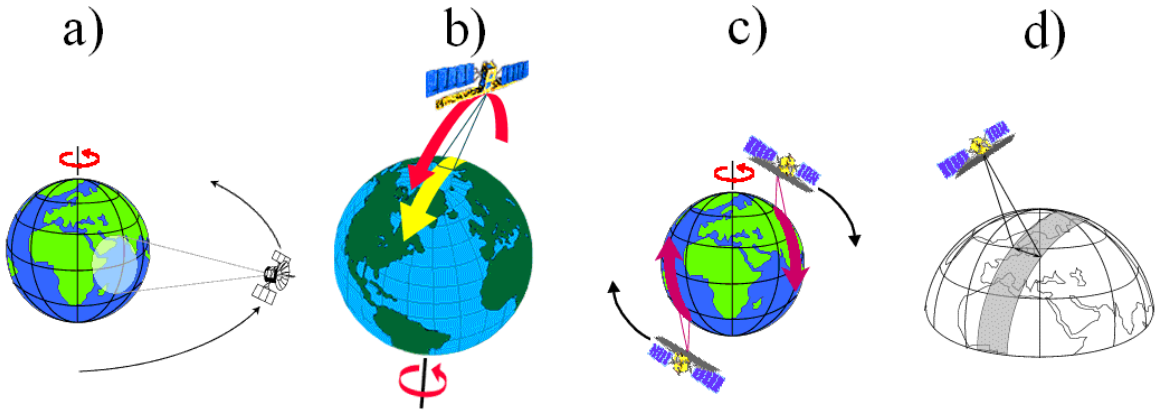

Figure A.4 : Caractéristiques d'un satellite (Soucre Centre Canadien de Télédétection, http://cct.rncan.gc.ca/)
a) Orbite géostationnaire
b) Orbite héliosynchrone
c) Orbite ascendante ou descendante
d) Fauchée

D'autres plates-formes spatiales suivent une orbite allant pratiquement du nord au sud ou vice versa (figure A.4.b). Cette configuration, combinée à la rotation de la terre (ouest- est), fait qu'au cours d'une certaine période, les satellites ont observé la presque totalité de la surface de la terre. Ce type d'orbite est appelé orbite quasi polaire à cause de l'inclinaison de l'orbite par rapport à une ligne passant par les pôles Nord et Sud de la terre. La plupart des satellites sur orbite quasi-polaires ont aussi une orbite héliosynchrone. (leur période de révolution est environ égale à celle de la rotation du soleil c.-à-d. 365 j 5 h 20 min). Ces satellites passent toujours à la même heure au-dessus du même point à une altitude d'environ 879 km (e.g. LANDSAT, SPOT). Ceci est un facteur important lorsqu'on compare deux images successives recueillies pour une saison particulière sur plusieurs années ou lorsqu'on produit une mosaïque avec des images adjacentes, puisque les images n'ont pas à être corrigées pour tenir compte des conditions similaires de l'illumination solaire.

De nos jours, la plupart des plates-formes satellitaires sont placées sur orbite quasi-polaires. Elles se déplacent donc vers le nord d'un côté de la terre, et vers le sud dans l'autre moitié de leur orbite. Ces deux types de passage du satellite se nomment respectivement orbite ascendante et orbite descendante (figure A.4.c). Si l'orbite est aussi héliosynchrone,

l'orbite ascendante du satellite se fait du côté ombragé de la terre, tandis que l'orbite descendante se fait du côté éclairé par le soleil. Les capteurs qui enregistrent l'énergie solaire réfléchie par la terre ne recueillent donc de l'information qu'au cours leur orbite descendante, lorsque le soleil illumine la terre. Par contre, les capteurs actifs qui possèdent leur propre source d'illumination ou les capteurs passifs qui enregistrent l'énergie émise par la planète (l'énergie infrarouge thermique par exemple) peuvent amasser des données autant lors des orbites ascendantes que descendantes de leurs satellites.

Le capteur d'un satellite en orbite autour de la terre, "observe" une certaine partie de la surface qui porte le nom de couloir couvert ou fauchée (figure A.4.d). Les capteurs sur plate-forme spatiale ont une fauchée dont la largeur varie généralement entre une dizaine et une centaine de kilomètres. Pour les satellites à orbite quasi-polaire, le satellite se déplace selon une trajectoire nord-sud. Cependant, vue de la terre, la trajectoire du satellite semble avoir une composante vers l'ouest à cause de la rotation de la terre. Ce mouvement apparent du satellite permet à la fauchée du capteur d'observer une nouvelle région à chacun des passages consécutifs du satellite. L'orbite du satellite et la rotation de la terre travaillent donc de concert, permettant une couverture complète de la surface de la planète après un cycle orbital complet.

A.2.3 Capteur

En télédétection, les capteurs sont des instruments qui permettent de mesurer le REM. Suivant le mode d'acquisition, on distingue:

(i) les capteurs passifs qui sont sensibles uniquement aux rayonnement d'origine naturelle usuellement réfléchie par le soleil ou l'énergie émise par un objet (e.g. radiomètres, caméras, etc.). Ils ne possèdent pas leur propre source d'éclairement et sont donc des récepteurs ;

(ii) les capteurs actifs qui possèdent leur propre source d'éclairement pour illuminer la cible et sont donc à la fois des émetteurs et des récepteurs. Les plus connus sont

les radar (Radio Detection and Ranging), lidar (Light Detection and Ranging), laser (Light Amplification by Stimulated Emited Radiation).

Ces capteurs sont classifiés en deux types : les imageurs et les non-imageurs (figure A.5).

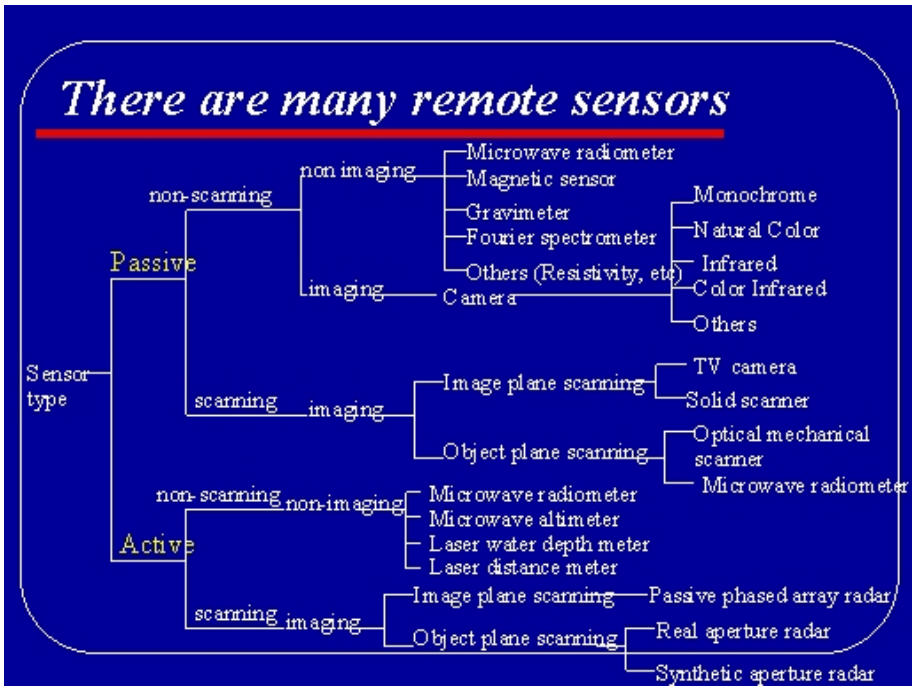

Figure A.5 : Différents types de capteurs (source NASA, http://rst.gsfc.nasa.gov/)

A.2.4 Résolution

Les données de télédétection sont décrites en terme de résolution. décrites par chaque capteur peuvent être décrites en termes de résolution radiométrique, spatiale, spectrale, temporelle.

La résolution radiométrique d'un système de télédétection décrit sa capacité de reconnaître de petites différences dans le spectre électromagnétique. Plus la résolution radiométrique d'un capteur est fine, plus le capteur est sensible à de petites différences dans l'intensité de l'énergie reçue.

La résolution numérique est très liée à la résolution radiométrique. Chaque quantité d'énergie mesurée en kilojoules est transformée en compte numérique ou valeurs radiométriques. Plus la gamme des comptes numériques est étendue, plus la résolution

numérique est grande. Par exemple, la gamme possible des valeurs enregistrée d'une image codée sur 8 bits est de 2^8=256 valeurs, ordonnées de 0 à 255.

Encore appelée taille du pixel, La résolution spatiale fait référence à la plus petite distance possible qui permette la discrimnation de deux objets voisins. Cette résolution s'exprime en mètres ou en kilomètres et mesure le coté d'un pixel. Ainsi une résolution grossière (plusieurs centaines de mètres voire plusieurs kilomètres) ne permettra pas de distinguer des objets spatiaux de petite taille (comme des bâtiments, des petites parcelles agricoles). A contrario, une résolution plus fine (quelques dizaines de mètres à un mètre) autorise par exemple l'identification des réseaux de communication les plus fins (routes, chemins...). Ainsi, pour une image à 20 mètres de résolution, chaque pixel représente un carré ayant 20 m de côté.

Encore appellée période de revisite, la résolution temporelle est la durée de révolution pour l'observation de la cible dans les mêmes conditions d'éclairement. Elle dépend des caractéistiques du capteur/satellite, de la largeur du recouvrement et de la latitude. Des caractéristiques spectrales de certains objets peuvent changer à travers le temps et ces changements peuvent être détectés par la collection et la comparasion d'images multitemporelles. Ceci permet par exemple de surveiller les changements d'origine naturelle (telles que la phénologie ou les inondations) ou anthropique (telles que l'urbanisation ou la déforestation). Le facteur temps est donc d'importance primordiale en imagerie lorsque

(i) l'ennuagement permanent limite les observations claires de la surface ;

(ii) des phénomènes de courte durée (inondations, marée noire) ont besoin d'être imagés ;

(iii) des comparaisons multitemporelles sont requises (e.g. progression d'une infection forestière d'une année sur l'autre) ;

(iv) le chanagement d'apparence d'un objet à trvers le temps peut être utilisé pour le distinguer d'un autre ayant des caratéristiques similaires très proches (mais versus mil).

La résolution spectrale d’un capteur est le nombre et la largeur des bandes spectrales utilisées par le système. L'eau ou la végétation par exemple ne nécessitent pas une résolution spectrale fine, à l'inverse des roches et minéraux réclament une résolution beaucoup plus grande. Si la résolution spectrale est trop grossière, il ne sera alors plus possible de bien différentier les différents minéraux. Les images panchromatiques sont thématiquement moins intéressantes que les images multispectrales. Effectivement, il est préférable d'avoir pluieurs bandes spectrales dont la combinaison de trois bandes peut conduire à des compositions colorées lors de la visualisation plutot qu’une seule bande spectrale qui ne peut qu’être réprésentée qu’en un seul niveau de gris. Il existe également des capteurs très perfectionnés capables de détecter des centaines de bandes spectrales, on les appelle capteurs hyperspectraux. Dans le domaine des micro-ondes, la résolution spectrale fait référence aux nombres de bandes disponibles, et à cause de leur capacité à accroître l’utilité des données, au nombre de combinaisons de polarisations qu’il est possible d’utiliser.

Annexe B : Evaluation d'une classification

Une carte est évaluée par comparaison des résultats classifiés avec des données indépendantes dites de référence. Il existe plusieurs méthodes de comparaion des classification. Une approche, très utilisée en télédétection, se base sur l'usage de la matrice de confusion ou matrice de contingence ou matrice d'erreur (voir tableau B 1). Lorsque les deux jeux de données possèdent le même nombre de clusters, les principaux éléménts issus d'une matrice de confusion comprennent l'indice de Cohen et les exactitudes selon l'utilisateur, le producteur, et totale.

Classe prédite \ Classe actuelle	A	B	C	D	Σ
A	n_{aa}	n_{ab}	n_{ac}	n_{ad}	n_{a+}
B	n_{ba}	n_{bb}	n_{bc}	n_{bd}	n_{b+}
C	n_{ca}	n_{cb}	n_{cc}	n_{cd}	n_{c+}
D	n_{da}	n_{db}	n_{dc}	n_{dd}	n_{d+}
Σ	n_{+a}	n_{+b}	n_{+c}	n_{+d}	n

Table B 1 Matrice de confusion

L'exactitude de l'utilisateur (user accuracy, UA) indique la précision de la classification pour une classe donnée:

$$UA = \frac{n_{ii}}{n_{i+}} \times 100\,, \qquad \text{(B1)}$$

où n_{i+} est le nombre total d'éléments qui ont été assignés à la catégorie *i* dans la validation, n_{ii} est le nombre d'éléments correspondant de la catégorie *i* entre les pixels classifiés et les pixels de référence.

L'exactitude du producteur (producer accuracy, PA) qualifie la précision avec laquelle une classe est bien classifiée dans le processus de classification :

$$PA = \frac{n_{ii}}{n_{+i}} \times 100 \,, \qquad \text{(B2)}$$

où n_{+i} est le nombre total d'éléments qui ont été assignés à la catégorie *i* dans la classification.

L'idéal serait d'abord que le UA et le PA soient identiques et par la suite que ces valeurs avoisinnent 1.

L'exactitude totale (overall agreement, OA) exprime la proportion des pixels pour lesquels les classes actuelles et les classes prédictes sont identiques :

$$OA = \frac{\sum_{i} n_{ii}}{n} \times 100 \,. \qquad \text{(B3)}$$

Une limitation de l'exactitude totale est la non-prise en compte de l'accord du à la chance. Cohen (1960) a défini le coefficient kappa encore appelé indice de Cohen de la façon suivante pour palier à cet inconvénient :

$$kappa = \frac{\sum_{k=1}^{q} n_{kk} - \sum_{k=1}^{q} \frac{n_{k+} \times n_{+k}}{n}}{n - \sum_{k=1}^{q} \frac{n_{k+} \times n_{+k}}{n}} \times 100 \qquad \text{(B4)}$$

L'équation A4 peut se réécrire sous la forme :

$$kappa = \frac{OA - Chance_{prob}}{1 - Chance_{prob}} \,, \qquad \text{(B5)}$$

où $Chance_{prob}$ est la probabilité selon laquelle l'accord est dû à la chance se calcule comme suit :

$$Chance_{prob} = \sum_{k=1}^{q} n_{k+} \times n_{+k} \qquad \text{(B5)}$$

Kappa indique donc la probabilité selon laquelle l'accord obtenu n'est pas du au hasard. Plus kappa est proche de 1, plus l'exactitude est forte.

Annexe C : Interpolation

Interpolation spatiale

La fonction $f(x,y,t)$ réprésente la valeur de la variable biophysique d'un pixel, localisé à la longitude *x* et à la latitude *y* et distribuée à l'instant t. On note $D_{C,t}\left((x,y);r\right)$ la boule fermée de rayon r centrée en (x,y) à l'instant *t* avec la distance de Manathan distance *d*, c.-à-d:

$$D_{C,t}\left[(x,y);r\right]=\left\{(\tilde{x},\tilde{y})\in C;\quad d\left[(\tilde{x},\tilde{y}),(x,y)\right]\le r,\ f\left(\tilde{x},\tilde{y},t\right)\in\Re\right\} \qquad \text{(C1)}$$

Nous cherchons le voisinage $\Omega_{C,t}$ contenu dans la classe d'occupation du sol *C* pour boucher les trous des cartes d'ECOCLIMAP-II de la variable $f(x,y,t)$, c-a-d:

$$\Omega_{C,t}\left(x,y\right)=\bigcap_{r\in\aleph^*} D_{C,t}\left[\,(x,y);r\,\right] \qquad \text{(C2)}$$

Nous définissons la valeur manquante comme étant une valeur pondérée des pixels valides:

$$f(x,y,t)=\frac{\sum_{\tilde{x},\tilde{y}\in\Omega_{C,t}} w(\tilde{x},\tilde{y})f\left(x,y,t\right)}{\sum_{\tilde{x},\tilde{y}\in\Omega_{C,t}} w(\tilde{x},\tilde{y})} \qquad \text{(C3)}$$

où

$$w(\tilde{x},\tilde{y})=\frac{1}{d\left[(\tilde{x},\tilde{y}),(x,y)\right]} \qquad \text{(C4)}$$

Interpolation temporelle

Initialement les paramètres LAI et albédo de MODIS ne sont pas distribués aux mêmes dates. En prenant *t* pour la date d'albédo, il est situé entre deux instants t_- et t_+ de production du LAI. Par exemple, la 1ère seizaine d'albédo de l'année est associée au jour julien 9. Durant cette période de composition, 2 valeurs de LAI correspondant aux jours

juliens 5 et 13 sont produites. Afin de faire correspondre le LAI avec l'albédo, une interpolation temporelle est systématiquement effectuée comme suit :

$$LAI(x,y,t) = LAI(x,y,t_-) + \frac{t-t_-}{t_+ - t_-}\left[LAI(x,y,t_+) - LAI(x,y,t_-)\right] \qquad \text{(C5)}$$

Annexe D : Bilans hydriques et énergétiques dans ISBA

D.1 Bilan énergétique

Dans ISBA, l'équation du bilan d'énergie s'écrit sous la forme :

$$R_n - H - LE = G\,, \qquad \text{(C.1)}$$

où R_n est le rayonnement net, H est le flux de chaleur sensible, LE est le flux de chaleur latente, et G est le flux de chaleur dans le sol.

Le rayonnement net ou quantité d'énergie radiative (de courte et grande longueur d'ondes) disponible à la surface de la terre dépend du rayonnement solaire (R_g) (dont une partie est réfléchie par la surface), de l'albédo de surface (A), du rayonnement atmosphérique (R_A), de l'émissivité de la surface (ε_s) et de la température de surface $\left(T_s\right)$:

$$R_n = R_g(1-A) + \varepsilon_s(R_A - \sigma T_s^4) \qquad \text{(C.2)}$$

Le flux de chaleur sensible (H) correspond au transfert de chaleur par turbulence entre la surface et l'air. Elle dépend de la densité de l'air $\left(\rho_a\right)$, de la capacité calorifique de l'air $\left(C_p\right)$, d'un coefficient de transfert turbulent $\left(C_H\right)$ fonction de la stabilité thermique de l'atmosphère, de la vitesse du vent $\left(V_a\right)$, du gradient de température (potentielle) près de la surface $\left(T_s - T_a\right)$:

$$H = \rho_a C_p R_a^{-1} V_a (T_s - T_a) \qquad \text{(C.3),}$$

où la résistance aérodynamique R_a est définie par

$$R_a = \left(C_H V_a\right)^{-1}, \qquad \text{(C.4)}$$

Le flux de chaleur latente $\left(LE\right)$ correspond à une quantité d'eau $\left(E\right)$ évaporée par unité de temps et unité de surface et convertie en énergie. La chaleur latente d'évaporation $\left(L\right)$

est la quantité de chaleur requise pour transformer une unité de liquide en vapeur à température constante (isotherme).

Le flux de chaleur par conduction dans le sol (*G*) dépend de la conductivité thermique du sol (C_T), du gradient de température entre une profondeur de référence (ici T_2) et la surface du sol (T_s) :

$$\frac{\partial T_s}{\partial t} = C_T G - \frac{2\pi}{\tau}\left(T_s - T_2\right), \qquad \text{(C.5)}$$

$$\frac{\partial T_2}{\partial t} = \frac{1}{\tau}\left(T_s - T_2\right) \qquad \text{(C.6)}$$

τ étant la constante temporelle d'un jour.

D.2 Bilan hydrique

Le bilan hydrique dans ISBA obéit à l'équation :

$$I_r = E + R_T + \frac{\Delta SoilMoist}{\Delta t}, \qquad \text{(C.7)}$$

où *E* est l'évaporation, *SoilMoist* l'humidité du sol et R_t le ruissellement total qui est égal à :

$$R_t = D + Q_s, \qquad \text{(C.8),}$$

où *D* est le Drainage et Q_s le ruissellement.

L'humidité superficielle SoilMoist s'obtient comme suit :

$$SoilMoist = \rho_w\left(w_2 d_2 + w_3\left(d_3 - d_2\right)\right), \qquad \text{(C.9)}$$

où ρ_w est la densité de l'eau liquide.

D. 2. 1. L'évaporation

L'évaporation $\left(E\right)$ est la somme des évaporations provenant du sol nu $\left(E_{soil}\right)$ et de la végétation $\left(E_v\right)$:

$$E = E_v + E_{soil} \qquad \text{(4.10)}$$

L'évaporation du sol nu E_{soil} est donnée par les expressions suivantes :

$$E_{soil} = (1 - veg)\rho_a R_a^{-1}(h_u q_{sat} - q_a), \qquad \text{(C.11)}$$

où ρ_a est la densité de l'air, q_a l'humidité spécifique de l'air, q_{sat} l'humidité spécifique à la saturation et h_u l'humidité relative de la surface, et

L'humidité relative de la surface h_u dépend du contenu en eau de la surface w_1 :

$$h_u = \begin{cases} \frac{1}{2}\left[1 - \cos\left(\frac{w_1}{w_{fc}}\pi\right)\right], & si \quad w_1 \prec w_{fc} \\ 1, & si \quad w_1 \geq w_{fc} \end{cases} \qquad \text{(C.12)}$$

L'évapotranspiration (E_v) est la somme de l'évaporation de la quantité d'eau retenue par les feuilles (E_{canop}) et de la transpiration (E_{transp}). Ainsi on a :

$$E_v = E_{canop} + E_{transp} \qquad \text{(C.13)}$$

Le flux d'évaporation de l'eau interceptée s'écrit :

$$E_{canop} = veg\ \rho_a \delta\ R_a^{-1}[q_{sat}(T_s) - q_a], \qquad \text{(C.14)}$$

où la fraction δ du feuillage recouverte par l'eau interceptée dépend du réservoir d'eau intercepté par la végétation en accord avec Deardorff (1977) :

$$\delta = \left(w_r / w_{r\max}\right)^{2/3}. \qquad \text{(C.15)}$$

Le flux de transpiration des plantes E_{transp} s'écrit :

$$E_{transp} = veg\ \rho_a \frac{1-\delta}{R_a + R_s}[q_{sat}(T_s) - q_a] \qquad \text{(C.16)}$$

où la résistance stomatique R_s est calculée suivant la méthode de Jarvis (1976) :

$$R_s = \frac{R_{s\min}}{LAI} F_1\ F_2\ F_3\ F_4 \qquad \text{(C.17)}$$

et les $(F_i)_{1\leq i\leq 4}$ sont des coefficients exprimant les dépendances vis-à-vis du stress hydrique, du rayonnement solaire, de la température et de l'humidité atmosphérique. Rsmin est constant et défini par type fonctionnel de plante.

D. 2. 2. Evolution du contenu en eau des différentes couches

Les équations suivantes permettent de décrire l'évolution temporelle des réservoirs d'eau :

$$\frac{\partial w_1}{\partial t} = \frac{C_1}{\rho_w d_1}\left(I_r - E_{soil}\right) - D_1 \quad \text{(C.18)}$$

$$\frac{\partial w_2}{\partial t} = \frac{1}{\rho_w d_2}\left(I_r - E_{soil} - E_{transp}\right) - K_2 - D_2 \quad \text{(C.19)}$$

$$\frac{\partial w_3}{\partial t} = \frac{d_2}{d_3 - d_2}\left(K_2 + D_2\right) - K_3 \quad \text{(C.20)}$$

où d_i est la profondeur de la couche *i* (*i*=1 couche de surface, *i*=2 zone racinaire, *i*=3 réservoir profond), w_i est le contenu volumétrique d'eau du sol pour la couche *i*, I_r est l'infiltration, K_i et D_i sont respectivement les coefficients de drainage et de diffusion de la couche *i*.

Ces coefficients peuvent s'écrire comme suit :

$$D_1 = \frac{C_2}{\tau}\left(w_1 - w_{eq}\right) \quad \text{(C.21)}$$

$$D_2 = \frac{C_4}{\tau}\left(w_2 - w_3\right) \quad \text{(C.22)}$$

$$K_3 = \frac{C_3}{\tau}\frac{d_3}{d_3 - d_2}\max\left(0; w_3 - w_{fc}\right) \quad \text{(C.23)}$$

$$K_2 = \frac{C_3}{\tau}\frac{d_3}{d_2}\max\left(0; w_2 - w_{fc}\right) \quad \text{(C.24)}$$

Les paramètres C_1, C_2, C_3 et C_4 sont les coefficients de la méthode « force-restore » qui dépendent de la texture du sol et de son contenu en eau (Boone et al., 1999).

Annexe E : Production scientifique

Articles parus dans des revues avec comité de lecture (1er auteur)

1. **Kaptué, T. A. T.**, Bégué, A., Roujean, J.-L, Garrigues, S., Sietse, O. L., Grey, W. M. F., Morisette, J. T., Boone, A., Mahfouf, J.-F. and Carrer, D. (2010), A revised land surface database (ECOCLIMAP-II) for use in climate and meteorological models. Part I: Generation of the biogeophisical parameters at the African continental scale, *Global Change Biology*, **submitted**.
2. **Kaptué, T. A. T.**, Boone, A. A., Faroux, S., Le Moigne, P., Kergoat, L., Alkama, M. R. and Roujean, J.-L (2010) Impact of a new MODIS-based physiography forcing on Land Surface Model simulations over the Western African Region, *Journal of Geophysical Research,* **submitted**.
3. **Kaptué, T. A. T.**, De Jong, S. M., Roujean, J.-L, Favier, C. and Mering, C. (2010), Ecosystem mapping at the African continent scale using a hybrid clustering approach based on 1-km resolution multi-annual data from SPOT/VEGETATION, *Remote Sensing of Environment*, **in press, doi:1016/j.rse.2010.09.015**.
4. **Kaptué, T. A. T**, Roujean, J.-L and De Jong, S. M. (2010), Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale, *International Journal of Applied Earth Observation and Geoinformation*, **accepted**.
5. **Kaptué, T. A. T**, Roujean, J.-L. and Faroux, S., (2010), ECOCLIMAP-II: an ecosystem classification and land surface parameter database of Western Africa at 1 km resolution for the Africa Monsoon Multidisciplinary Analysis (AMMA) project, *Remote Sensing of Environment*, **114**, 961-976.

Articles parus dans des actes de conférences

1. **Kaptué, A.**, Roujean, J.-L and Faroux, S. ECOCLIMAP-II programme: A new land cover classification at 1 km from MODIS and VEGETATION data time series over the western Africa in the frame of AMMA project, *International Geoscience And Remote Sensing Symposium*, **4**, 1023-1026, doi:10.1109/IGARSS.2009.5417555. Cape Town, 12-17 July 2009.
2. **Kaptué, A.**, Roujean, J.-L and Faroux, S. The ECOCLIMAP-II programme: A new land cover classification at 1 km from MODIS and VEGETATION data time series over the western Africa in the frame of AMMA project, 3rd *International Conference of African Monsoon Multidisciplinary Analyses, Ouagadougou 20-24 July 2009*.
3. **Kaptué, A.**, De Jong, S., Roujean, J.-L, Favier C and Mering, C. (2010) Ecosystem mapping at the African continental scale using a hybrid clustering approach based on 1-km resolution multi-annual data from SPOT/VEGETATION, *Remote Sensing and Photogrammetry Society Annual Conference*, Cork, August 31-September 3 2010.

PH.D DISSERTATION

AUTHOR : Armel Thibaut KAPTUE TCHUENTE

TITRE : Mapping ecosystem and biophysical variable satellite for the study of hydric flux over the African mainland

ADVISORE : Jean-Louis ROUJEAN

DISCIPLIN Remote sensing and physics of biosphere

LABORATORY ADRESS : Météo France/CNRM (CNRS)

42 Av. G. Coriolis 31057 Toulouse Cedex

PLACE AND DATE OF THESIS DEFENCE : CNRM, 10 november 2010

ABSTRACT :

In the context of climate change, the aim of this study is to characterize the heterogeneity of the African continent in order to provide some elements to better understand and quantify surfaces process acting on hydric fluxes. This work is intented to update the double ECOCLIMAP-I database which is constituted by a land cover map and a dassets of land biophysical parameters. To this end, we use remotely sensed data acquired by the latest generation sensors MODIS and SPOT/VEGETATION between 2000 and 2007.

During the first step, two methods of classifications has been developed for the mapping of land cover classes. The first method, which is supervised, is obtained by combining information provided by the both global land cover map GLC2000 and ECOCLIMAP-I using an interactive analysis of MODIS leaf area index (LAI). It has been performed in the framework of the AMMA project to discriminate ecosystems over the western African Region. The second method is hybrid in that it combines k-NN clustering, hierarchical principles and the Fast Fourier Transform (FFT) on the basis of multi-annual NDVI data from SPOT/VEGETATION to identify ecosystems at the whole African continent.

Then, methods for the estimation of land surface biophysical variables such as albedo, fractional vegetation cover and leaf area index has been developed and/or applied over the mainland. A statistical approach allow us to disentagle land surface albeo as a linear combination of bare soil albedo and vegetation albedo weighted by the fractional vegetation coverage as required in land surface models. After the application of the latter approach over the western african region, we demonstrate the robustness of the method by applying it over the entire mainland.

The sensitivity of two land surface scenarios was evaluated: scenarios with the ECOCLIMAP-I database and simulations with the new land surface boundary conditions developed over the western African Region. This was done to test the sensitivity of elements of the water cycle and energy balance to change in vegetation with the ISBA land surface model using the same atmospheric forcing for the both simulations. Heat and latent flux are mainly driven by the fractional vegetation coverage. The land surface model ISBA can be used to predict the impact of land cover change and accordingly the anthropic pressure on hydric fluxes.

KEY WORDS : Classification, ecosystem, Africa, leaf area index, bare soil albedo, vegetation albedo, fractional vegetation coverage, hydrologic cycle, energy budget, space remote sensing, MODIS, SPOT/VEGETATION

THESE DE DOCTORAT

AUTEUR : Armel Thibaut KAPTUE TCHUENTE

TITRE : Cartographie des écosystèmes et paramètres biophysiques satellitaires pour l'étude des flux hydriques sur le continent africain

DIRECTEUR DE THESE : Jean-Louis ROUJEAN

DISCIPLINE Télédétection et physique de la biosphère

INTITULE ET ADRESSE DU LABORATOIRE : Météo France/CNRM (CNRS)

42 Av. G. Coriolis 31057 Toulouse Cedex

LIEU ET DATE DE SOUTENANCE : CNRM, le 10 novembre 2010

RESUME :

Dans le contexte des changements climatiques, l'objectif du travail effectué est de caractériser l'hétérogénéité du continent africain afin de mieux comprendre et quantifier les processus de surface agissant sur les flux hydriques. Ce travail s'inscrit dans le cadre de la mise à jour de la base de données double ECOCLIMAP-I constituée d'une carte d'occupation des sols et des cartes de paramètres biophysiques. Pour cela, on s'appuie sur des données de télédétection acquises par les capteurs de dernière génération MODIS et SPOT/VEGETATION entre 2000 et 2007.

Dans un premier temps, Deux techniques de classifications ont ainsi été mises sur pied afin de cartographier les différentes classes d'occupation des sols. L'une, supervisée, a été conduite dans le contexte du programme AMMA afin de discriminer les écosystèmes sur la région ouest-africaine en combinant l'information complémentaire contenue dans les cartes d'occupation du sol GLC2000 et ECOCLIMAP-I par analyse interactive de l'indice foliare (LAI) MODIS. L'autre, hybride utilise les principes de regroupement hiérachique et dynamique de manière automatique en combinant l'usage du classificateur k-NN et celui de la transformée de Fourier Discrète sur la base des données d'indice de végétation normalisé (NDVI) SPOT/VEGETATION pour identifier les écosystèmes africains.

Puis des méthodes d'estimation des paramètres biophysiques tels que l'albédo, la fraction de végétation, l'indice foliaire onte été developpées et/ou appliquées sur le continent. Une approche statistique permet de décomposer l'albédo de surface comme combinaison linéaire de l'albédo du sol nu et de l'albédo de la végétation pondéré par la fraction de végétation tel que cela est requis dans les modèles de surface. La méthode a d'abord été appliquée sur la région ouest-africaine et sa robustesse a été prouvée lors de son application à l'intégralité du continent africain.

Ces conditions de surface ont ensuite été implémentées dans le modèle de surface ISBA pour reproduire les processus de surface. La sensibilité d'ISBA au forçage physiographique a été effectuée en analysant deux simulations avec le même forçage atmosphérique sur la région ouest-africaine : l'une en utilisant la classification ECOCLIMAP-I et l'autre en utilisant la nouvelle paramétrisation de la surface développée sur la région ouest-africaine. Les flux de chaleur latente et sensible sont principalement pilotés par la fraction de végétation. Le modèle ISBA peut être utilisé pour prédire l'impact d'un changement d'occupation du sol et par conséquent des actions anthropiques sur le bilan hydrique.

MOTS CLES : Classification, écosystèmes, Afrique, indice folaire, albédo du sol nu, albédo de la végétation, fraction de végétation, cycle hydrologique, bilan énergétique, télédétection spatiale, MODIS, SPOT/VEGETATION