THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT - AGROPARISTECH

N°: 2018AGPT0002

En Géomatique

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau – n°584
Portée par l'Université de Montpellier

Unité de recherche TETIS – Territoires, Environnement, Télédétection et Information Spatiale

Une approche multiscalaire par télédétection pour la cartographie et la caractérisation des systèmes agricoles à l'échelle régionale

Présentée par Beatriz BELLÓN DE LA CRUZ

Le 24 mai 2018

Sous la direction d'Agnès BÉGUÉ

Devant le jury composé de

Laurence HUBERT-MOY, Professeur des Universités, Université Rennes 2	Rapporteure
Laurent POLIDORI, Directeur de Recherche, CNRS	Rapporteur
Agnès BÉGUÉ, Cadre scientifique des EPIC, CIRAD	Directrice de thèse
Anne-Elisabeth LAQUES, Directeur de Recherche, IRD	Présidente du jury
David SHEEREN, Maître de Conférences, INPT/ENSAT	Examinateur
Florent MARAUX, Cadre scientifique des EPIC, CIRAD	Examinateur
Margareth SIMOES, Professeur des Universités, Université de Rio de Janeiro	Invitée

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

UNIVERSITÉ DE MONTPELLIER

Thèse réalisée au sein de :





UMR TETIS (Territoires, Environnement, Télédétection et Information Spatiale)
CIRAD - Département Environnements et Sociétés (ES)
Maison de la Télédétection
500 Rue Jean-François Breton
34093 Montpellier Cedex 5
France



« C'est en haut de l'arbre que l'on voit mieux le paysage »

V.LEBOURGEOIS, 2017.

Remerciements

De nombreuses personnes ont contribué, de façon directe ou indirecte, au bon aboutissement de cette thèse, c'est ainsi un grand plaisir de leur rendre hommage en leur dédiant ces quelques mots de remerciement.

En premier lieu, je tiens à remercier les personnes qui m'ont confié ce sujet de thèse, qui m'ont encadrée au cours de ces 3 dernières années et accompagnée dans son bon déroulement. Mes remerciements s'adressent tout d'abord à ma directrice de thèse, Agnès Bégué. Avec un admirable savoir faire scientifique, une véritable passion par la télédétection et des superbes qualités humaines, Agnès a rempli son rôle de directrice à la perfection et bien plus encore. Elle m'a toujours offert une écoute attentionnée, elle a prodigué de précieux conseils, réalisé des relectures assidues et des remarques toujours constructives, le tout dans une grande sérénité et avec un esprit toujours positif. Merci pour ton soutien, et pour la grande liberté d'action que tu m'as accordée. Merci également pour tout le reste, l'appui avec les nombreux casse-têtes administratifs, la mise en relation avec tant de personnes en France et à l'international et tous les bons moments partagés au bureau, en mission et lors des 'convites' conviviaux à la maison.

Je remercie également mes encadrants Danny Lo Seen et Valentine Lebourgeois pour leur soutien inestimable, leur amabilité et sympathie de tous les jours. Merci pour vos diverses relectures et vos excellents apports lors des réunions de travail qui m'ont aidé à prendre du recul sur les différentes questions de recherche soulevées et les choix méthodologiques. Votre esprit créatif et critique a stimulé l'ouverture de voies à explorer et a enrichi ce travail. Cela a été un véritable plaisir de travailler avec vous. Un grand merci Valentine pour m'avoir encouragé à me lancer dans cette thèse et m'avoir ouvert les portes de cette belle opportunité !

J'adresse aussi mes remerciements à toute l'équipe de l'UMR TETIS qui m'a chaleureusement accueilli au sein de la Maison de la Télédétection (MTD) et a participé de près ou de loin à cette thèse. Je tiens particulièrement à remercier Jean-Philippe Tonneau, pour la confiance qu'il m'a accordée dès mes premiers pas à la MTD et sa bienveillance. Merci pour ton implication avec les jeunes, pour tous les échanges pédagogiques sur les zonages mais aussi sur la vie. Je remercie également Raffaele Gaetano et Audrey Jolivot pour leur appui. Merci pour votre disponibilité et le temps que vous avez consacré à réaliser différents tests sur le Burkina, entre autres. Ma reconnaissance va également à Monsieur le Baron, mon fidèle camarade de cantine. Merci Christian pour toutes tes attentions et ton écoute, pour tes conseils avisés et pour ta bonne humeur au quotidien, source de décontraction neuronale indispensable pendant la thèse ! Merci aussi à Sandra Luque pour m'avoir donné les clés nécessaires pour aborder la notion de paysage. Enfin, je remercie infiniment tout l'appui logistique et administratif d'Alain Clopès, Annie Huguet, Jean Claude Lorente, Pierre-Yves Vyon, Véronique Jalabert, Frans Brouwers et Isabelle Altier, qui ont assuré des conditions de travail exceptionnelles. Merci Isa pour ton accueil chaleureux à la MTD et pour

tous ces moments sympathiques au travail et en dehors. Grand merci Annie pour ton amabilité et ton travail dévoué, y muchísimas gracias pour le coup de main le jour de la soutenance !

Ma reconnaissance va également aux rapporteurs de cette thèse, Laurence Hubert-Moy et Laurent Polidori pour leur relecture approfondie du manuscrit et la pertinence de leurs conseils, ainsi qu'aux autres membres du jury, Anne-Elisabeth Laques, David Sheeren et Florent Maraux pour leurs observations et suggestions constructives qui ont enrichi ce travail sur différents plans. Merci également aux membres du comité de suivi individuel, Damien Arvor, Claudio de Almeida, Delphine Leenhardt et Xavier Augusseau, pour l'intérêt qu'ils ont porté à ce travail et les avis éclairés qu'ils ont fournis à chaque comité.

Je tiens à remercier toutes les personnes qui m'ont accompagné d'une façon ou d'une autre lors des missions au Brésil et au Burkina Faso. Merci à Margareth Simões et Rodrigo Peçanha Demonte Ferraz d'EMBRAPA Solos pour leur implication dans cette thèse, leur aide précieuse lors de la prise de contact avec le terrain 'Tocantinense' et leur hospitalité. Je salue également l'accompagnement de Balbino Antônio Evangelista et Francelino Petenó de Camargo au nom de toute l'équipe du 'Núcleo temático de sistemas agrícolas' d'EMBRAPA Pesca e Aquicultura pendant les missions de terrain au Tocantins. Obrigada por seu apoio no campo, por sua calorosa acolhida, por todos os ensinamentos sobre a agricultura Tocantinense e a cultura brasileira, pela sua gentileza e todos os bons momentos compartilhados! Côté Burkina j'aimerais remercier Louise Leroux, ma partenaire de mission, pour l'organisation sans faille de la campagne de terrain du nord au sud et pour avoir assuré la prise de points lorsque la fièvre m'annulait ! Merci également à Théo Saunier pour l'appui logistique et à Ibra Touré, Denis Gautier et Patrice Sanou pour le temps que vous m'avez accordé.

C'est aussi à tous les amis co-thésards et collègues de la MTD que je souhaite apporter mes remerciements, responsables de la bonne ambiance au labo et sans lesquels ces années auraient été certainement moins réjouissantes. Merci à Jacques, Christian, Florian, Jean-Baptiste, Coralie, Sávio, Claudio, Patrick et Guilhem. Cela fût un vrai plaisir de partager tant de bons moments et de conversations enrichissantes au labo et en dehors ! Un énorme merci à Camille, Louise, Marc, Mathieu et Arthur, mes chers co-bureaux, pour les dépannages et coups de main multiples, les relectures du français et les pauses-détente au Lez. Merci pour tout ce que vous m'avez apporté et merci spécialement à toi Camille pour tous nos périples héraultais et pour le soutien moral pendant la thèse en m'apprenant entre autres qu'il faut « tout relativiser tout le temps » (difficile à maîtriser mais offrant tant de réconfort !).

Enfin, je tiens à remercier mes amis de Montpellier, d'Espagne, en diaspora autour du monde et ma famille. Je dédie ce travail à mes parents, Enriqueta y Juan et les remercie pour leur soutien affectif et inconditionnel pendant toutes mes années d'études. Gracias por todos los esfuerzos que habéis hecho por mí, por vuestro amor y dedicación. Gracias a ti, Iván, compañero, motivador e inspirador, por el camino recorrido estos años, tu generosidad y apoyo constante.

Titre : Une approche multiscalaire par télédétection pour la cartographie et la caractérisation des systèmes agricoles à l'échelle régionale.

Mots-clés : Systèmes agricoles, Utilisation des terres, Unités de paysage, Clustering, Série temporelle de NDVI, Analyse d'image orientée-objet.

Résumé : Dans un contexte de planification de l'utilisation des terres à l'échelle régionale, la cartographie des systèmes agricoles - espèces cultivées et pratiques culturales - permet de suivre ce qui est produit, où et comment, et constitue donc un élément essentiel d'évaluation régionale de la production et de son impact sur l'environnement. La production d'information sur les systèmes agricoles nécessite généralement beaucoup de données et d'expertise. Cette information est donc très hétérogène en quantité et en qualité dans l'espace et le temps, la disponibilité et les mises à jour étant extrêmement variables selon les pays et les régions. La télédétection, de par sa capacité à fournir une information spatiale synoptique sur l'état et la dynamique de la végétation à partir des images satellitaires, constitue un outil précieux pour le suivi de l'agriculture. Toutefois, la conversion des images en produits cartographiques à l'échelle régionale reste encore du domaine de la recherche pour de nombreuses applications.

Cette thèse propose des développements méthodologiques originaux dans une approche multiscalaire semi-automatique basée sur le traitement et l'analyse d'imagerie satellitaire optique pour la cartographie et la caractérisation des systèmes agricoles à l'échelle régionale. L'approche est composée de deux méthodes principales : (i) stratification régionale en unités de paysage et classification de ces unités pour produire une carte de systèmes d'utilisation agricole des terres ; (ii) segmentation à l'échelle de la parcelle et classification non supervisée des segments par une méthode de « landscape-clustering » pour produire une carte de systèmes de culture. Les méthodes ont été développées sur une région d'agriculture intensive, l'État brésilien du Tocantins, où le domaine cultivé, ainsi que les principaux systèmes d'utilisation agricole des terres et systèmes de culture ont été cartographiés avec succès à partir d'une série annuelle d'images NDVI-MODIS et d'une mosaïque d'images Landsat. La reproductibilité de l'approche a ensuite été évaluée au Burkina Faso, où les paysages sont façonnés par la petite agriculture familiale. Seul le domaine cultivé a pu être cartographié avec des résultats satisfaisants, mettant en évidence les limites de ces méthodes et de l'offre actuelle en imagerie satellitaire face aux spécificités contraignantes de ce type d'agriculture pour la télédétection. Les cartes résultantes ont été évaluées avec des données de vérité terrain et des statistiques agricoles, et comparées à d'autres produits cartographiques existants.

Les résultats de cette thèse montrent le potentiel de la nouvelle méthode de stratification régionale en unités de paysage qui, à partir de séries temporelles de NDVI et combinée à la méthode de classification de « landscape-clustering », contribue à améliorer de façon significative la discrimination des espèces cultivées et des pratiques agricoles, et permet de représenter les systèmes agricoles à différents niveaux d'organisation spatiale. L'originalité des méthodes développées réside principalement dans la simplicité de leur mise en œuvre. Elles sont presque exclusivement basées sur des données satellitaires et nécessitent peu d'intervention « experte » et de données externes, ce qui leur confère un fort potentiel de reproductibilité. Cette thèse contribue ainsi, avec ces nouvelles méthodes, au développement d'outils génériques pour soutenir le suivi de l'agriculture à grande échelle et fournir des produits d'aide à la décision pour une planification raisonnée de l'utilisation des terres.

Title: A multiscale remote sensing approach for agricultural systems mapping and characterization at regional scale.

Keywords: Agricultural systems, Land use, Landscape units, NDVI time series, Clustering, Object-based image analysis.

Abstract: In a context of regional land-use planning, agricultural systems' mapping - crop types and cropping practices – allows monitoring of what is being produced, where and how, and therefore represents a key element for regional assessment of the agricultural production and its environmental impact. The production of information on agricultural systems generally requires a lot of data and expertise. This information is thus very heterogeneous in quantity and quality in space and time, the availability and updates being extremely variable between countries and regions. Remote sensing, through its ability to retrieve synoptic spatial information on the state and dynamics of vegetation from satellite images, represents a valuable tool for agricultural monitoring. However, the conversion of images into regional-scale map products remains a field of research for many applications.

This thesis presents original methodological developments in a semi-automatic multiscale approach based on the processing and analysis of optical satellite imagery for the mapping and characterization of agricultural systems at regional scale. The approach is composed of two main methods: (i) regional stratification into landscape units and classification of these units to produce a map of agricultural land-use systems; (ii) field-level segmentation and unsupervised classification of the segments by a "landscape-clustering" method to produce a cropping systems' map. The methods were developed on a region of intensive agriculture, the Brazilian state of Tocantins, where the cultivated area, as well as the main agricultural land-use systems and cropping systems were successfully mapped from an annual NDVI-MODIS time series and a mosaic of Landsat images. The reproducibility of the approach was then assessed in Burkina Faso, where landscapes are shaped by the smallholder agriculture. Only the cultivated area could be mapped with satisfactory results, highlighting the limitations of these methods and the current offer in satellite imagery given the challenging specificities of this type of agriculture for remote sensing. The resulting maps were assessed with ground-truth data and agricultural statistics, and compared to other existing maps.

The results of this thesis show the potential of the new method of regional stratification into landscape units which, based on NDVI time series and combined to the unsupervised "landscape-clustering" classification method, contributes to significantly improve discrimination of crop types and agricultural practices, and allows representing the agricultural systems at different levels of spatial organization. The originality of the developed methods lies mainly in the simplicity of their implementation. They are almost exclusively based on satellite data and require little "expert" intervention and external data, which gives them a high reproducibility potential. Thereupon, this thesis contributes, with these novel methods, to the development of generic tools to support large-scale monitoring of agriculture and to provide decision-support products for reasoned land-use planning.

Table des matières

Remerciements i
Résumé : iii
Abstract: iv
Table des matières v
Liste des figures ix
Liste des tableaux xiv
Liste d'abréviations, sigles et acronymes xv
Introduction 2
CONTEXTE DE LA THÈSE 2
PRÉSENTATION GÉNÉRALE DE LA THÈSE 3
ORGANISATION DE LA THÈSE 5
Chapitre 1. Contexte et problématique de la thèse 8
1.1. L'AGRICULTURE FACE AUX ENJEUX GLOBAUX 8
1.2. LA CARTOGRAPHIE ET LA CARACTÉRISATION DE SYSTÈMES AGRICOLES 11
1.2.1. Les systèmes agricoles : concepts et définitions 11
1.2.2. L'intérêt de la cartographie et la caractérisation de systèmes agricoles 13
1.2.3. L'utilisation agricole des terres 15
1.2.4. La dimension spatio-temporelle dans la cartographie de systèmes agricoles : le concept d'échelle 15
1.2.5. Le défi de la cartographie de systèmes agricoles à l'échelle régionale 19
1.3. LA TÉLÉDÉTECTION ET LA CARTOGRAPHIE DE SYSTÈMES AGRICOLES 28
1.3.1. L'image satellitaire : source d'information sur les systèmes agricoles 28
1.3.2. État de l'art des approches de télédétection pour la cartographie de systèmes agricoles 34
1.4. OBJECTIF ET QUESTIONS DE RECHERCHE 42
1.5. APPROCHE GÉNÉRALE 44
Chapitre 2. Régions d'étude, données et outils 48
2.1. LES RÉGIONS D'ÉTUDE 48
2.1.1. L'État brésilien du Tocantins 48
2.1.2. Le Burkina Faso 54
2.2. LES DONNÉES 61
2.2.1. Données satellitaires 62
2.2.2. Carte d'occupation des terres 72
2.2.3. Données de validation 73
2.3. LES OUTILS 77
2.3.1. L'analyse en composantes principales 77
2.3.2. Les coefficients de relation entre variables 79

2.3.3. L'algorithme de segmentation multi-résolution 80
2.3.4. L'algorithme de classification non supervisée k-means 81
2.3.5. Les indices d'évaluation de classifications 82
Chapitre 3. Stratification et caractérisation des systèmes d'utilisation agricole des terres à l'échelle régionale 88
3.1. MÉTHODE ADOPTÉE 89
3.1.1. Délimitation des unités de paysage 90
3.1.2. Classification des unités de paysage 92
3.1.3. Évaluation des résultats de classification 94
3.2. RÉSULTATS 94
3.2.1. Les composantes principales 94
3.2.2. Les unités de paysage 99
3.2.3. Les systèmes d'utilisation agricole des terres 100
3.2.4. Résultats de l'évaluation de la classification 103
3.3. DISCUSSION 105
Chapitre 4. Classification des systèmes de culture à l'échelle régionale par la méthode «landscape-clustering» 110
4.1. MÉTHODE ADOPTÉE 111
4.1.1. Segmentation 112
4.1.2. Classification non supervisée 115
4.1.3. Évaluation des résultats de classification 117
4.2. RÉSULTATS 118
4.2.1. Résultats de classification 118
4.2.2. Résultats d'évaluation de la précision 120
4.2.3. Carte des systèmes de culture obtenue par « landscape-clustering » 123
4.3. DISCUSSION 124
Chapitre 5. Cas d'application au Burkina Faso 130
5.1. ÉTAT DE L'ART DES APPROCHES DE TÉLÉDÉTECTION POUR LA CARTOGRAPHIE DU DOMAINE AGRICOLE AU BURKINA FASO 130
5.1.1. Exemples de cartographie à l'échelle globale et continentale 130
5.1.2. Exemples de cartographie à l'échelle régionale et nationale 132
5.1.3. Défis pour la télédétection 133
5.2. ÉVALUATION DE LA REPRODUCTIBILITÉ DE L'APPROCHE DÉVELOPPÉE 134
5.2.1. La taille du parcellaire et la variabilité intra-parcellaire 134
5.2.2. La persistante couverture nuageuse 139
5.2.3. La difficile séparabilité des classes 139
5.2.4. Adaptation des méthodes 139
5.3. RÉSULTATS 142
5.3.1. Les composantes principales 142

5.3.2. Les unités de paysage 145
5.3.3. Résultats de classification 145
5.4. DISCUSSION 150
5.4.1. Perspectives 154
Conclusions et perspectives 158
CONTRIBUTIONS DE LA THÈSE 158
Apports méthodologiques 158
Enseignements tirés du cas d'application au Burkina Faso 160
Apports thématiques 161
PERSPECTIVES 163
Perspectives de recherche 163
Perspectives opérationnelles 166
Références bibliographiques 169
Annexe - Publications 193

Liste des figures

Figure 1. Exemples d'enjeux et défis de l'agriculture dans un contexte d'augmentation de la population mondiale. 9
Figure 2. Estimations du pourcentage de variation des rendements pour 11 grandes cultures entre la période de 1996-2005 et 2046-2050. Les valeurs ont été obtenues par une moyenne des résultats de cinq modèles climatiques et trois scénarios d'émission de gaz à effet de serre (Müller *et al.*, 2010; World Bank, 2010). 10
Figure 3. Hiérarchie de systèmes agricoles, et leur niveau d'organisation (adaptée de Fresco et Westphal (1988)). Les termes anglais sont mis en correspondance avec des termes français analogues, même si les concepts qui leur sont associés sont parfois employés différemment selon les groupes de recherche français et anglo-saxons (Fresco, 1984). 12
Figure 4. De la donnée spatiale à l'aménagement des territoires. 14
Figure 5. Exemple d'hiérarchie de niveaux d'organisation des systèmes agricoles (alignées à gauche) et unités spatiales ou échelles opérationnelles correspondantes (alignées à droite). 17
Figure 6. Systèmes d'utilisation des terres d'Afrique subsaharienne. ©FAO/UNEP GEF (source : LADA, 2008). 21
Figure 7. Classifications des systèmes de production du Burkina Faso (en haut à droite) et du Tocantins (en bas à droit) extraites des produits "*Major farming systems of Sub-Saharan Africa*" (en haut à gauche) et "*Major farming systems of South America*" (en bas à gauche) (adaptée de Dixon *et al.* (2001)). 23
Figure 8. Produit du zonage de moyens d'existence au Burkina Faso développé par l'USAID (source : FEWS NET/USAID, 2010). 24
Figure 9. Extrait des attributs associés aux strates du Tocantins (en bas à gauche) et du Burkina Faso (en bas à droite) du produit GAES niveau 4 (en haut) (adaptée de FAO (2017a); Mücher *et al.* (2016)). 26
Figure 10. Carte globale de la taille des parcelles agricoles (source : Fritz *et al.*, 2015). 33
Figure 11. Carte des quatre cultures commerciales et séquences culturales, de la saison de 2006-2007 sur les quatre régions agricoles principales du Mato Grosso (source : Arvor *et al.*, 2011). 36
Figure 12. Stratification spatiale de la France métropolitaine en unités de paysage à partir de données MODIS (source : Bisquert *et al.*, 2015). 40
Figure 13. Principales étapes de l'approche générale décrites dans les chapitres 3 et 4. 44
Figure 14. Localisation de l'État du Tocantins au centre-nord du Brésil et de sa capitale Palmas (point rouge), frontières avec les États limitrophes (trait gris) et limites des biomes de l'Amazonie et Cerrado (trait blanc) (sources : limites des états (IBGE, 2017) ; limites des biomes (IBGE, 2004) ; localisation capitale (SEPLAN, 2012) ; image de fond (mosaïque Landsat 8 OLI composée de scènes de juillet 2014 et 2015, USGS (s. d.)). 48
Figure 15. Localisation de la région de MATOPIBA (code de couleur orange), la nouvelle frontière agricole du Brésil, composée de parties des États du Maranhão (MA), Piauí (PI), Bahia (BA) et l'étendue complète du Tocantins (TO, rayé rouge) (sources : limites des états (IBGE, 2017) ; limites de MATOPIBA (de Miranda, Magalhães et de Carvalho, 2014)). 49

Figure 16. Exemples de différents types de pâturage au Tocantins : (a) pâturage naturel ; (b) pâturage planté ; (c) pâturage naturel dégradé ; (d) pâturage naturel dégradé par surpâturage (photographies réalisées par B.Bellón). 52
Figure 17. Clichés du Tocantins : (a) résidus de sorgho de couverture ; (b) système d'irrigation à pivot central sur une parcelle de canne à sucre ; (c) plantation d'hévéa (Hevea brasiliensis) ; (d) usine de traitement et raffinerie de soja ; (e) jeunes pousses de soja (cultivée par semis-direct) sur résidus de maïs ; (f) élevage de zébus à viande en confinement ; (g) bétail paissant dans un pâturage naturel ; (h) champ labouré et préparation du chaulage des sols (photographies réalisées par B.Bellón). 54
Figure 18. Localisation du Burkina Faso au centre de l'Afrique de l'Ouest, de sa capitale Ouagadougou (point rouge), des pays limitrophes et des zones climatiques (sources : limites des pays (FAO-UN, 1994), zones climatiques (Direction de la Météorologie du Burkina Faso), image de fond (mosaïque MODIS, produit MOD09A1 du 19/12/2014, NASA LP DAAC (s. d.-a)). 55
Figure 19. Calendrier saisonnier et configuration des pluies des zones du nord et sud-ouest du Burkina Faso, extraits du produit de 2010 des zonages de moyens d'existence au Burkina Faso (adaptée de FEWS NET/USAID, 2010). 57
Figure 20. Clichés du Burkina Faso : (a) troupeau composé de plusieurs espèces ; (b) sol latéritique impropre aux cultures ; (c) coton ; (d) mil ; (e) sorgo ; (f) sésame ; (g) parcelle d'un exploitant agro-éleveur ; (h) riz pluvial ; (i) récolte du niébé en association avec du mil (photographies réalisées par B.Bellón). 61
Figure 21. Emprise des tuiles MODIS couvrant Tocantins (TO) et Burkina Faso (BF) (source : limites administratives du Tocantins (IBGE, 2017) et du Burkina Faso (OCHA ROWCA, 2017) ; images de fond (NASA LP DAAC, s.d.-b)). 63
Figure 22. Illustration des bandes du produit MOD13Q1 téléchargées pour l'image composée 129 (correspondant au mois de mai) de 2015 pour le Tocantins, et l'image 129 de 2016 pour le Burkina Faso : NDVI ; bande de qualité (*Pixel reliability summary QA*) et ; jour julien (*Composite day of the year*). 66
Figure 23. Distribution spatiale des pixels considérés bruités dans la bande de qualité fournie avec le produit MOD13Q1 sur les séries annuelles de 23 dates de 2013-2014 (à gauche) et 2016 (à droite). 67
Figure 24. Profil de NDVI brut (avant lissage et interpolation) et lissée d'un pixel de la série annuelle couvrant la saison culturale de 2013-2014 du Tocantins. 68
Figure 25. Profil de NDVI lissé et corrigé (replacé sur des intervalles réguliers de 16 jours) d'un pixel de la série annuelle couvrant la saison culturale de 2013-2014 du Tocantins. 69
Figure 26. Emprise des tuiles Landsat 8 OLI couvrant Tocantins. L'identifiant des tuiles correspond au numéro de colonne et ligne de la grille de référence de Landsat (WRS2). La mosaïque libre de nuages obtenue à partir des 19 scènes est présentée en image de fond avec une composition colorée en fausses couleurs (RGB proche-infrarouge, vert, rouge) (source : image de fond (USGS, s.d.)). 70
Figure 27. Carte d'occupation des terres (OT) agricole du Tocantins extraite à partir de la carte TerraClass Amazônia 2014 (couvrant le nord-est du Tocantins) et de la carte TerraClass

Cerrado 2013 (couvrant le reste de l'État) (sources : TerraClass Amazônia 2014 (INPE, s.d.-a) ; TerraClass Cerrado 2013 (INPE, s.d.-b)). 73
Figure 28. Localisation et nombre de points GPS enregistrés dans la base de données spatiale du Tocantins lors de la campagne de terrain réalisée en octobre 2015. Les classes associées aux points sont celles utilisées pour la validation des cartes produites (source : limites de l'État, limites des municipalités et localités (SEPLAN, 2012)). 75
Figure 29. Localisation et nombre de points GPS enregistrés dans la base de données spatiale du Burkina Faso lors de la campagne de terrain réalisée en septembre 2016. Les classes associées aux points sont celles utilisées pour la validation des cartes produites (source : limites du pays et limites des provinces (OCHA ROWCA, 2017) ; localités (OCHA ROWCA, 2015)). 76
Figure 30. Représentation des étapes méthodologiques présentées dans le chapitre 3. 90
Figure 31. Règles de l'arbre de décision utilisé pour la classification des unités de paysage en quatre principaux types de systèmes d'utilisation agricole des terres (adaptée de de Almeida *et al.* (2016)). 93
Figure 32. Magnitude des vecteurs propres (axe d'ordonnées) des trois premières composantes principales en fonction des images de la série annuelle de 2013-2014 (représentées dans l'axe d'abscisses par leur code du jour julien). 95
Figure 33. Les trois premières composantes principales dérivées de la série temporelle de NDVI de l'année 2013-2014. Les images CP sont représentées en niveaux de gris, illustrant la gradation des valeurs numériques allant des valeurs basses (la couleur noire correspond à la valeur minimale) aux valeurs élevées (la couleur blanche correspond à la valeur maximale) pour chaque CP. 97
Figure 34. Corrélation (r) entre la CP1 et le NDVI cumulé sur la période annuelle 2013-2014. 98
Figure 35. Stratification de la région du Tocantins en 90 unités de paysage, obtenues par segmentation des 22 images de Composantes Principales (CP). Le fond d'image est une composition colorée des trois composantes principales CP2, CP3, CP4 (RGB). 100
Figure 36. Profils temporels de NDVI des 22 unités de paysage agricole appartenant au domaine cultivé. Les figures montrent la moyenne (courbe noire continue) et l'écart-type (courbe gris foncé ombrée) des valeurs NDVI des pixels à l'intérieur du masque de la classe « agriculture annuelle » de la carte d'occupation des terres TerraClass, pour l'année culturale 2013-2014. Un profil d'une unité de paysage agricole représentatif de chacun des trois groupes de profils temporels de NDVI (a), (b) et (c) est présenté en détail à gauche, et les profils des autres unités de paysage appartenant à chaque groupe sont représentés à droite. Les pics de croissance des cycles culturaux sont représentés par des bandes verticales grises ombrées (d'une durée de 32 jours) qui sont identiques entre figures d'un même groupe. 101
Figure 37. Carte des 6 principaux systèmes d'utilisation agricole des terres de l'État du Tocantins (saison culturale 2013-2014). 103
Figure 38. Proportion moyenne des cultures dans les municipalités du premier type (gauche), deuxième type (centre) et troisième type (droite), associés respectivement aux systèmes de monoculture de soja, double culture soja-céréale et double culture riz-soja. 104

Figure 39. Répartition spatiale des trois principaux systèmes de culture dans la carte de référence au niveau des municipalités dérivée des statistiques agricoles (a) et dans la classification réalisée par télédétection (b). Les limites des municipalités sont incluses dans les deux cartes pour faciliter leur comparaison. 105
Figure 40. Représentation des étapes méthodologiques présentées dans le chapitre 4. 112
Figure 41. Échantillons des résultats de la segmentation de la mosaïque Landsat 8 sur différents paysages agricoles. L'image de fond correspond à une composition colorée en fausses couleurs (RGB proche-infrarouge, vert, rouge) de la mosaïque Landsat (source : image de fond (USGS, s.d.)). 113
Figure 42. Résultat de la stratification régionale en unités de paysage suivant la méthode décrite dans le chapitre 3..... 115
Figure 43. Profils temporels des dix clusters finaux obtenus avec (a) la méthode hyperclustering et (b) la méthode « landscape-clustering », et les classes des systèmes de culture correspondantes. Les courbes continues représentent le NDVI moyen de tous les objets appartenant à chaque cluster en fonction du jour julien. Les profils des clusters qui semblent avoir capté les trois principaux systèmes de culture du Tocantins sont représentés dans des figures séparées (i), (ii) et (iii) avec leur écart-type (courbes discontinues), et les profils des clusters qui semblent avoir capté d'autres types d'occupation et utilisation des terres (OT et UT) sont représentés dans la même figure (iv). 119
Figure 44. Régression linéaire simple entre les estimations de la superficie de soja récoltée en 2015 par municipalité, issues des statistiques agricoles, et la superficie occupée par les objets affectés aux trois classes de systèmes de culture du résultat de classification de la méthode « landscape-clustering » par municipalité. 122
Figure 45. Carte des principaux systèmes de culture du Tocantins à l'échelle de la parcelle pour la saison culturale de 2014-2015 obtenue avec la méthode de classification non supervisée « landscape-clustering ». 123
Figure 46. Carte du domaine cultivé en 2014 au Burkina Faso (source : Knauer *et al.*, 2017). 133
Figure 47. Représentation d'une parcelle de taille moyenne d'environ 100 ha au Tocantins et 0,6 ha au Burkina Faso, sur un extrait de 9 km^2 d'images MODIS NDVI et Landsat 8. 135
Figure 48. Représentation d'une parcelle de taille moyenne du Tocantins et du Burkina Faso sous une grille correspondant aux pixels MODIS NDVI et Landsat 8. 135
Figure 49. Emprise spatiale des tuiles MODIS, Landsat 8, Sentinel-2 et SPOT 6/7 sur la région du Burkina Faso. 137
Figure 50. Extrait de la segmentation réalisée sur une image SPOT 7 dans le cadre du projet JECAM (*Joint Experiment of Crop Assessment and Monitoring*) (Raffaele Gaetano, communication personnelle 2017)..... 138
Figure 51. Représentation des étapes méthodologiques adoptées dans le chapitre 5. 141
Figure 52. Magnitude des vecteurs propres (axe d'ordonnées) des trois premières composantes principales en fonction des images de la série annuelle de 2016 (représentées en abscisse par leur code du jour julien). 143
Figure 53. Corrélation (r) entre la CP1 et le NDVI cumulé sur l'année 2016. 144

Figure 54. Résultat de la stratification régionale en unités de paysage obtenu suivant la méthode décrite dans le chapitre 3 à partir d'une série temporelle de données MOD13Q1 de l'année 2016. 145
Figure 55. Profils temporels des vingt clusters finaux obtenus avec la méthode « landscape-clustering » adaptée au Burkina Faso. Les courbes continues représentent le NDVI moyen de tous les pixels appartenant à chaque cluster en fonction du jour julien. Les profils des clusters qui semblent avoir capté les cultures annuelles sont représentés dans la figure (a) et les profils des clusters qui semblent avoir capté d'autres types d'occupation et utilisation des terres sont représentés dans les figures (b) et (c). 146
Figure 56. Carte du domaine cultivé (cultures annuelles) à 250 m de résolution spatiale pour la saison culturale de 2016 obtenue avec la méthode de classification non supervisée « landscape-clustering » et limites des unités de paysage (cf. Figure 54). 148
Figure 57. Extrait du résultat de classification du domaine cultivé de 2016 au nord du Burkina Faso..... 149
Figure 58. Extrait du résultat de classification du domaine cultivé de 2016 et limites des espaces protégés (source : limites des espaces protégés (IUCN/UNEP-WCMC, 2017)). 150
Figure 59 . Extraits de 9 km^2 des images à haute résolution Sentinel-2A et SPOT 7 pour illustrer les sources d'erreur de classification observées spatialement. La localisation des extraits et les limites des zones climatiques sont présentées à gauche. 152

Liste des tableaux

Tableau 1. Principales composantes des systèmes agricoles liées à la production végétale, détectables par télédétection à partir d'imagerie satellitaire (voir Bégué *et al.* (2018) pour plus de détails). 29
Tableau 2. Caractéristiques des principaux satellites opérationnels en 2017 fournissant des images à haute, moyenne et basse résolution spatiale dans le domaine optique pour le suivi de la végétation (entre autres applications). 30
Tableau 3. Exemple de limites de résolution spatiale auxquelles les systèmes agricoles peuvent être caractérisés à l'échelle de la parcelle, en se situant dans un "cas de haute résolution" pour différentes tailles de parcelle. 32
Tableau 4. Liste des 23 images composées téléchargées pour la série annuelle couvrant la saison culturale de 2014-2015 au Tocantins et 2016 au Burkina, dates de la période de composition de chaque image et code du jour julien utilisé pour désigner les images (la liste d'images pour la saison culturale de 2013-2014 du Tocantins a les mêmes fourchettes de dates et le même code du jour de l'année que les images pour la saison 2014-2015, ici montrées). .. 65
Tableau 5. Détails d'acquisition des scènes Landsat 8 OLI. La couverture nuageuse dans les scènes correspond à des zones situées à l'extérieur des limites de l'État du Tocantins. 71
Tableau 6. Exemple de matrice de confusion avec la notation utilisée pour la formulation des indices d'évaluation de classifications. 83
Tableau 7. Tableau récapitulatif des paramètres de segmentation utilisés pour la délimitation des parcelles et la stratification régionale en unités de paysage. *IR = infrarouge. 114
Tableau 8. Résultats de la précision globale de la classification et l'intervalle de confiance associé, le coefficient global de Kappa, et le désaccord sur la quantité et la répartition des objets, calculés à partir des 900 points GPS de vérité terrain pour chaque méthode de classification. 120
Tableau 9. Tableau de contingence résumant les résultats des deux méthodes de classification (nombre d'objets correctement classés ou mal classés) et résultats du test du χ^2 de McNemar. 121
Tableau 10. Résultats de la F-mesure pour les deux méthodes de classification. 121
Tableau 11. Ordres de grandeur des résultats de segmentation pour le Tocantins et le Burkina Faso sur une superficie équivalente de 240 000 km^2. Les valeurs ont été estimées à partir d'une segmentation réalisée sur le Tocantins (section 4.1.1.1) et une segmentation réalisée sur la commune de Koumbia. 139

Liste d'abréviations, sigles et acronymes

ABC	Agricultura de Baixa emissão de Carbono
(A)CP	(Analyse en) composantes principales
(GE)OBIA	(Geographic) Object-Based Image Analysis
AfDB	African Development Bank
CAPES	Coordination pour le Perfectionnement du personnel de l'Enseignement Supérieur au Brésil
CESBIO	Centre d'Études Spatiales de la BIOsphère
CIRAD	Centre de Recherche Agronomique pour le Développement
ClaPaS	Classement de Paysages et Segmentation
CNES	Centre National d'Études Spatiales
COFECUB	COmité Français d'Évaluation de la Coopération Universitaire et Scientifique avec le Brésil
CV-MVC	Constrained View angle-Maximum Value Composite
DGESS	Direction Générale des Études et des Statistiques Sectorielles du Ministère de l'agriculture et des aménagements hydrauliques du Burkina Faso
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
EPA	Enquête Permanente Agricole de la DGESS
EROS	Earth Resources Observation and Science
ESA	European Space Agency
ESPA	EROS Center Science Processing Architecture on demand interface
ESTARFM	Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model
EVI	Enhanced Vegetation Index
FAO	Food and Agricultural Organization
FEWS NET	Famine Early Warning Systems Network
GAES	Global Agro-Environmental Stratification
GAEZ	Global Agro-Ecological Zoning
GEOGLAM	Group on Earth Observations Global Agricultural Monitoring
HEA	Household Economy Approach
IIASA	Institut International pour l'Analyse des Systèmes Appliqués
IBGE	Instituto Brasileiro de Geografia e Estatística
IFPRI	International Food Policy Research Institute
INPE	Instituto Nacional de Pesquisas Espaciais
INSD	Institut National de la Statistique et de la Démographie
ISODATA	Iterative Self-Organizing Data Analysis Technique
JECAM	Joint Experiment of Crop Assessment and Monitoring
LADA	Land Degradation Assessment in Drylands
LDCM	Landsat Data Continuity Mission
LP DAAC	Land Processes Distributed Active Archive Center
MAAH	Ministère de l'Agriculture et des Aménagements Hydrauliques du Burkina Faso
MATOPIBA	Maranhão, Tocantins, Piauí et Bahia

MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NTSA	Núcleo Temático de Sistemas Agrícolas da EMBRAPA Pesca e Aquicultura
OLI	Operational Land Imager
ONU	Organisation des Nations Unies
OTB	Orfeo ToolBox
PAM	Produção Agrícola Municipal
RGB	Rouge, Vert, Bleu
SAR	Synthetic Aperture Radar
SIGMA	Stimulating Innovation for Global Monitoring of Agriculture
SPOT	Satellite Pour l'Observation de la Terre
SRTM	Shuttle Radar Topographic Mission
SVM	Support Vector Machine
TIRS	Thermal Infrared Sensor
TOSCA	Terre Solide, Océan, Surfaces Continentales et Atmosphère
UCL	Université Catholique de Louvain
USAID	United States Agency for International Development
USGS	United States Geological Survey
UTM	Universal Transverse Mercator

Introduction

CONTEXTE DE LA THÈSE 2
PRÉSENTATION GÉNÉRALE DE LA THÈSE 3
ORGANISATION DE LA THÈSE 5

Introduction

Contexte de la thèse

L'agriculture mondiale est confrontée à un double défi : elle doit fournir plus de produits alimentaires et non alimentaires (ex. fibres textiles, biocombustibles) pour répondre à la demande d'une population en croissance continue, tout en réduisant l'impact sur l'environnement. La capacité de l'agriculture mondiale à répondre à ce double défi est mise en difficulté par la dégradation et l'épuisement des ressources naturelles, les pressions socio-économiques et foncières et la saturation des terres, aggravés par les conséquences du changement climatique, qu'elle doit en même temps mitiger.

Ainsi, dans un contexte de planification de l'utilisation des terres à l'échelle des pays ou des régions, il y a un réel besoin d'indicateurs de suivi de l'agriculture, notamment des indicateurs permettant (i) d'évaluer les conditions de la sécurité alimentaire, spécialement dans les régions où la disponibilité alimentaire est irrégulière, et (ii) de suivre la gestion et l'état des ressources naturelles. Ces indicateurs concernent les estimations de la production agricole et des rendements des cultures, les estimations de l'extension des surfaces agricoles et de l'évolution des milieux naturels, ou encore les estimations de diversification et de l'intensification de la production.

La localisation des différents systèmes agricoles - des espèces cultivées et des pratiques agricoles adoptées - sont à la base de ces indicateurs. Il existe pourtant un manque général d'informations spatialement explicites et détaillées sur les systèmes agricoles des différentes régions du monde. Ces informations sont disponibles généralement sous forme de statistiques agricoles agrégées à des échelles administratives, ou de cartes issues de travaux sur l'utilisation agricole des terres à l'échelle globale ou régionale. La disponibilité et la qualité de ces données sont très hétérogènes à l'échelle globale et les informations sur les pratiques agricoles sont généralement omises. De plus, les moyens importants (techniques, humains et économiques) mobilisés pour leur production limitent leur mise à jour, empêchant le suivi régulier de l'évolution des systèmes agricoles.

Les outils de télédétection présentent un potentiel désormais reconnu pour la cartographie des systèmes agricoles sur des grandes étendues de par leur capacité d'extraction d'informations spatialisées synoptiques sur les dynamiques de la végétation résultant de l'activité agricole à partir de séries d'images satellitaires. Cette information spatiale est robuste et cohérente dans le temps lorsqu'on applique les mêmes traitements sur le même

type d'image satellitaire, les acquisitions périodiques permettant d'actualiser les informations régulièrement.

Toutefois, la cartographie des systèmes agricoles à l'échelle régionale par télédétection demeure un défi pour la communauté scientifique. En effet, à cette échelle, la variabilité de systèmes agricoles est souvent importante car liée à la variabilité régionale des conditions environnementales et socio-économiques. De plus, il reste difficile de dissocier l'information spectrale et temporelle associée aux différents systèmes agricoles de celle d'autres types d'occupation et d'utilisation des terres, notamment des milieux naturels végétalisés, également très diverses à cette échelle. En outre, il existe des contraintes géographiques de mise en œuvre des méthodes de télédétection liées principalement aux conditions climatiques locales (nébulosité) et aux types d'agriculture (de l'agro-industrie à la petite agriculture familiale). Ainsi, les études portant sur la cartographie des systèmes agricoles à l'échelle régionale par télédétection sont peu nombreuses, et s'appuient généralement sur des données externes, de l'expertise et des données de terrain pour assister les classifications, ce qui les rend coûteuses et qui limite leur reproductibilité.

Présentation générale de la thèse

Cette thèse est rattachée à trois projets qui partagent un objectif scientifique commun : développer des méthodologies innovantes pour améliorer le suivi des systèmes agricoles à grande échelle à partir de données d'observation de la Terre.

1. Le projet européen "FP7 SIGMA"[1] (2013-2017), qui a cofinancé ce travail de thèse avec le CIRAD (Centre de Recherche Agronomique pour le Développement), a été réalisé en appui à l'initiative mondiale GEOGLAM[2]. Ce projet vise à développer des méthodes et des indicateurs innovants de suivi et d'évaluation du progrès de l'agriculture mondiale vers une « agriculture durable » basées sur des données d'observation de la Terre et des données in situ. La thèse s'inscrit dans l'axe « évaluation de la productivité agricole » et plus spécifiquement elle s'insère dans le volet « analyse régionale et locale des systèmes agricoles » et répond à deux de ses objectifs : (i) caractériser les principales variables qui décrivent les systèmes de culture à l'échelle locale en utilisant des données de télédétection, et (ii) étudier comment ces variables sont transférées à l'échelle régionale et quels indicateurs issus

[1] *Stimulating Innovation for Global Monitoring of Agriculture – 7th Framework Programme for Research and Technological Development of the European Commission* (FP7 SIGMA).
[2] *Group on Earth Observations Global Agricultural Monitoring* (GEOGLAM).

d'images satellitaires à moyenne résolution peuvent être utilisés pour représenter les systèmes agricoles.

2. Le projet "Méthodes et innovations technologiques pour le suivi de systèmes agricoles bas-carbone en appui au programme ABC (*Agricultura de Baixa emissão de Carbono*)" (2015-2019) du programme franco-brésilien CAPES-COFECUB[3]. Ce projet a pour objectif d'évaluer l'intégration des méthodes et des innovations technologiques - nouveaux capteurs de télédétection, changement d'échelle et analyse spatiale - en vue d'élaborer des méthodes d'identification des systèmes agricoles et des changements de pratiques conduisant à des systèmes plus économes en matière d'émissions de carbone (ex. qui intègrent des pratiques telles que le semis direct, l'intégration agro-sylvopastoral et l'agroforesterie).
3. Le projet "AGRIZONE" TOSCA-CNES [4] (2016-2017) qui a pour objectif de développer et de tester des méthodes de zonage basées sur l'analyse paysagère et le traitement d'images satellitaires, pour caractériser les systèmes de production agricole à l'échelle régionale.

Cette thèse a pour objectif de développer et tester de nouvelles méthodes basées sur l'extraction des informations relatives aux systèmes agricoles - espèces cultivées et pratiques agricoles - par télédétection à partir des données satellitaires optiques en appui au suivi de l'agriculture à grande échelle. En particulier, elle explore le potentiel des séries temporelles d'indice de végétation par différence normalisé (NDVI, acronyme anglais de *Normalized Difference Vegetation Index*), des techniques d'extraction de « variables phénologiques » et de réduction de la variabilité spectrale, des méthodes de classification non supervisées et de transfert d'échelle, pour la cartographie et la caractérisation des systèmes agricoles à l'échelle régionale.

Les travaux sont menés sur deux régions du monde qui partagent une économie fortement basée sur la production agricole, mais avec des enjeux contrastés (l'une de « durabilité », et l'autre de sécurité alimentaire et nutritionnelle) : l'État brésilien du Tocantins, un nouveau front pionnier caractérisé par l'installation des grandes exploitations entrepreneuriales qui pratiquent une agriculture intensive, et le Burkina Faso, une région représentative de la petite agriculture familiale de l'Afrique de l'Ouest, caractérisée par des petites exploitations fragmentées qui pratiquent une culture manuelle et attelée avec peu d'intrants.

[3] Coordination pour le Perfectionnement du personnel de l'Enseignement Supérieur au Brésil (CAPES) ; COmité Français d'Évaluation de la Coopération Universitaire et scientifique avec le Brésil (COFECUB).
[4] Terre Solide, Océan, Surfaces Continentales et Atmosphère (TOSCA) ; Centre National d'Études Spatiales (CNES).

Organisation de la thèse

Le corps de la thèse est organisé en cinq chapitres qui donnent suite à cette section introductive. Le premier chapitre présente la thèse dans son contexte scientifique et parcourt l'état de l'art sur les thématiques et les concepts abordés, notamment les méthodes existantes, pour caractériser les systèmes agricoles de différentes régions du monde et les développements concernant l'utilisation de la télédétection dans ce domaine. Le contexte et la problématique abordés aboutissent à l'explicitation de l'objectif principal de la thèse et des questions de recherche soulevées. Le deuxième chapitre présente ensuite l'approche générale proposée, les régions d'étude, ainsi que les données et les outils mobilisés dans la thèse. Les chapitres trois et quatre s'articulent autour de deux articles scientifiques publiés, qui décrivent les différentes composantes méthodologiques de l'approche développée et les résultats obtenus pour le Tocantins. Le chapitre cinq est consacré à l'évaluation de la reproductibilité de l'approche appliquée au Burkina Faso. Enfin, une dernière section de conclusions et perspectives est consacrée au bilan des principales contributions de la thèse, principalement méthodologiques, et introduit des perspectives de recherche et d'utilisation des méthodes développées.

Chapitre 1.
Contexte et problématique de la thèse

1.1. L'AGRICULTURE FACE AUX ENJEUX GLOBAUX 8
1.2. LA CARTOGRAPHIE ET LA CARACTÉRISATION DE SYSTÈMES AGRICOLES 11
1.2.1. Les systèmes agricoles : concepts et définitions 11
1.2.2. L'intérêt de la cartographie et la caractérisation de systèmes agricoles 13
1.2.3. L'utilisation agricole des terres 15
1.2.4. La dimension spatio-temporelle dans la cartographie de systèmes agricoles : le concept d'échelle 15
1.2.4.1. Les unités de paysage agricole 18
1.2.5. Le défi de la cartographie de systèmes agricoles à l'échelle régionale 19
1.2.5.1. Exemples d'approches de cartographie des systèmes d'utilisation agricole des terres 20
1.2.5.2. Exemples d'approches de zonage agricole 25
1.2.5.3. Limites des approches présentées 26
1.3. LA TÉLÉDÉTECTION ET LA CARTOGRAPHIE DE SYSTÈMES AGRICOLES 28
1.3.1. L'image satellitaire : source d'information sur les systèmes agricoles 28
1.3.2. État de l'art des approches de télédétection pour la cartographie de systèmes agricoles 34
1.3.2.1. Exemples d'études menées à l'échelle régionale 35
1.3.2.2. Méthodes de classification 37
1.3.2.3. Techniques de réduction de la variabilité spectrale 38
1.3.2.4. La stratification régionale en unités de paysage 39
1.4. OBJECTIF ET QUESTIONS DE RECHERCHE 42
1.5. APPROCHE GÉNÉRALE 44

Chapitre 1. Contexte et problématique de la thèse

1.1. L'agriculture face aux enjeux globaux

Les dernières projections concernant l'évolution de la population mondiale faites par l'ONU (Organisation des Nations Unies) prévoient 9,8 milliards de personnes en 2050. Même si la croissance moyenne de la population à l'échelle mondiale semble ralentir par rapport aux dernières décennies, la population augmente toujours, avec un taux de croissance estimé de 1,24 % par an, soit 83 millions de nouvelles personnes chaque année (United Nations. Department of Economic and Social Affairs. Population Division, 2017).

Afin de répondre à la demande des produits issus de l'agriculture (alimentaires et énergétiques) que se profile pour 2050, la FAO (acronyme anglais pour l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture) estime que la production agricole mondiale devra augmenter d'environ 50 % par rapport à la production de 2012[5] (FAO, 2017). Cependant, malgré l'avancement rapide des innovations et développements technologiques, l'augmentation des rendements des surfaces cultivées ralentit (FAO, 2017). Ce constat est particulièrement alarmant étant donné que l'augmentation des rendements est la principale source d'augmentation de la production agricole mondiale[6], l'extension des surfaces cultivées étant limitée par la concurrence pour la ressource foncière et par la prise en compte de l'impact environnemental (Alexandratos et Bruinsma, 2012; Fischer, Hizsnyik et Wiberg, 2011).

Le ralentissement de l'augmentation des rendements est en partie dû à la dégradation des ressources naturelles par l'intensification des pratiques culturales (Matson *et al.*, 1997). Les pratiques intensives, bien qu'améliorant la production à court terme, entrainent un appauvrissement progressif des ressources, qui à long terme se traduit par des baisses des rendements. Par exemple, le labour des sols occasionne à long terme une perte importante de fertilité et augmente la vulnérabilité des sols à l'érosion.

[5] Les hypothèses d'évolution qui ressortent de différents travaux de prospective concernant la production agricole mondiale à l'horizon 2050 y compris les estimations de l'évolution de l'extension des surfaces cultivées et des rendements, présentent des résultats avec des divergences importantes qui proviennent des grandes différences dans la construction des scénarios utilisés pour estimer les tendances d'évolution futures (Even et Vert, 2011).

[6] D'après des estimations de la FAO datant de 2011, l'augmentation des rendements et l'intensification de pratiques culturales seraient responsables du 90 % de l'augmentation de la production agricole mondiale, et l'extension des surfaces cultivées du 10 % restant, pour la période de 2007–2050 (Bruinsma, 2011).

Outre l'appauvrissement des ressources et les conflits d'usage autour de la ressource foncière, l'agriculture mondiale doit également surmonter les conséquences du changement climatique (HLPE, 2017; Lobell et Field, 2007) qui s'intensifie au cours des années, tout en s'adaptant aux effets de la mondialisation sur les marchés internationaux et à l'évolution des politiques agricoles pour maintenir la rentabilité économique de la production (Lambin, Geist et Lepers, 2003). La Figure 1 présente à titre d'exemple certains enjeux et défis de l'agriculture mondiale.

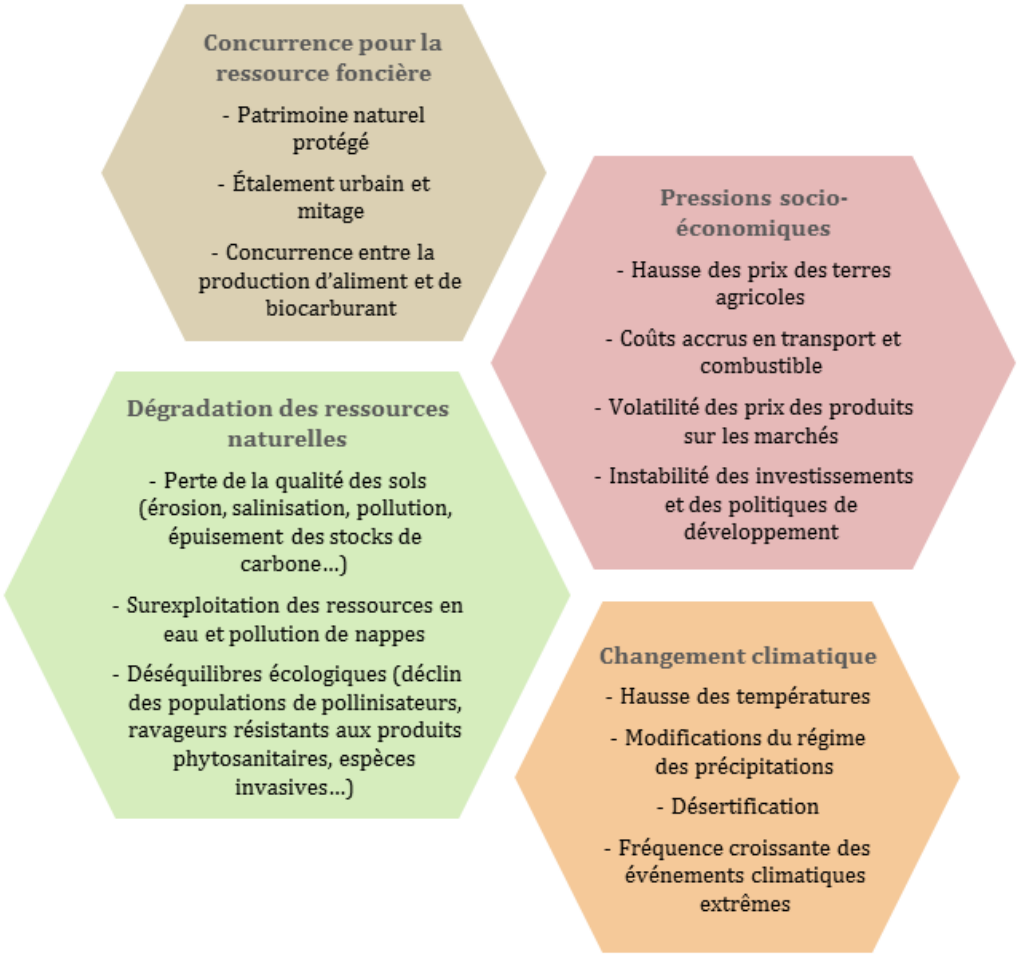


Figure 1. Exemples d'enjeux et défis de l'agriculture dans un contexte d'augmentation de la population mondiale.

Les enjeux globaux ne sont pas homogènes à l'échelle de la planète, ainsi que leur impact sur les agricultures locales (Collin, 2012). Les pays en développement étant particulièrement vulnérables sont confrontés à des défis de grande ampleur (World Bank, 2010). Par exemple, la Figure 2 montre que les conséquences du changement climatique pourraient altérer fortement la production des régions en situation d'insécurité alimentaire, avec une baisse des rendements importante pour la plupart des pays d'Afrique, de l'Asie du Sud et de l'Amérique latine entre la période actuelle et 2050 (Müller *et al.*, 2010; World Bank, 2010).

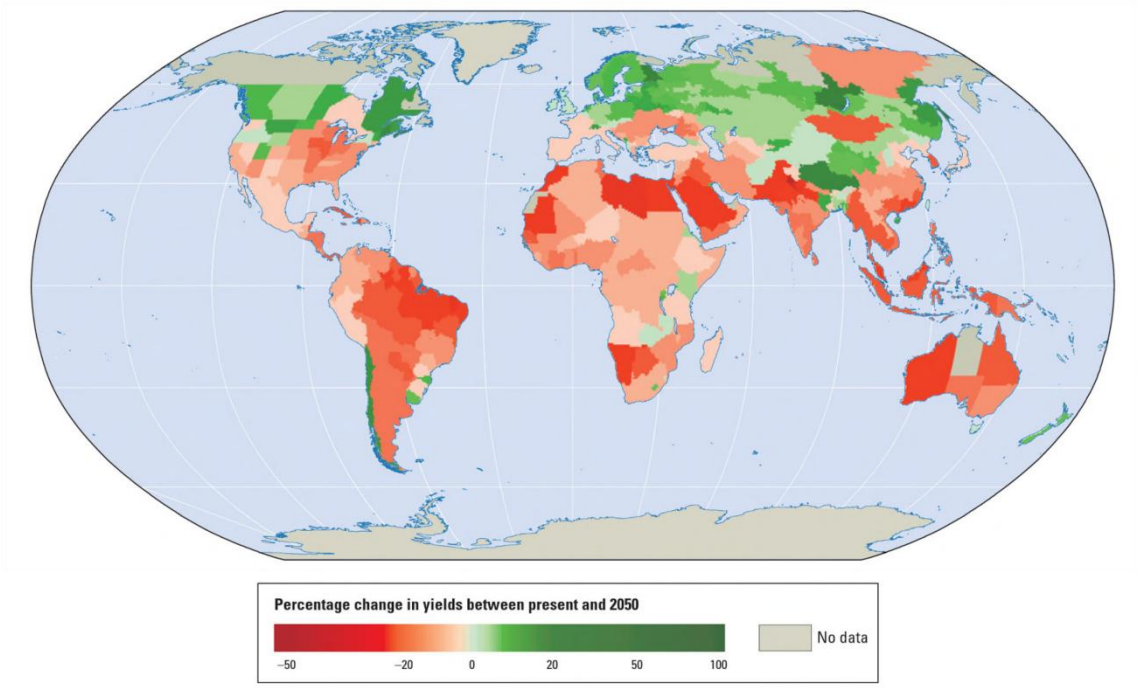


Figure 2. Estimations du pourcentage de variation des rendements pour 11 grandes cultures entre la période de 1996-2005 et 2046-2050. Les valeurs ont été obtenues par une moyenne des résultats de cinq modèles climatiques et trois scénarios d'émission de gaz à effet de serre (Müller *et al.*, 2010; World Bank, 2010).

Par conséquent, chaque région agricole doit développer sa propre résilience face aux changements globaux en employant des stratégies d'adaptation différentes. Toutefois, l'adoption de pratiques culturales à faible impact environnemental, telles que les pratiques associées à l'agro-écologie, se retrouvent largement dans les stratégies transrégionales pour évoluer vers une production agricole « durable » (de Schutter, 2010). L'agro-écologie regroupe différent types d'agriculture, avec des pratiques diversifiées, qui partagent un modèle commun de réduction de la dégradation des ressources naturelles (ex. agriculture de conservation, agriculture biologique, agroforesterie, agriculture raisonnée, permaculture) (Burger *et al.*, 2011).

L'amélioration de la fertilité du sol par l'adoption de pratiques agro-écologiques entraîne un accroissement de la productivité du sol (augmentation des rendements à l'hectare) et du travail, la diversification de la production faisant partie des stratégies qui contribuent à cet accroissement, en permettant au même temps aux agriculteurs d'être moins vulnérables aux fluctuations des prix sur les marchés mondiaux (Burger *et al.*, 2011). Même si la transition vers ces modèles de production à grande échelle n'est pas évident (Parmentier, 2014), ils continuent à prendre de l'ampleur au niveau mondial (FAO, 2011; Friedrich, Derpsch et Kassam, 2012; Willer et Lernoud, 2017).

Compte tenu de la forte évolution de l'agriculture mondiale et des grands contrastes régionaux vis-à-vis des défis économiques, sociaux et environnementaux, on s'attend à ce que les besoins en informations spatiales à l'échelle régionale sur les surfaces mises en culture et les pratiques agricoles ne cessent d'augmenter.

Bilan : L'agriculture face aux enjeux globaux

- La production agricole mondiale doit répondre à la demande d'une population en croissance continue, tout en s'adaptant aux conséquences du changement climatique, de la dégradation des ressources naturelles, des pressions socio-économiques et de la concurrence pour la ressource foncière.
- Les enjeux globaux ne sont pas homogènes à l'échelle de la planète, chaque région devant adopter des stratégies de développement adaptées à leurs potentiels et contraintes spécifiques.
- L'agriculture mondiale évolue vers une agriculture multifonctionnelle qui doit faire face au double défi de produire plus et de manière durable, notamment en adoptant des pratiques culturales à faible impact environnemental telles que les pratiques associées à l'agro-écologie.

1.2. La cartographie et la caractérisation de systèmes agricoles

1.2.1. Les systèmes agricoles : concepts et définitions

Les activités agricoles, composées de multiples éléments qui interagissent entre eux, sont communément conceptualisées sous la forme de systèmes. Le terme générique "système agricole" (équivalent du terme anglophone "*agricultural production system*") peut être considéré fondamentalement comme « l'ensemble d'activités humaines de gestion des ressources naturelles et d'application de la technologie pour générer certaines sorties de nourriture et de fibres souhaitées » (Geng, Hess et Auburn, 1990).

Le terme générique "système agricole" englobe ainsi plusieurs types de systèmes par lesquels on appréhende les différentes activités humaines en relation avec la production agricole. On peut concevoir une hiérarchie imbriquée entre ces différents systèmes associée à différents niveaux d'organisation (Figure 3) comme celle proposée par Fresco et Westphal (1988), où chaque système comprend des systèmes de niveau inférieur (sous-systèmes) et constitue en même temps un sous-système d'un système de niveau supérieur (supra-système).

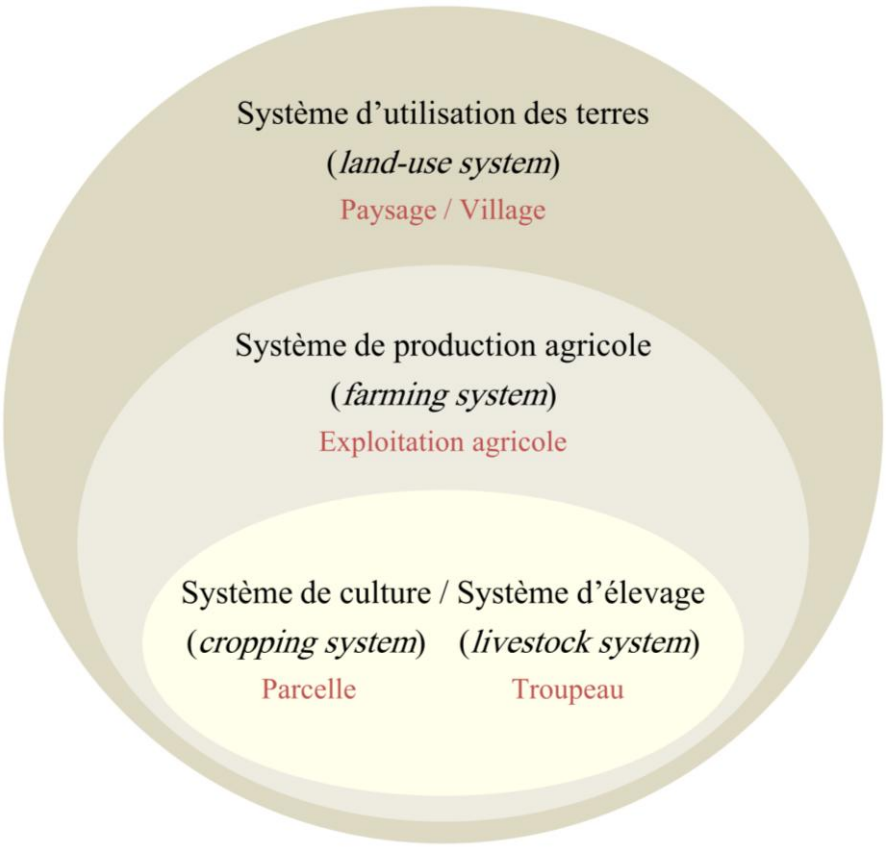


Figure 3. Hiérarchie de systèmes agricoles, et leur niveau d'organisation (adaptée de Fresco et Westphal (1988)). Les termes anglais sont mis en correspondance avec des termes français analogues, même si les concepts qui leur sont associés sont parfois employés différemment selon les groupes de recherche français et anglo-saxons (Fresco, 1984).

Concernant la production agricole végétale, le niveau d'organisation élémentaire correspond à la parcelle. Les activités liées à la production qui ont lieu au niveau de la parcelle se conceptualisent sous la forme de "système de culture". D'après Sebillotte (1990), un système de culture est défini comme « l'ensemble des modalités techniques mises en œuvre sur des parcelles cultivées de manière identique. Chaque système étant défini par :

- la nature des cultures et leur ordre de succession,
- les itinéraires techniques appliqués à ces différentes cultures, ce qui inclut le choix des variétés ».

À un niveau d'organisation supérieur, celui de l'exploitation agricole, on retrouve le concept de « système de production agricole ». Selon Sebillotte (1982) une exploitation agricole peut contenir un ou plusieurs systèmes de culture et se caractérise par son système de production qui peut être défini comme « un agencement des facteurs de production, des choix techniques et une combinaison des productions agricoles » (Brossier, 1987). Ce concept relie

les choix techniques associés à la production à une composante socio-économique en rapport aux facteurs de production (capital, foncier, travail) de l'exploitation agricole.

Finalement on retrouve le concept intégrateur de "système d'utilisation des terres", qui englobe les deux types de systèmes agricoles décrits auparavant, et se présente comme une combinaison de composantes socio-économiques et biophysiques qui interagissent à plusieurs niveaux (Stomph, Fresco et van Keulen, 1994). Le concept de système d'utilisation des terres est donc un concept multiniveau qui peut être abordé à différentes échelles selon les objectifs de l'analyse (Bouma, 1997; Lambin, Geist et Rindfuss, 2006; Stomph *et al.*, 1994).

L'étude des systèmes agricoles diffère, ainsi, entre différents travaux, selon leurs objectifs, le type de système agricole analysé et les différentes approches utilisées pour les analyser (ex. agronomique, économique, pluridisciplinaire) (Bouma, 1997; Brossier, 1987). Dans ce travail de thèse, nous nous positionnons dans un contexte cartographique, où l'on cherche à caractériser spatialement différents types de systèmes agricoles liés à la production végétale à plusieurs niveaux d'organisation.

1.2.2. L'intérêt de la cartographie et la caractérisation de systèmes agricoles

La localisation et la description des surfaces et des activités agricoles actuelles ou passées au travers de la cartographie et la caractérisation des systèmes agricoles permet de suivre ce qui est produit, où et comment. Elles répondent donc à un besoin de production d'information spatiale pour le suivi de l'agriculture (Leenhardt *et al.*, 2010), à la fois de l'impact sur l'environnement et les ressources naturelles et de la sécurité alimentaire et nutritionnelle.

Plus spécifiquement, les cartes de systèmes agricoles peuvent aider à localiser spatialement les différents types de production agricole y compris les types de produits et le mode de production (ex. extensive ou intensive). Cette information permet de suivre des processus tels que l'intensification économique de l'agriculture (ex. l'utilisation de variétés de cultures à haut rendement, l'irrigation, la mécanisation, la culture séquentielle ou encore l'utilisation de fertilisants et pesticides), l'intensification écologique (ex. l'association et la rotation culturale et les pratiques de conservation des sols) et l'extension des surfaces cultivées (ex. au détriment des surfaces naturelles) (Bonny, 2010; Matson *et al.*, 1997).

Ces informations spatiales contribuent simultanément :

- À dériver des indicateurs spatiaux sur l'impact environnemental et sur les ressources naturelles tels que des indicateurs locaux sur l'érosion, la baisse de la fertilité des sols et la réduction de la biodiversité, des indicateurs régionaux sur la

pollution, l'eutrophisation et l'appauvrissement des ressources en eau et même des indicateurs globaux sur les impacts sur la composition atmosphérique et le climat (Matson *et al.*, 1997) ;

- À dériver des indicateurs sur la quantité d'aliments produite, sur l'évolution des rendements et sur la diversité d'aliments, essentiels pour des analyses sur la sécurité alimentaire et la sécurité nutritionnelle (HLPE, 2017).

De plus, une production régulière de ce type d'information spatiale sert à définir la situation de référence sur laquelle se construisent les travaux de prospective qui simulent, sous différents scénarios, les tendances d'évolution futures des activités agricoles, de la production, et des impacts associés (Leenhardt *et al.*, 2010).

La cartographie de systèmes agricoles produit donc une information spatiale essentielle pour appuyer une planification raisonnée de l'utilisation des terres, et contribue ainsi à orienter les stratégies d'aménagement et de développement des territoires. La Figure 4 montre un schéma synthétique où la carte des systèmes agricoles est présentée comme une des données essentielles à la base du suivi de l'agriculture et de l'aménagement des territoires.

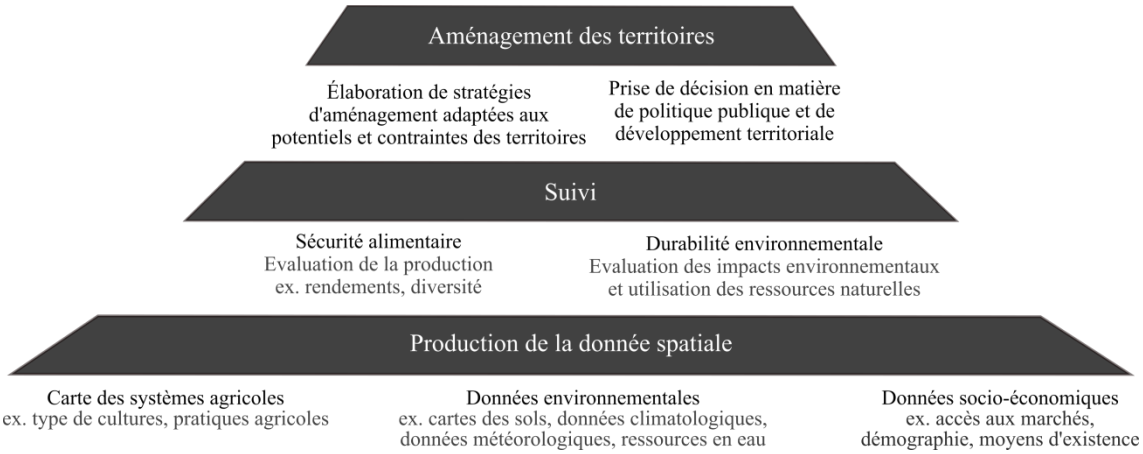


Figure 4. De la donnée spatiale à l'aménagement des territoires.

Le niveau de détail spatial et thématique des cartes produites, et par conséquent leur utilité, dépend néanmoins des données et des approches utilisées pour leur production, notamment de leur capacité pour extraire de l'information pertinente sur les activités agricoles, mais aussi de la dimension spatio-temporelle associée à la production cartographique.

1.2.3. L'utilisation agricole des terres

L'occupation et l'utilisation des terres sont deux concepts distincts mais fortement reliés entre eux (Fisher, Comber et Wadsworth, 2005; Loveland et DeFries, 2004). L'occupation des terres est définie comme « la couverture (bio) physique observée sur la surface terrestre » (Di Gregorio et Jansen, 2000), y compris la végétation naturelle, la surface cultivée, la surface en eau, le sol nu, le bâti, etc. Le concept d'utilisation des terres, quant à lui, fait référence d'une part à « la séquence d'opérations effectuée dans le but d'obtenir des biens et des services de la terre » (Nachtergaele et Petri, 2013), et d'autre part « aux fins pour lesquelles les humains exploitent la couverture terrestre » (Lambin *et al.*, 2006).

De ce fait, l'utilisation agricole des terres pour obtenir des produits végétaux concerne aussi bien les pratiques culturales (ex. semis, récolte, rotation culturale), que le but pour lequel elles sont appliquées, qu'il soit obtenir des produits (ex. culture céréalière, culture légumineuse), ou des services (ex. culture de couverture).

La caractérisation spatialisée des systèmes agricoles peut être faite, donc, à partir d'une cartographie et une classification de l'utilisation agricole des terres, décrite par les espèces cultivées et les pratiques culturales associées. Même si la caractérisation des systèmes agricoles au travers de l'utilisation des terres reste une caractérisation partielle, celle-ci est une composante essentielle pour l'analyse de leur production et leur impact sur l'environnement (de Bie, 2000; Nachtergaele et Petri, 2013).

1.2.4. La dimension spatio-temporelle dans la cartographie de systèmes agricoles : le concept d'échelle

Le concept d'échelle est défini par Gibson, Ostrom et Ahn (2000) comme « les dimensions spatiales, temporelles, quantitatives ou analytiques utilisées pour mesurer et étudier des phénomènes ». Ces composantes d'échelle, et notamment les composantes spatiale et temporelle, sont utilisées dans la cartographie et la caractérisation de systèmes agricoles comme décrit ci-dessous.

L'échelle temporelle est déterminée par la durée de la période étudiée. Concernant la caractérisation des systèmes agricoles, la période d'étude correspond généralement à la saison culturale ou à plusieurs saisons, ce qui permet de caractériser les systèmes par leurs calendriers culturaux et les différentes pratiques culturales utilisées, telles que le mode de culture sur une année ou saison culturale (ex. culture séquentielle), ou les successions culturales sur plusieurs années ou saisons (ex. rotations culturales).

Dans un contexte cartographique, l'échelle spatiale est sous-divisée en quatre types (Lam, 2008) : l'échelle cartographique, l'échelle géographique, la résolution spatiale et l'échelle opérationnelle. L'échelle cartographique désigne le rapport entre la distance représentée sur une carte et la distance réelle sur la surface terrestre. L'échelle géographique ou étendue spatiale, évoque la taille de la dimension spatiale (Gibson *et al.*, 2000) et va donc être déterminée par la taille de la surface cartographiée (c.-à-d. la taille de la région d'étude). La résolution spatiale (désignée dans la littérature en écologie comme le "grain"), va être quant à elle déterminée par la taille des unités individuelles d'observation (Wiens, 1989); c'est-à-dire les plus petites entités qui peuvent être distinguées. Enfin, l'échelle opérationnelle fait référence à l'échelle spatiale à laquelle les processus étudiés ont lieu (Lam, 2008).

Ce dernier type d'échelle repose sur le principe que les processus qui ont lieu dans un environnement donné présentent une organisation spatiale qui peut être observée au travers de motifs caractéristiques dans l'espace, et sont, en quelque sorte, délimités par une étendue spatiale (Lam, 2008; Sayre, 2005). L'échelle opérationnelle va être donc définie en fonction des processus à cartographier et va déterminer l'étendue et la résolution spatiale à utiliser pour révéler au mieux les processus étudiés (Lam, 2008).

Comme introduit précédemment dans la section 1.2.1, les systèmes agricoles auxquels on s'intéresse s'organisent de façon hiérarchique sur plusieurs niveaux d'organisation (Bouma, 1997; Lambin *et al.*, 2006; Stomph *et al.*, 1994). On peut assumer dans ce modèle de représentation de systèmes hiérarchique, comme ceux utilisés en écologie suivant la théorie hiérarchique (O'Neill *et al.*, 1986), que les processus de production agricole – comprenant l'utilisation des terres –, à l'image des processus écologiques au sein des écosystèmes naturels, présentent des motifs spatiaux d'étendue variable selon les niveaux d'organisation.

Ainsi, on peut concevoir des unités spatiales (unités cartographiques) avec des étendues proportionnelles aux différents niveaux d'organisation des systèmes agricoles comme illustré dans la Figure 5. Chaque unité spatiale serait caractérisée par une organisation spatiale entre les éléments physiques de la surface, plus ou moins complexe selon si l'on monte ou descend dans la hiérarchie.

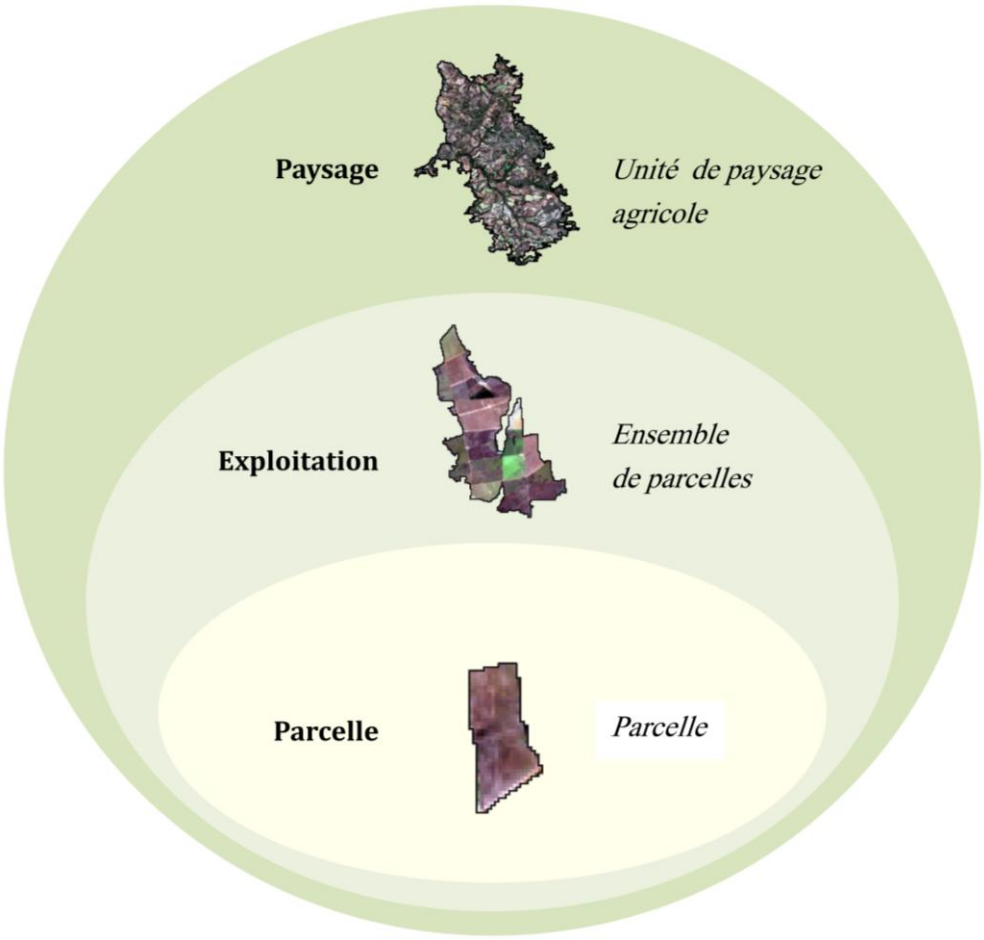


Figure 5. Exemple d'hiérarchie de niveaux d'organisation des systèmes agricoles (alignées à gauche) et unités spatiales ou échelles opérationnelles correspondantes (alignées à droite).

En ce qui concerne la cartographie des systèmes agricoles liés à la production végétale, la parcelle agricole serait ainsi l'unité spatiale ou échelle opérationnelle élémentaire, permettant d'accéder aux informations sur la nature des cultures et les pratiques culturales, à partir de laquelle on peut réaliser une carte de systèmes de culture. Les cartes de systèmes de culture représentent la donnée de base à partir de laquelle on peut calculer la surface occupée par les differents types de culture et extraire des indicateurs sur la production et les impacts environnementaux.

Or, la représentation des systèmes agricoles avec des unités spatiales situées à des niveaux supérieurs de la hiérarchie peut apporter des informations complémentaires sur les systèmes, notamment sur leur organisation spatiale, pertinentes pour les évaluations de la production agricole et de l'impact environnemental à grande échelle. Jouve (2006) souligne

qu'une bonne compréhension de l'organisation de l'espace rural[7] passe par l'appréhension de la dimension spatiale et territoriale des systèmes de culture (i.e. leur organisation spatiale au sein des territoires). Plus particulièrement, il constate qu'à l'échelle des territoires, il existe une organisation spatiale des systèmes de culture, en correspondance avec une différentiation des conditions biophysiques et des modes d'exploitation agricole, qui devrait être représentée et analysée pour mieux orienter la gestion de l'espace cultivé et des ressources naturelles (Jouve, 2006).

1.2.4.1. Les unités de paysage agricole

Ces mêmes observations ont mené à des agronomes et des géographes à proposer des façons de représenter l'organisation spatiale des systèmes agricoles au sein des territoires, notamment au niveau du paysage[8] (ex. Deffontaines, 1973; Deffontaines, 1997; Rizzo *et al.*, 2013; Veldkamp *et al.*, 2001). Ceci a conduit à la conception d'entités cartographiques, dites "unités de paysage agricole" qui délimitent spatialement des "agropaysages"[9]. Ces concepts reposent sur l'hypothèse selon laquelle un espace rural serait composé de différents paysages agricoles, qui sont un produit de la configuration spatiale des parcelles, ainsi que des activités agricoles et des conditions biophysiques et environnementales, et qui peuvent être donc directement liés aux systèmes agricoles sur place (Deffontaines, 1997; Deffontaines, Thenail et Baudry, 1995; Girard et Girard, 2010; Rizzo *et al.*, 2013).

Il peut être considéré, par conséquent, qu'une unité de paysage agricole particulière présente une certaine homogénéité dans sa physionomie et utilisation des terres qui diffère des paysages adjacents, tout en étant hétérogène dans sa structure, puisqu'elle est composé de multiples éléments (Deffontaines, 1973; Galochet, Godard et Hotyat, 2001). Compte tenu de leur nature hétérogène, il n'existe pas de frontière claire entre deux unités de paysage[10], mais

[7] Plet (2003) décrit l'espace rural comme un espace caractérisé par une organisation spatiale multifonctionnelle et une faible densité relative de population, avec des paysages marqués par l'agriculture.

[8] La notion de "paysage" désigne selon la Convention Européenne du Paysage « une partie de territoire telle que perçue par les populations, dont le caractère résulte de l'action de facteurs naturels et/ou humains et de leurs interrelations » (Conseil de l'Europe, 2000).

[9] Girard et Girard (2010) en présentant la notion d'agropaysage précisent qu'elle « repose sur l'axiome suivant : il existe une organisation spatiale, non aléatoire, des facteurs du milieu. L'organisation spatiale d'un milieu rural est en lien avec l'utilisation par l'homme de ce milieu et les actions de l'homme sont inscrites dans l'espace géographique. D'autre part, les facteurs du milieu constituent la base de l'organisation spatiale. En conséquence, certaines actions de l'homme […] peuvent être déchiffrées, au moins en partie, dans l'agropaysage au moyen de l'analyse spatiale ».

[10] Le terme de "*land unit*" est parfois employé par les groupes de recherche anglo-saxons comme synonyme du terme "*landscape unit*", étant dans ces cas l'équivalent du terme français "unité de paysage".

plutôt des transitions relativement discontinues provoquées par les éléments naturels et/ou l'action humaine (Zonneveld, 1989). En outre, les différents facteurs qui configurent les paysages agricoles (ex. facteurs biophysiques, pratiques agricoles) agissent à différentes échelles, et donc on peut assumer que les unités de paysage n'ont pas de dimension fixe, mais varient plutôt en forme et en extension selon l'échelle d'analyse.

D'après ces réflexions, le paysage semble être un niveau d'organisation approprié pour la cartographie de systèmes d'utilisation agricole des terres (concept présenté dans la section 1.2.1), telles que l'exploitation agricole le serait pour le système de production agricole ou la parcelle pour le système de culture, et les unités de paysage agricole seraient les unités cartographiques correspondantes.

1.2.5. Le défi de la cartographie de systèmes agricoles à l'échelle régionale

Puisque les politiques d'aménagement et de développement des territoires interviennent généralement sur des divisions administratives de grande taille (ex. nations, divisions sous-nationales), il est important de produire des cartes à ces échelles géographiques. Ainsi, dans ce travail de thèse on s'intéresse à la cartographie à l'échelle régionale.

La région peut être définie comme « certaines parties de la surface terrestre où la population, organisée dans certaines limites sociales, politiques ou administratives, fait face à des défis naturels, sociaux, politiques et économiques » (Nir, 1987). L'échelle régionale semble donc l'échelle appropriée pour la production de la donnée spatiale au service de la planification du développement et l'aménagement des territoires. Bien que la région fasse aussi référence à des espaces qui se caractérisent par une similitude physiographique ou encore par des valeurs historiques ou culturelles, on va réserver le terme par la suite à la région en tant qu'échelon politico-administratif avec des contours fixes et relativement pérennes (Girard, 2004).

Une carte de systèmes agricoles à l'échelle régionale peut relever ainsi de l'information sur des superficies allant des régions sous-nationales à des régions multinationales. Cependant, plus l'étendue considérée est extensive, plus elle est susceptible de présenter une grande variabilité de systèmes agricoles, augmentant la difficulté technique liée à la production cartographique et suscitant un majeur effort de production.

Malgré l'adéquation des cartes à l'échelle régionale pour les propos mentionnés, il est commun donc de retrouver à cette échelle une représentation spatiale simplifiée des activités agricoles et de l'utilisation des terres avec un manque général de détail thématique et une basse résolution spatiale (Jouve, 2006; Leenhardt *et al.*, 2010). Plusieurs travaux pertinents

relatifs à la planification d'utilisation agricole des terres à grande échelle, dont on présente ci-dessous quelques exemples, rencontrent ces limites.

1.2.5.1. Exemples d'approches de cartographie des systèmes d'utilisation agricole des terres

Driessen et Konijn (1992) en consolidant les concepts précurseurs présentés par la FAO dans le "Cadre pour l'évaluation des terres" (FAO, 1976), approchent le concept de système d'utilisation des terres d'un point de vue cartographique, en le définissant comme « la combinaison d'une unité de terre (unité cartographique) et un type d'utilisation de terre ». À l'origine, la cartographie de systèmes d'utilisation des terres consistait à délimiter des zones physiographiques relativement homogènes (les unités de terre), et de les relier à un type spécifique d'utilisation des terres (actuel ou potentiel).

Le travail cartographique le plus complet issu de ce type d'approche a été réalisé dans le cadre du projet d'évaluation de la dégradation des terres dans les zones arides (LADA, acronyme de *Land Degradation Assessment in Drylands*) de la FAO (Nachtergaele et Petri, 2013). Il s'agit d'une cartographie de 40 systèmes d'utilisation des terres à l'échelle globale pour l'analyse de l'évaluation de la dégradation des terres (voir extrait de l'Afrique subsaharienne en Figure 6) (Nachtergaele et Petri, 2013).

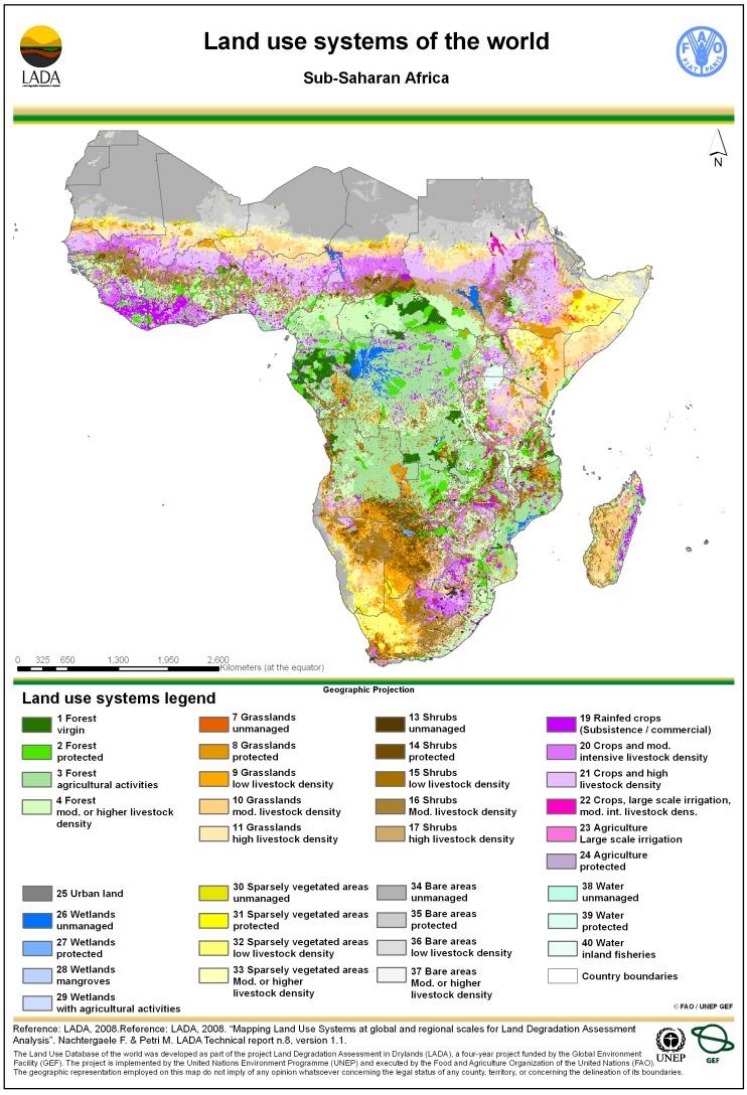


Figure 6. Systèmes d'utilisation des terres d'Afrique subsaharienne. ©FAO/UNEP GEF (source : LADA, 2008).

Les systèmes d'utilisation agricole des terres représentés dans les cartes du projet LADA sont caractérisés par des informations sur la présence de bétail, la culture ou groupe de culture dominant, présence d'irrigation et un indice de gestion des cultures que caractérise le niveau d'intensité des pratiques culturales (Nachtergaele et Petri, 2013). Ces informations proviennent de différentes sources statistiques et sont agrégées à une résolution spatiale d'environ 9 km au niveau de l'équateur.

Six pays pilotes ont élaboré leurs propres cartes à une résolution spatiale plus fine en adaptant la méthode originelle du projet LADA à l'échelle nationale, mais seulement trois parmi eux (Tunisie, Chine et Cuba) ont inclus des informations spatiales sur le type de culture et les pratiques culturales dans la caractérisation des systèmes d'utilisation agricole des terres (Biancalani *et al.*, 2013).

Une approche à l'échelle globale, établie par la FAO et la Banque Mondiale a été utilisée pour la production de cartes des principaux systèmes de production agricole[11] de six grandes régions du monde (échelle multinationale) (Dixon, Gulliver et Gibbon, 2001). Un total de 72 systèmes de production agricole a été cartographié en utilisant des multiples données biophysiques et socio-économiques comprenant : « (i) les ressources naturelles disponibles ; (ii) le modèle dominant des activités agricoles et des moyens d'existence des ménages, y compris les relations avec les marchés et (iii) l'intensité des activités de production » (Dixon *et al.*, 2001).

Les frontières des systèmes de production agricoles ont été délimitées par des experts de façon subjective sur des grandes étendues, en dimensionnant les systèmes de production de façon à respecter l'échelle d'analyse multinationale qui est privilégiée (Dixon *et al.*, 2001). Par conséquent, la résolution spatiale des cartes finales n'est pas définie et la plupart des pays sont caractérisés par moins de quatre classes. À titre d'exemple, la Figure 7 montre la classification du Burkina Faso extraite du produit "*Major farming systems of Sub-Saharan Africa*"[12], et la classification du Tocantins extraite du produit "*Major farming systems of South America*".

[11] Le concept de "système de production agricole" (équivalent du terme anglais "*Farming system*") utilisé par Dixon *et al.* (2001) se rapproche du concept de "système d'utilisation agricole des terres" présenté dans la section 1.2.2. dès lors qu'ils sont cartographiés et caractérisés à une échelle supérieure à celle de l'exploitation agricole. La définition exacte de "système de production agricole" donnée par Dixon *et al.* (2001) fait référence à « un groupe d'exploitations agricoles qui ont des bases de ressources largement similaires, des modèles d'entreprise, des moyens d'existence et des contraintes des ménages, et pour lesquels des stratégies de développement et des interventions similaires seraient appropriées ».

[12] Le produit "*Major farming systems of Sub-Saharan Africa*" a été actualisé pour l'année 2010 en ajoutant des nouvelles données (Garrity, Dixon, et Boffa, 2012).

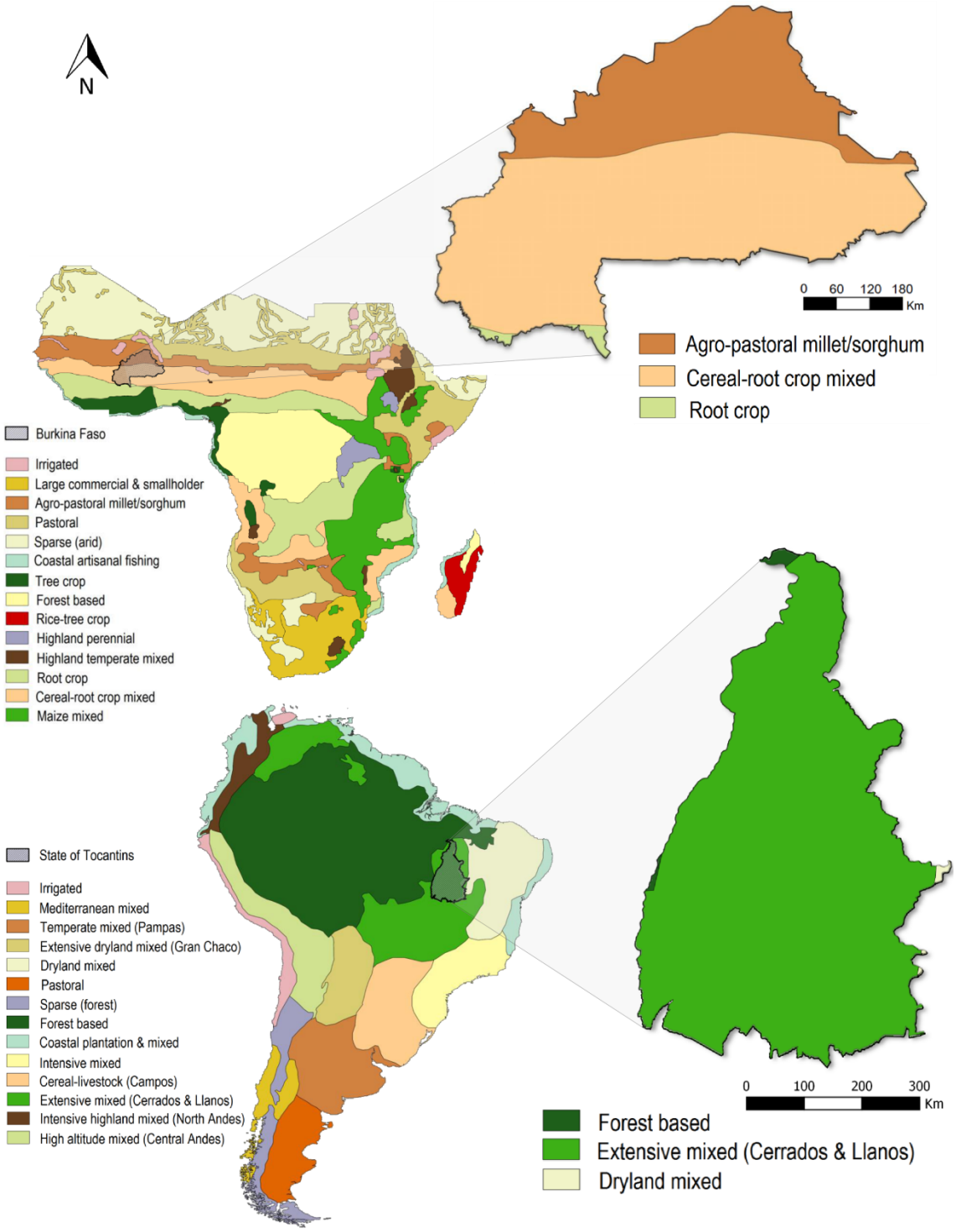


Figure 7. Classifications des systèmes de production du Burkina Faso (en haut à droite) et du Tocantins (en bas à droit) extraites des produits "*Major farming systems of Sub-Saharan Africa*" (en haut à gauche) et "*Major farming systems of South America*" (en bas à gauche) (adaptée de Dixon *et al.* (2001)).

Les zonages de moyens d'existence, développés par l'Agence des États-Unis pour le Développement International (USAID) pour le Réseau du Système d'Alerte Précoce (FEWS NET), peuvent également servir d'exemple d'approche de cartographie des systèmes

d'utilisation agricole des terres. Ces zonages, élaborés à une échelle nationale dans le cadre du suivi de la sécurité alimentaire des pays à risque, délimitent, pour 37 pays, des zones où la population partage approximativement les mêmes options pour obtenir des aliments y compris les revenus, les moyens d'accès aux marchés et globalement le même système de production agricole, entre autres facteurs (Grillo et Holt, 2009).

L'élaboration de ces zonages suit une « approche de l'économie des ménages » (HEA, acronyme anglais de *Household Economy Approach*), la première étape de cette approche consistant à réaliser une cartographie participative et à dire d'expert à partir de multiples données spatiales (ex. occupation des terres, accès aux marchés, conditions agro-climatiques, accès aux ressources en eau), des enquêtes de terrain et des statistiques socio-économiques (FEG Consulting and Save the Children, 2008). Les cartes de zonage résultantes présentent globalement des classes représentatives des principaux systèmes de production agricole des pays concernés. À titre d'exemple la Figure 8 présente le zonage de moyens d'existence du Burkina Faso (les profils associés à chaque zone peuvent être consultés dans FEWS NET/USAID (2010)).

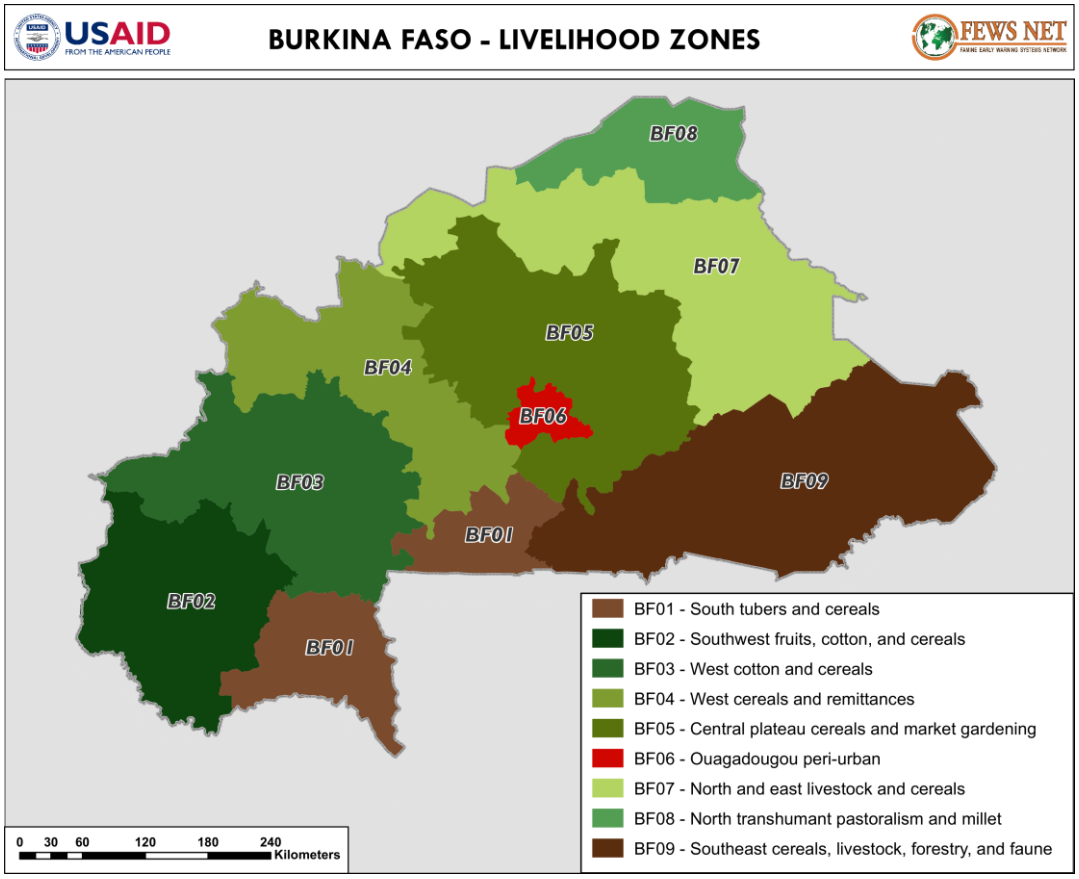


Figure 8. Produit du zonage de moyens d'existence au Burkina Faso développé par l'USAID (source : FEWS NET/USAID, 2010).

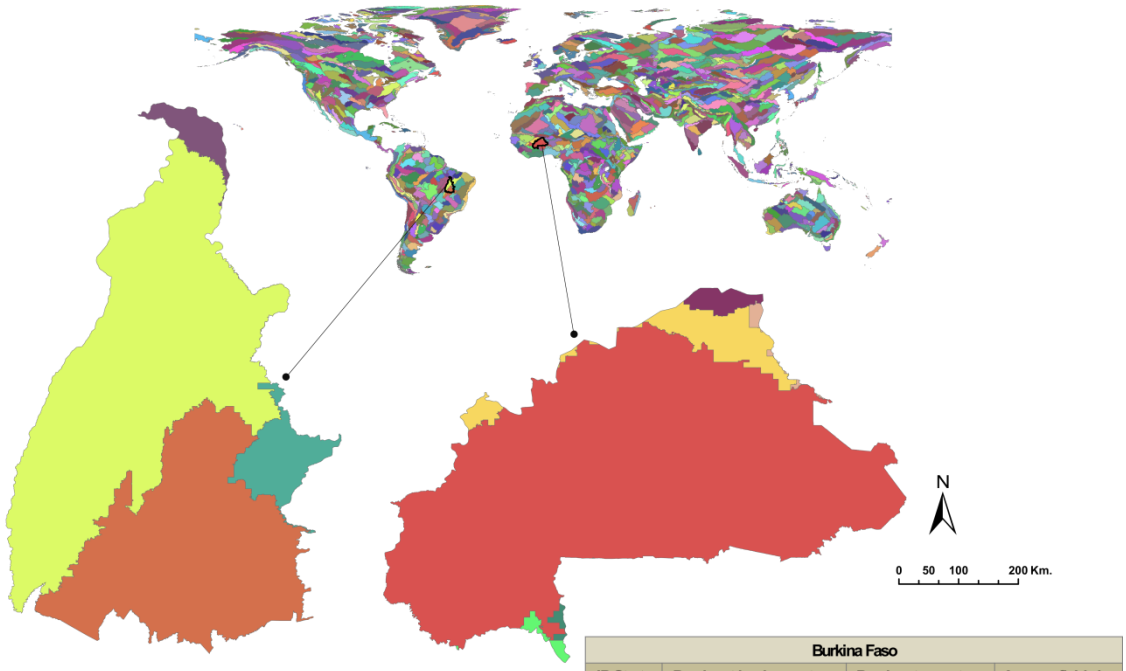
1.2.5.2. Exemples d'approches de zonage agricole

Les approches de zonage agricole font aussi partie des exemples de travaux qui visent à délimiter des unités spatiales et à les relier à un type d'utilisation des terres ; plus spécifiquement, ils évaluent le potentiel agricole associé à chaque zone analysée. Ce type d'approche largement décliné (ex. zonages agro-écologiques, agro-climatiques, agro-environnementaux, agro-écologique-économiques) est mené à plusieurs échelles, participant à la planification de l'utilisation des terres de plusieurs régions du monde.

Un des produits les plus connus est le "*Global Agro-Ecological Zones*" (GAEZ), conçu par la FAO en collaboration avec l'Institut International pour l'Analyse des Systèmes Appliqués (IIASA) (Fischer *et al.*, 2012). Depuis l'année 2000, de nombreux produits cartographiques à l'échelle globale ont été publiés, avec des bases de données associées qui contiennent un grand nombre de variables[13].

Les bases de données produites par le projet GAEZ sont utilisées dans différentes études dont la plus récente, le "*Global Agro-Environmental Stratification*" (GAES), a été menée par l'Institut de recherche environnementale de l'Université de Wageningen (Alterra) dans le cadre du projet SIGMA (Mücher *et al.*, 2016). Cette étude a réalisé un ré-échantillonnage de 13 couches spatiales de données agro-environnementales globales (ex. régimes climatiques, sols, élévation, ressources en eau et occupation des terres) à différentes résolutions spatiales à une résolution commune de 1 km, sur lesquelles une segmentation à quatre niveaux a été ensuite réalisée pour obtenir une hiérarchie de stratifications globales des « zones de production agricole » (Mücher *et al.*, 2016). Ces zones ou strates sont ensuite caractérisées en termes de phénologie (début et pic de la saison de croissance des cultures), de modes de gestion de l'eau, de taille des parcelles agricoles et de l'accès aux marchés, entre autres (Mücher *et al.*, 2016). La Figure 9 présente les attributs associés aux strates du Tocantins et du Burkina extraits à partir du produit GAES niveau 4 (l'échelle la plus fine sur les 4 niveaux de stratification disponibles).

[13] Les bases de données associées aux produits cartographiques de GAEZ comprennent des variables liées aux ressources en terres et en eau, aux ressources agro-climatiques, à l'aptitude à différentes cultures et utilisations des terres, et à la production et rendements réels, potentiels, et l'écart entre les deux, des principales cultures (Fischer *et al.*, 2012).



Tocantins			
ID Strata	Dominant land cover type	Dominant crop type	Average field size
299	Tree Covered Areas	Other crops	Medium (2 - 20 ha)
280	Tree Covered Areas	Other crops	Large (> 20 ha)
269	Shrubs Covered Areas	Dry beans	Medium (2 - 20 ha)
308	Tree Covered Areas	Soybean	Large (> 20 ha)

Burkina Faso			
ID Strata	Dominant land cover type	Dominant crop type	Average field size
188	Grassland	Other crops	Small (0.5 - 2 ha)
177	Grassland	Other crops	Small (0.5 - 2 ha)
163	Grassland	Other crops	Small (0.5 - 2 ha)
162	Cropland	Other crops	Small (0.5 - 2 ha)
141	Tree Covered Areas	Coffee	Small (0.5 - 2 ha)
148	Shrubs Covered Areas	Maize	Small (0.5 - 2 ha)

Figure 9. Extrait des attributs associés aux strates du Tocantins (en bas à gauche) et du Burkina Faso (en bas à droite) du produit GAES niveau 4 (en haut) (adaptée de FAO (2017a); Mücher *et al.* (2016)).

1.2.5.3. Limites des approches présentées

Ces exemples de travaux destinés à la planification de l'utilisation des terres à grande échelle sont fondés sur le rassemblement de multiples données historiques hétérogènes. La superposition spatiale de couches de données, dont la qualité, la résolution spatiale et la date d'acquisition sont variables, et la pondération subjective de ces couches d'information, remet en question la précision et fiabilité des produits cartographiques résultants et limite inévitablement la mise à jour des résultats (Fischer *et al.*, 2002; Nachtergaele et Petri, 2013; Van Wart *et al.*, 2013).

Le plus souvent, les informations sur l'utilisation agricole des terres, notamment le type de cultures et les pratiques culturales, qui sont utilisées pour dériver d'autres variables (ex. l'écart entre le rendement potentiel et réel de différentes cultures en tenant compte des pratiques culturales du produit GAEZ), proviennent des statistiques agricoles agrégées à des échelles administratives peu représentatives de la répartition spatiale des surfaces cultivées.

La plupart des données statistiques portent sur les surfaces occupées par différentes cultures, sur les rendements des cultures et sur les surfaces irriguées à l'échelle des divisions administratives nationales ou sous-nationales qui proviennent de grandes bases de données comme "FAOSTAT" (FAO, 2017a), "Agro-Maps" (FAO, s.d.-a) ou "AQUASTAT" (FAO, s.d.-b). Les données statistiques dans ces bases de données peuvent être incomplètes, dû aux contraintes de disponibilité dans certaines régions du monde, et contiennent des informations de qualité hétérogène d'une région à l'autre (AfDB, 2011; Dixon *et al.*, 2001; Nachtergaele et Petri, 2013).

En outre, les divisions administratives ayant des étendues variables dans les différents pays, les statistiques agricoles agrégées à des échelles administratives sous-nationales présentent un niveau de détail variable selon le pays considéré. Par conséquent, la plupart de ces travaux ont recours à des techniques de ré-échantillonnage pour ramener les différentes couches d'information à un quadrillage matriciel avec une résolution spatiale (réduite) commune au travers de complexes modèles d'allocation spatiale (ex. Monfreda, Ramankutty et Foley 2008; Siebert *et al.*, 2015; Siebert, Hoogeveen et Frenken, 2006; You et Wood, 2006; You *et al.*, 2014). Les approches comme celui des "Systèmes de production" de la FAO et la Banque Mondiale (Dixon *et al.*, 2001) qui, à l'image des approches de zonage traditionnel, extrapolent les données d'entrée sur des larges unités cartographiques synthétiques, se retrouvent avec des résultats trop généralisés qui s'avèrent peu utiles au sein des pays (Nachtergaele et Petri, 2013).

Dans tous les cas, le manque général de données spatialement explicites sur les différents systèmes agricoles à l'échelle nationale et régionale, ainsi que des méthodes opérationnelles permettant un suivi régulier des surfaces cultivées et des pratiques culturales, affectent la précision et la mise à jour des cartes et analyses issues de ces approches (Leenhardt *et al.*, 2010; Monfreda *et al.*, 2008). À terme, ces limites freinent leur usage par les gestionnaires lors de la planification territoriale et la prise de décision concernant l'aménagement et l'utilisation des terres (Fischer *et al.*, 2002; Van Wart *et al.*, 2013).

Les nouveaux progrès technologiques en matière d'observation spatiale de la Terre offrent un potentiel encore peu exploité pour complémenter les statistiques agricoles régionales avec des informations détaillées et spatialement explicites sur les systèmes agricoles, d'autant plus précieuses dans les régions où les statistiques agricoles sont manquantes ou peu fiables. Dans ce contexte, sont présentés dans la section suivante les développements concernant l'utilisation de la télédétection pour la cartographie de systèmes agricoles et des exemples d'approches existantes.

Bilan : La cartographie et la caractérisation de systèmes agricoles

- La cartographie de systèmes agricoles fournit des informations spatiales sur la localisation des activités agricoles, qui sont nécessaires pour toute évaluation agronomique (en termes de production agricole) et environnementale, actuelle ou future.
- La caractérisation spatialisée de systèmes agricoles peut être faite à partir de la cartographie et de l'analyse de l'utilisation agricole des terres, décrite par les espèces cultivées et les pratiques culturales associées, qui est une composante essentielle des systèmes agricoles.
- Les systèmes agricoles peuvent être cartographiés à plusieurs échelles temporelles et spatiales, chaque échelle apportant des informations complémentaires sur les systèmes.
- Les approches de cartographie des systèmes d'utilisation agricole des terres et de zonage agricole à l'échelle globale et régionale présentent plusieurs limites, liées à la grande hétérogénéité des données utilisées, qui affectent la précision et la mise à jour des produits. Ceci entraîne un manque général d'informations spatialement explicites et régulières sur les systèmes agricoles des différentes régions du monde.

1.3. La télédétection et la cartographie de systèmes agricoles

1.3.1. L'image satellitaire : source d'information sur les systèmes agricoles

La télédétection aérospatiale, qui se définit comme « l'ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance » (Ministère de l'éducation nationale et ministère de l'enseignement supérieur et de la recherche, 1980), permet, à partir de l'acquisition et le traitement des images satellitaires, d'extraire des informations sur l'état et l'évolution de la végétation de façon régulière sur l'ensemble de la surface terrestre (les principes de base de la télédétection des couverts végétaux peuvent être consultés dans l'ouvrage de Girard et Girard (2010)). Depuis les années 70, les rapides progrès technologiques ont permis la multiplication du nombre d'images satellitaires d'observation de la Terre disponibles, offrant une large gamme de résolutions temporelles, spatiales et spectrales.

Grâce à cette grande variété d'images disponibles et la continuité des missions satellitaires, la télédétection permet aujourd'hui d'estimer et cartographier sur une ou plusieurs années, diverses composantes des systèmes agricoles liées à la production végétale (Tableau 1)[14].

Tableau 1. Principales composantes des systèmes agricoles liées à la production végétale, détectables par télédétection à partir d'imagerie satellitaire (voir Bégué *et al.* (2018) pour plus de détails).

Sur une année	Sur plusieurs années
Surfaces cultivées	**Trajectoires des surfaces cultivées**
Type de culture	Évolution des assolements
Groupe de cultures	Extension des cultures
Pratiques culturales	**Évolution des pratiques culturales**
Culture séquentielle	Monoculture
Association culturale	Succession et rotation culturales
Irrigation	Jachère
Travail du sol	Intensification des pratiques
Production végétale	**Tendances de la production**
Biomasse	Évolution de la production de biomasse
Rendement	Évolution du rendement

Les satellites d'observation de la Terre et leurs capteurs à bord ont des caractéristiques particulières qui font que les images sont plus ou moins adaptées à la cartographie de ces différentes composantes. Les caractéristiques spectrales correspondent au domaine et au nombre de bandes spectrales (richesse spectrale) et à leur largeur (résolution spectrale), à la fréquence d'acquisition des images (ou répétitivité de passage du satellite) qui détermine sa résolution temporelle, à la résolution spatiale et à la fauchée (surface « observée » par le capteur) (voir Girard et Girard (2010) pour plus de détails).

Le Tableau 2 montre les caractéristiques des principales missions satellitaires pour le suivi de la végétation qui sont actuellement opérationnelles.

[14] Des exemples d'applications de la télédétection dans le domaine de l'agriculture sont répertoriés dans les articles de synthèse d'Atzberger (2013), Bégué *et al.* (2015) et Bégué *et al.* (2018).

Tableau 2. Caractéristiques des principaux satellites opérationnels en 2017 fournissant des images à haute, moyenne et basse résolution spatiale dans le domaine optique pour le suivi de la végétation (entre autres applications).

Satellite/Constellation	Début d'acquisition	Capteur	Bandes spectrales VIS et IR	Résolution temporelle (jour)	Résolution spatiale (m)	Fauchée (km)	Organisation(s)	Politique de données
Basse résolution spatiale (kilométrique)								
NOAA 18 ; 19	2005 ; 2009	AVHRR	R, PIR, MIR, IRT	Journalière	1100	2399	NOAA	Libre accès
Metop-A ; B	2006 ; 2012					2893	EUMESAT	
Moyenne résolution spatiale (hectométrique)								
PROBA-V	2013	VGT-P	B, R, PIR, MIR	Journalière	100 ; 300 ; 1000	2285	ESA	Libre accès
Terra ; Aqua	2000 ; 2002	MODIS	B, V, R, PIR, MIR, IRT	Journalière	250 ; 500 ; 1000	2330	NASA	Libre accès
Haute résolution spatiale (décamétrique et métrique)								
Landsat 8	2013	OLI	B, V, R, PIR, MIR	16 jours	15* ; 30	185	USGS/NASA	Libre accès
		TIRS	IRT		100			
Landsat 7	1999	ETM+	B, V, R, PIR, MIR, IRT		15* ; 30 ; 60			
Sentinel-2A ; 2B	2015 ; 2017	MSI	B, V, R, PIR, MIR	5 jours	10 ; 20	290	ESA	Libre accès
Venµs	2017	VSSC	B, V, R, PIR	2 jours	5	27,5	CNES/ISA	Projets scientifiques sélectionnés
ResourceSat-2A	2016	LISS-3	V, R, PIR, MIR	5 jours	23,5	140	ISRO	Commerciale
		LISS-4	V, R, PIR		5,8	23,9		
		AWiFS	V, R, PIR, MIR		56	740		
SPOT 6 ; 7	2012 ; 2014	NAOMI	B, V, R, PIR	Journalière	1,5* ; 6	60	CNES/Airbus	Commerciale

* Correspond aux résolutions spatiales des bandes panchromatiques. Les autres valeurs de résolution spatiale correspondent aux bandes multi-spectrales dans les domaines visible (VIS) et Infra-Rouge (IR) : Bleu (B), Vert (V), Rouge (R), Proche Infra-Rouge (PIR), Moyen Infra-Rouge (MIR) et Infra-Rouge Thermique (IRT).

Les images issues des satellites présentés dans le Tableau 2 réunissent les conditions nécessaires pour la cartographie et le suivi des surfaces cultivées, des pratiques culturales et de la production agricole. Ces conditions sont d'une part, des canaux spectraux ou bandes spectrales dans les domaines visible et infrarouge, nécessaires pour le calcul des indices de végétation[15], et d'autre part, une haute résolution temporelle.

Les indices de végétation, dont le plus connu est l'indice de végétation par différence normalisé ou NDVI (acronyme anglais de *Normalized Difference Vegetation Index*) (Rouse *et al.*, 1973), sont liés aux variables biophysiques du couvert végétal telles que le taux de couverture, la biomasse, l'indice de surface foliaire ou l'indice de surface foliaire verte (surface foliaire photosynthétiquement active) (Girard et Girard, 2010). Les types de culture, des indices de production et certaines pratiques peuvent être donc estimés à partir des indices de végétation.

Les images acquises à une fréquence temporelle suffisamment élevée pour capter les variations intra-saisonnières de la végétation constituent, quant à elles, des séries annuelles qui sont essentielles pour le suivi de la phénologie des surfaces cultivées. Les différents cycles culturaux et stades de croissance des cultures peuvent être détectés à partir de ces séries d'images, et certaines pratiques telles que la culture séquentielle (plusieurs espèces cultivées les unes après les autres pendant une année) ou les dates de semis et de récolte, peuvent être ensuite déduites de ces informations (des exemples sont présentés par Bégué *et al.* (2018)).

Les images radar à synthèse d'ouverture ou SAR[16] (acronyme anglais de *Synthetic Aperture Radar*), avec des canaux dans le domaine des micro-ondes, permettent aussi d'obtenir des informations sur la végétation (ex. indice de surface foliaire, teneur en eau, biomasse), ainsi que sur les états de surface (ex. humidité et rugosité du sol). Ces images peuvent être ainsi utilisées pour la détection des types de culture (ex. Corgne, Hubert-Moy et Betbeder, 2016) et des pratiques culturales notamment l'irrigation et le travail au sol (ex. Hadria *et al.*, 2009; Pacheco, McNairn et Merzouki, 2010).

[15] Les indices de végétation peuvent être calculés à partir des bandes spectrales fournies dans les images originelles ou bien récupérées directement sous forme de produits de synthèse d'images qui sont mis à disposition dans le cadre de certains programmes satellitaires. "MOD13Q1" et "ENDVI10" sont deux exemples de produits d'indices de végétation actuellement disponibles.

[16] Quelques satellites opérationnels en 2017 fournissant des images SAR sont Sentinel-1A, 1B, RadarSat-2, ALOS-2 et TerraSAR-X.

En outre, la durée des missions satellitaires pluriannuelles permet le suivi des changements d'utilisation de terres agricole ainsi que l'estimation des tendances de production sur plusieurs années consécutives. De plus, grâce à la continuité des missions satellitaires[17] on trouve aujourd'hui une disponibilité accrue d'images d'archive qui permet de réaliser des études rétrospectives et un suivi à long terme de ces dynamiques.

En fonction de l'objectif recherché, toutes les images n'ont pas le même intérêt. Par exemple, si l'on cherche à caractériser les systèmes agricoles à l'échelle de la parcelle, nous ne nous intéresserons qu'aux images ayant une résolution inférieure à la taille des parcelles, de sorte à avoir au moins un pixel avec de l'information spectrale « pure » sur les parcelles ciblées. Strahler, Woodcock et Smith (1986) définissent cette situation comme un "cas de haute résolution" (*H-resolution case*) et la situation inverse, lorsque la taille de l'unité spatiale à cartographier (ex. la parcelle) est plus petite que la taille du pixel, comme un "cas de basse résolution" (*L-resolution case*).

Woodcock et Strahler (1987) montrent que les pixels d'une taille supérieure aux ¾ de la taille des unités spatiales d'intérêt, contiennent des informations spectrales mixtes (issues de plusieurs éléments de la surface). Ainsi, on pourrait utiliser le facteur d'échelle de ¾ pour situer les limites de résolution spatiale dans lequel on se trouverait sous un "cas de haute résolution" pour différentes tailles de parcelle (cf. Tableau 3):

Tableau 3. Exemple de limites de résolution spatiale auxquelles les systèmes agricoles peuvent être caractérisés à l'échelle de la parcelle, en se situant dans un "cas de haute résolution" pour différentes tailles de parcelle.

	Taille de parcelle		
	Petite (< 1 ha)	Moyenne (1 – 12 ha)	Grande (> 12 ha)
Ordre de grandeur de la résolution spatiale	< 90 m	90 – 300 m	> 300 m

Les besoins en résolution spatiale diffèrent donc selon la région à cartographier compte tenu de la grande variabilité de la taille des parcelles agricoles dans le monde (Figure 10). Pour cartographier les régions avec des parcelles de petite taille (ex. régions d'Afrique subsaharienne ou d'Asie du Sud), l'on aura besoin d'images dotées d'une résolution spatiale fine, tandis que pour les régions avec des parcelles de grande taille l'on pourra se servir d'un

[17] À titre d'exemple, les satellites NOAA, Landsat et SPOT présentés dans le Tableau 2 assurent la continuité de décennies d'acquisitions d'images avec de séries de satellites qui ont débuté en 1979, 1972 et 1986, respectivement.

choix plus large d'images, avec des résolutions spatiales allant de la basse à la haute résolution.

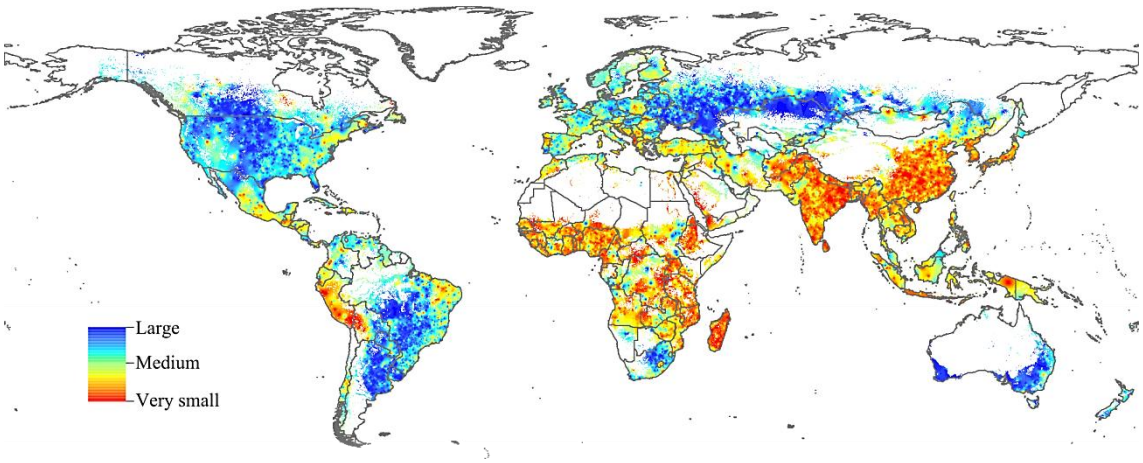


Figure 10. Carte globale de la taille des parcelles agricoles (source : Fritz *et al.*, 2015).

Par ailleurs, de façon générale il existe une relation inverse entre la résolution spatiale et l'emprise des images (déterminée par la fauchée du satellite), i.e. plus la résolution spatiale des images satellitaires est fine, plus petite est l'emprise de celles-ci. Ainsi, pour obtenir une couverture à haute résolution spatiale sur une grande étendue on a généralement besoin d'un nombre élevé d'images. Un nombre élevé d'images compliquent les traitements postérieurs (ex. les images peuvent présenter différentes dates d'acquisition) et les rend coûteux en termes de ressources informatiques et de temps de calcul (ex. plus de pixels à traiter et plus d'étapes à inclure dans la chaîne de traitement). De plus, si les images sont commerciales, le prix peut résulter très élevé.

Outre la taille des parcelles agricoles, la localisation de la région d'étude introduit aussi des contraintes dans le choix des images. Plus spécifiquement, les régions tropicales sont affectées par une forte nébulosité pendant de longues périodes, ce qui rend l'information des séries temporelles d'images difficilement utilisable. De plus, la saison culturale correspond généralement à la saison de pluies ou saison humide, ce qui limite l'information sur la phénologie des surfaces cultivées. Pour augmenter les chances de récupérer de l'information spectrale non bruitée par les effets atmosphériques pendant la saison culturale, on va ainsi privilégier sur ces régions des images acquises avec une fréquence journalière. Le couplage des images radar avec les images optiques est aussi utilisé dans ces cas, l'information des images radar étant peu affectée par la nébulosité (ex. Hadria *et al.*, 2009; McNairn *et al.*, 2009; Panigrahy *et al.*, 2011; Skakun *et al.*, 2016).

Finalement, la politique de données des différentes missions satellitaires limite ou facilite l'accès aux images, ce qui a une forte influence sur la cartographie et le suivi de systèmes agricoles, notamment sur la fréquence de production des cartes et leur mise à jour. La mise à disposition des images d'observation de la Terre en libre accès a commencé avec l'ouverture des archives Landsat en 2008 et 2009 par l'Institut d'Études Géologiques des États-Unis (USGS, acronyme anglais de *United States Geological Survey*), et continue aujourd'hui avec des programmes comme Copernicus de l'Agence Spatiale Européenne (ESA, acronyme anglais de *European Space Agency*), qui met à disposition publiquement les images d'archive des constellations Sentinel. Concernant la programmation des acquisitions d'images commerciales, les prix varient selon les caractéristiques de l'acquisition (ex. taille de l'emprise), le niveau de prétraitements appliqués et l'utilisation des images (des tarifs préférentiels étant appliqués par certains programmes pour les projets de recherche), mais sont généralement prohibitifs pour des applications agricoles, celles-ci étant très gourmandes en données temporelles.

1.3.2. État de l'art des approches de télédétection pour la cartographie de systèmes agricoles

Les approches de télédétection appliquées au suivi de l'agriculture se sont développées au cours des dernières décennies en parallèle au développement de la technologie satellitaire et des capacités de traitement d'images. Cependant, l'article de synthèse réalisé par Bégué *et al.* (2015) montre que moins de 10% des publications scientifiques concernant la télédétection appliquée à l'agriculture prennent en considération les pratiques culturales dans leurs approches, et que ces études sont généralement menées à l'échelle locale (ex. Daughtry *et al.*, 2006; Galford *et al.*, 2008; Hadria *et al.*, 2009; Maus *et al.*, 2016; Pacheco et McNairn, 2010; Viña *et al.*, 2003)[18].

À l'échelle régionale, relativement peu d'approches de cartographie de systèmes agricoles par télédétection existent, et les informations sur les systèmes sont souvent limitées. À cette échelle, les approches ont tendance à simplifier la caractérisation des systèmes agricoles en se focalisant uniquement sur les principaux groupes de cultures et des espèces cultivées, omettant généralement les informations sur les pratiques culturales (Leenhardt *et al.*, 2010; Monfreda *et al.*, 2008).

[18] Un état de l'art exhaustif des approches de cartographie des systèmes de culture par télédétection, avec des exemples des méthodes utilisés pour détecter les différents types de pratiques culturales est présenté dans l'article de synthèse réalisé par Bégué *et al.* (2018).

La contrainte pour extraire des informations détaillées sur les systèmes agricoles à l'échelle régionale par télédétection ne provient pas seulement de la limite de données appropriées et disponibles, présentée dans la section 1.3.1, mais aussi de la complexité des systèmes en soi. En effet, les systèmes agricoles résultent d'une combinaison de facteurs humains et environnementaux qui varient naturellement dans l'espace, et plus l'espace considéré est grand, plus la variabilité spatiale des systèmes est importante. Ainsi, un même système agricole va présenter des spécificités locales qui conduisent à des réponses spectrales et temporelles très variables au sein d'une même région. De plus, à l'échelle régionale, les systèmes agricoles font partie d'une mosaïque de paysages complexe, composée d'une grande diversité de types d'occupation et d'utilisation des terres desquels il est difficile de les dissocier.

1.3.2.1. Exemples d'études menées à l'échelle régionale

Des études de télédétection récentes ont toutefois surmonté ces défis méthodologiques en adoptant des approches originales pour cartographier des systèmes agricoles à l'échelle régionale. En effet, au cours des dix dernières années, les systèmes agricoles de différentes régions du monde[19] ont été caractérisés au travers de l'utilisation agricole des terres, en classifiant les différentes espèces cultivées et certaines pratiques culturales associées, notamment la culture séquentielle (ex. Arvor *et al.*, 2011; Brown *et al.*, 2013; Cheema et Bastiaanssen, 2010; Guan *et al.*, 2016; Gumma *et al.*, 2016; Mingwei *et al.*, 2008; Wardlow et Egbert, 2008) et l'irrigation (ex. Bridhikitti et Overcamp, 2012; Cheema et Bastiaanssen, 2010; Dheeravath *et al.*, 2010; Guan *et al.*, 2016; Gumma *et al.*, 2016; Wardlow et Egbert, 2008). Le résultat cartographique obtenu par Arvor *et al.* (2011) est présenté dans la Figure 11 à titre d'exemple.

[19] L'échelle régionale considérée dans ces travaux concerne des régions administratives et des régions biophysiques (ex. le bassin versant de l'Indus) de plus de 100 000 km^2.

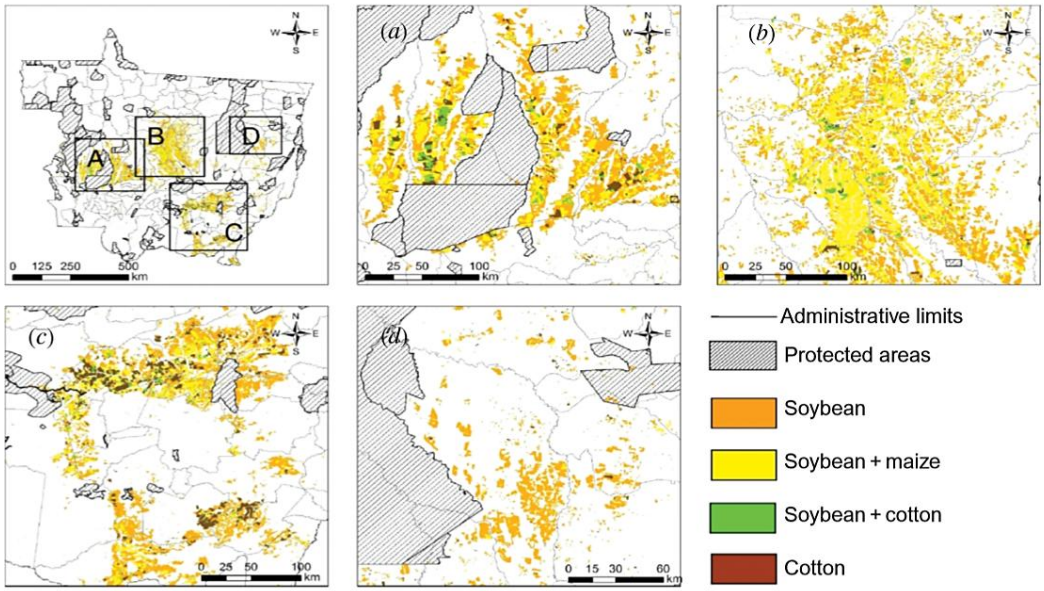


Figure 11. Carte des quatre cultures commerciales et séquences culturales, de la saison de 2006-2007 sur les quatre régions agricoles principales du Mato Grosso (source : Arvor *et al.*, 2011).

Des pratiques sur plusieurs années, comme les rotations culturales et la monoculture, ont pu être aussi détectées (ex. Plourde *et al.*, 2013; Sahajpal *et al.*, 2014) à partir de cartes d'espèces cultivées obtenues principalement de données satellitaires (Boryan *et al.*, 2011). Cependant, il y a un manque général d'approches sur des pratiques pluriannuelles puisqu'elles reposent entièrement sur une production régulière de cartes d'espèces cultivées à l'échelle régionale. Les pratiques comme l'association de cultures, dont la culture intercalaire et l'agroforesterie, qui sont détectées à une échelle intra-parcellaire, sont aussi rarement abordées par télédétection satellitaire, et les peu d'études sont menées à l'échelle locale puisqu'elles nécessitent des images à très haute résolution spatiale, métrique ou même infra-métrique[20] (ex. Karlson *et al.*, 2016; Lelong, Alexandre et Dupuy, 2014).

Les études qui sont menées à l'échelle régionale, comme celles citées précédemment, privilégient plutôt l'utilisation d'images acquises à une fréquence journalière et sur des grandes étendues. Les séries temporelles d'indices de végétation issues du capteur MODIS (*Moderate Resolution Imaging Spectroradiometer*) ont été ainsi jusqu'à aujourd'hui largement les plus utilisées dans la cartographie de systèmes agricoles par télédétection. Dans la plupart des cas, l'information des indices de végétation est utilisée directement, mais quelques études en extraient de nouveaux indices qui apportent des informations

[20] Les images satellitaires à très haute résolution spatiale issues de satellites commerciaux actuellement opérationnels (ex. Pléiades, Worldview, KOMPSAT, SkySat) ont des emprises inférieures à 500 km^2 et sont composées d'un grand nombre de pixels, ce qui rendrait à ce jour les méthodes trop coûteuses en termes de ressources informatiques pour qu'elles soient adaptées aux applications à l'échelle régionale.

complémentaires. Quelques techniques utilisées à cette fin sont la transformée de Fourier (ex. Mingwei *et al.*, 2008), la transformation en ondelettes (ex. Qiu *et al.*, 2014) et le calcul de textures (ex. Vintrou *et al.*, 2012). D'autres études complètent l'information des séries temporelles optiques avec celle des données radar (ex. Panigrahy *et al.*, 2011).

1.3.2.2. Méthodes de classification

Les méthodes de traitement de l'information des séries temporelles pour la production de cartes concernent généralement des classifications statistiques basées sur des algorithmes d'apprentissage automatique, supervisées ou non supervisées[21]. Quelques exemples de méthodes de classification supervisées sont le maximum de vraisemblance, les arbres de décision, les réseaux de neurones et les machines à vecteurs de support (ex. Arvor *et al.*, 2011; Brown *et al.*, 2013; Massey *et al.*, 2017; Mingwei *et al.*, 2008; Wardlow et Egbert, 2008). Les méthodes supervisées reposent entièrement sur des données de référence, habituellement sous forme d'observations de terrain pour entrainer les modèles de classification. Or, la collecte de données de terrain une ou plusieurs fois par an s'avère une tâche difficile, qui demande beaucoup de temps et de ressources. Cet exercice est particulièrement ardu quand il doit être réalisé de façon extensive pour assurer la représentativité de la variabilité spectrale liée à la forte variabilité spatiale des systèmes agricoles au sein d'une région (Franklin et Wulder, 2002). Par conséquent, la reproductibilité des méthodes supervisées est assez limitée dans le temps et entre différentes régions du monde.

En revanche, les méthodes de classification non supervisées, basées uniquement sur l'information spectrale et temporelle des séries d'images, sont plus opérationnelles, sans pour autant affecter la précision des cartes résultantes (ex. Cheema et Bastiaanssen, 2010; Gumma *et al.*, 2016). Des exemples de méthodes non supervisées sont les algorithmes de partitionnement ou *clustering* itératives tels que le *k*-means, l'ISODATA (*Iterative Self-Organizing Data Analysis Technique*) ou la classification ascendante hiérarchique. Les techniques d'"hyperclustering" sont généralement utilisées, produisant un grand nombre initial de clusters qui est ensuite réduit par des étapes de traitements successifs. Cette technique garantit que les classes d'intérêt sous-représentées soient correctement isolées et présente l'avantage de réduire l'impact des paramètres définis par l'utilisateur (ex. le nombre de clusters) sur la classification finale (Cihlar, 2000). Cependant, les techniques de raffinement de classification utilisées dans ces études, telles que la fusion ou la sous-division

[21] Pour plus de détails sur les différents algorithmes de classification supervisée et non supervisée consulter Richards et Jia (2006).

de clusters, sont fortement axées sur des connaissances expertes et reposent souvent sur des données externes (ex. masques du domaine cultivée, données de référence) qui limitent leur application dans d'autres régions du monde.

Certaines études combinent classification supervisée et non supervisée dans des approches mixtes, où les étapes de classification s'enchaînent de façon à sous-diviser successivement les classes ; par exemple le domaine cultivé est classé en premier et ensuite sous-divisé en différents systèmes de culture (ex. Mingwei *et al.*, 2008; Panigrahy *et al.*, 2011; Wardlow et Egbert, 2008). Ces approches de classification par étapes, parfois hiérarchiques, aident à réduire progressivement la variabilité spectrale initiale de l'image, en facilitant la séparabilité finale des classes d'intérêt, mais ces études ont recours à des données externes et des données de référence collectées sur le terrain pour guider les étapes de classification.

1.3.2.3. Techniques de réduction de la variabilité spectrale

Afin de diminuer la variabilité spectrale élevée qui existe sur de grandes étendues sans dépendre de données externes, les approches de cartographie de systèmes agricoles à l'échelle régionale bénéficieraient largement d'étapes de réduction de la variabilité spatiale avant la classification.

Les techniques de segmentation dérivées de l'analyse d'image orientée-objet, connues comme OBIA ou GEOBIA (acronymes anglais de *(Geographic) Object-Based Image Analysis*), sont maintenant largement utilisées par la communauté de télédétection pour réduire la variabilité spectrale en regroupant des pixels individuels en objets thématiquement significatifs qui résument leurs propriétés spectrales (Blaschke *et al.*, 2014; Bunker *et al.*, 2016; Hay et Castilla, 2008). Or, les exemples d'études travaillant à l'échelle régionale répertoriés dans la section 1.3.2.1 privilégient tous des approches de classification par pixel.

La stratégie de réduction de la variabilité spatiale, dite stratification[22], consiste à partitionner les régions en sous-régions, ou strates, en amont des classifications (ex. Cai *et al.*, 2011; Cano *et al.*, 2017; Franklin et Wulder, 2002; Massey *et al.*, 2017; Vintrou *et al.*, 2012; Wardlow et Egbert, 2008; Xiong *et al.*, 2017). Traditionnellement, des cartes externes sont utilisées (ex. écorégions, zonages agro-écologiques, régions agro-climatiques), mais quelques

[22] La communauté de télédétection n'a pas encore formulée une définition consensuelle du terme stratification. Ce terme, à l'origine statistique, a été pourtant communément employé en référence à un découpage spatial des territoires en sous-régions géographiques relativement homogènes à partir de variables environnementales et/ou climatiques (ex. Hazeu *et al.*, 2011; Mücher *et al.*, 2016).

approches innovantes ont commencé à explorer le potentiel des stratifications issues d'images satellitaires en utilisant des techniques de segmentation OBIA pour les classifications de l'occupation des terres sur de grandes étendues (ex. Cano *et al.*, 2017; Vintrou *et al.*, 2012).

Ces premiers travaux de stratification par télédétection, délimitant des objets géographiques extensifs à partir des données à moyenne résolution spatiale, divergent de l'utilisation courante des techniques de segmentation OBIA, qui généralement concerne la délimitation des objets géographiques à une échelle fine (ex. arbres, bâtiments, routes parcelles) à partir de données à haute ou très haute résolution spatiale.

1.3.2.4. La stratification régionale en unités de paysage

On constate un manque général dans la littérature d'approches de télédétection pour la cartographie et la caractérisation des systèmes agricoles à des niveaux d'organisation supérieurs à celui de la parcelle. En effet, les cartes de systèmes agricoles par télédétection sont généralement produites à l'échelle de la parcelle, ou ensemble de parcelles pour les études qui travaillent sur des régions avec un parcellaire de taille inférieure à la taille des pixels (ex. Cheema et Bastiaanssen, 2010; Guan *et al.*, 2016; Gumma *et al.*, 2016).

La stratification spatiale aurait dans ce cas un potentiel peu exploré pour la mise à l'échelle des classifications obtenues à l'échelle de la parcelle, permettant de les représenter à des niveaux d'organisation supérieurs. Par exemple, Panigrahy *et al.* (2011) ont caractérisé les systèmes agricoles de la plaine indo-gangétique au niveau de régions agro-climatiques à partir d'informations issues d'enquêtes auprès des producteurs sur les successions culturales, le niveau de fertilisation et la qualité des sols. Ils ont relié cette information à des indices sur la diversité des cultures et l'intensité des pratiques qu'ils ont calculés à partir des classifications des cultures séquentielles par télédétection et mise à l'échelle de régions administratives (Sushma Panigrahy *et al.*, 2011).

Les techniques d'OBIA avec lesquelles on peut segmenter automatiquement l'imagerie satellitaire en objets relativement homogènes à plusieurs échelles (Hay et Castilla, 2008), semblent particulièrement adaptées pour la stratification des régions, pouvant générer des unités cartographiques (strates) de taille supérieure à la parcelle et variable selon l'échelle utilisée. Bien que ces techniques aient été largement utilisées pour la segmentation de la végétation à une échelle fine à partir d'imagerie à haute résolution spatiale (Blaschke, Johansen et Tiede, 2011), leur potentiel pour la stratification de régions en unités de paysage commence à peine à être exploré (ex. Bisquert, Bégué et Deshayes, 2015).

Bisquert *et al.* (2015) ont introduit une approche innovante basée sur la segmentation de variables spectrales et texturales dérivées de données MODIS (trois images d'EVI, *Enhanced Vegetation Index,* et de texture correspondant aux moyennes d'avril, juillet et décembre sur 5 ans), pour produire une stratification spatiale de la France métropolitaine en 300 unités de paysage (présentée dans la Figure 12). Les unités de paysage résultantes, non classées, présentent une cohérence avec l'organisation spatiale de l'occupation des terres (Bisquert *et al.*, 2017). Cependant, les variables utilisées représentant seulement les informations de 3 mois de l'année et synthétisant ces informations sur 5 ans perdent le détail temporel des variations intra-saisonnières qui caractérisent les systèmes agricoles, et donc la segmentation n'est pas représentative de l'organisation spatiale de ceux-ci. En outre, dans cette approche, l'échelle de segmentation est déterminée par un critère d'optimisation statistique (i.e. l'échelle à laquelle l'hétérogénéité entre unités de paysage et l'homogénéité au sein des unités sont maximisées est retenue). Or, le concept d'unité de paysage ne s'attache pas à une échelle optimale, mais est plutôt conçu comme un concept multiscalaire (Bourget et Le Dû-Blayo, 2010; Zonneveld, 1989).

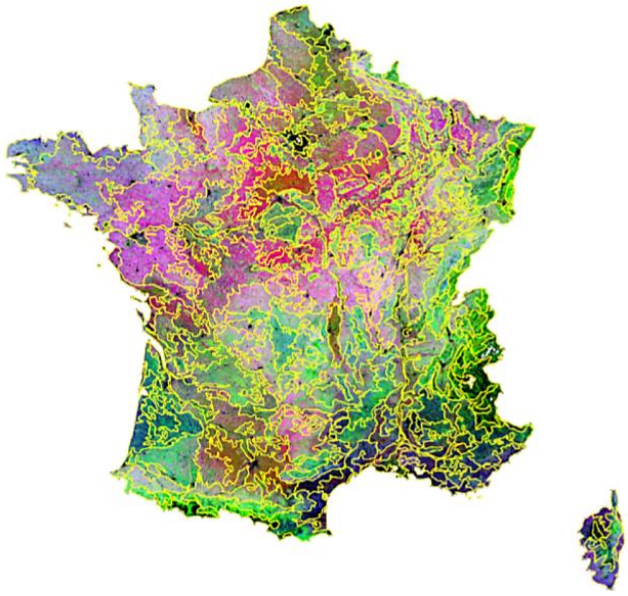


Figure 12. Stratification spatiale de la France métropolitaine en unités de paysage à partir de données MODIS (source : Bisquert *et al.*, 2015).

Bourget et Le Dû-Blayo (2010) proposent une approche alternative de stratification régionale en unités de paysage. Leur approche ascendante part de données locales issues de la télédétection qui sont synthétisées progressivement jusqu'à obtenir des « regroupements en entités homogènes » représentatives des paysages de Bretagne. Les données locales de départ sont une classification des types de paysage (typologie qui correspond à l'occupation des terres) de la région à partir d'une série temporelle d'images MODIS et des données de pente et d'altitude issues d'un modèle numérique de terrain. La résolution spatiale de MODIS est

considérée comme « très pertinente pour les études à l'échelle régionale » puisqu'elle représente « une première synthèse des paysages » (Bourget et Le Dû-Blayo, 2010).

L'approche consiste dans un premier temps à extraire des unités de paysage suivant trois méthodes : une délimitation visuelle, un regroupement des informations locales à l'aide du logiciel ClaPaS (CLAssement de PAysages et Segmentation) (Robbez-Masson *et al.*, 1999), et une méthode basée sur la morphologie mathématique. Les résultats des trois méthodes sont ensuite synthétisés visuellement et par croisement avec des Atlas départementaux de paysages, jusqu'à la délimitation et la classification finale d'unités de paysage synthétiques (Bourget et Le Dû-Blayo, 2010).

Cette approche est innovante du fait de l'utilisation de la télédétection et de certaines méthodes automatiques pour la stratification régionale en unités de paysage. Cependant, des méthodes dites holistiques sont aussi mobilisées (travail de synthèse visuelle, apport de connaissances expertes, croisement avec des données externes), résultant en une approche très « expertisée », qui limite sa reproductibilité et applicabilité sur d'autres régions. En outre, le découpage final en unités de paysage ne représente pas spécialement l'organisation spatiale des systèmes agricoles, mais délimite plutôt des espaces naturels (ex. bassins versants, massifs montagneux, plateaux, marais) avec une utilisation des terres diversifiée au sein des unités extraites.

À partir de ces exemples, on peut néanmoins percevoir le fort potentiel de la stratification régionale à partir d'imagerie satellitaire et des techniques d'OBIA pour la délimitation d'unités de paysage agricole dans le but de représenter les systèmes agricoles à des niveaux d'organisation spatiale supérieurs à celui de la parcelle.

Bilan : La télédétection et la cartographie de systèmes agricoles

- Les progrès technologiques ont contribué au déploiement de nombreux satellites d'observation de la Terre qui permettent la détection de diverses composantes des systèmes agricoles liées à l'utilisation et la productivité des terres. La cartographie de systèmes agricoles à l'échelle régionale par télédétection s'avère cependant un exercice complexe, dû, entre autres, à la forte variabilité spectrale des surfaces liée à la grande diversité des conditions environnementales et des pratiques agricoles au sein d'une région.
- La majorité des études sont ainsi menées à l'échelle locale et utilisent des classifications supervisées pour cartographier les espèces cultivées, sans incorporer des informations sur les pratiques agricoles. À l'échelle régionale, la classification supervisée est aussi privilégiée, et les approches reposent souvent sur des connaissances expertes et des données externes, ce qui limite la mise à jour régulière des cartes et le transfert entre régions du monde.
- Afin de diminuer la variabilité spectrale sans dépendre de données externes, la cartographie de systèmes agricoles à l'échelle régionale par télédétection bénéficierait de techniques de réduction de la variabilité spatiale telles que la segmentation et la stratification spatiale.
- La stratification spatiale par télédétection aurait également un potentiel non exploré pour réaliser un transfert d'échelle des informations sur les systèmes agricoles obtenues à l'échelle de la parcelle jusqu'à des niveaux d'organisation supérieurs, au travers de la délimitation d'unités de paysage agricole.

1.4. Objectif et questions de recherche

Dans le cadre de la production de données spatiales en appui au suivi de l'agriculture à grande échelle, on constate que la production de cartes de systèmes agricoles à l'échelle régionale s'avère nécessaire pour compléter le manque général d'informations sur les activités agricoles et leur localisation avec des informations détaillées, régulières, objectives et spatialement explicites. La production de ces informations de façon opérationnelle nécessite des approches de cartographie innovantes et génériques qui intègrent des méthodes simples et pragmatiques, non figées en termes de données, en équilibrant la performance technique et la facilité de mise en œuvre.

La télédétection offre un potentiel avéré dans la production de cartes de systèmes agricoles à l'échelle régionale de façon opérationnelle, mais aujourd'hui la plupart des approches de télédétection concernent la classification de l'occupation des terres (ex. espèces cultivées) à l'échelle locale avec des méthodes supervisées. Ces approches reposent fortement sur des données externes et les informations sur les pratiques agricoles sont généralement omises.

Nous émettons l'hypothèse que les images satellitaires contiennent des informations synoptiques sur les principaux éléments constitutifs des paysages agricoles et sur les facteurs humains qui modèlent ces paysages tels que l'organisation spatiale des surfaces cultivées, la taille et la forme du parcellaire, les pratiques agricoles et les changements d'occupation et utilisation des terres, et reflètent en même temps les facteurs environnementaux tels que le climat, les ressources en eau, le relief et les propriétés des sols qui déterminent l'aptitude à la production agricole, la phénologie des surfaces agricoles et leur productivité. Ainsi, on part du principe que ces informations peuvent être extraites des images satellitaires sous forme d'indicateurs temporels et spectraux suffisamment descriptifs des systèmes agricoles pour leur localisation et caractérisation à plusieurs échelles sans avoir recours à de données externes.

> L'objectif principal de cette thèse est de développer et d'évaluer une approche multiscalaire, basée sur le traitement et l'analyse d'imagerie satellitaire optique, pour la cartographie et la caractérisation de systèmes agricoles à l'échelle régionale.

Pour atteindre l'objectif général, ce travail de thèse s'attache à répondre aux questions de recherche suivantes :

- Comment délimiter des unités de paysage agricole à partir des données satellitaires pour la cartographie des systèmes d'utilisation agricole des terres ?
- Comment aborder la forte variabilité spectrale associée à la grande diversité de surfaces et d'utilisation des terres au sein d'une région pour cartographier les espèces cultivées et les pratiques agricoles ?
- Quelle est la reproductibilité de l'approche développée face au changement de type d'agriculture (de l'agriculture intensive à la petite agriculture familiale) ?

1.5. Approche générale

L'approche développée est basée sur le traitement et l'analyse d'images satellitaires optiques avec des techniques de réduction de la variabilité spectrale et des méthodes de classification non supervisée pour la cartographie et la caractérisation de systèmes agricoles à l'échelle régionale. C'est une approche multiscalaire qui se décompose en deux méthodes principales, visant chacune à produire un type de carte régionale de systèmes agricoles :

(i) Cartographie des systèmes d'utilisation agricole des terres par une méthode de stratification régionale en unités de paysage et classification des unités à partir d'une série annuelle de NDVI à moyenne résolution spatiale ;

(i) Cartographie des systèmes de culture par une méthode de classification orientée-objet non supervisée à l'échelle de la parcelle à partir d'une mosaïque à haute résolution spatiale et d'une série annuelle de NDVI à moyenne résolution spatiale.

L'approche est mise en œuvre sur l'État brésilien du Tocantins, une région de plus de 200 000 km^2 avec une grande diversité de surfaces et d'utilisation des terres, et est décrite dans les chapitres 3 et 4. Ces deux chapitres fournissent des réponses aux deux premières questions de recherche. La Figure 13 montre sous forme de schéma synthétique les principales étapes de l'approche générale décrites dans les chapitres 3, 4 au cœur de cette thèse.

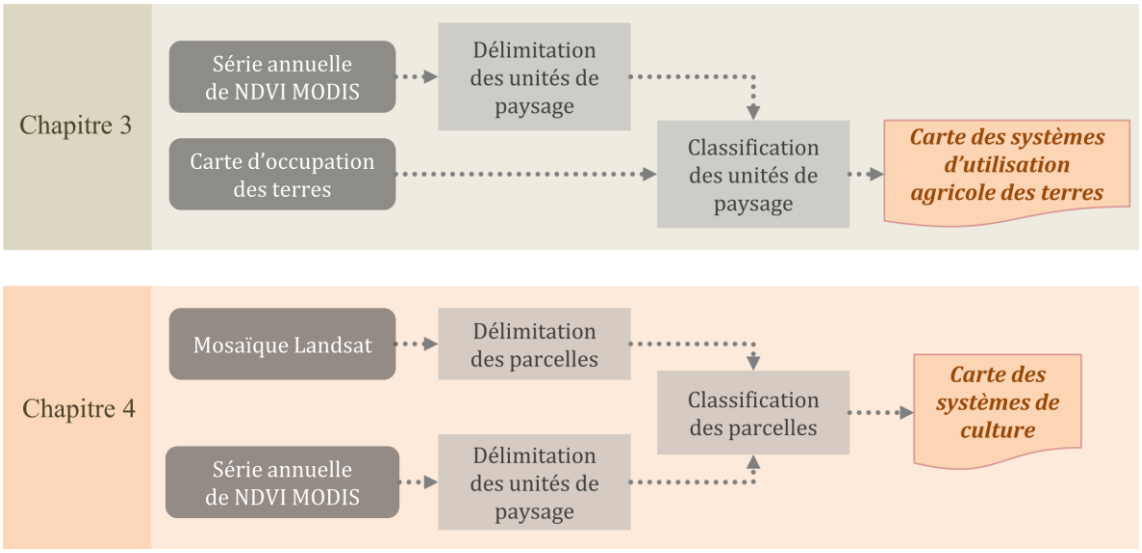


Figure 13. Principales étapes de l'approche générale décrites dans les chapitres 3 et 4.

Ensuite, la reproductibilité des méthodes est évaluée en Afrique de l'Ouest au Burkina Faso, une région de superficie semblable, en réponse à la dernière question de recherche. Les résultats de cette évaluation sont présentés et discutés dans le chapitre 5.

Chapitre 2.
Régions d'étude, données et outils

2.1. LES RÉGIONS D'ÉTUDE 48
2.1.1. L'État brésilien du Tocantins 48
2.1.1.1. Présentation générale du Tocantins 48
2.1.1.2. Les systèmes agricoles du Tocantins 51
2.1.2. Le Burkina Faso 54
2.1.2.1. Présentation générale du Burkina Faso 54
2.1.2.2. Les systèmes agricoles du Burkina Faso 59
2.2. LES DONNÉES 61
2.2.1. Données satellitaires 62
2.2.1.1. Les séries temporelles d'images de NDVI du capteur MODIS 62
2.2.1.2. La mosaïque Landsat 8 69
2.2.1.3. Les images Sentinel-2A et SPOT 7 71
2.2.2. Carte d'occupation des terres 72
2.2.3. Données de validation 73
2.2.3.1. Données de vérité terrain 73
2.2.3.2. Statistiques agricoles 76
2.3. LES OUTILS 77
2.3.1. L'analyse en composantes principales 77
2.3.2. Les coefficients de relation entre variables 79
2.3.3. L'algorithme de segmentation multi-résolution 80
2.3.4. L'algorithme de classification non supervisée k-means 81
2.3.5. Les indices d'évaluation de classifications 82

Chapitre 2. Régions d'étude, données et outils

2.1. Les régions d'étude

2.1.1. L'État brésilien du Tocantins

2.1.1.1. Présentation générale du Tocantins

L'État du Tocantins, situé au centre-nord du Brésil (5°10' - 13°27'S, 45°41' - 50°44'O), s'étend sur 277 621 km^2 dans une zone en transition entre les biomes de l'Amazonie et du Cerrado (Figure 14). C'est une région agricole à forte influence anthropique qui a connu une évolution importante de l'utilisation des terres au cours des dernières décennies.

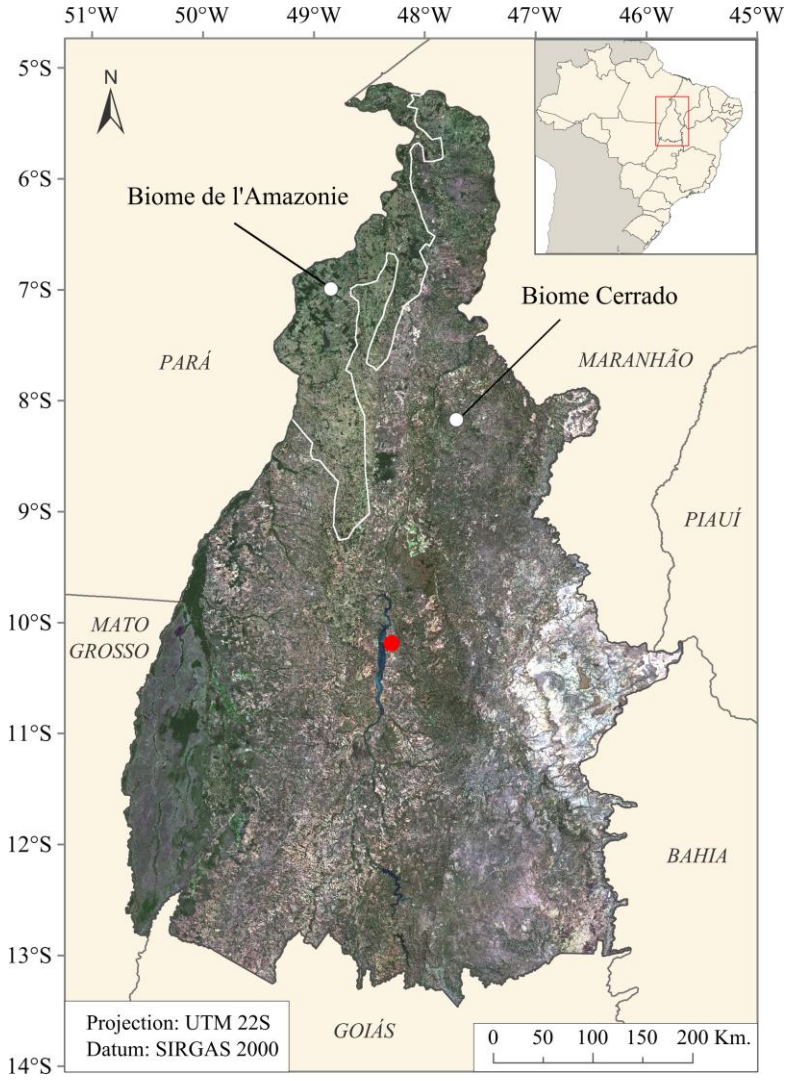


Figure 14. Localisation de l'État du Tocantins au centre-nord du Brésil et de sa capitale Palmas (point rouge), frontières avec les États limitrophes (trait gris) et limites des biomes de l'Amazonie et Cerrado (trait blanc) (sources : limites des états (IBGE, 2017) ; limites des biomes (IBGE, 2004) ; localisation capitale (SEPLAN, 2012) ; image de fond (mosaïque Landsat 8 OLI composée de scènes de juillet 2014 et 2015, USGS (s. d.)).

Depuis sa création en 1988, le Tocantins connaît une forte croissance démographique et son économie est en pleine évolution, basée sur un secteur agricole en développement (Fornaro, 2012; Pedroso da Silva et Almeida, 2007). Outre le développement extensif des systèmes d'élevage, induit par l'extension du réseau de transport depuis les années 70, qui dominent actuellement la plupart des terres en dehors des limites des aires protégées, cette région subit, depuis la dernière décennie, une extension des terres cultivées principalement au détriment de pâturages dégradés, mais aussi de zones de végétation naturelle (Barona *et al.*, 2010; Fornaro, 2012; Pedroso da Silva et Almeida, 2007; Rudorff *et al.*, 2015).

Cette tendance est fortement influencée par la demande mondiale de soja brésilien, responsable de l'extension des superficies cultivées de soja dans la nouvelle, et sans doute dernière, frontière agricole du Brésil : la région MATOPIBA (Figure 15), nommée d'après les États qui la composent : Maranhão, Tocantins, Piauí et Bahia (Bolfe *et al.*, 2016; Fornaro, 2012; Spangler, Lynch et Spera, 2017). Pour le Tocantins seul, la superficie cultivée de soja a augmenté de 900 000 hectares entre l'année 2000 et 2018 (CONAB, 2018).



Figure 15. Localisation de la région de MATOPIBA (code de couleur orange), la nouvelle frontière agricole du Brésil, composée de parties des États du Maranhão (MA), Piauí (PI), Bahia (BA) et l'étendue complète du Tocantins (TO, rayé rouge) (sources : limites des états (IBGE, 2017) ; limites de MATOPIBA (de Miranda, Magalhães et de Carvalho, 2014)).

Les incitations qui ont motivé l'arrivée de producteurs agricoles au Tocantins, notamment des régions du sud du Brésil, et des investisseurs, brésiliens et étrangers, sont nombreuses. D'une part, bien que prédominent des sols à faible fertilité et forte acidité qui requièrent des pratiques de fertilisation et chaulage (Lima, Oliveira et de Aquino, 2000), les conditions environnementales et climatiques sont généralement très favorables à la production agricole. Celles-ci comprennent un relief dominé par des plaines, un climat tropical avec une température et une précipitation moyennes annuelles qui oscillent entre 25°C et 29°C et 1200 mm et 2100 mm selon la zone (Barros de Sousa, Teixeira Borges et Ribeiro Dias, 2012), une saison de pluies qui a une durée d'environ 6 mois (d'octobre à mars), ainsi que des abondantes ressources hydriques.

Une étude sur l'aptitude à la production agricole (cultures, pâturages, foresterie) au Tocantins réalisée par l'Institut brésilien de recherche agronomique (EMBRAPA, acronyme portugais de *Empresa Brasileira de Pesquisa Agropecuária*), estime que seulement 5,8 %[23] de la superficie totale de l'État ne présentent pas d'aptitude à la production agricole, et 35,3 % serait apte aux cultures (Hirano *et al.*, 1989).

D'autre part, l'arrivée de producteurs agricoles et des investisseurs est facilité par des incitations fiscales du gouvernement fédéral destinées à encourager la production agricole, notamment du soja et du riz dans la région (ex. au travers de programmes comme le projet Rio Formoso) (Sauer et Leite, 2012). De plus, l'investissement en infrastructures publiques soutient également la production agricole, notamment, le potentiel d'exportation de l'État a considérablement augmenté grace à l'ouverture des écluses de Tucuruí sur la rivière Tocantins en 2010, et est largement favorisé par l'axe routier Belém-Brasília qui permet la circulation de la production vers d'autres États et de ports côtiers.

Enfin, la législation environnementale en vigueur pour la protection du biome Cerrado - qui occupe environ 91 % du Tocantins -, est moins sévère que celle des autres biomes. Par exemple, suivant les directives du Code forestier brésilien, les propriétaires de terres au Cerrado sont tenus de ne conserver que 20 à 35 % de leurs propriétés en tant que réserve légale, contre 80 % en Amazonie.

[23] L'étude d'aptitude à la production agricole ne prend pas en compte la superficie des zones protégées, les aires de conservation ni les réserves indigènes dans leurs estimations. Les estimations sont réalisées sur la base des multiples données environnementales et climatiques (ex. propriétés chimiques et physiques des sols, végétation, relief, pente, érosion, profondeur effective, drainage, température minimale et maximale, précipitations) (Hirano *et al.*, 1989).

Le biome Cerrado, caractérisé par une savane arbustive (*cerrado sensu stricto,* dominée par des arbustes et des arbres clairsemés sur une strate herbacée), alternée avec des formations fermées (*cerradão*) qui se connectent à des forêts de galerie le long des cours d'eau (*mata de galeria*), est pourtant considéré comme l'un des 34 *hotspots* de la biodiversité mondiale (Klink et Machado, 2005; Silva, 2007). Cependant, les efforts de conservation ont été modestes et il est fortement menacé par la déforestation, entre autres pressions anthropiques (ex. incendies et exploitations minières) (Klink et Machado, 2005; Silva, 2007; Spera, 2017).

Au Tocantins, les aires de conservation représentent 25,8 % de la superficie totale, réparties entre unités de conservation, corridors écologiques et réserves indigènes, mais seulement 5,7 % de la superficie a un statut de "protection integral"[24] (Barros de Sousa *et al.*, 2012). Les aires de conservation au Tocantins sont considérées insuffisantes en nombre et en extension et peu représentatives de l'hétérogénéité régionale d'espaces naturels, le paysage étant décrit comme des îles de Cerrado insérées dans une matrice d'agro-écosystèmes (Silva, 2007).

2.1.1.2. Les systèmes agricoles du Tocantins

Traditionnellement, l'agriculture au Tocantins était limitée à des sites de terres fertiles et l'élevage extensif était pratiqué sur les grandes étendues de savanes naturelles (Silva, 2007). À partir des années 70, l'arrivée des techniques agricoles modernes telles que le chaulage des sols, et des plantes fourragères telles que le Brachiaria, a fortement transformé les systèmes agricoles qui aujourd'hui utilisent des pratiques mécanisées et intensives (Silva, 2007).

Le Tocantins intègre à ce jour différents systèmes d'élevage, culture et agroforesterie. Les systèmes d'élevage concernent principalement des systèmes d'élevage extensif bovin pour la production de viande. Le bétail est alimenté avec les pâturages des formations végétales ouvertes (*campo limpo*), ou des zones déboisées. Les pâturages naturels dégradés abondent, avec différents degrés de dégradation (perte du couvert végétal et recolonisation par espèces « invasives ») et surpâturage, qui sont éventuellement abandonnés dû à leur faible productivité. Les pâturages cultivés avec des espèces fourragères africaines telles que le Brachiaria ou l'Andropogon représentent une alternative largement adoptée pour augmenter la production de viande dans la région (Borghi *et al.*, 2015). Le modèle d'élevage intensif par confinement du bétail est une autre alternative, qui commence à prendre de l'ampleur dans la

[24] Les "unités de conservation de protection intégrale" ne peuvent pas être habitées par l'homme. Seules sont autorisées l'utilisation indirecte de ses ressources naturelles au travers d'activités comme la recherche scientifique et le tourisme écologique.

région pour satisfaire la toujours croissante demande de viande. La Figure 16 présente une illustration de différents pâturages au Tocantins.



Figure 16. Exemples de différents types de pâturage au Tocantins : (a) pâturage naturel ; (b) pâturage planté ; (c) pâturage naturel dégradé ; (d) pâturage naturel dégradé par surpâturage (photographies réalisées par B.Bellón).

La production laitière est mineure par rapport à la production de viande, et se développe principalement dans les municipalités[25] de Taguatinga, Combinado et Lavandeira au sud-est, avec une productivité en dessous de la moyenne nationale (c'est un système d'élevage majoritairement extensif et de faible technicité) (Rodrigues de Melo *et al.*, 2010).

Les systèmes de culture sont dominés par trois systèmes de production de soja, caractérisés par des pratiques de gestion hautement mécanisées et des grandes parcelles, d'environ 100 hectares (Borghi *et al.*, 2015). L'un de ces systèmes est basé sur la monoculture de soja, caractérisée par un seul cycle de culture (le semis a lieu généralement entre octobre et novembre, avec le début des pluies, et la récolte en avril), et les deux autres sont caractérisés par une culture séquentielle avec un double cycle cultural au cours de la même saison de croissance. La culture séquentielle est une forme d'intensification qui est couramment adoptée dans cette région car elle permet d'augmenter les rendements tout en réduisant le besoin d'extension des terres cultivées (Spangler *et al.*, 2017).

[25] Au Brésil, une municipalité (*município*) correspond à la plus petite division du pays, après la division par États (n-2).

Dans l'un de ces systèmes en double culture, une variété de soja à cycle court est cultivée pendant l'été (semis entre octobre et novembre, et récolte[26] entre mars et avril), suivie d'une culture céréalière (principalement maïs) cultivée entre la fin de l'été et début de l'automne (récolte durant les mois de juin et juillet). Le deuxième cycle cultural est principalement destiné à la production de paille sur laquelle le soja est replanté en semis-direct[27] lors de la saison culturale suivante. Dans l'autre système en double culture, localisé sur les plaines inondables des rivières Formoso et Javaés dans le sud-ouest, du riz irrigué par inondation est cultivée pendant la saison de pluies (semis entre octobre et novembre et récolte entre mars et avril), suivie d'une culture tardive de soja, cultivée entre mai et septembre en utilisant une pratique de "sous-irrigation"[28].

La culture de la canne à sucre irriguée, localisée dans la municipalité de Pedro Afonso au centre-nord de l'État, se développe rapidement dû à la croissante demande de biocarburants, remplaçant les cultures annuelles dans les propriétés de la municipalité (Borghi *et al.*, 2015; de Souza *et al.*, 2017). Des systèmes de cultures pérennes sont également présents dans la région et comprennent les plantations d'eucalyptus et d'hévéa au nord de l'État.

[26] Pour distinguer la période culturale des différents produits, la période de récolte des produits du premier cycle est connue dans la région comme "*safra*", et celle des produits du deuxième cycle comme "*safrinha*", ainsi on trouve souvent des appellations des différents produits telles que "*soja-safra* " ou "*milho-safrinha*".

[27] La pratique du semis-direct sur résidus végétaux suit les principes de l'agriculture de conservation, et se caractérise par le non-travail et le couvert permanent du sol par des résidus de culture, ce qui permet de ne pas dégrader la structure du sol et le protéger de l'érosion. Le maïs, qui a une forte valeur économique est aujourd'hui privilégiée pour la production du paillage au Tocantins, le grain étant récolté et commercialisé, mais le mil et le sorgho sont aussi utilisés (le dernier est également produit pour l'alimentation du bétail), et parfois des graminées comme le Braquiaria (Borghi *et al.*, 2015).

[28] Cette pratique d'irrigation consiste à irriguer le système racinaire par le contrôle de la nappe phréatique de sorte à maintenir les eaux souterraines à une profondeur optimale par la gestion de l'eau d'irrigation et du drainage.

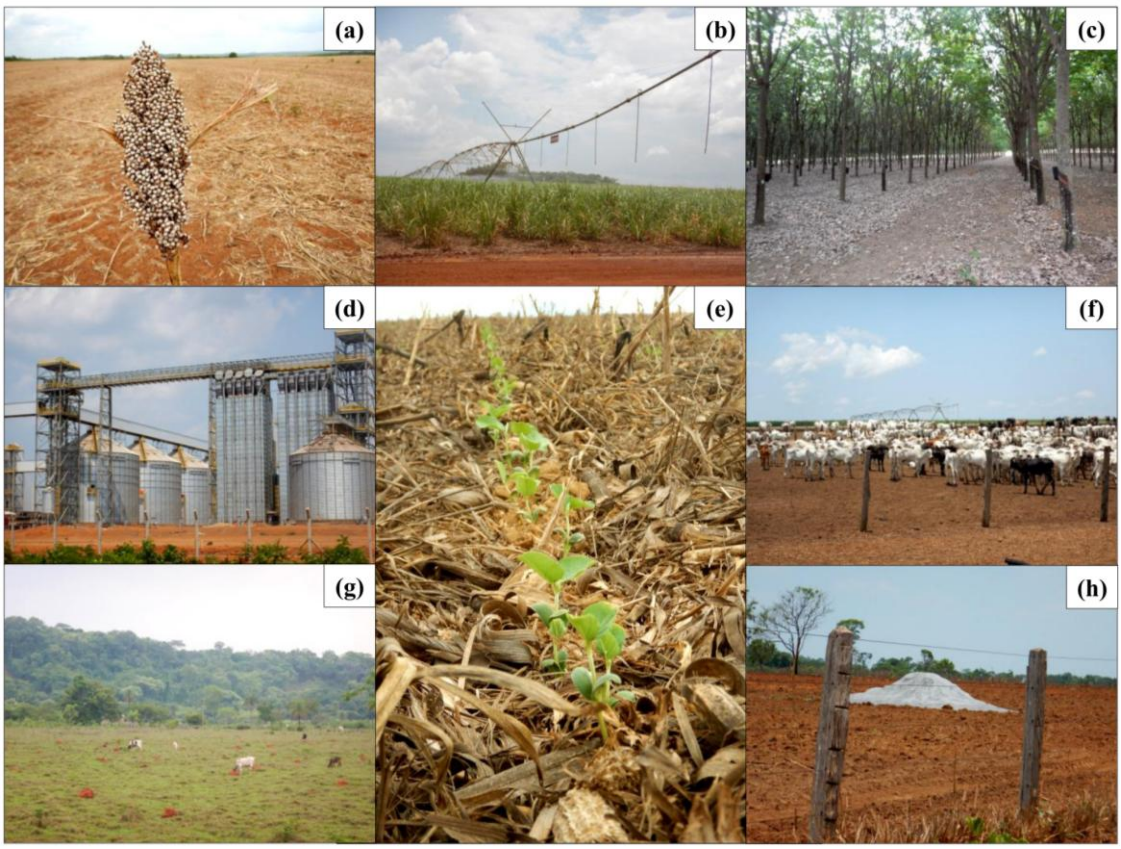


Figure 17. Clichés du Tocantins : (a) résidus de sorgho de couverture ; (b) système d'irrigation à pivot central sur une parcelle de canne à sucre ; (c) plantation d'hévéa (Hevea brasiliensis) ; (d) usine de traitement et raffinerie de soja ; (e) jeunes pousses de soja (cultivée par semis-direct) sur résidus de maïs ; (f) élevage de zébus à viande en confinement ; (g) bétail paissant dans un pâturage naturel ; (h) champ labouré et préparation du chaulage des sols (photographies réalisées par B.Bellón).

2.1.2. Le Burkina Faso

2.1.2.1. Présentation générale du Burkina Faso

Le Burkina Faso, enclavé au cœur de l'Afrique de l'Ouest (9°24' - 15°05'N, 2°22'E - 5°29'O), s'étend sur 274 122 km^2 couvrant trois zones climatiques qui suivent un gradient pluviométrique latitudinal (Figure 18) : la zone sahélienne au nord du pays, avec une précipitation annuelle inférieure à 600 mm, la zone de transition soudano-sahélienne avec une précipitation annuelle entre 600 et 900 mm et la zone soudano-guinéenne au sud avec une précipitation annuelle supérieure à 900 mm (Dembélé, 2010).

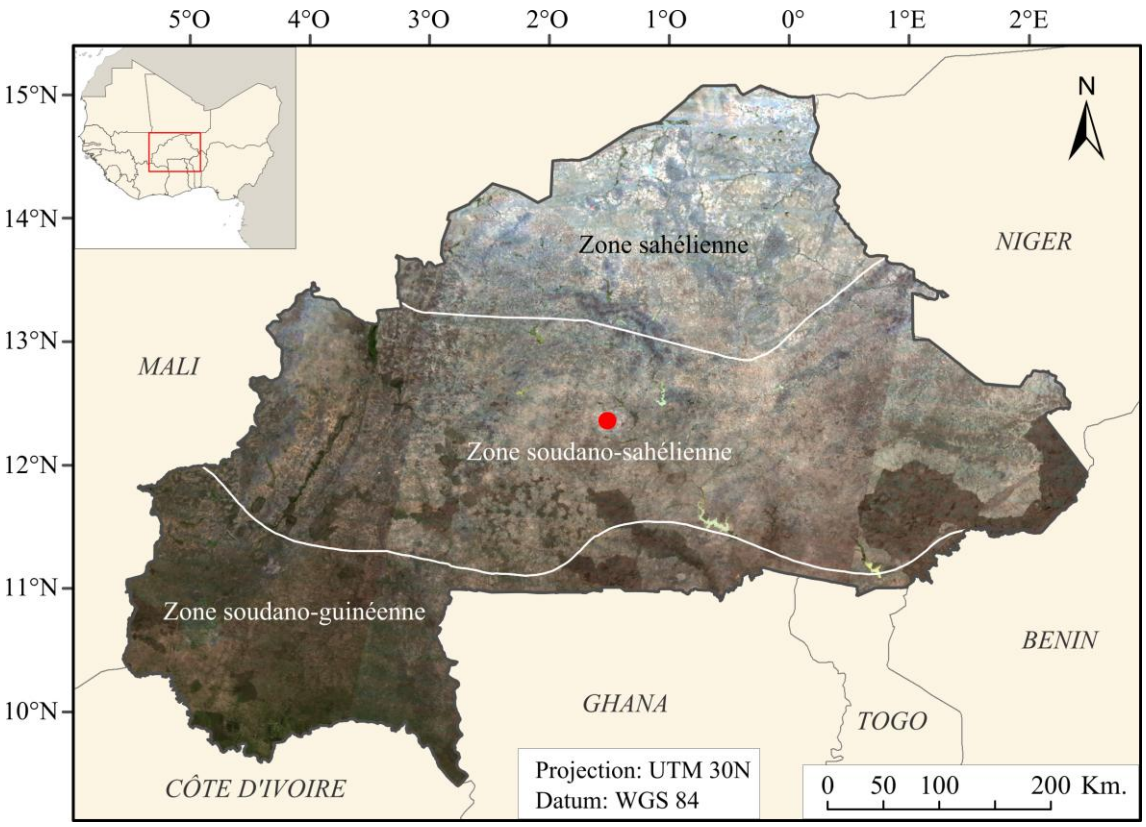


Figure 18. Localisation du Burkina Faso au centre de l'Afrique de l'Ouest, de sa capitale Ouagadougou (point rouge), des pays limitrophes et des zones climatiques (sources : limites des pays (FAO-UN, 1994), zones climatiques (Direction de la Météorologie du Burkina Faso), image de fond (mosaïque MODIS, produit MOD09A1 du 19/12/2014, NASA LP DAAC (s. d.-a)).

L'économie du Burkina Faso est peu diversifiée et repose largement sur l'agriculture qui représente la principale activité de la région, contribuant à 36 % du produit intérieur brut[29] et occupant environ 80 % de la population active (Dembélé, 2010; FEWS NET/USAID, 2010; World Bank, 2017).

Bien que l'agriculture soit le moteur économique du pays, avec quelques cultures de rente dédiées à l'exportation, elle demeure essentiellement une agriculture extensive adressée principalement à l'autoconsommation et au marché local (OSS, 2015). L'agriculture pluviale domine, avec plus de 85 % de la superficie cultivée (Dembélé, 2010). La pluviosité est ainsi l'élément dominant qui détermine la production et la distribution spatiale des systèmes agricoles (FEWS NET/USAID, 2010). En effet, au-delà des conditions climatiques, les conditions environnementales au Burkina Faso présentent une variabilité spatiale peu marquée, avec des sols généralement peu profonds et pauvres en matières organiques et un

[29] Estimation de la Banque Mondiale qui inclut la valeur ajoutée de l'agriculture, de l'exploitation forestière, de la chasse et de la pêche. Elle a été calculée sur une moyenne sur la période de 2000 à 2016 (World Bank, 2017).

relief relativement plat, dominé par une pénéplaine (appelée le « Plateau Central ») qui couvre plus des trois quarts du pays (CILSS, 2016; OSS, 2015).

La saisonnalité pluviométrique d'une année moyenne se caractérise par une courte saison de pluies (généralement de mai à septembre) et une longue saison sèche (d'octobre à avril), mais varie spatialement : dans la zone sahélienne la saison de pluies dure entre 2 et 4 mois, dans la zone soudano-sahélienne entre 4 et 5 mois et dans la zone soudano-guinéenne entre 6 et 7 mois (CILSS, 2016). Ainsi, dans le nord aride du pays, les activités d'élevage prédominent, dans le Plateau Central les cultures vivrières, principalement céréalières, et dans le sud-ouest plus de terres sont dédiées aux cultures de rente et à la diversification de la production avec des zones maraîchères et de vergers (CILSS, 2016; FEWS NET/USAID, 2010).

Les zonages de moyens d'existence au Burkina Faso développés par USAID pour FEWS NET (présentés dans la section 1.2.5.1) mettent en évidence la variabilité spatiale de la production et du calendrier agricole par rapport à la saisonnalité pluviométrique (un extrait des zones du nord et sud-ouest du produit de 2010 est présenté dans la Figure 19) (FEWS NET/USAID, 2010).

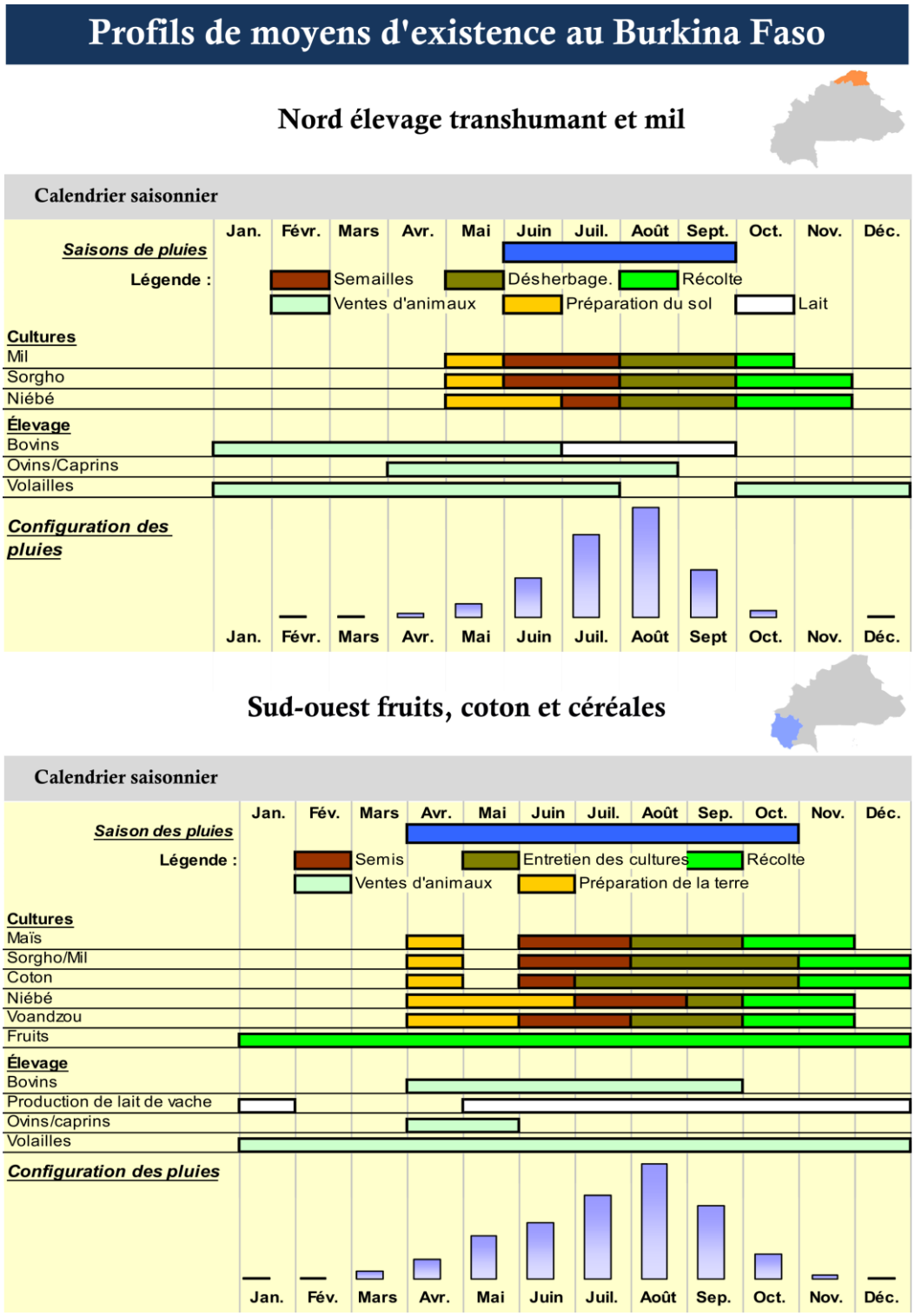


Figure 19. Calendrier saisonnier et configuration des pluies des zones du nord et sud-ouest du Burkina Faso, extraits du produit de 2010 des zonages de moyens d'existence au Burkina Faso (adaptée de FEWS NET/USAID, 2010).

Outre une forte variabilité spatiale, la pluviométrie au Burkina se caractérise par des fluctuations interannuelles importantes (Dembélé, 2010). Le secteur agricole subit ainsi les effets des années de sécheresse, où la production est déficitaire, entraînant des périodes de

soudure plus ou moins longues selon l'année, et une persistance de l'insécurité alimentaire et économique (Dembélé, 2010). De plus, une tendance à l'aridification du climat burkinabé est observée, avec une réduction sensible des volumes de précipitation et une élévation des températures (avec les plus fortes températures relevées autour de 40° à 45°C), ce qui augmente progressivement les taux d'évaporation (Dembélé, 2010).

L'aridification constitue un enjeu majeur pour les ressources en eau du pays. Les plans d'eau diminuent en volume, et les cours d'eau, la plupart avec un écoulement temporaire, présentent des débits avec des tendances à la baisse (Dembélé, 2010). Afin de réduire la vulnérabilité à la sécheresse et les pénuries d'eau, nombreux barrages et digues ont été construits le long des fleuves (CILSS, 2016). Ces réservoirs approvisionnement les espaces urbains, et servent à l'irrigation de zones maraîchères, contribuant à la diversification de la production agricole (CILSS, 2016).

La dégradation des sols représente aussi un enjeu environnemental majeur pour la production agricole burkinabé. Les sols du Burkina, qui naturellement présentent une vulnérabilité à l'érosion hydrique et éolienne, sont exposés à une baisse de la fertilité et une érosion accélérée liée à la déforestation et les pratiques culturales inadaptées (Dembélé, 2010; OSS, 2015). Afin de compenser la faible de productivité des sols et sa dégradation, les surfaces cultivées s'étendent rapidement et les périodes de jachère se raccourcissent (Knauer *et al.*, 2017).

De plus, la population augmente rapidement, avec un taux de croissance annuelle de 2.39 %[30] (World Bank, 2018) ce qui accélère l'ouverture de nouvelles zones agricoles, au détriment des surfaces naturelles, et la dégradation des sols (Knauer *et al.*, 2017). En conséquence, il est estimé que le pays atteindra ses limites de terres cultivables en 2030 (Mathys et Gardner, 2009).

À l'exception du nord, où la surface couverte par des steppes demeure relativement stable dû aux contraintes climatiques qui ne favorisent pas l'installation de cultures, les formations végétales naturelles du reste du pays ont été fortement réduites (une réduction estimée de 39 % entre 1975 et 2013) afin de faire place aux cultures et se trouvent confinées dans les aires protégées (CILSS, 2016). Les ressources forestières, également exploitées pour la satisfaction des besoins énergétiques de la population, se raréfient (CILSS, 2016). Aujourd'hui, même les réserves naturelles voient leur surface diminuer considérablement dans un milieu naturel hautement fragmenté (CILSS, 2016; Knauer *et al.*, 2017).

[30] Estimations de la Banque Mondiale calculée sur une moyenne sur la période de 2001 à 2016 (World Bank, 2018).

2.1.2.2. Les systèmes agricoles du Burkina Faso

Les systèmes d'élevage au Burkina Faso, concernent essentiellement des systèmes pastoraux du type transhumant ou sédentaire extensif (Dembélé, 2010; OSS, 2015). Les troupeaux peuvent être composés de plusieurs espèces (ex. bovins, ovins, caprins etc.), mais les bovins zébus dominent (Dembélé, 2010). L'alimentation des troupeaux est fournie pour l'essentiel par le pâturage naturel, principalement des steppes et des savanes arbustives (Akoudjin *et al.*, 2016).

Pendant la saison de pluies, les pâturages maintiennent leur valeur nutritive, qui diminue considérablement en fin de saison, ainsi que les points d'eau (Dembélé, 2010). Généralement, les éleveurs transhumants du nord se déplacent en début de saison sèche vers le sud, où les animaux peuvent s'alimenter de l'herbe des bas-fonds et des résidus des cultures (Dembélé, 2010). Les troupeaux représentent une source importante de fumier pour les terres cultivées, mais les dégâts provoqués par les animaux dans les champs sont sources de nombreux conflits entre éleveurs et agriculteurs, ce qui a limité les parcours des animaux à des zones de sols incultivables (Akoudjin *et al.*, 2016; BAfD, 2015; Dembélé, 2010).

L'élevage sédentaire et les systèmes agro-pastoraux se sont rapidement développés à partir des années 80, avec deux origines : l'installation des migrants venus du nord sahélien dans le sud du Burkina et la diversification des activités dans les exploitations agricoles (Dembélé, 2010; Gonin et Tallet, 2012).

Les exploitations agricoles sont majoritairement familiales et peuvent être divisées en trois types en fonction de leur taille et leurs niveaux de technification : (i) les petites exploitations, avec des superficies entre 3 et 6 hectares qui pratiquent l'agriculture vivrière de façon manuelle et sans utilisation d'intrants ; (ii) les exploitations de taille moyenne, avec des superficies d'au moins 6 hectares qui dédient une partie de leurs terres aux cultures de rente, pratiquent la culture attelée et utilisent quelques intrants ; (iii) les grandes exploitations, d'au moins 10 hectares qui consacrent plus de terres à la production de cultures de rente qu'à la production vivrière et qui pratiquent la culture attelée voire mécanisée et utilisent plus d'intrants (Dembélé, 2010).

Les grandes exploitations sont concentrées dans le bassin cotonnier dans l'ouest et sont essentiellement dédiées à la culture du coton, qui est la principale culture de rente (suivie des cultures oléagineuses telles que le sésame et l'arachide) et le deuxième produit exporté du Burkina Faso après l'or (OSS, 2015). Les assolements de ces exploitations consacrent de plus

en plus de superficie au maïs, cultivé en système de rotation avec le coton (Guenot et Huchet-Bourdon, 2014).

En dehors du bassin cotonnier les petites exploitations prédominent, avec un parcellaire de petite taille (d'environ 0.6 hectares) et une forte hétérogénéité intra-parcellaire, exacerbée au sud par la présence des arbres des systèmes agro-forestiers. Leur production est dominée par les céréales, notamment du mil et du sorgho, et d'autres cultures vivrières légumineuses comme le niébé (cultivé en association avec les céréales), l'igname ou le voandzou (Dembélé, 2010). Le défrichement sélectif lors de la création des champs est une pratique courante au Burkina pour conserver les espèces arborées à valeur alimentaire, énergétique, ou encore culturelle, en association avec les cultures (Thiombiano et Kampmann, 2010). La jachère est aussi une pratique agro-forestière ancestrale utilisée dans la région pour reconstituer la fertilité des sols (Thiombiano et Kampmann, 2010).

Les systèmes de culture irriguée occupent environ 12% des superficies cultivées et sont principalement situés autour des réservoirs d'eau (Dembélé, 2010; Knauer *et al.*, 2017). Les principales cultures irriguées sont le riz, la canne à sucre, les cultures maraîchères et fruitières. La filière fruticole, installée principalement dans le sud-ouest où les conditions climatiques sont favorables, a connu une augmentation significative de la superficie de plantation au même temps qu'une augmentation et de la valeur des exportations (ex. mangues, agrumes et noix de karité), contribuant significativement à l'amélioration des revenus locaux (Dembélé, 2010; Knauer *et al.*, 2017).

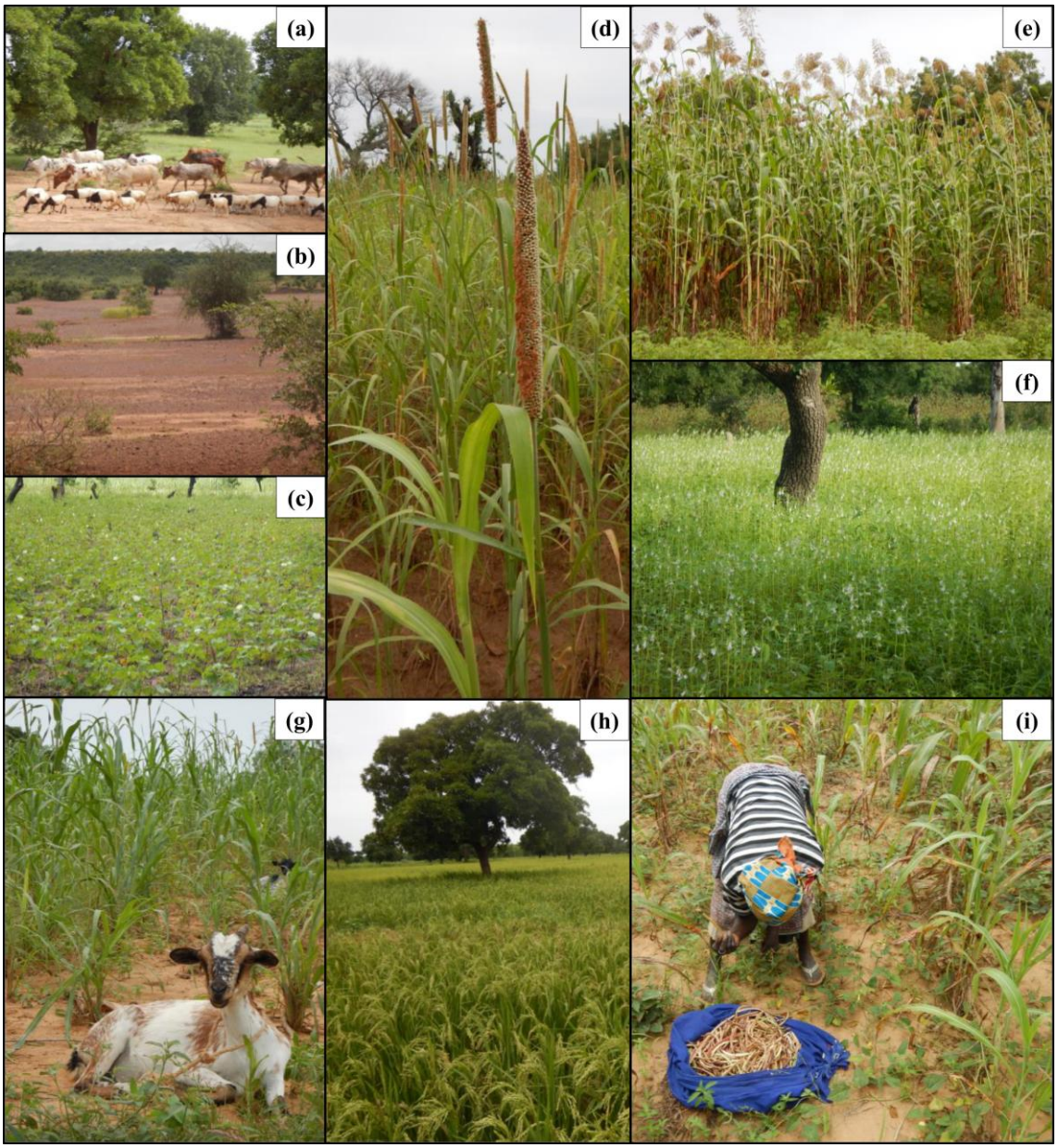


Figure 20. Clichés du Burkina Faso : (a) troupeau composé de plusieurs espèces ; (b) sol latéritique impropre aux cultures ; (c) coton ; (d) mil ; (e) sorgo ; (f) sésame ; (g) parcelle d'un exploitant agro-éleveur ; (h) riz pluvial ; (i) récolte du niébé en association avec du mil (photographies réalisées par B.Bellón).

2.2. Les données

Différentes données ont été utilisés au cours de ce travail : des données satellitaires (des images MODIS, Landsat 8, Sentinel-2A et SPOT 7), une carte d'occupation des terres du Tocantins et des données de validation (des statistiques agricoles et des données de terrain). Les caractéristiques de ces données ainsi que les prétraitements appliqués sont décrits dans les sous-sections suivantes.

2.2.1. Données satellitaires

2.2.1.1. Les séries temporelles d'images de NDVI du capteur MODIS

Comme présenté précédemment (dans la section 1.3.2.1), les séries temporelles d'indices de végétation issues du capteur MODIS, à bord des satellites Terra et Aqua de la NASA (*National Aeronautics and Space Administration*), sont à ce jour les données privilégiées pour la classification de systèmes agricoles à l'échelle régionale par télédétection. Ce choix est motivé par les caractéristiques de ces données, particulièrement appropriées :

- la couverture globale à haute répétitivité temporelle (journalière) permet le suivi du développement de la végétation à pas de temps régulier sur n'importe quelle région du monde ;
- la grande emprise spatiale des scènes (fauchée de 2330 km) permet le suivi de grandes superficies, tout en minimisant le nombre de scènes à traiter ;
- l'accès gratuit et la mise à disposition en temps quasi-réel de données ortho-rectifiées et en réflectance de surface, facilitent leur utilisation directe et minimisent les étapes de prétraitement ;
- la mise à disposition de produits de synthèse bimensuelle d'indices de végétation offre un gain de temps aux utilisateurs, évitant les étapes de calcul de néocanaux ;
- la profondeur temporelle des archives (disponibles depuis l'année 2000) permet le suivi de la dynamique saisonnière et interannuelle de la végétation et de l'utilisation des terres.

Pour ce travail de thèse, nous avons utilisé le produit "*MODIS Vegetation Indices 16-Day L3 Global 250m*" (MOD13Q1, version 5) (Solano *et al.*, 2010). Le produit MOD13Q1 fournit des synthèses (« images composées ») de 16 jours des indices de végétation NDVI et EVI (*Enhanced Vegetation Index*) à 250 m de résolution. Ces synthèses sont dérivées des bandes spectrales dans le domaine du rouge et proche-infrarouge des produits de réflectance de surface corrigés des effets atmosphériques de la série journalière "MOD09" (Vermote, El Saleous et Justice, 2002).

Le processus de synthèse sur 16 jours du produit MOD13Q1 utilise l'algorithme CV-MVC (*Constrained View angle-Maximum Value Composite*), qui vise à minimiser les effets atmosphériques et directionnels en sélectionnant (pixel par pixel) l'observation avec la valeur de NDVI la plus élevée et le plus petit angle de visée parmi toutes les observations

effectuées pendant la période de 16 jours (Solano *et al.*, 2010). Le jour de l'année de l'observation retenue pour chaque pixel est enregistré dans la bande "*composite day of the year*", fournie avec le produit MOD13Q1.

Parmi les deux indices de végétation fournis, on a privilégié l'utilisation du NDVI. Bien que le NDVI soit théoriquement plus affecté par le bruit (ex. effets des aérosols) et par la saturation radiométrique dans les régions à forte biomasse (ex. fôrets) que l'EVI (Huete *et al.*, 2002), il est hautement sensible aux légères variations de la quantité de végétation photosynthétique (Soudani *et al.*, 2008), et donc semble bien adapté pour capturer la dynamique de la végétation des systèmes agricoles.

Nous avons récupéré trois séries temporelles de données, deux années pour le Tocantins et une année pour le Burkina Faso, auprès du centre d'archives LP DAAC de la NASA (*NASA Land Processes Distributed Active Archive Center*) au travers de la librairie "MODIStsp v.1.3.2" du logiciel R (Busetto et Ranghetti, 2016; R Core Team, 2016). Les données couvrent l'ensemble des régions d'étude en 5 tuiles (Figure 21). Trois bandes ont été récupérées pour chaque image : la bande NDVI ; la bande de qualité "*pixel reliability summary QA*" ; la bande du jour julien (jour de l'année) "*composite day of the year*".

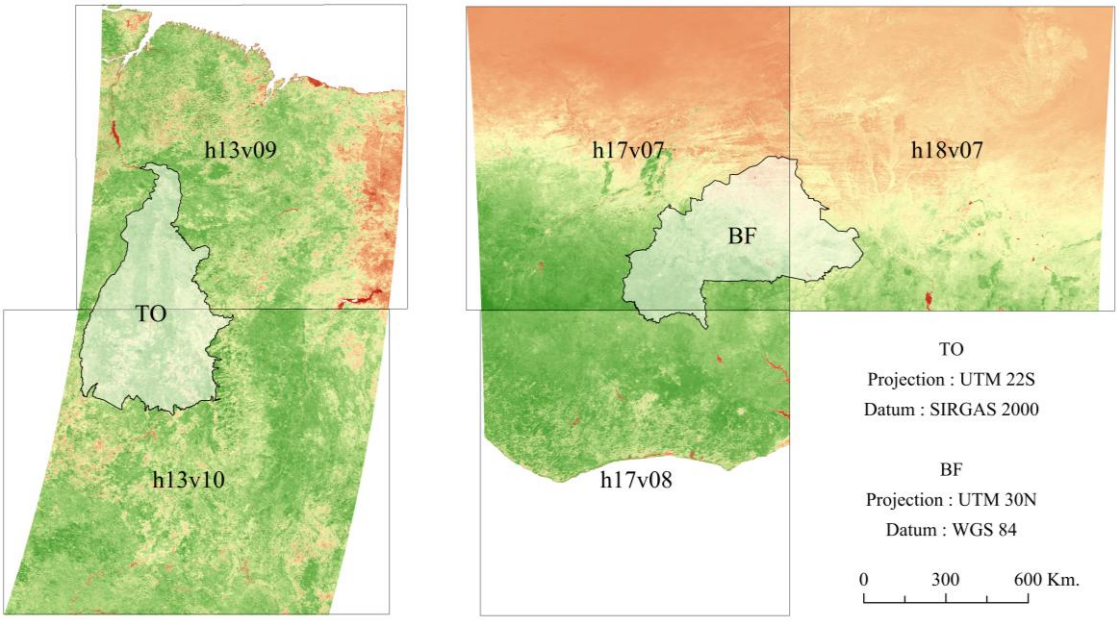


Figure 21. Emprise des tuiles MODIS couvrant Tocantins (TO) et Burkina Faso (BF) (source : limites administratives du Tocantins (IBGE, 2017) et du Burkina Faso (OCHA ROWCA, 2017) ; images de fond (NASA LP DAAC, s.d.-b)).

Chaque série temporelle est composée d'une série annuelle d'images composées (23 images composées par an) qui couvre les saisons culturales d'intérêt, et de vingt images supplémentaires[31], 10 images avant la première image de la série annuelle et 10 images après la dernière image de la série annuelle. Un total de 43 images composées par série temporelle est donc téléchargé avec la période d'intérêt annuelle au milieu de chaque série. Les séries annuelles pour le Tocantins couvrent les saisons culturales 2013-2014 et 2014-2015, avec des images allant du début de la saison culturale en octobre jusqu'à septembre de l'année suivante. La série annuelle pour le Burkina Faso couvre la saison culturale de 2016, avec des images allant de janvier jusqu'à décembre. Au total, 261 images (132 images pour le Tocantins et 129 images pour le Burkina) et 783 bandes ont été téléchargées pour les cinq tuiles.

Le Tableau 4 présente la liste des images composées des séries annuelles avec les dates de la période de 16 jours par image et le code du jour de l'année désormais utilisé pour désigner chaque image des séries annuelles.

[31] Les images supplémentaires sont utilisées lors du lissage des séries annuelles (décrit ultérieurement dans cette section) pour éviter des résultats biaisés aux extrémités des séries annuelles d'intérêt.

Tableau 4. Liste des 23 images composées téléchargées pour la série annuelle couvrant la saison culturale de 2014-2015 au Tocantins et 2016 au Burkina, dates de la période de composition de chaque image et code du jour julien utilisé pour désigner les images (la liste d'images pour la saison culturale de 2013-2014 du Tocantins a les mêmes fourchettes de dates et le même code du jour de l'année que les images pour la saison 2014-2015, ici montrées).

Tocantins				Burkina Faso			
Code de l'image	Dates		Jour julien	Code de l'image	Dates		Jour julien
MOD13Q1.A2014273	30/09/2014	15/10/2014	273	MOD13Q1.A2016001	01/01/2016	16/01/2016	1
MOD13Q1.A2014289	16/10/2014	31/10/2014	289	MOD13Q1.A2016017	17/01/2016	01/02/2016	17
MOD13Q1.A2014305	01/11/2014	16/11/2014	305	MOD13Q1.A2016033	02/02/2016	17/02/2016	33
MOD13Q1.A2014321	17/11/2014	02/12/2014	321	MOD13Q1.A2016049	18/02/2016	05/03/2016	49
MOD13Q1.A2014337	03/12/2014	18/12/2014	337	MOD13Q1.A2016065	06/03/2016	21/03/2016	65
MOD13Q1.A2014353	19/12/2014	03/01/2015	353	MOD13Q1.A2016081	22/03/2016	06/04/2016	81
MOD13Q1.A2015001	01/01/2015	16/01/2015	1	MOD13Q1.A2016097	07/04/2016	22/04/2016	97
MOD13Q1.A2015017	17/01/2015	01/02/2015	17	MOD13Q1.A2016113	23/04/2016	08/05/2016	113
MOD13Q1.A2015033	02/02/2015	17/02/2015	33	MOD13Q1.A2016129	09/05/2016	24/05/2016	129
MOD13Q1.A2015049	18/02/2015	05/03/2015	49	MOD13Q1.A2016145	25/05/2016	09/06/2016	145
MOD13Q1.A2015065	06/03/2015	21/03/2015	65	MOD13Q1.A2016161	10/06/2016	25/06/2016	161
MOD13Q1.A2015081	22/03/2015	06/04/2015	81	MOD13Q1.A2016177	26/06/2016	11/07/2016	177
MOD13Q1.A2015097	07/04/2015	22/04/2015	97	MOD13Q1.A2016193	12/07/2016	27/07/2016	193
MOD13Q1.A2015113	23/04/2015	08/05/2015	113	MOD13Q1.A2016209	28/07/2016	12/08/2016	209
MOD13Q1.A2015129	09/05/2015	24/05/2015	129	MOD13Q1.A2016225	13/08/2016	28/08/2016	225
MOD13Q1.A2015145	25/05/2015	09/06/2015	145	MOD13Q1.A2016241	29/08/2016	13/09/2016	241
MOD13Q1.A2015161	10/06/2015	25/06/2015	161	MOD13Q1.A2016257	14/09/2016	29/09/2016	257
MOD13Q1.A2015177	26/06/2015	11/07/2015	177	MOD13Q1.A2016273	30/09/2016	15/10/2016	273
MOD13Q1.A2015193	12/07/2015	27/07/2015	193	MOD13Q1.A2016289	16/10/2016	31/10/2016	289
MOD13Q1.A2015209	28/07/2015	12/08/2015	209	MOD13Q1.A2016305	01/11/2016	16/11/2016	305
MOD13Q1.A2015225	13/08/2015	28/08/2015	225	MOD13Q1.A2016321	17/11/2016	02/12/2016	321
MOD13Q1.A2015241	29/08/2015	13/09/2015	241	MOD13Q1.A2016337	03/12/2016	18/12/2016	337
MOD13Q1.A2015257	14/09/2015	29/09/2015	257	MOD13Q1.A2016353	19/12/2016	03/01/2017	353

Tout comme le Tocantins, le Burkina s'étend sur plusieurs zones UTM (*Universal Transverse Mercator*). Ainsi, toutes les images téléchargées avec leur projection d'origine Sinusoïdale ont été reprojetées au système de coordonnées de la zone UTM la plus étendue sur chaque région : UTM zone 22S pour le Tocantins et UTM zone 30N pour le Burkina Faso. Le système géodésique (*datum*) choisi pour le Burkina est le "WGS 84" et pour le Tocantins "SIRGAS 2000" (le système de référence géodésique brésilien). Un aperçu des trois bandes téléchargées après reprojection, mosaïquage et découpage à l'emprise des régions d'étude[32] est présenté dans la Figure 22.

[32] Trois librairies du logiciel R sont utilisées pour les opérations basiques des prétraitements des images (ex. reprojection, mosaïquage, découpage, interpolations linéaires) : "raster v. 2.5.8", "maptools v.0.9.2" et "rgdal v.1.2.7" (Bivand, Keitt, et Rowlingson, 2016; Bivand et Lewin-Koh, 2016; Hijmans, 2016).

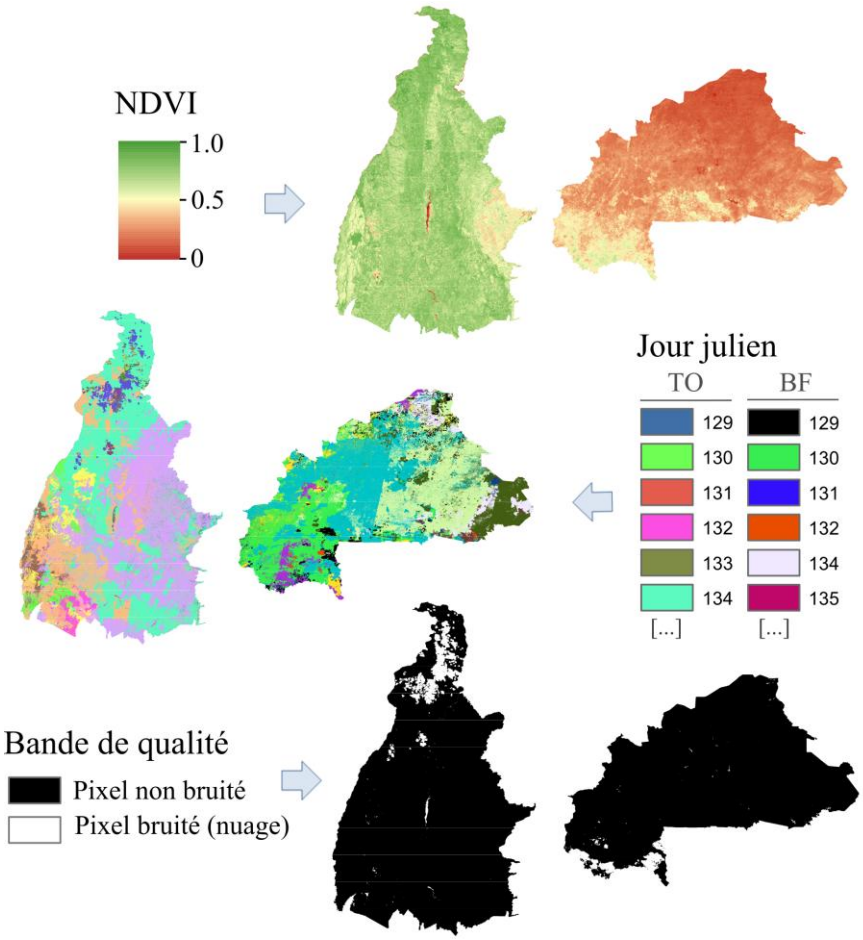


Figure 22. Illustration des bandes du produit MOD13Q1 téléchargées pour l'image composée 129 (correspondant au mois de mai) de 2015 pour le Tocantins, et l'image 129 de 2016 pour le Burkina Faso : NDVI ; bande de qualité (*Pixel reliability summary QA*) et ; jour julien (*Composite day of the year*).

La bande de qualité a été utilisée pour évaluer la quantité de bruit présente dans les séries annuelles de 23 images. La répartition spatiale des pixels bruités pour la série de 2013-2014 au Tocantins et 2016 au Burkina est présentée dans la Figure 23. On s'aperçoit que malgré l'utilisation de l'algorithme CV-MVC, il y a du bruit résiduel dans les séries temporelles, notamment dans les zones à forte nébulosité (nord du Tocantins et sud du Burkina). Cependant, la bande de qualité s'avère peu fiable puisqu'on observe des surestimations systématiques de présence de nuages sur les parcelles agricoles, qui sont particulièrement évidentes pendant les périodes de récolte. Cela peut être expliqué par la baisse des valeurs de NDVI pendant ces périodes qui a dû être confondue avec une baisse des valeurs par la présence de nuages et donc masqués en conséquence.

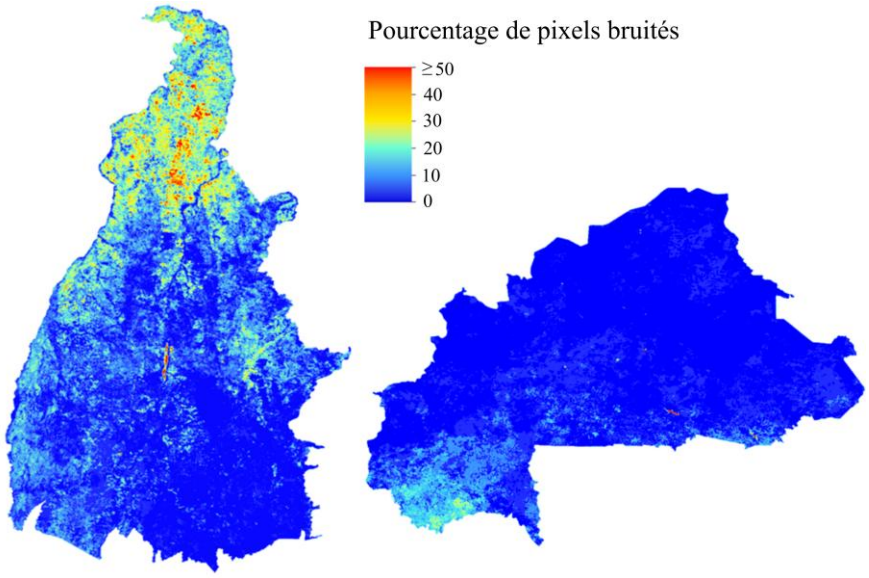


Figure 23. Distribution spatiale des pixels considérés bruités dans la bande de qualité fournie avec le produit MOD13Q1 sur les séries annuelles de 23 dates de 2013-2014 (à gauche) et 2016 (à droite).

Afin de réduire le bruit résiduel des séries temporelles, on a appliqué une méthode de lissage basée sur l'utilisation de l'algorithme de Savitzky-Golay (Savitzky et Golay, 1964) en adaptant la méthode présentée par Chen *et al.* (2004). Cette méthode consiste à réaliser une interpolation linéaire des pixels affectés par les nuages, suivie de l'utilisation du filtre de Savitzky-Golay par un processus itératif (Chen *et al.*, 2004).

Prenant en compte les surestimations de présence de nuages de la bande de qualité sur les zones cultivées, nous évitons son utilisation lors de l'interpolation linéaire des pixels bruités. L'interpolation est réalisée uniquement sur les pixels qui présentent une variation de NDVI « anormale », soit des pixels présentant un écart de valeurs de 0,4 entre deux images composées ($t_1 - t_0 = 0{,}4$), le pixel de l'image t_0 étant dans ce cas considéré comme bruité et donc interpolé. Le filtre Savitzky-Golay est ensuite appliqué à l'aide de la librairie R "signal v.0.7.6" (Ligges, Short et Kienzle, 2015) pour réduire d'autres effets résiduels tels que la nébulosité consécutive sur plusieurs images composées.

Deux paramètres du filtre doivent être définis : la demi-largeur de la fenêtre temporelle (m) et le degré du polynôme utilisé pour le lissage (d) (Savitzky et Golay, 1964). Suivant la méthode de Chen *et al.* (2004), le filtre de Savitzky-Golay est appliqué en deux temps. On l'applique dans un premier temps pour obtenir une courbe de changement à long terme représentant la variation graduelle de la végétation au cours de l'année (Chen *et al.*, 2004). Les deux critères qualitatifs pour obtenir une courbe de changement à long terme performante tels que définis par Chen *et al.* (2004) sont : (i) elle devrait suivre la variation graduelle de la végétation sans provoquer une perte trop importante des détails temporels dans la série de

NDVI ; (ii) les points bruités par les nuages devraient se retrouver par-dessous de la courbe. Cette première application du filtre va corriger les chutes de valeurs de NDVI correspondant aux points bruités, mais au même temps aura un effet de lissage sur les pics de valeurs de NDVI. Ainsi, le filtre est appliqué dans un deuxième temps de façon itérative jusqu'à que la courbe s'ajuste à l'enveloppe supérieure du profil de NDVI original (Chen *et al.*, 2004).

Après plusieurs tests, les paramètres retenus sont m=2 et d=2 pour la courbe de changement à long terme, et m=4 et d=6 pour l'itération d'ajustement. Cette méthode de lissage a largement amélioré la qualité des trois séries temporelles en réduisant le bruit du signal tout en préservant le détail des profils temporels de NDVI. Le profil NDVI brut (avant lissage) et lissé d'un pixel représentant un double cycle cultural est présenté dans la Figure 24. On remarque que les points bruités sont corrigés, que le profil lissé est ajusté à l'enveloppe supérieure du profil brut et que la forme du profil de NDVI pendant la période de récolte entre les deux cycles de culture distincts est respectée.

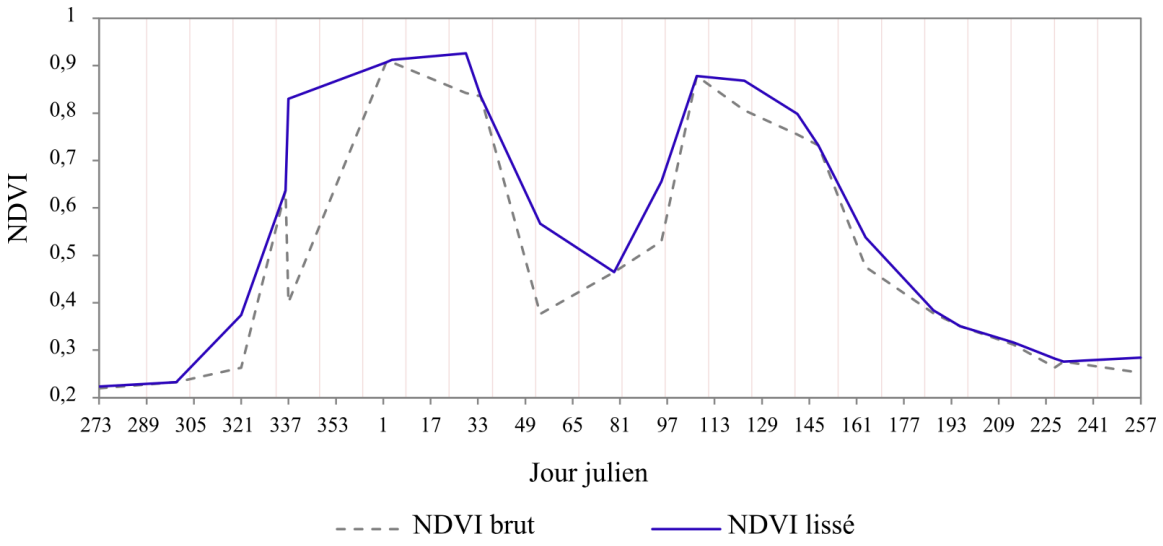


Figure 24. Profil de NDVI brut (avant lissage et interpolation) et lissée d'un pixel de la série annuelle couvrant la saison culturale de 2013-2014 du Tocantins.

Enfin, pour éviter les discontinuités temporelles dans les séries, générées par le processus de synthèse des images composées sur 16 jours, les valeurs des séries temporelles lissées ont été interpolées linéairement, en utilisant leurs dates d'acquisition réelles fournies dans la bande du jour julien, pour les placer à des intervalles réguliers de 16 jours. Le profil NDVI lissé et corrigé (replacé sur des intervalles réguliers de 16 jours) est présenté dans la Figure 25.

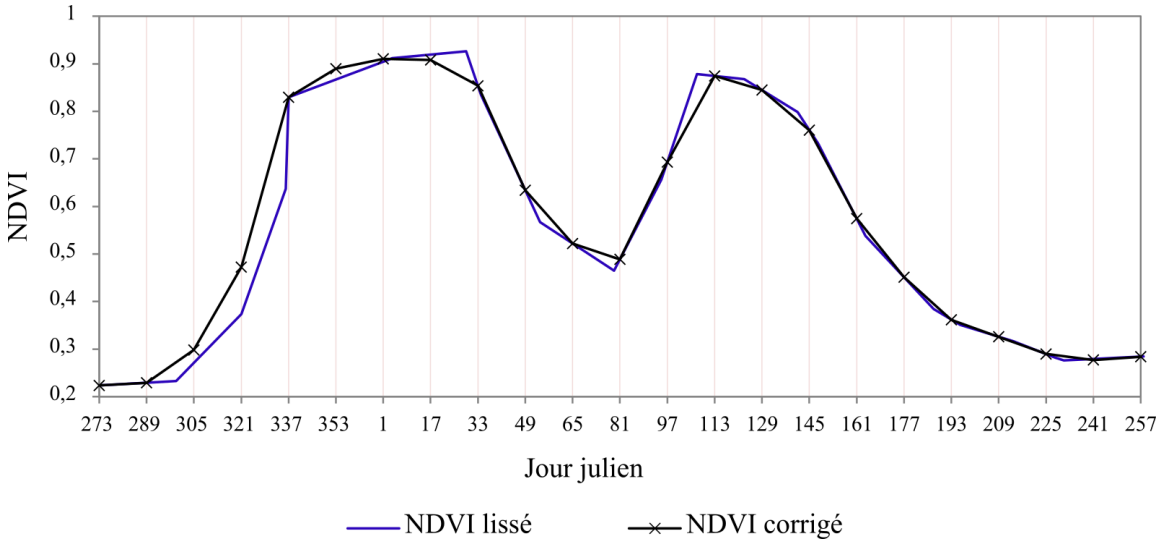


Figure 25. Profil de NDVI lissé et corrigé (replacé sur des intervalles réguliers de 16 jours) d'un pixel de la série annuelle couvrant la saison culturale de 2013-2014 du Tocantins.

2.2.1.2. La mosaïque Landsat 8

Des données Landsat 8, avec une haute résolution spatiale, ont été récupérées en complémentarité de la haute résolution temporelle des données MODIS. Ces images, mieux résolues spatialement ont été utilisées pour générer des objets à l'échelle de la parcelle pour la deuxième étape de l'approche (présentée dans la section 1.5 et décrite dans le chapitre 4).

Le satellite Landsat 8, aussi connu comme LDCM (*Landsat Data Continuity Mission*) est le dernier satellite opérationnel de la série Landsat de la NASA. Il a été lancé en 2013 assurant la continuité de la mission qui a débuté en 1972, poursuivant les observations dans les domaines du visible et infrarouge (proche, moyen et thermique) (USGS, 2016). Il transporte deux capteurs : l'OLI (*Operational Land Imager*) qui fournit une bande panchromatique à 15 m de résolution et des bandes multispectrales à 30 m, et le TIRS (*Thermal Infrared Sensor*) qui fournit des bandes thermiques à 100 m de résolution. Tout comme les données MODIS, les données Landsat 8 sont aujourd'hui mises à disposition gratuitement en temps quasi-réel, et l'accès aux données est facilité par les interfaces fonctionnelles "Earth Explorer" et "ESPA" (*EROS Center Science Processing Architecture on demand interface*).

Afin de générer une mosaïque à 30 m de résolution sur l'ensemble du Tocantins, nous avons récupéré dix-neuf scènes OLI (la mosaïque et les tuiles des scènes sont présentées dans la Figure 26). Pour assurer une couverture entièrement libre de nuages du Tocantins, nous avons sélectionné des scènes acquises en juillet 2014 et juillet 2015. Le Tableau 5 présente les dates d'acquisition des scènes et leur identifiant dans la grille de référence de Landsat

(i.e. les numéros de colonne et ligne). Les images (ortho-rectifiées et en réflectance de surface) ont été récupérées via l'interface ESPA (Jenkerson, 2013). Parmi les 11 bandes fournies avec le produit Landsat 8, nous avons utilisé les bandes 2 à 6 correspondant aux domaines du bleu (B2), vert (B3), rouge (B4), proche infrarouge (B5) et moyen infrarouge (B6). La mosaïque générée a été projetée dans le système de coordonnées UTM 22S et découpée suivant l'emprise du Tocantins.

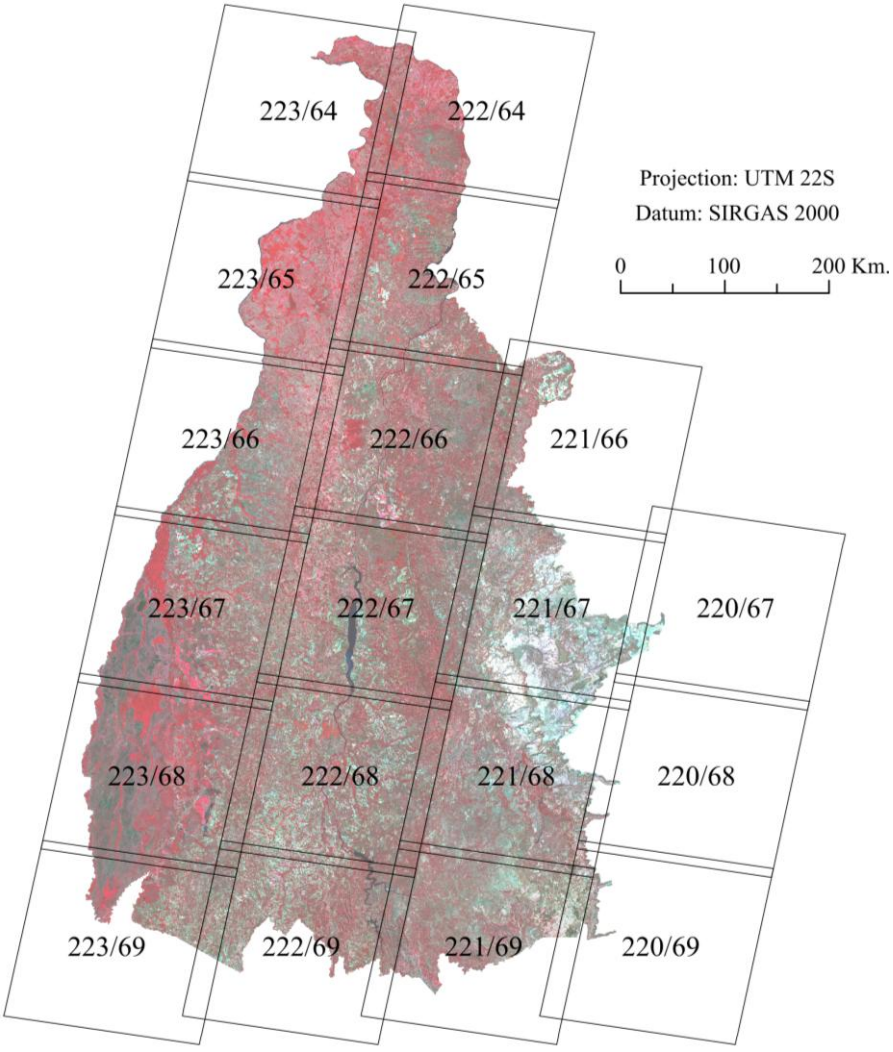


Figure 26. Emprise des tuiles Landsat 8 OLI couvrant Tocantins. L'identifiant des tuiles correspond au numéro de colonne et ligne de la grille de référence de Landsat (WRS2). La mosaïque libre de nuages obtenue à partir des 19 scènes est présentée en image de fond avec une composition colorée en fausses couleurs (RGB proche-infrarouge, vert, rouge) (source : image de fond (USGS, s.d.)).

Tableau 5. Détails d'acquisition des scènes Landsat 8 OLI. La couverture nuageuse dans les scènes correspond à des zones situées à l'extérieur des limites de l'État du Tocantins.

Date d'acquisition	Colonne	Ligne	Couverture nuageuse de la scène (%)
17/07/2015	223	64	0.06
		65	0.05
		66	0.00
		67	0.00
		68	0.00
		69	0.00
07/07/2014	222	64	0.04
		65	0.00
		66	0.00
		67	0.00
		68	0.00
		69	0.21
19/07/2015	221	66	0.00
		67	0.00
		68	0.00
		69	0.44
25/07/2014	220	67	0.68
		68	0.00
		69	0.00

2.2.1.3. Les images Sentinel-2A et SPOT 7

Des images supplémentaires Sentinel-2A et SPOT 7 ont été acquises pour appuyer l'acquisition de points GPS au Burkina Faso (présentée dans la section 2.2.3.1) et la photo-interprétation des résultats de classification du Burkina (présentés dans la section 5.3.3).

Des acquisitions SPOT 6/7 ont été programmées afin d'obtenir des images à très haute résolution spatiale sur les zones à enquêter, la résolution spatiale de 6 m de ces images est particulièrement adaptée à la petite taille du parcellaire burkinabé, facilitant le positionnement des points sur les surfaces ciblées et la photo-interprétation des surfaces cultivées. Deux zones au nord et au sud ont bénéficié des données SPOT 7, et deux images Sentinel-2A à 10 m de résolution ont été récupérées sur deux autres zones (au centre et à l'ouest), les acquisitions SPOT étant très affectées par la nébulosité sur ces zones.

Les deux images SPOT 7 ortho-rectifiées datant du 24/08/2016 et 12/09/2016 ont été fournies par la Station de Réception Directe "GEOSUD/Airbus DS" 2016, et les deux images Sentinel-2A datant du 14/06/2016 et 02/09/2016 ont été fournies au niveau 2A par le "pôle Surfaces Continentales THEIA", prétraitées en réflectance de surface via le "MACCS-

ATCOR Joint Algorithm" (Hagolle *et al.*, 2015) développé par le Centre National d'Études Spatiales (CNES). Les images Sentinel-2 ont été découpées à l'emprise des acquisitions SPOT (~3 600 km^2). L'emprise finale des images est présentée dans la Figure 29.

2.2.2. Carte d'occupation des terres

Pour la mise en place de l'approche, nous avons utilisé deux produits cartographiques issus du projet "TerraClass", dirigé par l'Institut national brésilien de recherche spatiale (INPE, acronyme portugais d'*Instituto Nacional de Pesquisas Espaciais*) en partenariat avec l'EMBRAPA.

Le projet TerraClass, conçu pour suivre les récentes transitions d'utilisation des terres au Brésil par des techniques d'observation de la Terre, a mis en place des méthodes de cartographie de l'occupation des terres basées sur le traitement de données Landsat et MODIS par des techniques de télédétection et photo-interprétation (de Almeida *et al.*, 2016). Les produits TerraClass comprennent des cartes de l'occupation des terres dans le biome de l'Amazonie, produites tous les deux ans depuis 2004, ainsi qu'une première version pour le biome Cerrado qui a été publiée récemment pour l'année 2013 (INPE, s.d.-a, s.d.-b). Ces données nous permettent de connaître la répartition spatiale des activités agricoles dans la région du Tocantins à l'échelle locale, avec une résolution spatiale de 30 m.

Nous avons ainsi combiné les résultats des derniers produits TerraClass ("TerraClass Amazônia 2014" et "TerraClass Cerrado 2013") pour générer une carte d'occupation des terres du Tocantins, à partir de laquelle l'on a extrait les classes agricoles : pâturages et agriculture annuelle (Figure 27).

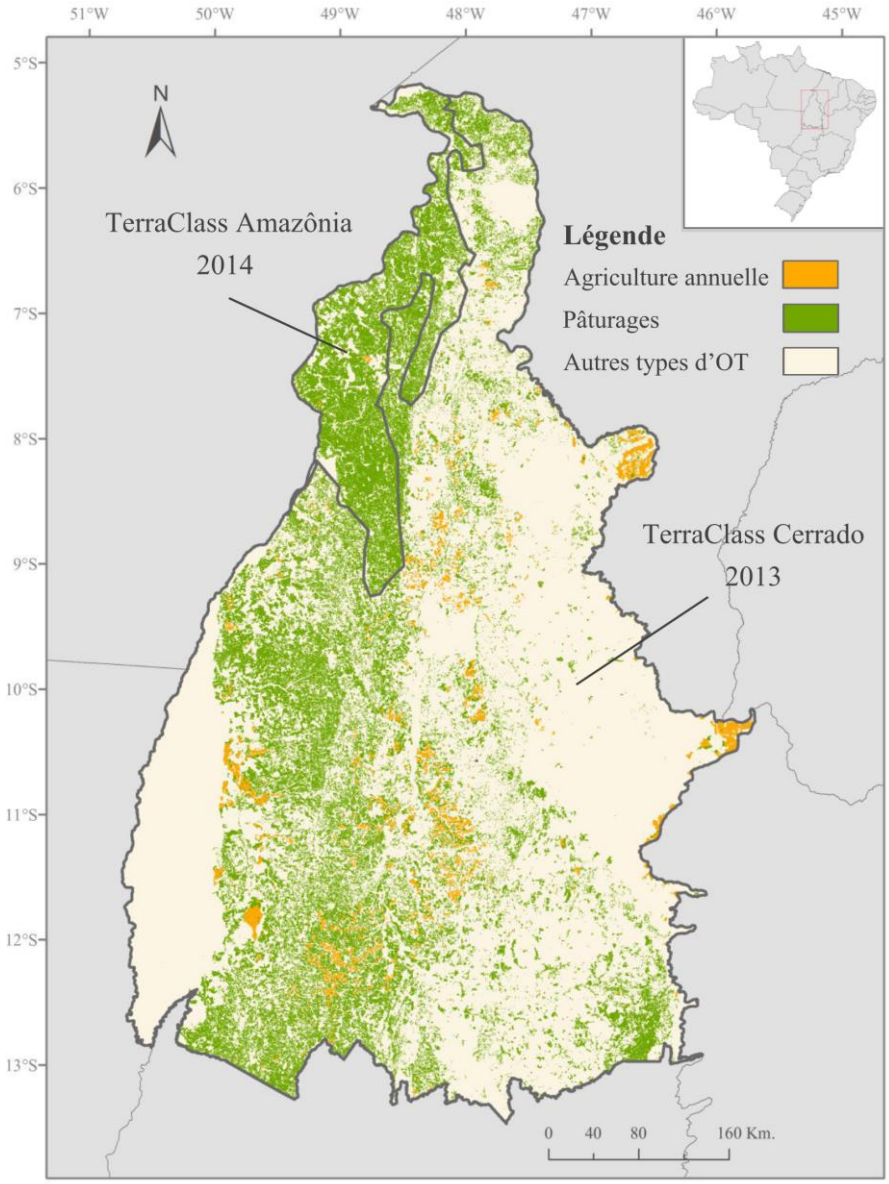


Figure 27. Carte d'occupation des terres (OT) agricole du Tocantins extraite à partir de la carte TerraClass Amazônia 2014 (couvrant le nord-est du Tocantins) et de la carte TerraClass Cerrado 2013 (couvrant le reste de l'État) (sources : TerraClass Amazônia 2014 (INPE, s.d.-a) ; TerraClass Cerrado 2013 (INPE, s.d.-b)).

2.2.3. Données de validation

2.2.3.1. Données de vérité terrain

Deux campagnes de terrain ont été réalisées afin de collecter des points GPS qui ont été utilisés comme vérité de terrain pour la validation des cartes produites : une campagne au Tocantins en octobre 2015, à la fin de la saison culturale 2014-2015, et une campagne au Burkina Faso en septembre 2016, au milieu de la saison culturale.

La stratégie d'échantillonnage utilisée visait à parcourir de façon extensive différents paysages agricoles des deux régions afin de collecter des points GPS représentatifs de la variabilité spatiale des systèmes agricoles. Les campagnes de terrain ont été ainsi planifiées différemment pour les deux régions. La sélection des zones à échantillonner au Tocantins a été assistée par les experts locaux de différentes institutions en lien avec l'agriculture et l'élevage dans la région[33]. Des séances d'enquêtes cartographiques ont été organisées à cette fin, où les acteurs ont localisé sur des spatiocartes les différentes zones agricoles à enquêter. Dans le cas du Burkina Faso, la prise de points GPS a été réalisée sur trois zones distribuées sur un transect nord-sud pour capturer la variabilité spatiale de systèmes agricoles sur le gradient latitudinal, et une zone sur la région cotonnière dans l'ouest du Burkina Faso.

L'équipement pour la collecte de points GPS comprenait une tablette durcie "Trimble Yuma 2", avec un GPS intégré, et le logiciel de système d'information géographique "Quantum GIS" (QGIS) qui permettait la navigation en temps réel et l'enregistrement direct des points GPS avec les attributs associés dans des bases de données spatiales. Pour la mission au Tocantins, la tablette a été équipée avec la mosaïque multispectrale Landsat 8 OLI (présentée dans la section 2.2.1.2). Pour la mission au Burkina, la tablette a été équipée avec les images Sentinel-2A et SPOT 7 (présentées dans la section 2.2.1.3).

Les attributs associés à chaque point GPS correspondent au type de système de culture annuel pour les surfaces cultivées, et le type d'occupation de terres pour les autres types de surface (végétation herbacée et pâturages, jachères, cultures pérennes, végétation arbustive, forêts, surfaces bâties, sols nus et corps d'eau). Des photographies géolocalisées ont été également associées à chaque point. La pratique courante du semis-direct dans les systèmes de cultures annuelles au Tocantins nous a permis d'identifier les cultures de la saison culturale 2014-2015 à partir des résidus qui couvrent les parcelles après la récolte. L'identification des systèmes était assistée par un responsable technique de l'unité thématique sur les systèmes agricoles (NTSA, *Núcleo Temático de Sistemas Agrícolas*) de l'"EMBRAPA Pesca e Aquicultura" et confirmée dans certains cas par les propriétaires des différentes exploitations visitées.

Les chemins secondaires traversant des zones agricoles étaient privilégiés pour maximiser le nombre de points associés à des systèmes de culture. Les points associés à d'autres types d'occupation et utilisation des terres (ex. corps d'eau, espace bâti, forêt), principalement

[33] Les institutions qui ont participé aux séances d'enquêtes cartographiques sont : "Agência de Defensa AgroPECuária do Estado do Tocantins (ADAPEC Tocantins)" ; "Companhia Nacional de Abastecimento (CONAB)" ; "Secretaria da Agricultura e Pecuária - Governo do Estado do Tocantins (SEAGRO)" ; "EMBRAPA Pesca e Aquicultura" (unité thématique sur les systèmes agricoles (NTSA)).

enregistrées suivant les routes principales lors des déplacements entre zones agricoles, ont été complétés a posteriori par photo-interprétation des images satellitaires à haute résolution spatiale (Landsat 8 pour le Tocantins et Sentinel-2 et SPOT 7 pour le Burkina Faso), ainsi que des points associés au système de monoculture de soja très représenté dans une zone isolée dans l'est du Tocantins qui n'a pas pu être visitée.

Au total, 900 points répartis sur 35 municipalités ont été enregistrées pour le Tocantins (Figure 28), et 3663 points répartis sur 11 provinces pour le Burkina (Figure 29). Les classes associées aux points représentées dans les Figures 28 et 29 sont celles utilisées pour la validation des cartes produites (les cartes et les étapes de validation sont présentées dans les chapitres 4 et 5).

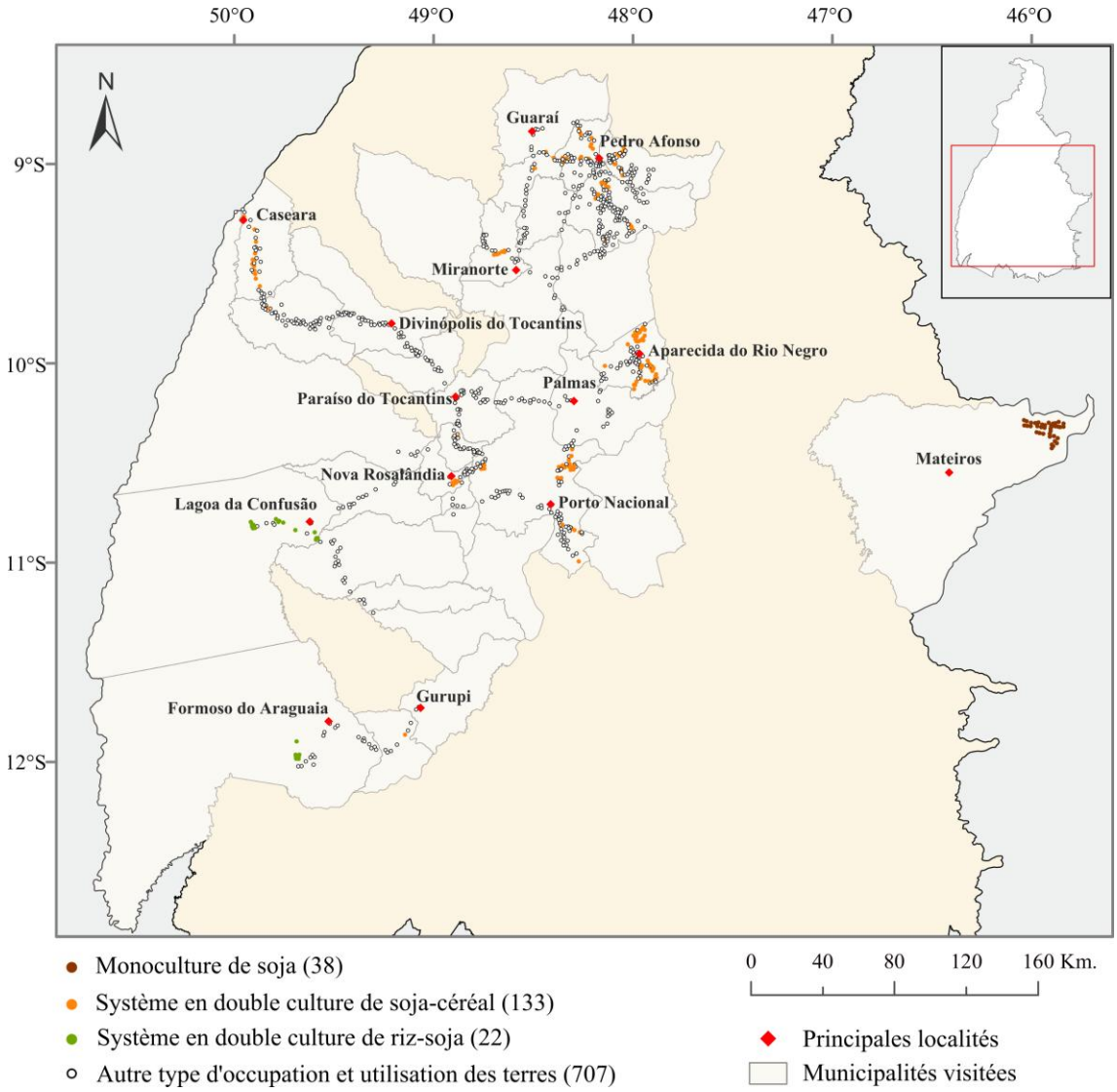


Figure 28. Localisation et nombre de points GPS enregistrés dans la base de données spatiale du Tocantins lors de la campagne de terrain réalisée en octobre 2015. Les classes associées aux points sont celles utilisées pour la validation des cartes produites (source : limites de l'État, limites des municipalités et localités (SEPLAN, 2012)).

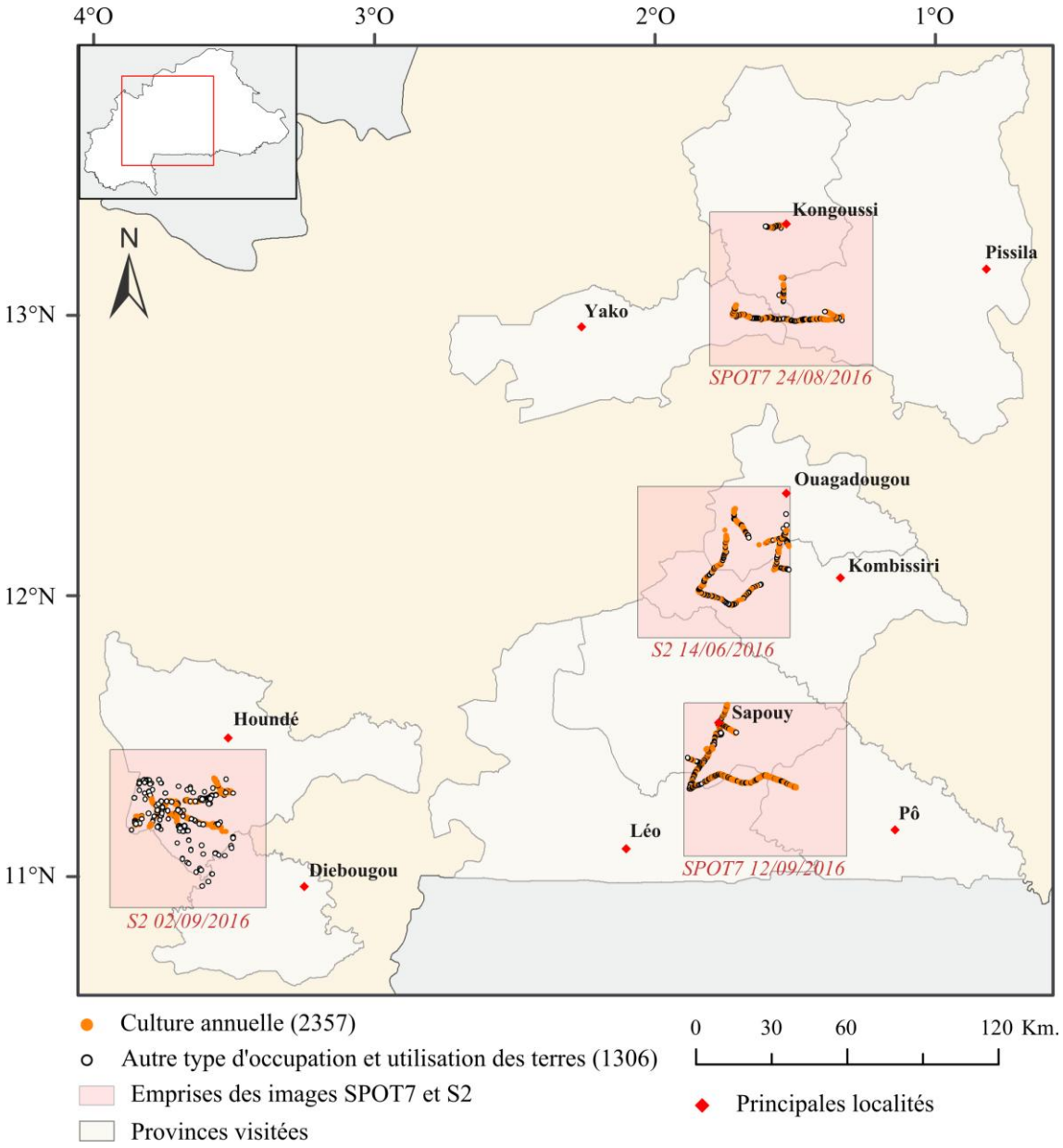


Figure 29. Localisation et nombre de points GPS enregistrés dans la base de données spatiale du Burkina Faso lors de la campagne de terrain réalisée en septembre 2016. Les classes associées aux points sont celles utilisées pour la validation des cartes produites (source : limites du pays et limites des provinces (OCHA ROWCA, 2017) ; localités (OCHA ROWCA, 2015)).

2.2.3.2. Statistiques agricoles

Des statistiques agricoles sur les principales cultures annuelles du Tocantins ont été récupérées à partir de la base de données de la production agricole municipale (PAM, *Produção agrícola municipal*) de l'IBGE (acronyme portugais pour l'Institut brésilien de géographie et statistique) (IBGE, 2014, 2015). Les estimations de la superficie de soja, maïs et riz récoltée en 2014 et 2015 pour les 139 municipalités du Tocantins ont été utilisées pour évaluer nos résultats de classification (présentés dans les chapitres 3 et 4).

Des statistiques agricoles au Burkina Faso ne pouvant pas être accessibles à l'échelle sous-nationale, nous avons récupéré les statistiques agricoles officielles agrégées à l'échelle nationale. Ces statistiques concernent les estimations de la superficie des principales cultures vivrières et des principales cultures de rente du résultat de l'"Enquête Permanente Agricole (EPA)" de 2016 (Campagne 2016/2017) réalisée auprès des ménages ruraux[34] par la Direction Générale des Études et des Statistiques Sectorielles (DGESS) du Ministère de l'Agriculture et des Aménagements Hydrauliques du Burkina Faso (MAAH) (DGESS/MAAH, 2017).

Les estimations des superficies des différentes cultures recensées[35] ont été agrégées pour dériver une estimation de la superficie occupée par des cultures annuelles en 2016 qui a été utilisée pour évaluer notre résultat de classification (présenté dans le chapitre 5). Cependant, ces estimations ne précisent pas la superficie occupée par les cultures associées (ex. l'association du niébé avec les cultures céréalières), l'estimation finale de la superficie occupée par des cultures annuelles surestimant ainsi la superficie « réelle ».

2.3. Les outils

Différents outils ont été utilisés pour traiter et analyser les données : un outil statistique pour la transformation et l'extraction d'informations des images (l'analyse en composantes principales), deux indicateurs de la relation entre variables (le coefficient de corrélation de Pearson et le coefficient de détermination), un outil de segmentation, un outil de classification et six indices statistiques pour évaluer les résultats des classifications. Les caractéristiques de ces outils sont présentées dans les sous-sections suivantes.

2.3.1. L'analyse en composantes principales

Formulée à l'origine par Pearson (1901), l'analyse en composantes principales (ACP) consiste en une transformation orthogonale linéaire (ou rotation) des variables originelles en un ensemble de variables non corrélées connues sous le nom de composantes principales (CP).

[34] Le dispositif de l'EPA couvre les quarante-cinq provinces du pays avec un échantillon d'environ 887 villages (soit 6 ménages au plus par village) et 5163 ménages enquêtés pour la campagne 2016/2017. La méthode d'obtention des estimations est décrite comme suit : « Les superficies […] sont obtenues par la mesure objective, à l'aide de matériel topographique (boussole, ruban métrique, jalons), de l'ensemble des parcelles de tous les ménages échantillons. Ces observations sont pondérées selon les formules statistiques appropriées pour obtenir la superficie totale pour chaque spéculation par province » (DGESS/MAAH, 2017).

[35] Les cultures incluses dans les estimations sont les suivantes : « Cultures céréalières» mil, maïs, riz, fonio, sorgo blanc, sorgo rouge ; « Cultures de rente » coton, sésame, arachide, soja ; « Autres cultures vivrières » Niébé, Voandzou, Igname, Patate.

Lorsqu'elle est appliquée à des séries temporelles d'images satellitaires, l'ACP calcule de nouvelles valeurs pour chaque pixel afin de produire un ensemble de « nouvelles » images CP (également appelées *eigenchannels*). Les CP résultantes ne sont pas corrélées entre elles et sont ordonnées par la proportion de variance retenue de la variance totale initiale (suivant un ordre décroissant). Ainsi, les premières CP (généralement les trois premières CP) concentrent la plus grande proportion de la variance de la série temporelle et captent donc la redondance des informations parmi les images originales, tandis que les dernières CP captent des informations moins redondantes, associées à des variations des valeurs ponctuelles et localisées (les aspects mathématiques et les démonstrations de l'ACP appliquée aux images satellitaires, peuvent être consultés dans l'ouvrage de Richards et Jia (2006)).

Deux méthodes peuvent être utilisées : l'ACP simple, basée sur la matrice de covariances ou l'ACP standard, basée sur la matrice de corrélations. Cette dernière a le même effet que centrer-réduire les valeurs des images originales avant la transformation et est conseillée lorsque les images utilisées ont une plage de dynamique (résolution radiométrique[36]) très différente (ex. quand les images proviennent de différents capteurs).

Les images de la série temporelle de NDVI du capteur MODIS, sur lesquelles on applique l'ACP, ayant une résolution radiométrique commune (16-bits), et afin de préserver la dynamique des valeurs originelle des données dans l'analyse, nous avons choisi d'utiliser une ACP simple. L'outil "rasterPCA" de la librairie R de traitement d'imagerie satellitaire "RStoolbox v.0.1.8" (Leutner et Horning, 2017) a été utilisé pour réaliser l'ACP (la méthode utilisée est décrite dans la section 3.1.1.1).

L'interprétation des résultats de l'ACP peut être basée sur l'examen de la distribution temporelle des vecteurs propres (*eigenvectors* ou *loadings* en anglais) (sortie temporelle de l'ACP) et l'examen des valeurs numériques (*digital numbers* en anglais) des images composantes principales (sortie spatiale de l'ACP) (ex. de Almeida *et al.*, 2015; Eastman et Fulk, 1993; Gupta *et al.*, 2013; Young et Anyamba, 1999).

La distribution temporelle des vecteurs propres illustre le degré auquel chacune des images originales de la série temporelle est corrélée aux images CP résultantes, facilitant ainsi l'interprétation des portions d'informations sur la série temporelle qui ont été captées par

[36] La résolution radiométrique d'une image satellitaire correspond au nombre maximum de niveaux d'intensité (de l'énergie électromagnétique) ou de valeurs numériques disponibles, qui est déterminé par le nombre de bits utilisés pour enregistrer l'image (Richards et Jia, 2006). Le nombre de valeurs numériques disponibles dans une image est égal au nombre 2 à la puissance du nombre de bits. Par exemple les images MODIS avec une résolution radiométrique de 16-bits peuvent stocker 2^{16} valeurs soit 65536 valeurs.

chaque image CP (mise en évidence notamment par les fortes corrélations positives ou négatives avec les images de la série temporelle).

Les valeurs numériques (sans unité) des pixels des images CP résultent de la somme des produits (par pixel) des vecteurs propres par les valeurs radiométriques des images originelles (Gupta *et al.*, 2013). Ainsi, les vecteurs propres estiment les informations « partagées » entre les images originales et les CP, et représentent au même temps la « pondération » de la contribution relative de chaque image originale à la dérivation de chaque CP (i.e. un vecteur propre, de signe positif ou négatif, élevé entre une image originelle et une CP données, indique que la radiométrie de l'image originelle contribue de façon significative à la génération de la CP, et inversement, un vecteur propre faible indique une faible contribution).

En plus de l'interprétation des vecteurs propres et des valeurs des images CP, nous avons calculé et interprété la corrélation entre la première CP dérivée de l'ACP et le NDVI cumulé sur la période annuelle (somme des valeurs de NDVI par pixel des 23 images de la série) à l'aide du coefficient de corrélation linéaire de Pearson présenté dans la sous-section suivante (section 2.3.2). Pour un aperçu synthétique des résultats de l'ACP seuls les trois premières composantes sont présentées. Les résultats de l'ACP et leur interprétation sont présentés dans les sections 3.2.1 et 5.3.1.

2.3.2. Les coefficients de relation entre variables

Le coefficient de corrélation linéaire de Pearson (noté r) permet de mesurer l'intensité de la relation entre deux variables, et le sens de la relation (positif si les variables varient dans le même sens, ou négatif quand elles varient en sens inverse). Ce coefficient est compris entre -1 et 1, plus la valeur de r est proche de -1 ou 1, plus la relation linéaire est forte entre les deux variables, et une valeur de 0 signifie qu'il n'y a pas de relation linéaire.

Ce coefficient a été utilisé lors de l'interprétation des résultats de l'ACP (présentée dans les sections 3.2.1 et 5.3.1) et il a été également utilisé pour mesurer la relation entre les statistiques agricoles et les résultats de classification (présentés dans le chapitre 4).

Le coefficient de détermination linéaire (r^2), qui mesure la proportion de la variation expliquée par un modèle sur la variation totale des valeurs à prédire, est utilisé en complémentarité du coefficient de corrélation pour évaluer la qualité prédictive du modèle de classification utilisé (décrit dans le chapitre 4). Il a été calculé à partir des régressions linéaires simples, où les résultats de classification ont été utilisés comme variable indépendante et les statistiques agricoles comme variable dépendante. Le coefficient de détermination se calcule comme le carré du coefficient de corrélation linéaire (r), et varie

entre 0 et 1, plus la valeur est proche de 1, meilleure est la qualité prédictive du modèle de classification.

2.3.3. L'algorithme de segmentation multi-résolution

L'algorithme de segmentation multi-résolution introduit par Baatz et Schäpe (2000) est basé sur une segmentation itérative de type « croissance de régions » qui a été spécialement conçue pour la segmentation multi-niveau d'imagerie satellitaire[37]. Ce type de segmentation consiste à agglomérer des pixels voisins en « régions » (segments) à partir de pixels de départ (connus comme « graines » ou « *seeds* » en anglais), par des critères de croissance définis par l'utilisateur.

Dans la phase d'initialisation de l'algorithme développé par Baatz et Schäpe (2000), chaque pixel de l'image forme un segment. À chaque itération, les segments voisins sont fusionnés en fonction du degré de similitude entre leurs caractéristiques, déterminé par un « critère d'homogénéité » ("homogeneity criterion"), jusqu'à ce qu'un seuil défini par l'utilisateur, le « paramètre d'échelle » ("scale parameter"), soit atteint.

La croissance des segments, et donc leur aspect final, est ainsi régulée par le critère d'homogénéité, qui est défini par l'opérateur en fixant le poids attribué au paramètre de forme ("shape") en opposition au poids attribué au paramètre de « couleur » ("color"), ces paramètres déterminant l'influence de la forme et de la radiométrie dans la conception des segments, respectivement. Lorsqu'un poids minimum est attribué au paramètre de forme, deux paramètres supplémentaires doivent être pondérés : la compacité de la forme des segments ("compactness") en opposition à une forme aux contours irréguliers et sinueux ("smoothness"). La somme des poids attribués aux paramètres de forme et couleur doit être égale à 1, ainsi que la somme des paramètres de « compacité » et « sinuosité » (0 et 1 étant respectivement les poids minimum et maximum pour chacun des quatre paramètres). Enfin, les différentes bandes spectrales utilisées doivent être aussi pondérées pour contrôler leur influence sur la génération des segments. Pour une contribution égale de toutes les bandes, les poids sont fixés à 1.

La croissance des segments est à terme limitée par le paramètre d'échelle, ceci étant un seuil sans unité qui détermine « l'hétérogénéité maximale autorisée », et donc la taille finale des segments. En conséquence, les segments relativement homogènes (selon le critère

[37] Les détails sur l'algorithme de segmentation multi-résolution avec des exemples d'application à l'imagerie satellitaire peuvent être consultés dans Baatz et Schäpe (2000) et dans Benz *et al.* (2004).

d'homogénéité défini) atteindront une taille plus grande que les segments plus hétérogènes pour un paramètre d'échelle donné. Tout de même, l'algorithme assure une croissance relativement régulière des segments sur toute l'emprise de l'image, et donc généralement une valeur faible de paramètre d'échelle conduit à de nombreux segments de petite taille et inversement, un paramètre d'échelle élevé permettra la croissance de segments plus larges.

L'algorithme de segmentation multi-résolution est utilisé dans le cadre de notre travail pour tester les deux techniques de réduction de la variabilité spatiale, introduites dans la section 1.3.2.3, dans nos régions d'étude. Ainsi, il est mobilisé d'une part pour la stratification des régions d'étude en unités de paysage à partir des données MODIS à moyenne résolution spatiale (correspondant à la première étape de l'approche développée, présenté dans la section 1.5, et décrite dans le chapitre 3), et d'autre part, il est utilisé pour la segmentation de la mosaïque Landsat à haute résolution à l'échelle de la parcelle (correspondant à la deuxième étape de l'approche, qui est décrite dans le chapitre 4).

Les segmentations ont été réalisées avec l'outil "Multiresolution Segmentation" implémenté dans le logiciel d'analyse d'images orientée-objet "eCognition® Developer 9.0" (Trimble Germany GmbH, 2014).

2.3.4. L'algorithme de classification non supervisée *k*-means

L'algorithme de *k*-means (traduit *k*-moyennes en français) a été brièvement présenté dans la section 1.3.2.2, lors de l'introduction des méthodes de classification non supervisées, comme un algorithme de partitionnement ou *clustering* itératif. Cet algorithme vise à partitionner les données en *k* groupes (appelés « clusters ») de façon à minimiser la somme des carrés intra-cluster[38] (c.-à-d. regroupant les individus similaires et séparant les individus dissimilaires). Lorsqu'il est utilisé pour la classification d'imagerie satellitaire, les individus, qui correspondent aux pixels ou aux segments dans le cas d'OBIA, sont regroupés en clusters, avec des caractéristiques spectrales similaires, qui peuvent être ensuite classés par l'utilisateur (ex. en associant une classe d'occupation ou d'utilisation des terres) (Richards et Jia, 2006).

L'algorithme de *k*-means est à ce jour le plus largement utilisé des techniques de fouille de données par partitionnement du fait de la simplicité de sa mise en œuvre, sa rapidité et sa performance avec tout type de données quantitatives (Jain, 2010). Il existe un certain nombre

[38] La somme des carrés intra-cluster est calculée comme la somme des carrés des distances (généralement euclidiennes) entre les individus et la moyenne des individus du cluster d'appartenance. Minimiser la somme des carrés intra-cluster équivaut à minimiser l'inertie (dispersion) des clusters.

de variations de c*et al*gorithme dont les plus connus sont les algorithmes de MacQueen (1967), Forgy (1965), Lloyd (1982) et Hartigan et Wong, (1979).

Pour la classification non supervisée (correspondant à la deuxième étape de l'approche développée, présenté dans la section 1.5, et décrite dans le chapitre 4), nous avons utilisé la version proposée par Hartigan et Wong (1979), implémentée dans l'outil "kmeans" de la librairie R "stats v.3.3.2" du package (R Core Team, 2016). Cette version s'avère moins sensible à la partition aléatoire initiale des données que d'autres versions de l'algorithme du fait que les centres des partitions, aléatoirement distribués lors de l'initialisation de l'algorithme, sont recalculés de façon itérative à chaque réaffectation d'un individu d'un cluster à un autre (Slonim, Aharoni et Crammer, 2013).

Ci-dessous est présenté un pseudo-code de l'algorithme de Hartigan et Wong (1979) utilisé :

1. Choix du nombre de clusters (k)
2. Tirage aléatoire de k individus qui constituent les « centres » de cluster initiaux
3. Calcul de la distance euclidienne entre les individus et chaque centre
4. Affectation de chaque individu au centre le plus proche
5. Calcul des moyennes des individus par cluster, qui deviennent les nouveaux centres, et de la somme des carrés par cluster
6. Si la moyenne d'un cluster a été mise à jour, la somme des carrés de chaque cluster pour chaque individu étant affecté dans un autre cluster est calculée
7. Si l'une des sommes des carrés calculées dans l'étape 6 pour un cluster donné est inférieure à la somme des carrés précédente (étape 5), l'individu responsable du changement est affecté au cluster
8. Boucle itérative des étapes 5, 6 et 7 jusqu'à ce que la somme des carrés intra-cluster soit minimisée pour chacun des clusters.

2.3.5. Les indices d'évaluation de classifications

Pour évaluer les différents résultats de classification (présentés dans les chapitre 4 et 5), nous avons mobilisé six indices : la précision globale et l'intervalle de confiance associé (Rosenfield et Fitzpatrick-Lins, 1986), le coefficient Kappa de Cohen (1960), le désaccord sur la quantité et le désaccord sur la répartition (ou *quantity and allocation disagreement* en anglais) (Pontius et Millones, 2011), le test du χ^2 de McNemar (1947) et le F-mesure (aussi connu comme F-score) (Labatut et Cherifi, 2012). Les bases de données spatiales de points GPS (présentés dans la section 2.2.3.1) ont été utilisées comme vérité-terrain pour le

calcul de ces indices afin d'avoir une appréciation globale de la précision des résultats des classifications.

Tous ces indices peuvent être dérivés d'une matrice de confusion, construite à partir du nombre de points classés (en ligne) et du nombre de points de référence (en colonne). Ainsi, la diagonale de la matrice correspond aux points bien classés. Le Tableau 6 présente une matrice de confusion simplifiée. La notation utilisée dans le Tableau 6 est reprise dans la formulation des indices d'évaluation détaillée ci-après.

Tableau 6. Exemple de matrice de confusion avec la notation utilisée pour la formulation des indices d'évaluation de classifications.

		Référence			
		Classe 1	…	Classe n	Total
Classification	Classe 1	a_{11}	$\cdots$	a_{1n}	$L_1 = \sum_{i=1}^{n} a_{1i}$
	⋮	⋮	⋱	⋮	⋮
	Classe n	a_{n1}	$\cdots$	a_{nn}	$L_n = \sum_{i=1}^{n} a_{ni}$
	Total	$C_1 = \sum_{i=1}^{n} a_{i1}$	$\cdots$	$C_n = \sum_{i=1}^{n} a_{in}$	$N = \sum_{i=1}^{n} L_i = \sum_{i=1}^{n} C_i$

La précision globale (PG) correspond au nombre total de points (objets ou pixels selon le type de classification : orientée-objet ou pixellaire) correctement classés divisé par le nombre total de points évalués et s'exprime en pourcentage :

$$PG = \frac{\sum_{i=1}^{n} a_{ii}}{N}$$

Le coefficient Kappa mesure la différence entre l'accord (entre les points de référence et les points classés) observé et la probabilité d'un accord dû uniquement au hasard, et est exprimé entre 0 (totalement en désaccord ou pas plus d'accord que prévu par hasard) et 1 (totalement en accord) :

$$Kappa = \frac{N \sum_{i=1}^{n} a_{ii} - \sum_{i=1}^{n} L_i C_i}{N^2 - \sum_{i=1}^{n} L_i C_i}$$

Le coefficient Kappa est à ce jour omniprésent dans la communauté de télédétection pour mesurer la précision globale des classifications, mais son utilisation est remise en question par de nombreux auteurs (Foody, 2002; Pontius et Millones, 2011; Strahler *et al.*, 2006). Les mesures de désaccord sur la quantité (DQ) et sur la répartition (DR) ont été ainsi incluses en tant que mesures complémentaires (Pontius et Millones, 2011; Warrens, 2015). Ces deux mesures décomposent le désaccord global (qui est le complément de la précision globale) en deux types de désaccord et fournissent ainsi un aperçu supplémentaire des erreurs de classification. Le désaccord sur la quantité mesure la différence entre les points de référence et les points classés qui est due à la moins que parfaite correspondance dans les proportions des classes, et le désaccord sur la répartition quantifie la différence entre les points de référence et les points classés qui est due à la correspondance moins que parfaite dans la répartition spatiale des classes compte tenu des proportions des classes dans les points classés et de référence (Pontius et Millones, 2011). Ces deux mesures sont exprimées en pourcentage.

$$DQ = \frac{\sum_{i=1}^{n} |l_i - c_i|}{2} \qquad DR = \sum_{i=1}^{n} \min(l_i, c_i) - PG$$

$$\text{où} \quad l_i = \frac{L_i}{N} \quad \text{et} \quad c_i = \frac{C_i}{N}$$

Pour mesurer la précision de la classification pour chacune des classes (i), nous avons utilisé la F-mesure, qui correspond à la moyenne harmonique de la précision du réalisateur (PR) et de l'utilisateur (PU)[39] et varie entre 0 (précision la plus basse) et 1 (précision la plus élevée).

$$F_i = 2\left(\frac{PU_i PR_i}{PU_i + PR_i}\right)$$

$$\text{où} \quad PR_i = \frac{a_{ii}}{C_i} \quad \text{et} \quad PU_i = \frac{a_{ii}}{L_i}$$

Enfin, le test du χ^2 de McNemar a été utilisé pour déterminer s'il existe une différence statistiquement significative entre les précisions de deux résultats de classification (la notation b est utilisée pour représenter les valeurs d'une deuxième matrice de confusion, en équivalence à la notation a dans la première matrice de confusion présentée dans le Tableau 6). La matrice de confusion est résumée dans un tableau de contingence à deux

[39] La précision pour le réalisateur (ou *producer's accuracy* en anglais), équivalent au « rappel » en statistiques, correspond au pourcentage de points de référence d'une classe correctement affectés à cette même classe par la classification. La précision pour l'utilisateur (ou *user's accuracy* en anglais), équivalent à la « précision » en statistiques, exprime quant à elle le pourcentage de points d'une classe qui sont correctement classés.

dimensions qui inclut le nombre de points correctement classés dans les deux classifications (m_{11}), le nombre de points mal classés dans les deux classifications (m_{00}), le nombre de points correctement classés dans la première classification mais mal classés dans la deuxième classification (m_{10}) et le nombre de points correctement classés dans la deuxième classification mais mal classés dans la première classification (m_{01}).

$m_{11} = \min(a_{ii}, b_{ii})$	$m_{10} = \sum_{\substack{i=1 \\ a_{ii} \geq b_{ii}}}^{n} a_{ii} - b_{ii}$
$m_{01} = \sum_{\substack{i=1 \\ a_{ii} \leq b_{ii}}}^{n} b_{ii} - a_{ii}$	$m_{00} = N - \sum_{i=1}^{n} \max(a_{ii}, b_{ii})$

$$\chi^2 = \frac{(m_{01} - m_{10})^2}{m_{01} + m_{10}}$$

Chapitre 3.
Stratification et caractérisation des systèmes d'utilisation agricole des terres à l'échelle régionale

3.1. MÉTHODE ADOPTÉE 89
3.1.1. Délimitation des unités de paysage 90
3.1.1.1. Extraction des « variables phénologiques » 90
3.1.1.2. Segmentation des composantes principales 91
3.1.2. Classification des unités de paysage 92
3.1.2.1. Classification par règles de décision 92
3.1.2.2. Classification par analyse des profils temporels de NDVI 93
3.1.3. Évaluation des résultats de classification 94
3.2. RÉSULTATS 94
3.2.1. Les composantes principales 94
3.2.2. Les unités de paysage 99
3.2.3. Les systèmes d'utilisation agricole des terres 100
3.2.4. Résultats de l'évaluation de la classification 103
3.3. DISCUSSION 105

Chapitre 3. Stratification et caractérisation des systèmes d'utilisation agricole des terres à l'échelle régionale

Comme cela a été présenté précédemment dans le chapitre introductif (chapitre 1), les systèmes agricoles peuvent être cartographiés à plusieurs niveaux d'organisation, chaque niveau apportant des informations complémentaires sur les systèmes, mais les approches de cartographie concernent généralement la cartographie des systèmes de culture menée à l'échelle de parcelle. On s'intéresse dans ce troisième chapitre à la cartographie des systèmes agricoles à un niveau d'organisation supérieur notamment celui du paysage, au travers de la délimitation d'unités de paysage agricole.

Ainsi, dans ce chapitre on s'attache à apporter quelques éléments de réponse à la question suivante (présentée dans la section 1.4) :

- Comment délimiter des unités de paysage agricole à partir des données satellitaires pour la cartographie des systèmes d'utilisation agricole des terres ?

Une piste méthodologique avancée dans le chapitre introductif concerne le potentiel non exploré de la stratification spatiale par télédétection pour la délimitation d'unités de paysage agricole. L'hypothèse de départ qui sous-tend cette piste considère que les images satellitaires contiennent des informations synoptiques sur les principaux éléments constitutifs des paysages agricoles et sur les facteurs humains et environnementaux qui modèlent ces paysages et permettraient donc de les localiser et délimiter sans avoir recours à des données externes.

Néanmoins, cette piste méthodologique soulève d'autres questions :

- Quels indicateurs issus des images satellitaires seraient pertinents pour la délimitation d'unités de paysage agricole ? Quelles méthodes de télédétection permettraient d'automatiser la délimitation des unités spatiales de façon relativement objective à partir des indicateurs ? Comment transférer les informations locales à l'échelle des unités de paysage ?

Ce chapitre apporte des réponses méthodologiques à ces questions, en proposant des méthodes originales de délimitation d'unités de paysage à partir de séries temporelles de NDVI du capteur MODIS, et de classification de ces unités de paysage à partir d'une carte d'occupation des terres.

Le travail présenté dans ce chapitre a été valorisé par une publication dans une revue internationale de rang A en 2017 (présentée en annexe) :

> Bellón, B., Bégué, A., Lo Seen, D., de Almeida, C.A., et Simões, M. (2017). "A Remote Sensing Approach for Regional-Scale Mapping of Agricultural Land-Use Systems Based on NDVI Time Series". *Remote Sensing*, 9(6), 600. http://doi.org/10.3390/rs9060600

3.1. Méthode adoptée

La méthode adoptée dans ce chapitre se décompose en deux étapes : (i) une étape de délimitation des unités de paysage à partir de la détection des variations saisonnières des surfaces cultivées à partir d'une série annuelle d'images MODIS-NDVI et une segmentation OBIA et (ii) une étape de classification des unités de paysage à partir d'une carte d'occupation des terres par règles de décision et analyse visuelle des profils temporels de NDVI.

La Figure 30 présente un diagramme de flux avec les principales étapes méthodologiques qui sont décrites dans les sous-sections suivantes. Les données en entrée correspondent à la série annuelle de NDVI de la saison 2013-2014 au Tocantins (présentée dans la section 2.2.1.1) et la carte d'occupation des terres TerraClass (présentée dans la section 2.2.2).

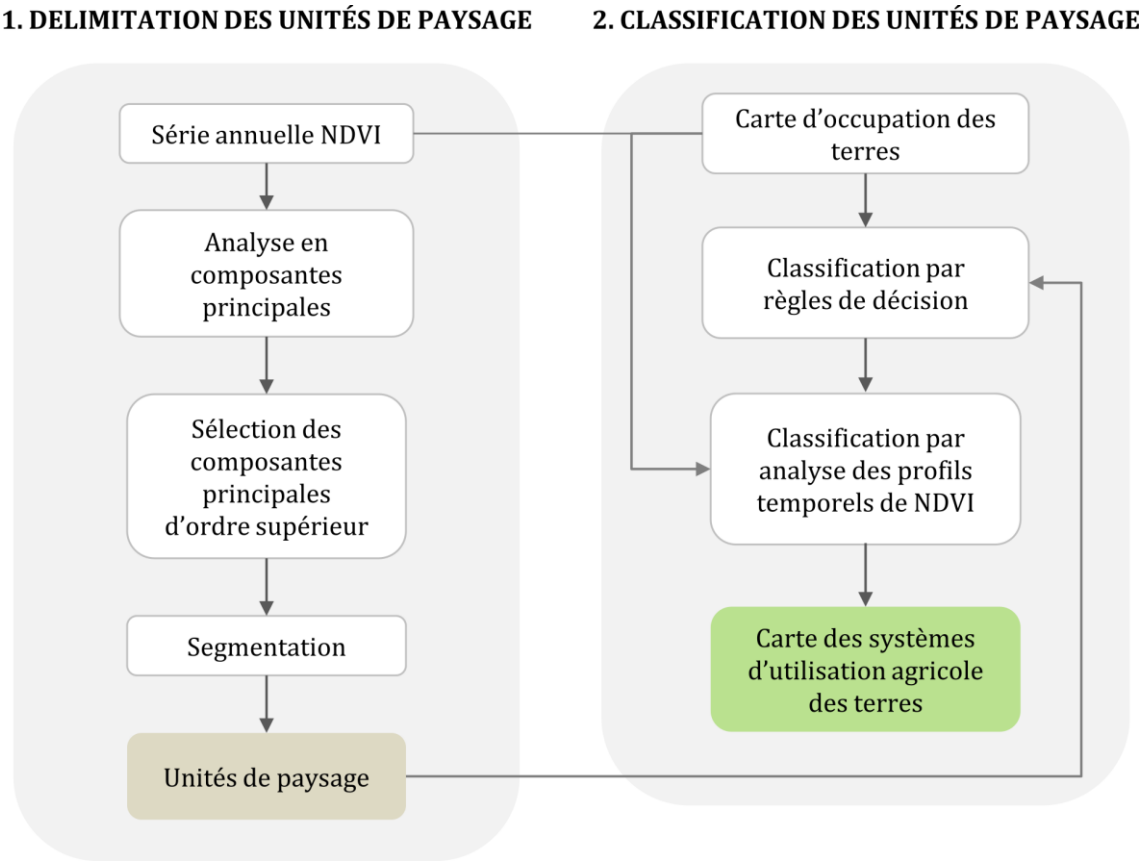


Figure 30. Représentation des étapes méthodologiques présentées dans le chapitre 3.

3.1.1. Délimitation des unités de paysage

La délimitation des unités de paysage se décompose en deux étapes : (i) extraction des « variables phénologiques » à partir de la série annuelle de NDVI par analyse en composantes principales (section 3.1.1.1) et (ii) délimitation des unités de paysage à l'aide d'une segmentation des composantes principales d'ordre supérieur (section 3.1.1.2).

3.1.1.1. Extraction des « variables phénologiques »

L'analyse en composantes principales (ACP) est particulièrement performante lorsqu'elle est appliquée aux séries temporelles de NDVI pour identifier les variations saisonnières des surfaces terrestres et révéler leur relation avec différents facteurs, principalement la variabilité climatique et les facteurs liés à l'activité humaine (Eastman et Fulk, 1993; Gurgel et Ferreira, 2003; Hall-Beyer, 2003; Hirosawa, Marsh et Kliman, 1996; Wang *et al.*, 2010; Young et Anyamba, 1999). Le type d'informations captées par chaque composante principale, ou CP, varie en fonction des caractéristiques des séries temporelles de NDVI en entrée (résolution temporelle et longueur des séries temporelles, résolution spatiale et étendue).

Cependant, des résultats communs sont observés lorsque la zone analysée est assez large pour inclure différents types de surface : la première CP capte l'élément de majeure variabilité de la série temporelle associé à la variabilité spatiale du NDVI, et s'avère corrélée au NDVI intégré sur la période considérée (c.-à-d. capte la redondance des informations parmi les images de la série), tandis que les CP d'ordre supérieur captent la variabilité saisonnière et intra-saisonnière du NDVI, chacune étant liée à un facteur particulier ou à une combinaison de facteurs de changement (Eastman et Fulk, 1993; Gurgel et Ferreira, 2003; Hall-Beyer, 2003; Hirosawa *et al.*, 1996; Wang *et al.*, 2010; Young et Anyamba, 1999).

Compte tenu de la superficie du Tocantins, supérieure à 200 000 km^2, et de la résolution spatiale des images issues du capteur MODIS à 250 m, nous nous attendons à ce que la majeure proportion de la variance de la série temporelle de NDVI soit également associée à la variabilité spatiale des valeurs de NDVI, et qu'elle soit conformément exprimée dans la première CP. Pareillement, nous prévoyons que les variations temporelles localisées associées aux systèmes agricoles, configurées par le développement saisonnier de la végétation et la variabilité intra-saisonnière induite par les pratiques agricoles, soient exprimées dans les CP d'ordre supérieur.

Ainsi, nous avons réalisé une ACP sur la série annuelle 23 images de NDVI et nous avons écarté la première CP des traitements postérieurs, après avoir vérifié qu'elle est liée à la variabilité spatiale du NDVI cumulé sur la période annuelle, en cohérence avec les résultats d'études précédentes (ex. Gurgel et Ferreira, 2003; Hall-Beyer, 2003) (les résultats de l'ACP sont présentés dans la section 3.2.1). Les 22 dernières CP (les images CP d'ordre supérieur), qui contiennent 27 % de la variance totale, ont été retenues pour la suite, et sont utilisées en tant que « variables phénologiques » du fait de leur relation avec les variations temporelles de la végétation.

3.1.1.2. Segmentation des composantes principales

Les images CP retenues ont été segmentées en utilisant l'algorithme de segmentation multi-résolution présenté dans la section 2.3.3, l'objectif ultime de la segmentation étant la stratification de la région en unités de paysage avec une phénologie relativement homogène.

Le paramètre de « forme » lors de la définition du critère d'homogénéité pour la conception des strates (correspondant aux unités de paysage) devient ainsi superflu, et nous attribuons donc le poids total à la radiométrie (paramètre « couleur ») en le fixant à 1. En pratique, cela signifie que l'homogénéité relative des strates résultantes est entièrement basée sur les valeurs des pixels des images CP et donc sur les informations implicites liées à

la variabilité temporelle du couvert végétal. En outre, chacune des 22 images CP contribue de façon égale à la conception des strates (le poids de chaque image est fixé à 1).

Une échelle, selon Benz *et al.* (2004), est définie comme le niveau d'agrégation et d'abstraction auquel un objet peut être clairement décrit. Puisque nos objets géographiques d'intérêt, les unités de paysage, ne se rapportent pas à une échelle particulière mais peuvent plutôt être décrits à différentes échelles d'analyse selon l'objectif d'étude, nous considérons qu'il n'y a pas d'échelle de segmentation optimale et que différents paramètres d'échelle doivent être testés dans chaque cas pour approcher des unités de paysage thématiquement significatives. Nous partons de l'hypothèse que l'algorithme de segmentation multi-résolution basé sur le critère d'homogénéité délimite des objets relativement homogènes (d'un point de vue statistique) quel que soit le paramètre d'échelle utilisé.

Cependant, quelques principes directeurs simples ont été suivis pour choisir parmi différents paramètres d'échelle : les objets doivent être plus grands que l'unité cartographique minimale d'un système agricole (c.-à-d. la parcelle) et il doit y avoir un compromis judicieux entre le nombre de strates et leur degré d'hétérogénéité. Après avoir testé plusieurs échelles de segmentation, le seuil du paramètre d'échelle (sans unité) a été empiriquement fixé à 850.

3.1.2. Classification des unités de paysage

La classification des unités de paysage se développe en deux temps : (i) classification basée sur des règles de décision (section 3.1.2.1) et (ii) classification par analyse des profils temporels de NDVI (section 3.1.2.2).

3.1.2.1. Classification par règles de décision

La carte d'occupation agricole des terres extraite du produit TerraClass (présentée dans la section 2.2.2) a été utilisée pour classer les unités de paysage associées à des activités agricoles (culture ou élevage) en quatre principaux types de systèmes d'utilisation agricole des terres : système de culture dominant ; système mixte (système de culture-élevage) ; système d'élevage semi-intensif ; système d'élevage intensif[40].

[40] Dans l'arbre de décision utilisé, la notion d'intensification des systèmes d'élevage est liée à l'intensification de la production (de viande ou de lait) dans les systèmes d'élevage extensifs traditionnels brésiliens et donc liés à l'extension de la superficie des pâturages (i.e. une unité de paysage classée avec un type de système d'élevage intensif correspond ainsi à une unité de paysage caractérisée par une grande extension de pâturages) (de Almeida, 2016).

Cette classification préliminaire des systèmes d'utilisation des terres a été réalisée suivant les règles de l'arbre de décision développé par de Almeida *et al.* (2016) pour la localisation des systèmes agricoles de l'État brésilien du Rondônia au niveau des municipalités, présenté dans la Figure 31. Cet arbre de décision semble bien adapté pour la classification des systèmes d'utilisation des terres de l'État du Tocantins considérant sa proximité avec l'État du Rondônia en termes géographique, et d'occupation et d'utilisation des terres.

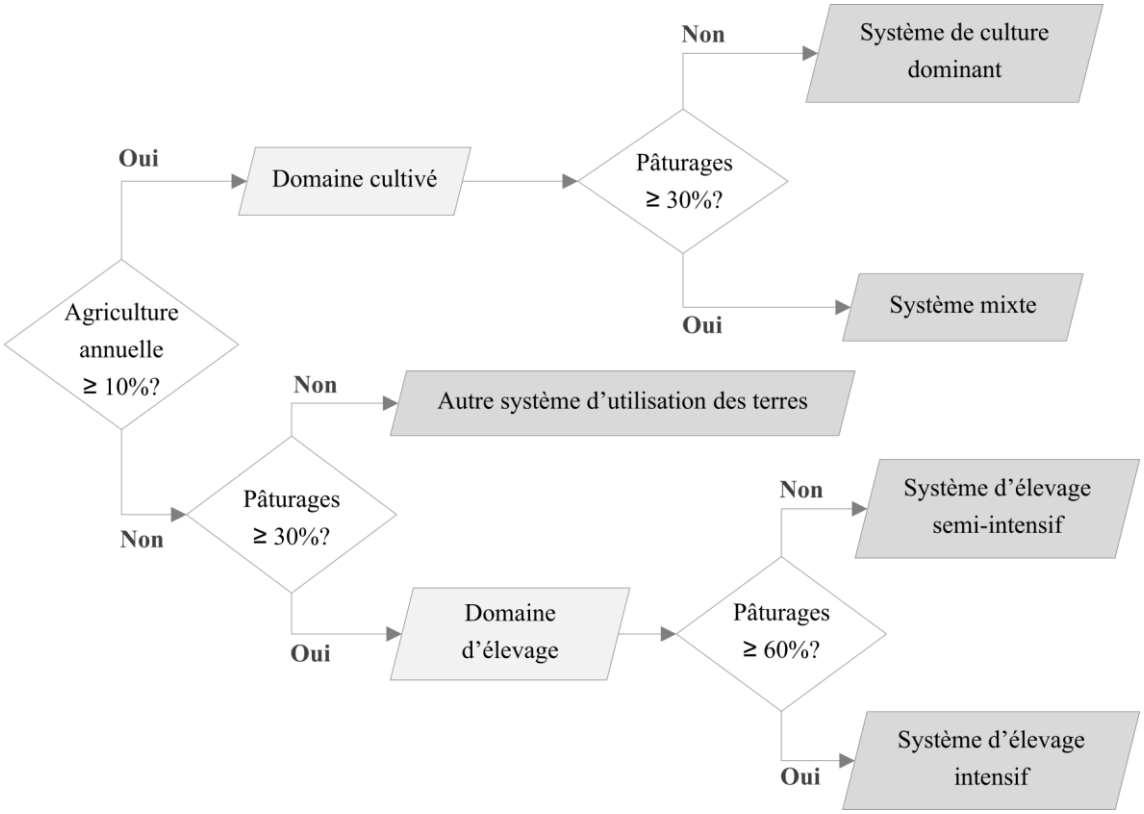


Figure 31. Règles de l'arbre de décision utilisé pour la classification des unités de paysage en quatre principaux types de systèmes d'utilisation agricole des terres (adaptée de de Almeida *et al.* (2016)).

3.1.2.2. Classification par analyse des profils temporels de NDVI

Les unités de paysage appartenant au « domaine cultivé » ont été ensuite classées plus finement en système de culture majoritaire par analyse de leurs profils temporels de NDVI. Le profil moyen de NDVI des surfaces cultivées à l'intérieur de chaque unité de paysage a été extrait de la série temporelle de NDVI en utilisant le masque « agriculture annuelle » de la carte TerraClass.

Les variations générales du NDVI au cours de la saison culturale ont été analysées par interprétation visuelle de la forme des profils temporels moyens du NDVI. La forme des profils temporels de NDVI a révélé les principaux stades de développement du domaine cultivé au sein de chaque unité de paysage. Les profils de NDVI avec des formes

visuellement équivalentes ont été regroupés et classés en différents systèmes de culture par mise en correspondance avec le calendrier cultural des principaux systèmes de culture du Tocantins.

3.1.3. Évaluation des résultats de classification

Compte tenu de l'absence de cartes thématiques de référence pour valider la précision de la classification de la carte finale de systèmes d'utilisation agricole des terres, nous avons réalisé une évaluation non supervisée de la carte finale à partir des statistiques agricoles sur les superficies des principales cultures annuelles récoltées en 2014 par municipalité (présentées dans la section 2.2.3.2). Pour élaborer une « pseudo-carte » de référence, nous avons d'abord calculé la proportion de la superficie occupée par chaque type de culture dans les municipalités avec plus de 2000 hectares de superficie cultivée. Les types de cultures prédominants (représentant la plus grande partie de la superficie cultivée par municipalité) ont été différenciés des cultures mineures et utilisés pour caractériser le système de culture principal dans chaque municipalité et réaliser ainsi une typologie de municipalités.

La carte de systèmes d'utilisation agricole des terres produite a finalement été évaluée sur la base de l'accord spatial entre les systèmes de culture classés et la typologie des systèmes de culture de la carte de référence.

L'évaluation a été réalisée uniquement pour les résultats de classification des systèmes de culture, car aucune statistique officielle n'était disponible sur les pâturages avec lesquelles évaluer les résultats de la classification des systèmes d'élevage.

3.2. Résultats

3.2.1. Les composantes principales

La Figure 32 montre les vecteurs propres des trois premières composantes principales en fonction des images de la série annuelle de 2013-2014 (représentées par leur code du jour julien).

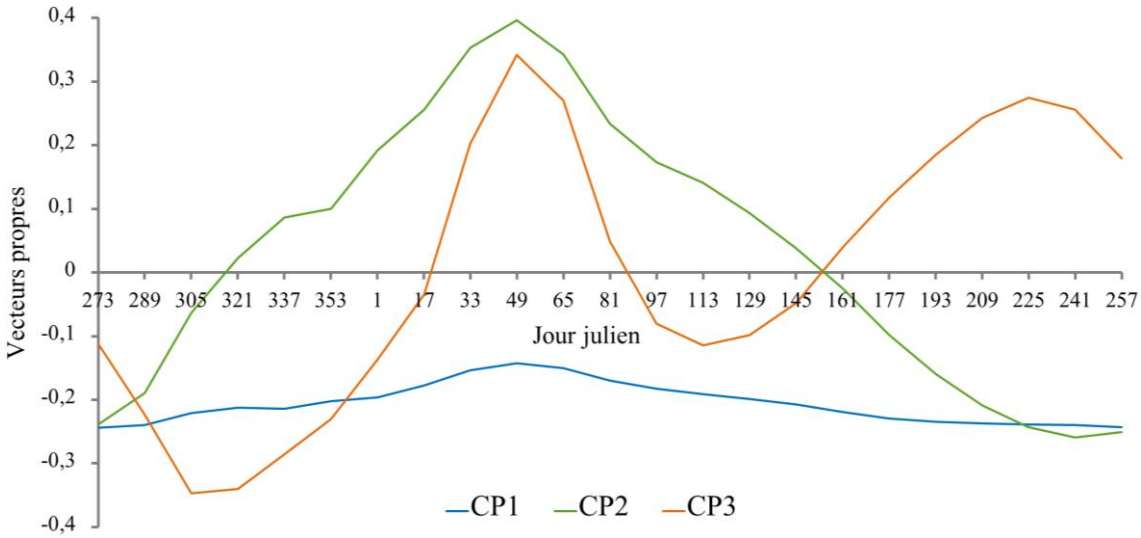


Figure 32. Magnitude des vecteurs propres (axe d'ordonnées) des trois premières composantes principales en fonction des images de la série annuelle de 2013-2014 (représentées dans l'axe d'abscisses par leur code du jour julien).

On observe dans la Figure 32 une distribution temporelle relativement constante des vecteurs propres de la première CP sur toute la période annuelle avec le même signe négatif. Ceci indique que la CP1 est corrélée de façon négative avec toutes les images de la série temporelle, ne représentant pas de variation temporelle particulière. Exprimé autrement, toutes les images de la série annuelle ont contribué de façon proportionnelle à la dérivation de la première CP. Ce résultat implique que le principal facteur de variabilité dans la série temporelle capté par la première composante (représentant 73 % de la variance totale) est celui qui se produit dans l'espace (i.e. la variabilité spatiale de valeurs de NDVI associée à la radiométrie des différents types de surface) (Eastman et Fulk, 1993).

Le profil des vecteurs propres de la deuxième CP (Figure 32 CP2) montre un cycle annuel associé au deuxième facteur majeur de variabilité des valeurs de NDVI dans la série temporelle (après la variabilité spatiale). On observe que les images allant du jour 321 au jour 145 ont une corrélation positive avec les vecteurs propres du CP2 et inversement, les images aux extrêmes de la série annuelle (allant du jour 273 à 305 et 161 à 257) présentent une corrélation négative. Ce résultat suggère que le deuxième CP a capté le facteur de variation saisonnière dominant dans la série annuelle qui est lié au contraste de valeurs entre la saison sèche et la saison de pluies.

Le pic maximum des vecteurs propres positifs (jour 49) serait ainsi associé au pic du maximum de croissance de la végétation pendant la saison de pluies (mois de février) et les pics minimums des vecteurs propres négatifs (jour 273 et 241) seraient associés aux plus faibles valeurs de NDVI pendant le mois de septembre et première quinzaine d'octobre, à la

fin de la saison sèche et avant les premières pluies. Le changement de signe autour des jours 321 (novembre) et 161 (juin) seraient ainsi associés respectivement aux processus de reverdissement et de sénescence des surfaces végétalisées (i.e. l'intersection avec la valeur zéro dans l'axe d'ordonnées est comprise comme une limite entre les deux saisons, lorsqu'une situation de la végétation, sénescence ou reverdissement, commence à dominer sur l'autre), et entre les pics minimums et maximum les vecteurs propres suggèrent une diminution ou augmentation des valeurs de NDVI graduelle suivant le régime pluviométrique.

Le profil des vecteurs propres de la CP3 (Figure 32 CP3), comme celui de la CP2, montre un cycle annuel qui indique que cette CP a capté des variations de la végétation dans la série annuelle. Les variations temporelles, mises en évidence par les oscillations du profil des vecteurs propres en fonction du temps et les inversions de signe, semblent être moins graduelles que celles représentées par le profil de la CP2, indiquant que la CP3 a capté des variations des valeurs de NDVI plus rapides dans le temps (variations intra-saisonnières). La CP3 et les CP ultérieures (de la CP4 jusqu'à la CP23) distinguent ainsi des épisodes de changement isolés qui s'écartent de la variation saisonnière dominante capté dans la CP2, leur interprétation étant progressivement plus complexe. Les variations intra-saisonnières de NDVI captés par ces CP d'ordre supérieur peuvent être liées ainsi à divers facteurs ou combinaison de facteurs naturelles (ex. feux, inondations) et anthropiques (ex. déforestation) y compris les pratiques agricoles.

La Figure 33 montre les trois premières images CP. L'interprétation de ces images CP suit le comportement et les motifs des profils temporels des vecteurs propres respectifs, les valeurs des pixels dépendant directement de la magnitude de vecteurs propres et de leur signe (Gupta *et al.*, 2013).

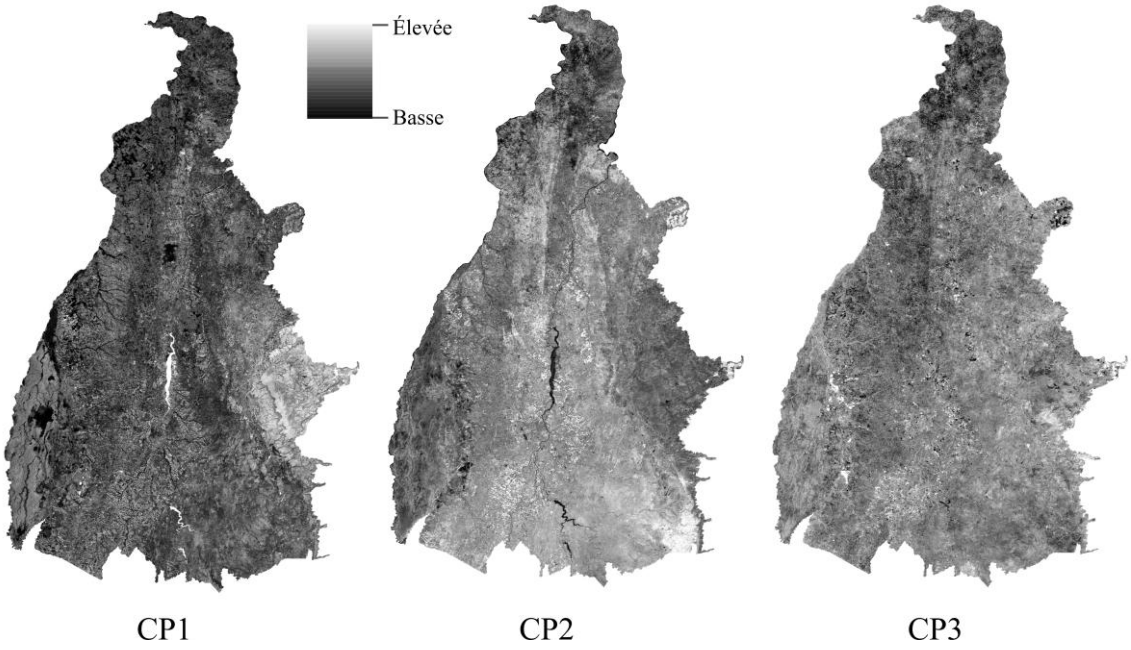


Figure 33. Les trois premières composantes principales dérivées de la série temporelle de NDVI de l'année 2013-2014. Les images CP sont représentées en niveaux de gris, illustrant la gradation des valeurs numériques allant des valeurs basses (la couleur noire correspond à la valeur minimale) aux valeurs élevées (la couleur blanche correspond à la valeur maximale) pour chaque CP.

Dans la première image CP, les zones correspondantes aux formations végétales denses avec une couverture végétale stable au cours de l'année (ex. formations boisées permanentes comme les forêts ou la végétation ripicole) sont représentées par des pixels foncés (correspondant à des valeurs numériques faibles) en accord avec la corrélation négative du profil des vecteurs propres du CP1 avec toutes les images de la série annuelle. Inversement, les surfaces non végétalisées (ex. les terres dénudées, les affleurements rocheux, les corps d'eau, les espaces bâties) sont représentées par des pixels clairs (correspondant à des valeurs numériques élevées).

En calculant la corrélation entre la CP1 et le NDVI cumulé sur toute la période annuelle à l'aide du coefficient de corrélation linéaire de Pearson, on observe une forte corrélation linéaire globale négative de r = - 0,99. Le résultat de la corrélation par pixel est présenté dans la Figure 34.

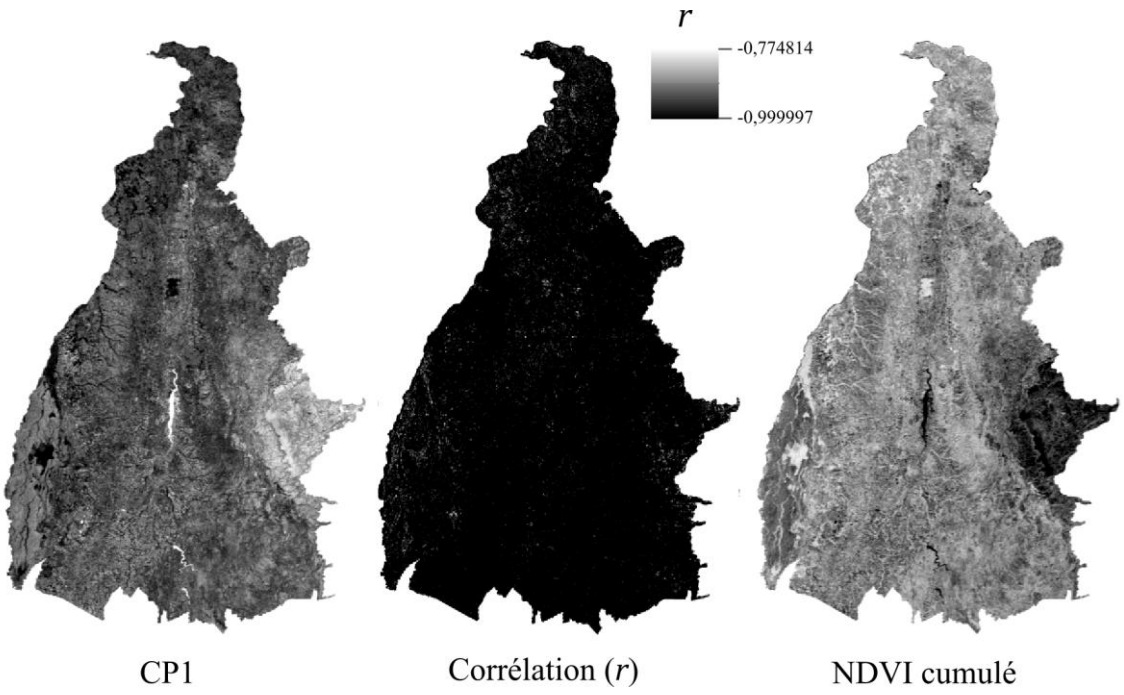


Figure 34. Corrélation ($\boldsymbol{r}$) entre la CP1 et le NDVI cumulé sur la période annuelle 2013-2014.

On observe ainsi, une inversion entre les valeurs élevées de la CP1 (pixels clairs) et les valeurs basses du NDVI cumulé (pixels sombres) et vice-versa (Figure 34). La corrélation est proche à 1 pour la plupart des pixels sauf pour quelques pixels isolés. Ce résultat suggère que la CP1 a capté la répartition spatiale du NDVI cumulé sur l'année.

Dans l'image CP2 (Figure 33) les surfaces qui présentent une variabilité saisonnière liée au régime pluviométrique sont soulignées par des pixels clairs (valeurs numériques élevées). Étant donné que les valeurs numériques des images CP résultent de la somme des produits des vecteurs propres par les valeurs radiométriques des images originelles (Gupta *et al.*, 2013), les valeurs numériques élevées associées à ces surfaces s'expliquent comme suit : les vecteurs propres positifs multipliés par des valeurs de NDVI élevées pendant la saison de pluies et les valeurs propres négatives multipliées par des valeurs de NDVI faibles pendant la saison sèche, résultent tous les deux en valeurs numériques élevées dans l'image CP2.

Dans l'image CP3 (Figure 33) on observe que les pixels les plus clairs et sombres sont répartis de façon éparse et isolée. Ce résultat implique que la troisième CP (et les CP ultérieures) ne capte pas seulement la variabilité de valeurs de NDVI « localisée » dans le temps (i.e. les changements rapides liés ou variations intra-saisonnières) mais aussi variabilité localisée spatialement (Eastman et Fulk, 1993).

Les résultats obtenus suivent ainsi la logique de l'ACP, qui ordonne les CP par la proportion de variance retenue de la variance totale initiale de manière décroissante, la CP1 captant les informations plus redondantes dans la série temporelle (associées à la variabilité spatiale des valeurs de NDVI) et les CP d'ordre supérieur captant progressivement des variations des valeurs plus ponctuelles et localisées (associées à la variabilité temporelle des valeurs de NDVI). Ces résultats sont cohérents avec les études précédentes appliquant des ACP à partir des séries temporelles d'images d'indice de végétation sur des grandes étendues, présentées dans la section 3.1.1.1.

3.2.2. Les unités de paysage

La segmentation des « variables phénologiques » a produit un total de 90 unités de paysage (Figure 35), avec une superficie moyenne de 3 090 km^2, 44 km^2 pour la plus petite et 15 069 km^2 pour la plus large.

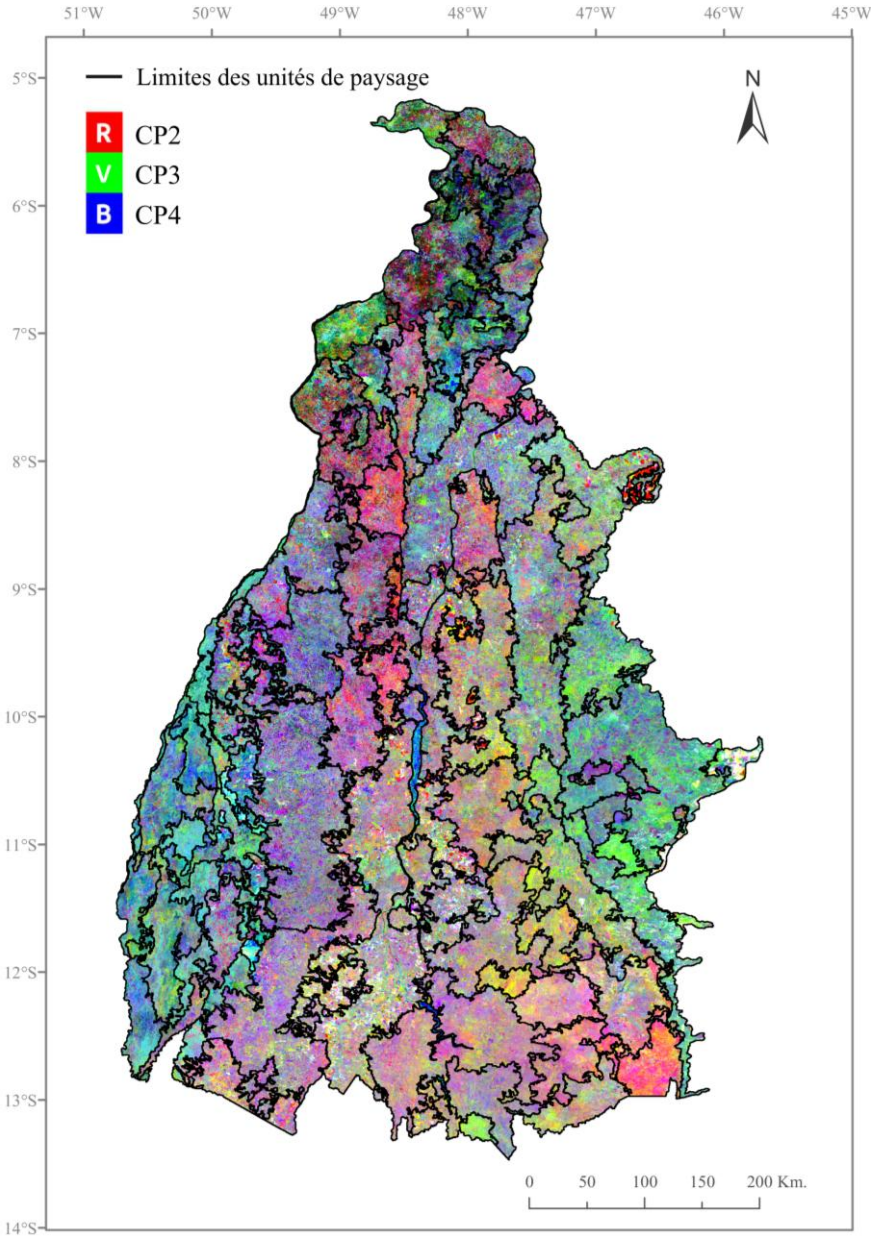


Figure 35. Stratification de la région du Tocantins en 90 unités de paysage, obtenues par segmentation des 22 images de Composantes Principales (CP). Le fond d'image est une composition colorée des trois composantes principales CP2, CP3, CP4 (RGB).

3.2.3. Les systèmes d'utilisation agricole des terres

Les règles de décision utilisées pour classer les unités de paysage à partir de la carte d'occupation des terres agricole TerraClass (Figure 31) ont permis d'identifier un total de 44 unités de paysage agricole (114 114 km^2) parmi les 90 unités de paysage délimitées. Sur les 44 unités de paysage agricole, 21 ont été classées comme « système de culture dominant » (11 193 km^2), 15 comme « système d'élevage semi-intensif » (82 576 km^2), 7comme « système d'élevage intensif » (18 329 km^2), et 1 a été classée comme « système mixte » (2 016 km^2).

Trois groupes de profils temporels de NDVI avec des formes distinctes ont été identifiés par analyse visuelle des profils temporels moyens de NDVI des surfaces cultivées au sein des 22 unités de paysage agricole classées comme appartenant au domaine cultivé (21 unités classées « système de culture dominant » et 1 unité classée « système mixte »). Les 22 profils temporels de NDVI, ainsi que leur groupe, sont présentés dans la Figure 36.

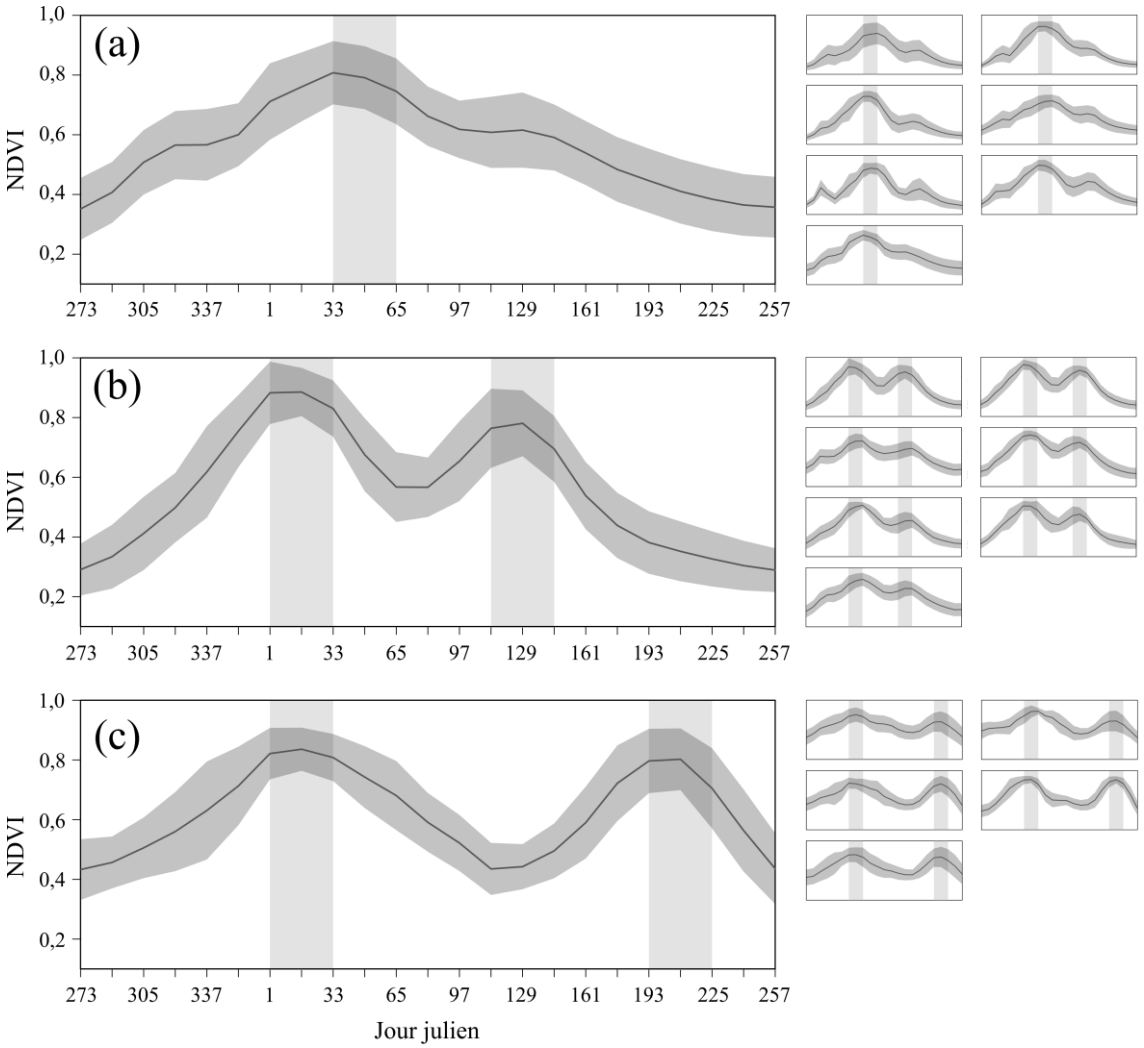


Figure 36. Profils temporels de NDVI des 22 unités de paysage agricole appartenant au domaine cultivé. Les figures montrent la moyenne (courbe noire continue) et l'écart-type (courbe gris foncé ombrée) des valeurs NDVI des pixels à l'intérieur du masque de la classe « agriculture annuelle » de la carte d'occupation des terres TerraClass, pour l'année culturale 2013-2014. Un profil d'une unité de paysage agricole représentatif de chacun des trois groupes de profils temporels de NDVI (a), (b) et (c) est présenté en détail à gauche, et les profils des autres unités de paysage appartenant à chaque groupe sont représentés à droite. Les pics de croissance des cycles culturaux sont représentés par des bandes verticales grises ombrées (d'une durée de 32 jours) qui sont identiques entre figures d'un même groupe.

Les calendriers culturaux des principaux systèmes de culture du Tocantins (présentés dans la section 2.1.1.2) nous ont aidé à interpréter les différents groupes de profils temporels de NDVI. Dans le premier groupe de profils (Figure 36(a)), un cycle de culture unique prédomine, avec un pic de croissance (mis en évidence par le pic de NDVI) autour de février (du jour 33 au jour 65). Étant donné que seules des variations mineures du NDVI apparaissent avant et après le cycle de croissance principal, on peut considérer qu'il n'y a pas de deuxième cycle de développement significatif du couvert végétal et que, par conséquent, les unités de paysage appartenant au premier groupe sont principalement associées à un seul système de culture. Étant donné que la culture de soja est la culture dominante du Tocantins, le premier groupe d'unités de paysage semble représentatif du système de monoculture de soja.

Les deuxième et troisième groupes de profils (Figure 36(b)(c)) présentent tous les deux un double pic de NDVI représentatif de deux cycles de culture distincts. Bien que le premier cycle des deux groupes présente une période de pic de NDVI commune autour de janvier (du jour 1 au jour 33), ces deux groupes sont facilement dissociables par leur deuxième cycle, qui apparaît avec plus d'un mois de différence entre les deux groupes.

- Les profils phénologiques du deuxième groupe (Figure 36(b)) sont représentatifs du système de double culture soja-céréale, caractérisé par une culture de soja d'été, cultivée pendant la saison de pluies, suivie d'une culture de céréale d'hiver. Ce profil présente de faibles valeurs de NDVI autour de mars (du jour 65 au jour 81) en raison de la récolte de soja, la céréale d'hiver est généralement immédiatement semée puis récoltée en juin et juillet (mis en évidence par les faibles valeurs de NDVI après le jour 161).

- Les profils temporels de NDVI du troisième groupe (Figure 36(c)) semblent représentatifs du système de double culture riz-soja. Les faibles valeurs de NDVI courant mai (du jour 113 au jour 145) suivies d'une augmentation jusqu'à la deuxième période de pic de NDVI autour de juillet (du jour 193 au jour 225), révèlent le deuxième cycle de culture de soja tardif, qui est unique au système de double culture riz-soja.

Les 22 unités de paysage appartenant au « domaine cultivé » ont été donc classées selon la correspondance entre leurs profils temporels et le calendrier cultural de l'un des trois systèmes de cultures identifiés. La carte des systèmes d'utilisation agricole des terres du Tocantins, à 250 m de résolution, avec la classification finale des unités de paysage est présentée dans la Figure 37.

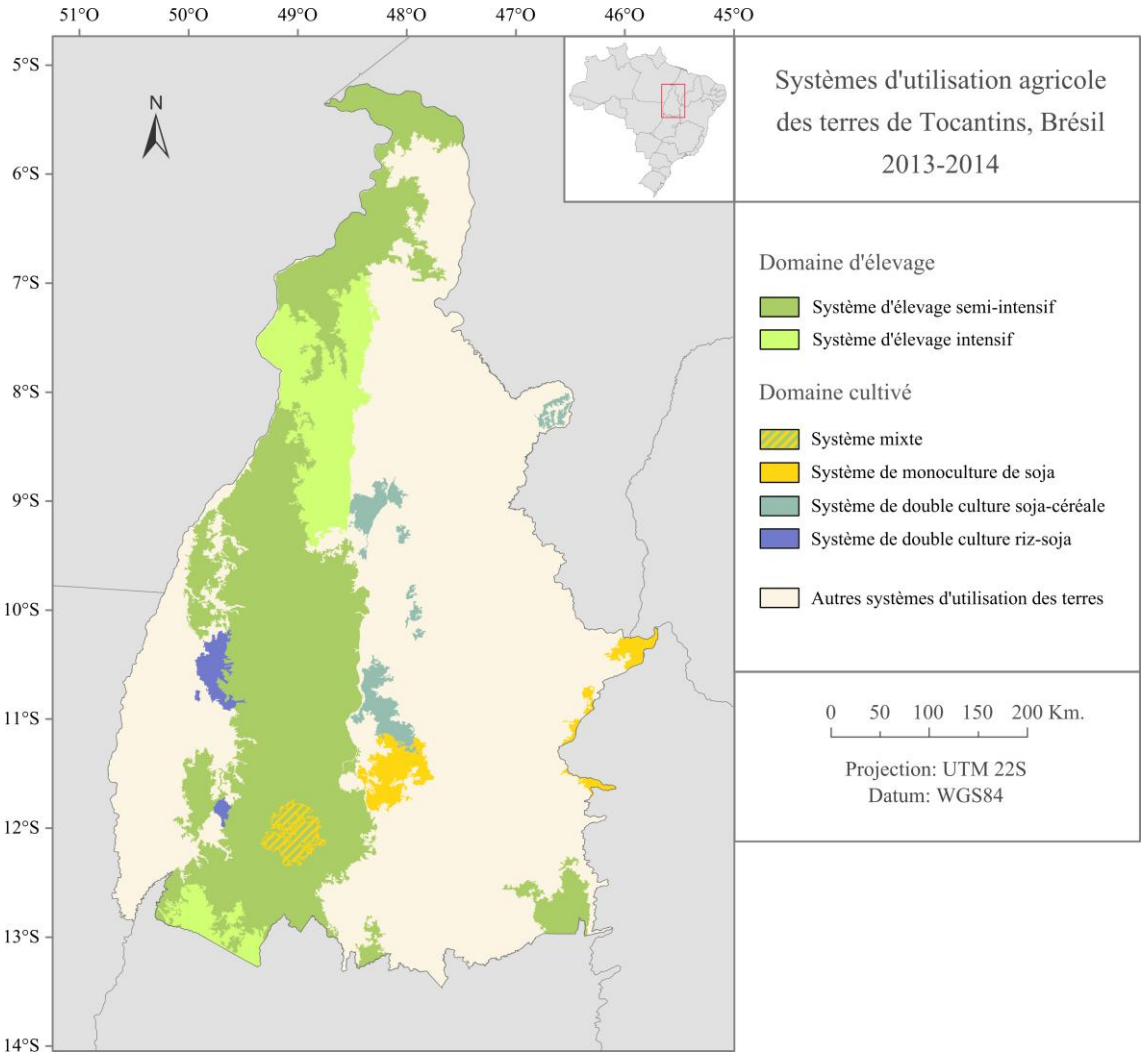


Figure 37. Carte des 6 principaux systèmes d'utilisation agricole des terres de l'État du Tocantins (saison culturale 2013-2014).

Comme l'on peut l'apprécier dans la Figure 37, les systèmes d'élevage sont plus extensifs que les systèmes de culture, et sont principalement répartis le long d'un transept longitudinal situé au centre-ouest de l'État. Les systèmes de culture sont concentrés dans des zones localisées qui sont réparties de manière éparse sur différentes parties de l'État, à l'exception du tiers nord, où les systèmes d'élevage prédominent. Les activités d'élevage et d'agriculture sont, pour la plupart de cas, séparées spatialement, à l'exception de la seule zone de systèmes mixtes située dans la région sud du transept longitudinal dominé autrement par les activités d'élevage.

3.2.4. Résultats de l'évaluation de la classification

Les statistiques agricoles au niveau des municipalités ont révélé trois cultures prédominantes qui représentent plus de 94 % de la superficie totale cultivée : le soja (75 %), le riz (11,1 %) et le maïs de deuxième cycle (*milho-safrinha* en portugais) (8,6%). Toutes les

municipalités avec plus de 2000 hectares de surfaces cultivées, dédient une partie importante des surfaces à la production de soja, celle-ci représentant entre 29 % et 100 % de leur surface cultivée totale, ce qui confirme la prédominance de ce type de culture au Tocantins.

Cependant, trois types de municipalité distincts ont été observés. Un premier type, composé de 15 municipalités, où la surface cultivée est presque exclusivement consacrée au soja, a été donc associé au système de monoculture de soja. Un deuxième type, composé de 7 municipalités, où la proportion de maïs de deuxième cycle est considérablement représentée, a été associé au système de double culture soja-céréale. Enfin, un troisième type de municipalité, composé de 3 municipalités, où la proportion de riz est presque équivalente à la proportion de soja, était liée au système de double culture riz-soja. La proportion moyenne des différents types de cultures pour chaque type de municipalité est présentée dans la Figure 38.

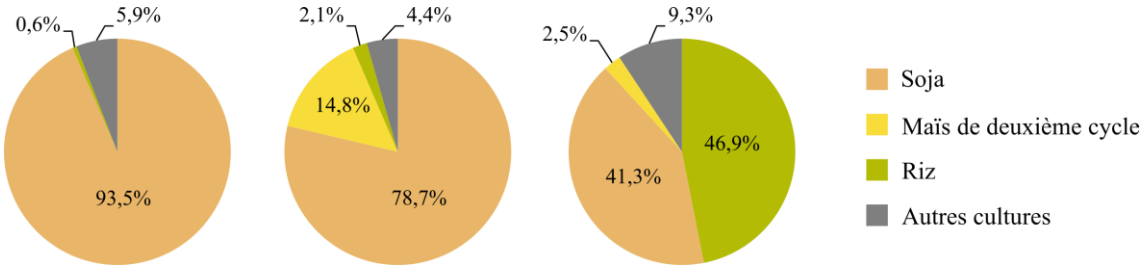


Figure 38. Proportion moyenne des cultures dans les municipalités du premier type (gauche), deuxième type (centre) et troisième type (droite), associés respectivement aux systèmes de monoculture de soja, double culture soja-céréale et double culture riz-soja.

La typologie de systèmes de culture par municipalité est ensuite cartographiée (Figure 39(a)). La carte de référence ainsi produite est comparée à la carte de systèmes d'utilisation des terres produite par télédétection dans la Figure 39(b).

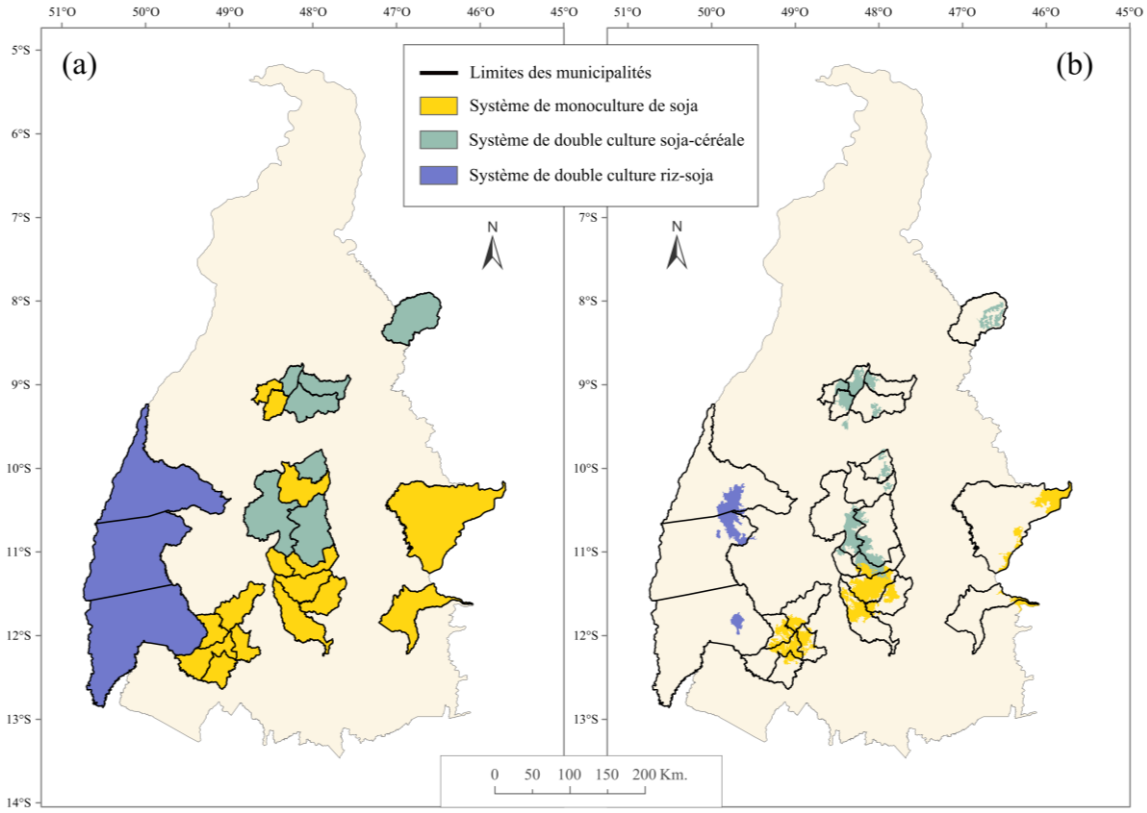


Figure 39. Répartition spatiale des trois principaux systèmes de culture dans la carte de référence au niveau des municipalités dérivée des statistiques agricoles (a) et dans la classification réalisée par télédétection (b). Les limites des municipalités sont incluses dans les deux cartes pour faciliter leur comparaison.

Bien que la carte de référence et la classification par télédétection soient produites à différentes échelles, on observe une bonne cohérence spatiale des systèmes de culture, notamment des systèmes de monoculture de soja et de double culture riz-soja. La répartition spatiale du système de double culture soja-céréale semble être surestimée localement pour certaines municipalités qui sont classées en système de monoculture de soja.

3.3. Discussion

Nous avons réalisé une stratification de la région d'étude en unités de paysage homogènes en termes de variation temporelle de la végétation. Les variations temporelles de la végétation étant, pour certaines unités, liées à l'utilisation agricole des terres, nous avons pu produire une carte de systèmes d'utilisation agricole des terres à l'échelle régionale.

Deux systèmes d'élevage ont pu être différenciés par la proportion de superficie occupée par des pâturages et les trois principaux systèmes de culture ont pu être différenciés par la proportion de superficie occupée par des cultures annuelles et les profils temporels de NDVI associés à chaque unité de paysage. Ceci a permis non seulement de caractériser les unités de

paysage agricole en termes de types de culture mais aussi par les pratiques culturales associées, notamment le mode de culture (ex. monoculture ou culture séquentielle).

Les résultats ont montré une cohérence spatiale générale entre les trois principaux systèmes de culture de la classification par télédétection et la carte de référence, à l'exception de quelques surestimations localisées de la présence du système de double culture de soja-céréale sur des municipalités associées au système de monoculture de soja dans la carte de référence. Cette surestimation pourrait être due à l'exclusion des cultures de couverture d'été non commerciales telles que le mil, le sorgho ou le braquiaria dans les statistiques agricoles et donc dans la carte de référence. Ces cultures de couverture sont parfois semées après la récolte de soja au lieu du maïs dans le cadre de la pratique de semis-direct largement adoptée dans la région. Ces cultures de couverture pourraient contribuer au deuxième pic de NDVI qui apparaît dans les profils temporels de NDVI des unités de paysage qui recoupent partiellement les municipalités où seule la superficie du soja a été officiellement recensée.

La carte de référence utilisée pour l'évaluation des résultats de la classification montre que des cartes des systèmes d'utilisation des terres peuvent être produites à partir de statistiques agricoles. Cependant, la résolution thématique de ce type de carte est limitée par les données disponibles dans les statistiques, ne tenant pas compte de certains systèmes (par exemple les systèmes d'élevage) et la précision spatiale est fortement limitée par le niveau d'agrégation des données), les limites administratives n'étant pas représentatives de la distribution spatiale réelle des systèmes d'utilisation des terres. L'approche par télédétection présentée produit une carte des systèmes d'utilisation agricole des terres qui surmonte ces limites, en traçant automatiquement des unités avec des limites qui représentent des véritables séparations physiques liées aux variations temporelles de la végétation.

Bien que la capacité de l'ACP pour découvrir des événements de changements intra-annuels significatifs à partir de séries temporelles de NDVI ait été prouvée précédemment (ex. Gurgel et Ferreira, 2003; Hall-Beyer, 2003), les études concentrent leurs analyses généralement sur les CP qui capturent la plus grande proportion de la variance de la série, généralement les trois premières, en écartant le reste de CP. En revanche, toutes les images CP d'ordre supérieur ont été incluses dans notre approche, car elles captent une proportion mineure de la variance totale de la série qui peut être associée aux variations de valeurs localisées spatialement et temporellement, telles que les variations intra-saisonnières et les changements rapides induits par les pratiques agricoles.

Dans l'ensemble, les résultats confirment que les images CP d'ordre supérieur (qui contiennent que 27 % de la variance totale) ont saisi l'information relative à la variabilité temporelle de la couverture végétale, qui était la partie essentielle de l'information de la série temporelle de NDVI nécessaire pour la stratification de la région d'étude en unités de paysage avec une phénologie relativement homogène. De plus, les profils temporels de NDVI présentés dans la Figure 36 reflètent les variations intra-saisonnières liées aux pratiques culturales telles que les dates de plantation et de récolte des différentes cultures.

L'extraction automatique d'unités de paysage avec une phénologie relativement homogène à travers d'une ACP et une segmentation automatique est une approche simple et innovante qui peut être reproduite dans d'autres régions. Cette approche surmonte la propagation d'erreurs potentielles des méthodes basées sur la délimitation des unités par des experts et le choix subjectif de variables hétérogènes des approches traditionnelles de cartographie des systèmes d'utilisation des terres. De plus, l'approche présentée est indépendante de l'échelle et permet donc une analyse multi-niveau, ce qui pourrait aider à améliorer la compréhension des systèmes d'utilisation des terres en les caractérisant à différents niveaux organisationnels.

En outre, l'ensemble de règles de l'arbre de décision utilisé pour la classification de quatre grands systèmes d'utilisation agricole des terres (présenté dans la Figure 31), adapté de de Almeida *et al.* (2016), est conçu avec des règles de seuil simples, basées sur le pourcentage d'occupation des terres agricoles, qui peut être facilement adapté ou directement appliqué, à d'autres régions. Cette classification dépend néanmoins de l'existence d'une carte de l'occupation des terres. Même si ces cartes sont maintenant produites et disponibles pour de nombreuses régions, leur précision conditionne les résultats des classifications, qui en plus dépendraient entièrement de leur mise à jour régulière et leur mise à disposition.

Par conséquent, l'approche présentée bénéficierait de l'inclusion d'une méthode de classification automatique entièrement basée sur des données satellitaires pour la cartographie des systèmes agricoles à l'échelle de la parcelle. Cette méthode correspond ainsi à la deuxième étape de l'approche générale (section 1.5) et est présentée en détail dans le prochain chapitre.

Bilan : Stratification et caractérisation des systèmes d'utilisation agricole des terres à l'échelle régionale

- Ce chapitre présente une approche simple et potentiellement reproductible pour la cartographie des systèmes d'utilisation agricole des terres à l'échelle régionale basée sur une approche objet (OBIA) et l'analyse d'une série annuelle de NDVI.
- L'approche présentée se décompose en deux étapes principales : (i) la stratification de la région d'étude en unités de paysage relativement homogènes en termes de variations temporelles de la végétation, réalisée à l'aide d'une ACP et d'une segmentation, et (ii) la caractérisation des unités de paysage à l'aide d'une classification par règles de décision à partir d'une carte d'occupation des terres et l'analyse des profils temporels de NDVI associés à l'utilisation agricole des terres dans la région d'étude.
- Une carte des principaux systèmes d'utilisation agricole des terres du Tocantins a été produite avec l'approche développée pour la saison culturale de 2013-2014. On observe une bonne cohérence spatiale entre les systèmes de culture ainsi cartographiés, et ceux de la carte de référence produite à partir des statistiques agricoles officielles.

Chapitre 4.
Classification des systèmes de culture à l'échelle régionale par la méthode « landscape-clustering »

4.1. MÉTHODE ADOPTÉE 111
4.1.1. Segmentation 112
4.1.1.1. Délimitation des parcelles 112
4.1.1.2. Stratification régionale en unités de paysage 113
4.1.2. Classification non supervisée 115
4.1.2.1. Méthode hyperclustering 116
4.1.2.2. Méthode « landscape-clustering » 116
4.1.2.3. Affectation de classes aux clusters 117
4.1.3. Évaluation des résultats de classification 117
4.2. RÉSULTATS 118
4.2.1. Résultats de classification 118
4.2.2. Résultats d'évaluation de la précision 120
4.2.3. Carte des systèmes de culture obtenue par « landscape-clustering » 123
4.3. DISCUSSION 124

Chapitre 4. Classification des systèmes de culture à l'échelle régionale par la méthode «landscape-clustering»

Les cartes de systèmes de culture, produites à l'échelle de la parcelle, représentent (comme signalé dans la section 1.2.4) la donnée de base à partir de laquelle on peut extraire de nombreux indicateurs tels que la superficie occupée par différents types de culture, l'extension d'une pratique agricole déterminée et des indicateurs liés à la production agricole et aux impacts environnementaux associés. Cependant, cet exercice s'avère difficile en raison de la grande variabilité spectrale liée à la grande diversité spatiale d'utilisation des terres et la forte variabilité de conditions environnementales lorsque l'on considère une région extensive.

Ainsi, dans ce quatrième chapitre on s'attache à traiter la question suivante (présentée dans la section 1.4) :

- Comment aborder la forte variabilité spectrale associée à la grande diversité de surfaces et d'utilisation des terres au sein d'une région pour cartographier les espèces cultivées et les pratiques agricoles ?

D'un point de vue méthodologique on s'intéresse plus particulièrement aux stratégies qui vont permettre d'y parvenir sans avoir recours ni à des données externes (y compris les données de terrain), ni à des étapes de classification guidées par des connaissances expertes, pour pouvoir opérationnaliser autant que possible les méthodes et faciliter ainsi leur reproductibilité. Une piste présentée dans le chapitre introductif pour aborder ce défi méthodologique porte sur l'inclusion d'étapes de réduction de la variabilité spatiale au préalable des classifications (cf. section 1.3.2.3). Cette piste repose sur l'hypothèse suivante :

- L'utilisation d'une stratification au préalable d'une classification non supervisée améliore la précision des cartes de systèmes de culture à l'échelle régionale en réduisant la variabilité spatiale et spectrale des données satellitaires.

Nous présentons dans ce chapitre des nouveaux développements méthodologiques qui explorent cette piste et testent l'hypothèse en comparant la précision des résultats de deux variantes d'une méthode de classification non supervisée : une méthode hyperclustering, et une méthode dite « landscape-clustering » qui comporte une stratification en préalable à la classification.

Le travail présenté dans ce chapitre a été valorisé par une publication dans une revue internationale de rang A en 2018 (présentée en annexe) :

> Bellón, B., Bégué, A., Lo Seen, D., Lebourgeois, V., Evangelista, B. A., Simões, M., et Demonte Ferraz, R.P. (2018). "Improved regional-scale Brazilian cropping systems' mapping based on a semi-automatic object-based clustering approach". *International Journal of Applied Earth Observations and Geoinformation*, 68C, 127-138. http://doi.org/10.1016/j.jag.2018.01.019

4.1. Méthode adoptée

La méthode adoptée dans ce chapitre se décompose en deux étapes : (i) une étape de segmentation à l'échelle de la parcelle à partir d'une mosaïque Landsat à haute résolution spatiale et de délimitation des unités de paysage par la méthode présentée dans le chapitre précédent (section 3.1.1) et (ii) une étape de classification non supervisée des segments à l'échelle de la parcelle par une méthode hyperclustering et par une méthode « landscape-clustering » à partir des informations d'une série annuelle MODIS-NDVI.

La Figure 40 présente un diagramme de flux avec les principales étapes méthodologiques décrites dans les sous-sections suivantes. La donnée en entrée pour la délimitation des parcelles correspond à la mosaïque Landsat 8 (présentée dans la section 2.2.1.2) et pour la stratification, la série annuelle de NDVI de la saison 2014-2015 sur l'État du Tocantins (présentée dans la section 2.2.1.1). Pour éviter la confusion, les segments issus de la délimitation des parcelles sont désignés dans ce chapitre comme « les objets », et les segments issus de la stratification sont désignés comme « les strates ».

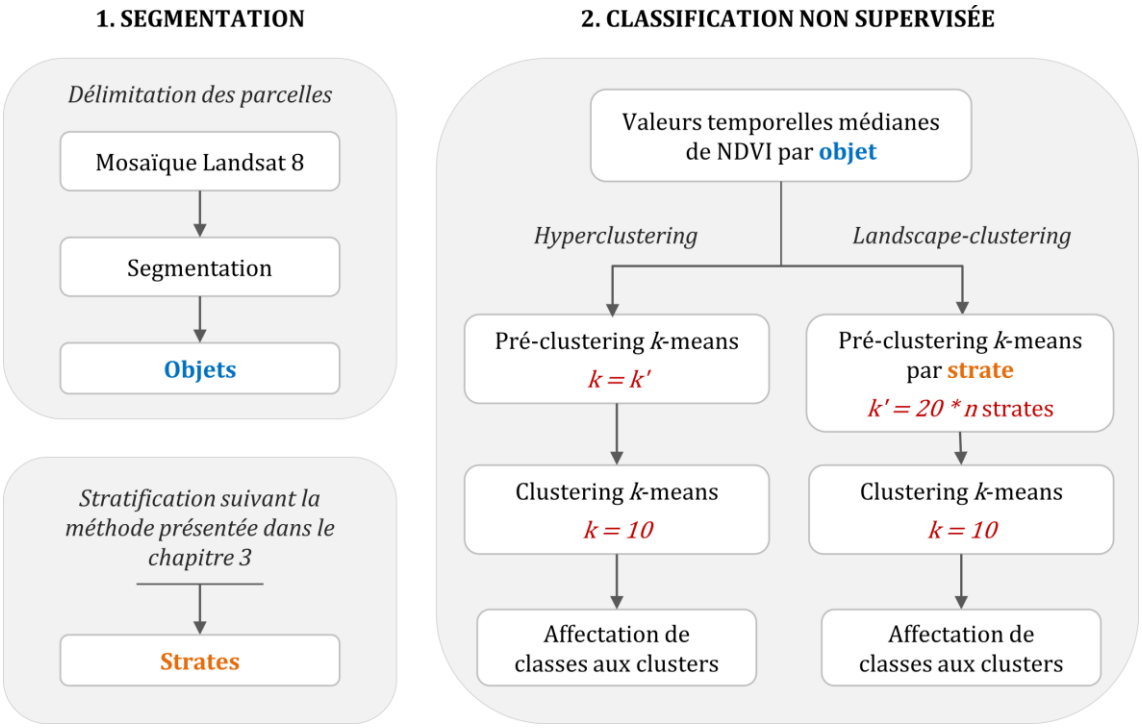


Figure 40. Représentation des étapes méthodologiques présentées dans le chapitre 4.

4.1.1. Segmentation

4.1.1.1. Délimitation des parcelles

Nous avons segmenté la mosaïque Landsat 8, en utilisant l'algorithme de segmentation multi-résolution présenté dans la section 2.3.3, pour délimiter les parcelles agricoles et d'autres éléments de paysage. Les cinq bandes spectrales de la mosaïque ont été pondérées de manière égale et les paramètres de segmentation ont été empiriquement fixés après des tests séquentiels de différentes combinaisons de paramètres et une évaluation visuelle des résultats de chaque combinaison. L'objectif de la segmentation étant la délimitation des parcelles agricoles, qui sont généralement de forme rectangulaire, nous avons fortement pondéré en faveur du paramètre de « forme » par rapport à la radiométrie (paramètre de « couleur »), et en faveur du paramètre de « compacité » par rapport au paramètre de « sinuosité » (les valeurs des paramètres de segmentation retenues sont présentées dans le Tableau 7).

Les objets qui en résultent, présentent des formes compactes et relativement rectangulaires qui délimitent efficacement les parcelles agricoles de la région d'étude (la Figure 41 présente des échantillons des résultats de la segmentation sur différents paysages agricoles). Un total de 187 877 objets à une résolution spatiale de 30 m a été obtenu sur l'ensemble de la région d'étude.

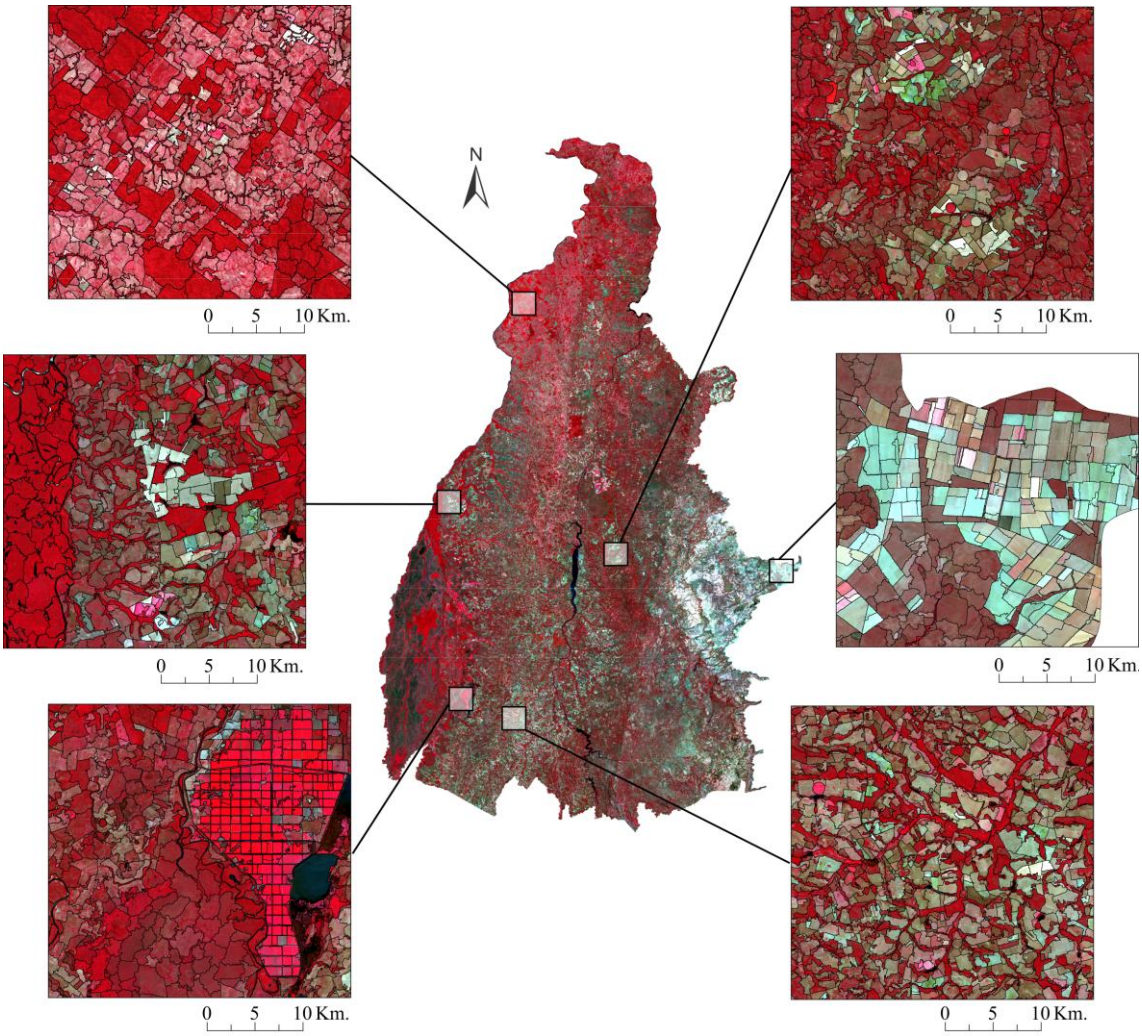


Figure 41. Échantillons des résultats de la segmentation de la mosaïque Landsat 8 sur différents paysages agricoles. L'image de fond correspond à une composition colorée en fausses couleurs (RGB proche-infrarouge, vert, rouge) de la mosaïque Landsat (source : image de fond (USGS, s.d.)).

4.1.1.2. Stratification régionale en unités de paysage

La stratification a été réalisée suivant la méthode introduite dans le chapitre précédent (section 3.1.1) à partir de la série annuelle d'images de NDVI du capteur MODIS couvrant l'année culturale 2014-2015.

La méthode de stratification de la région d'étude en unités de paysage, utilisée dans le chapitre précédent pour délimiter les unités cartographiques de la carte de systèmes d'utilisation agricole des terres, est cette fois-ci utilisée pour réduire la variabilité spectrale de la série temporelle de NDVI dans le cadre de la méthode de classification dite « landscape-clustering » présentée ultérieurement (section 4.1.2.2).

Les « variables phénologiques » issues de l'ACP ont été calculées pour la nouvelle série temporelle de NDVI et les paramètres de segmentation originaux ont été conservés (cf. Tableau 7).

Tableau 7. Tableau récapitulatif des paramètres de segmentation utilisés pour la délimitation des parcelles et la stratification régionale en unités de paysage. *IR = infrarouge.

Type de segmentation	Données	Résolution spatiale	Bandes spectrales	Paramètre d'échelle	Paramètre de couleur	Paramètre de forme
Délimitation des parcelles	Mosaïque Landsat 8	30 m	Bleue (B2) Verte (B3) Rouge (B4) Proche IR* (B5) Moyen IR (B6)	110	0,2	0,8 (compacité =1 ; sinuosité=0)
Stratification en unités de paysage	Images CP calculées à partir de la série annuelle de NDVI de 2014-2015	250 m	22 CP d'ordre supérieur	850	1	0

Un total de 120 unités de paysage, présentées dans la Figure 42, a été obtenu à partir de la stratification, avec une superficie moyenne de 2 318 km^2, 19 km^2 pour la plus petite et 15 052 km^2 pour la plus large.

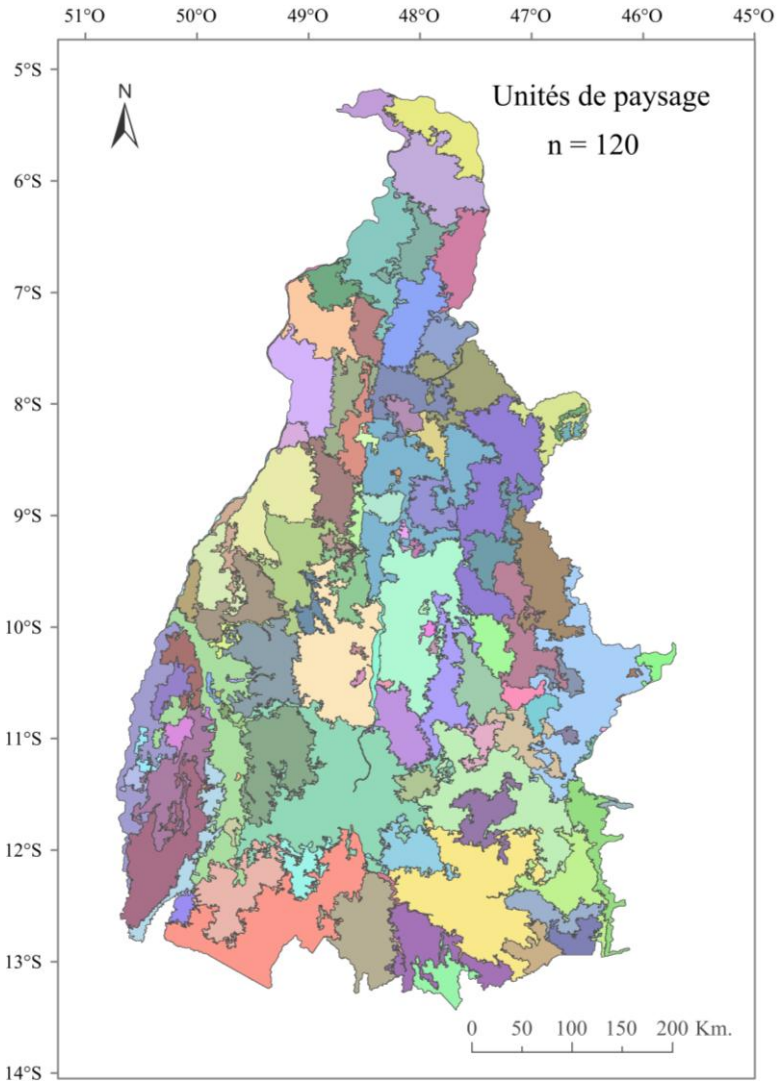


Figure 42. Résultat de la stratification régionale en unités de paysage suivant la méthode décrite dans le chapitre 3.

4.1.2. Classification non supervisée

Les deux variantes de la méthode de classification non supervisée ont été basées sur les informations temporelles de NDVI contenues dans la série annuelle de NDVI. Un jeu de données commun a été généré en extrayant la valeur médiane de NDVI par objet pour chacune des 23 dates de la série annuelle. En raison de la différence entre la résolution spatiale des pixels MODIS à 250 m et les objets à 30 m, seules les valeurs des pixels MODIS dont le centre est situé à l'intérieur d'un objet sont utilisées pour le calcul de la valeur médiane NDVI par objet. La valeur médiane a été choisie comme mesure de tendance centrale car elle est moins sensible à l'asymétrie dans la distribution des valeurs que la moyenne, et donc moins affectée par l'influence des pixels mixtes situés en bordure des objets.

Les deux variantes de la méthode de classification non supervisée ont utilisé l'algorithme k-means pour partitionner le jeu de données d'objets (contenant les valeurs temporelles médianes de NDVI en deux temps) en deux étapes : (i) une première étape de « pré-clustering » a partitionné les objets en un grand nombre de sous-clusters (sous-groupes) initiaux, et (ii) une deuxième étape de clustering a regroupé les sous-clusters résultants dans le nombre désiré de clusters finaux.

4.1.2.1. Méthode hyperclustering

Dans la méthode hyperclustering, le nombre de sous-clusters k a été fixé à 2400 pour correspondre au nombre de sous-clusters k' de la méthode « landscape-clustering » présentée dans la section suivante (section 4.1.2.2).

Ainsi, 2400 profils temporels moyens de NDVI par sous-cluster[41] ont été produits et repartitionnés (regroupés) en dix clusters finaux. Le nombre de clusters finaux a été empiriquement fixé, en s'attendant à ce que dix clusters discriminent adéquatement les trois classes ciblées (i.e. les trois principaux systèmes de culture du Tocantins) d'autres types d'occupation et utilisation des terres, tout en restant facilement interprétables.

Même si la version de k-means proposé par Hartigan et Wong (1979) paraît être moins sensible à la partition aléatoire initiale des données que d'autres versions de l'algorithme, nous avons effectué dix répétitions du clustering final, et le résultat maximisant la dissimilitude entre clusters (présentant le rapport le plus élevé entre la somme des carrés intra-cluster et la somme totale des carrés (R^2)) et la similitude au sein des clusters (présentant le rapport le plus élevé entre la somme des carrés inter-cluster et la somme totale des carrés ($R^2/(1-R^2)$)) a été retenu.

4.1.2.2. Méthode « landscape-clustering »

Dans la méthode « landscape-clustering », les strates ont été d'abord superposées spatialement avec les objets, et tous les objets dont le centroïde était situé à l'intérieur des limites d'une strate donnée ont été regroupés, aboutissant à un groupe d'objets par strate. Un pré-clustering a été ensuite effectué sur chaque groupe d'objets, faisant autant de partitions que de nombre strates (n = 120). Le nombre exact de classes spectrales par strate étant inconnu, le nombre de sous-groupes par groupe d'objets a été empiriquement fixé à 20, ce qui

[41] Chaque profil temporel moyen de NDVI par sous-cluster correspond aux valeurs moyennes de NDVI pour les 23 dates, calculées à partir de tous les objets groupés dans le même sous-cluster.

représente un total de 2400 sous-clusters pour toute la région d'étude (k'= 20 sous-clusters × 120 strates).

Le profil temporel moyen de NDVI par sous-cluster a ensuite été calculé, résultant en un total de 2400 profils temporels qui ont été repartitionnés en dix clusters finaux. Les résultats de dix répétitions du clustering final ont été évalués et le résultat avec les rapports R^2 et $R^2/(1-R^2)$ les plus élevés a été retenu.

4.1.2.3. Affectation de classes aux clusters

Les profils temporels moyens de NDVI des dix clusters finaux de chaque méthode de clustering ont été calculés et tracés. Comme cela a été mis en évidence dans le chapitre précédent, en général les profils temporels de NDVI associés aux systèmes de culture présentent des augmentations et des diminutions de valeurs relativement rapides liées aux périodes de croissance des cultures et des périodes de récolte, respectivement. La forme des profils nous a permis d'identifier et d'affecter une classe aux clusters finaux par analyse visuelle de leurs variations de NDVI au cours de la saison culturale.

Ainsi, les clusters dont la forme du profil temporel de NDVI suit le calendrier cultural des principaux systèmes de culture du Tocantins sont affectés à leur classe correspondante, tandis que ceux qui ne présentent pas d'oscillations remarquables sont affectés à la classe « autre type d'occupation et utilisation des terres ».

4.1.3. Évaluation des résultats de classification

La précision du résultat de classification de la méthode hyperclustering a été estimée et comparée à celle de la méthode « landscape-clustering », en utilisant six indices d'évaluation de classifications (présentés dans la section 2.3.5) calculés à partir des données de vérité terrain (présentés dans la section 2.2.3.1).

En plus de ces indices nous avons calculé la corrélation entre les résultats de classification et les estimations faites par les statistiques agricoles des superficies des principales cultures annuelles récoltées en 2015 par municipalité (présentées dans la section 2.2.3.2) à l'aide des coefficients de corrélation (r^2) et de détermination linéaire (r) (présentés dans la section 2.3.2).

4.2. Résultats

4.2.1. Résultats de classification

Les profils temporels moyens de NDVI des dix clusters finaux sont présentés dans la Figure 43 pour la méthode hyperclustering (Figure 43(a)) et la méthode « landscape-clustering » (Figure 43(b)).

Trois profils temporels représentatifs des principaux systèmes de culture du Tocantins peuvent être identifiés parmi l'ensemble des profils. Ces profils sont particulièrement nets pour un des clusters obtenus avec la méthode hyperclustering (Figure 43(a)(iii)), et pour trois clusters obtenus avec la méthode « landscape-clustering » (Figure 43(b)(i)(ii)(iii)), grâce aux variations remarquables des valeurs de NDVI.

Parmi les profils des clusters obtenus avec la méthode hyperclustering, deux profils présentent de légères oscillations des valeurs de NDVI (Figure 43(a)(i)(ii)), ce qui suggère que ces clusters ont capturé les variations temporelles liées aux systèmes de culture et peuvent donc être associés à leur classe correspondante, même si la platitude générale de leur forme suggère un mélange de classes dans ces deux clusters.

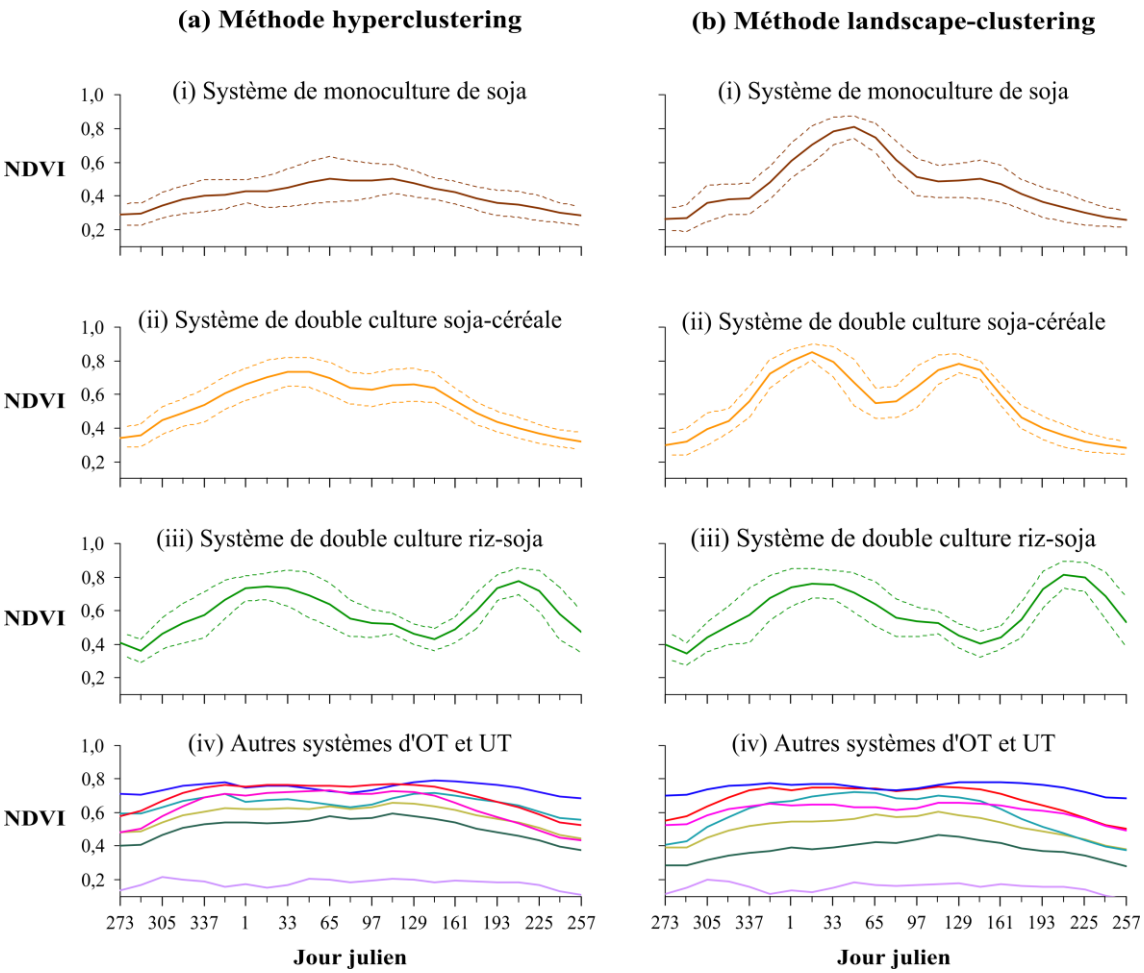


Figure 43. Profils temporels des dix clusters finaux obtenus avec (a) la méthode hyperclustering et (b) la méthode « landscape-clustering », et les classes des systèmes de culture correspondantes. Les courbes continues représentent le NDVI moyen de tous les objets appartenant à chaque cluster en fonction du jour julien. Les profils des clusters qui semblent avoir capté les trois principaux systèmes de culture du Tocantins sont représentés dans des figures séparées (i), (ii) et (iii) avec leur écart-type (courbes discontinues), et les profils des clusters qui semblent avoir capté d'autres types d'occupation et utilisation des terres (OT et UT) sont représentés dans la même figure (iv).

La forme du profil temporel de NDVI de la Figure 43(b)(i) montre un seul cycle de culture significatif qui présente un pic des valeurs autour de février, suggérant que le cluster correspondant a regroupé les objets appartenant au système de monoculture de soja. La forme de profil de la Figure 43(b)(ii) semble représentative du système de double culture soja-céréale, les valeurs de NDVI étant faibles courant mars reflétant la récolte de la culture de soja du premier cycle, après quoi le NDVI augmente suivant la croissance de la culture céréalière du deuxième cycle jusqu'à qu'elle atteigne son maximum vers le mois de mai. Enfin, la forme du profil de la Figure 43(b)(iii) semble représenter le système de double culture riz-soja, avec un long cycle cultural pendant la saison de pluies et un cycle court plus

tardif, de mai à septembre. Ce dernier cycle cultural du soja est mis en évidence par la période du pic de NDVI autour de juillet.

4.2.2. Résultats d'évaluation de la précision

Le Tableau 8 présente les estimations de la précision globale et son intervalle de confiance, du coefficient Kappa de Cohen et du désaccord sur la quantité et sur la répartition des résultats de classification en quatre classes (les trois principaux systèmes de culture et la classe « autre type d'occupation et utilisation des terres »), issus des méthodes hyperclustering et « landscape-clustering ».

Tableau 8. Résultats de la précision globale de la classification et l'intervalle de confiance associé, le coefficient global de Kappa, et le désaccord sur la quantité et la répartition des objets, calculés à partir des 900 points GPS de vérité terrain pour chaque méthode de classification.

Méthode de classification	Précision globale (%)	Intervalle de confiance 95 % (%)	Kappa	Désaccord sur la quantité (%)	Désaccord sur la répartition (%)
Hyperclustering	93,2	91,4 - 94,8	0,81	4,5	2,3
Landscape-clustering	77,9	75,0 – 80,6	0,52	15,3	6,8

La précision globale de la classification de la méthode « landscape-clustering » (93,2 %) est significativement plus élevée que celle de la méthode hyperclustering (77,9 %), i.e. la méthode « landscape-clustering » a classé correctement un plus grand nombre d'objets que la méthode hyperclustering. Par conséquent, le désaccord global (composé du désaccord sur la quantité et du désaccord sur la répartition) est plus faible pour la classification par « landscape-clustering » (6,8%) que pour la classification par hyperclustering (22,1 %). Le désaccord sur la quantité est plus élevé que le désaccord sur la répartition pour les deux résultats de classification. Ce résultat indique que les erreurs de classification sont davantage liées à un déséquilibre dans la proportion d'objets de référence et d'objets classés pour chaque classe (ex. pour une classe donnée le nombre d'objets classés est plus grand ou plus petit que le nombre d'objets de référence pour cette classe) qu'à une erreur de répartition spatiale des classes par rapport à la répartition spatiale des classes dans les objets de référence.

Le résultat du coefficient Kappa de la classification par « landscape-clustering » est également plus élevé que celui de la classification par hyperclustering, i.e. la différence entre l'accord observé entre les objets classés et les points de référence et la probabilité d'accord par hasard est plus élevée que pour la classification par hyperclustering, ce qui suggère que la fiabilité de la classification par « landscape-clustering » est plus élevée.

De plus, les résultats du χ^2 de McNemar (présentés dans le Tableau 9) ont montré une amélioration statistiquement significative des résultats de la méthode « landscape-clustering » par rapport aux résultats de la méthode hyperclustering.

Tableau 9. Tableau de contingence résumant les résultats des deux méthodes de classification (nombre d'objets correctement classés ou mal classés) et résultats du test du χ^2 de McNemar.

	Landscape-clustering (correctement classés)	Landscape-clustering (mal classés)	
Hyperclustering (correctement classés)	684	17	$\chi^2 = 110{,}72$ avec 1 degré de liberté (*valeur-p* < 0.0001)
Hyperclustering (mal classés)	155	44	

La précision de classification pour chaque classe est estimée avec la *F*-mesure pour les classifications par hyperclustering et par « landscape-clustering » (Tableau 10). La classe « système de monoculture de soja » obtient la plus faible valeur de *F*-mesure dans les deux cas (*F*-mesure < 0,7) et la classe « système de double culture riz-soja » est correctement classée par les deux méthodes, obtenant la valeur la plus élevée de *F* -mesure (*F*-mesure = 1). Les deux autres classes « système de double culture soja-céréale » et « autre type d'occupation et utilisation des terres » obtiennent des valeurs de *F*-mesure élevées (*F*-mesure > 0,8) pour la classification par « landscape-clustering ». Pour la classification par hyperclustering, les classes « système de monoculture de soja » et « système de double culture soja-céréale » présentent les plus faibles scores (respectivement 0,46 et 0,59).

Ainsi, à l'exception de la classe « système de double culture riz-soja » qui a obtenu la même valeur de *F*-mesure pour les deux classifications, les résultats indiquent une meilleure performance de la méthode « landscape-clustering ».

Tableau 10. Résultats de la *F*-mesure pour les deux méthodes de classification.

Classe	Hyperclustering	Landscape-clustering
Système de monoculture de soja	0,46	0,65
Système de double culture soja-céréale	0,59	0,81
Système de double culture riz-soja	1,00	1,00
Autre type d'occupation et utilisation des terres	0,85	0,96

Enfin, la régression linéaire simple entre les estimations de la superficie de soja récoltée en 2015 par municipalité, issues des statistiques agricoles, et la superficie occupée par les objets affectés aux trois classes de systèmes de culture du résultat de classification de la méthode « landscape-clustering » par municipalité est présentée dans la Figure 44.

La superficie totale de soja (en considérant les trois systèmes de culture) estimée par la méthode « landscape-clustering » représente 73 % de la superficie totale déclarée par l'IBGE, obtenant une forte corrélation positive (r) de 0,87 (r^2= 0,75) à l'échelle des municipalités.

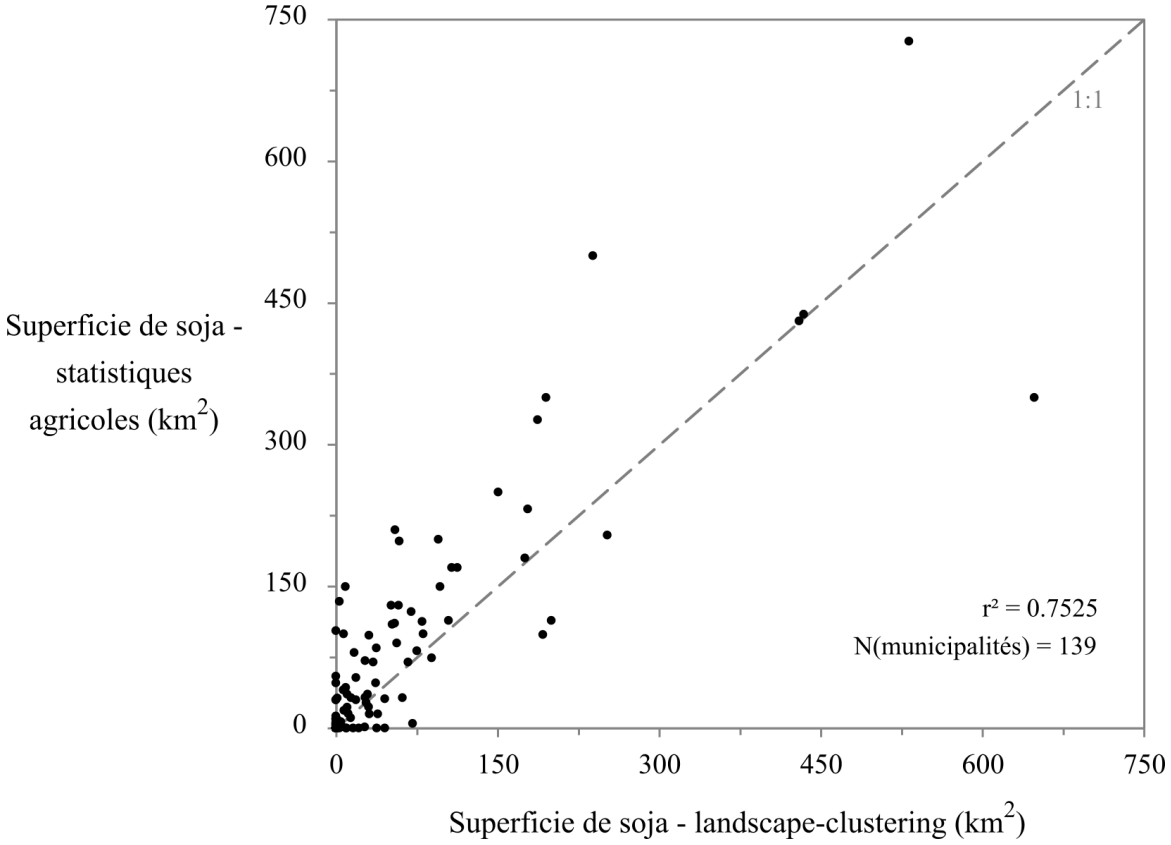


Figure 44. Régression linéaire simple entre les estimations de la superficie de soja récoltée en 2015 par municipalité, issues des statistiques agricoles, et la superficie occupée par les objets affectés aux trois classes de systèmes de culture du résultat de classification de la méthode « landscape-clustering » par municipalité.

De même, les estimations des superficies récoltées de maïs de deuxième cycle montrent une forte corrélation avec les résultats de la classification du système de double culture soja-céréale du résultat de classification de la méthode « landscape-clustering », avec un coefficient de corrélation à l'échelle des municipalités de 0,87 (r^2 = 0,76). Enfin, les estimations des superficies récoltées de riz montrent également une forte corrélation avec la classification du système de double culture riz-soja, avec un coefficient de corrélation de 0,99 (r^2 = 0,97).

4.2.3. Carte des systèmes de culture obtenue par « landscape-clustering »

La Figure 45 présente la carte obtenue à 30 m de résolution spatiale des principaux systèmes de culture du Tocantins à l'échelle de la parcelle obtenue avec la méthode de classification « landscape-clustering », avec une précision globale de 93,2 % (Tableau 8).

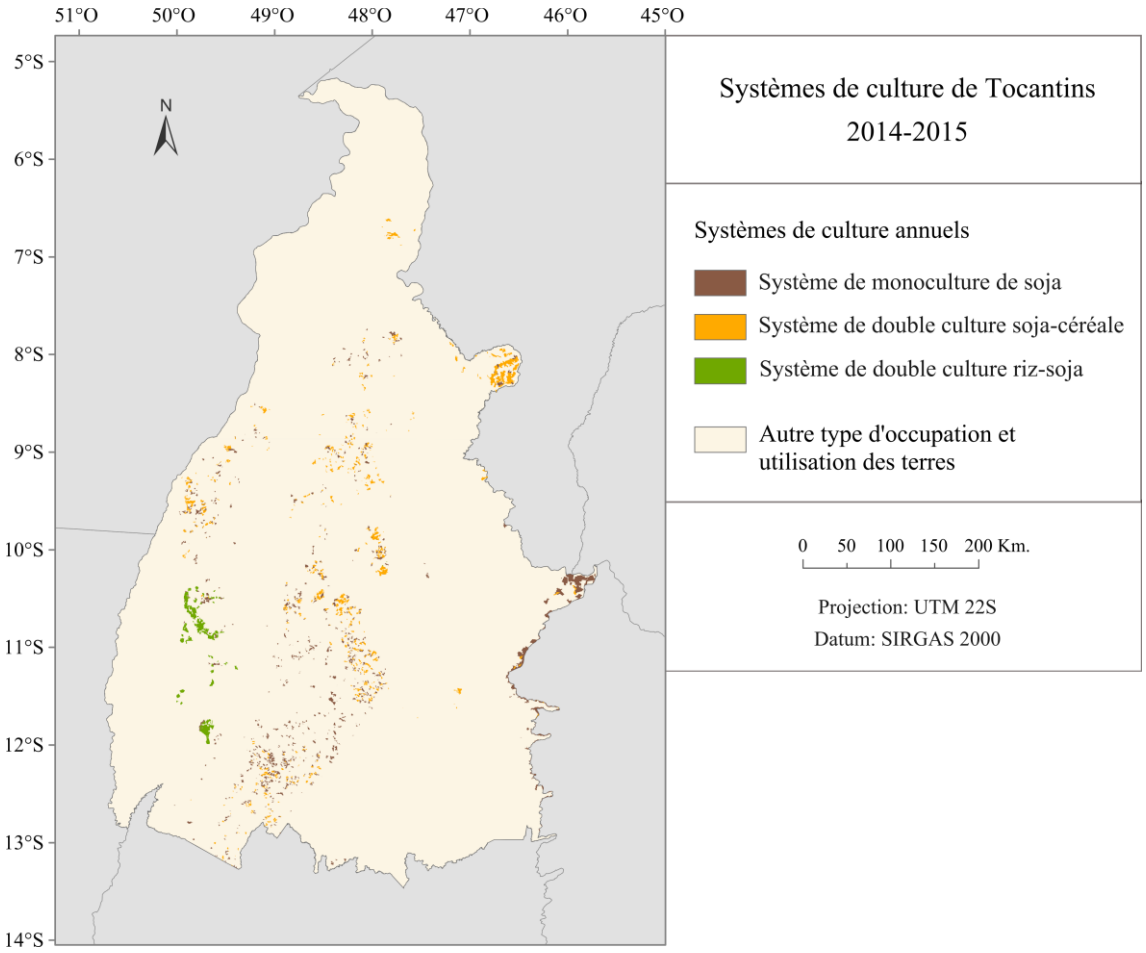


Figure 45. Carte des principaux systèmes de culture du Tocantins à l'échelle de la parcelle pour la saison culturale de 2014-2015 obtenue avec la méthode de classification non supervisée « landscape-clustering ».

Comme le montre la carte, le système de double culture riz-soja est exclusivement localisé dans le sud-ouest du Tocantins, tandis que les systèmes de monoculture de soja et double culture soja-céréale présentent une distribution éparse sur le reste de l'État, avec cependant deux zones de forte concentration à la frontière de l'État (respectivement dans le sud-est et le nord-est).

4.3. Discussion

Globalement, les résultats ont montré que la nouvelle méthode de classification non supervisée orientée-objet, basée exclusivement sur des données satellitaires, peut être utilisée pour produire des cartes de systèmes de culture (à l'échelle de la parcelle), décrits par le type de culture et le mode de culture, à l'échelle régionale.

Les résultats de l'évaluation de la précision ont montré que les deux variantes de la méthode ont produit de bonnes classifications, avec des précisions globales de plus de 75 %, même si la méthode de « landscape-clustering » a clairement surpassé la performance de la méthode hyperclustering. Ce constat est particulièrement évident dans les résultats des deux systèmes de culture spatialement dispersés, le système de monoculture de soja et le système de double culture de soja-céréale, la méthode de « landscape-clustering » obtenant une *F*-mesure d'environ 0,2 plus élevée que la méthode hyperclustering pour ces deux classes. Les coefficients Kappa suggèrent en outre une grande fiabilité de la classification par « landscape-clustering », qui est plus bas pour la classification par hyperclustering.

L'amélioration significative de la précision de la classification lors de l'utilisation d'une stratification au préalable de la classification mise en évidence par les résultats obtenus, est en cohérence avec les résultats de Cano *et al.* (2017). Cette amélioration suggère que le découpage de la région d'étude en sous-régions, dans ce cas par une stratification en unités de paysage à partir des « variables phénologiques » issues d'une série annuelle de NDVI, est une stratégie efficace de réduction de la variabilité spectrale associée à la variabilité spatiale des surfaces et de l'utilisation des terres pour améliorer la précision des classifications des systèmes de culture à l'échelle régionale. Ce résultat confirme ainsi l'hypothèse présentée en introduction de ce chapitre.

Cette stratégie de réduction de la variabilité spectrale semble particulièrement bénéfique aux systèmes de culture qui sont spatialement dispersés et sous-représentés par rapport à d'autres types d'occupation et d'utilisation des terres plus étendus puisqu'elle maximise la possibilité de les dissocier correctement avec le pré-clustering par strate (unité de paysage). Au contraire, la méthode hyperclustering sans stratification, même avec un nombre élevé de clusters initiaux (2400 dans ce cas), n'a pas évité le mélange des classes, rendant ainsi difficile l'affectation de classes aux clusters finaux et obtenant une plus faible précision de classification.

Le système de monoculture de soja a obtenu les précisions les plus faibles avec les deux variantes de la méthode (F -mesure de 0,65 et 0,46 pour la méthode de « landscape-clustering » et hyperclustering, respectivement). Ceci est principalement dû à la confusion avec les surfaces de pâturages qui ont un fort développement végétatif (avec un pic de NDVI) pendant la saison de pluies qui correspond également au pic de croissance du soja. En revanche, tous les points de terrain du système de double culture riz-soja ont été correctement classés par les deux variantes de la méthode. Cet excellent résultat peut être lié à la phénologie singulière de ce système de culture, avec deux cycles de culture distincts et temporairement éloignés, qui ne ressemble pas à la phénologie des autres systèmes d'utilisation des terres. Par ailleurs, la méthode de « landscape-clustering » n'a pas montré d'amélioration dans la classification de ce système de culture particulier, ce qui peut être expliqué par la localisation enclavée des parcelles associées à ce système dans la même microrégion (la microrégion Rio Formoso) qui présentent ainsi une variabilité spatiale limitée. La correcte localisation de ce système de culture pourrait en partie expliquer la plus faible valeur de désaccord sur la répartition que sur la quantité observée entre les classifications des deux méthodes et le jeu de données de référence.

La combinaison des deux étapes de réduction de la variabilité spectrale (segmentation à l'échelle de la parcelle et stratification de la région en unités de paysage) et une classification non supervisée en deux étapes semble être un moyen efficace de réduire la variabilité spectrale de l'ensemble de données sans avoir recours à des étapes de classification complexes, guidées par des connaissances expertes, ni à des données externes. La couverture d'une région d'une superficie proche de celle du Tocantins (277 621 km^2) avec des données à moyenne résolution spatiale (pixels MODIS à 250 m), représente un jeu de données initial d'environ 10 millions de pixels par image. Suivant la méthodologie présentée, cette dimensionnalité spatiale élevée a été successivement réduite à 187 877 objets, puis à 2400 pré-clusters, et finalement à 10 clusters. Outre les paramètres de segmentation définis par l'utilisateur et le nombre de clusters, les traitements sont entièrement automatiques et le faible nombre de clusters finaux permet d'affecter une classe par analyse visuelle des profils temporels de NDVI, tout en exigeant une connaissance de base du calendrier cultural pour les systèmes de culture d'intérêt.

La comparaison des résultats de la classification par « landscape-clustering » avec les estimations de l'IBGE concernant la superficie récoltée pour les cultures de soja, riz et maïs de deuxième cycle a montré des fortes corrélations positives à l'échelle des municipalités (r~0,9). La légère sous-estimation de la superficie de soja dans les résultats de classification est due à la confusion avec les surfaces de pâturages. Cette sous-estimation ainsi que certaines

surestimations sont responsables de la dispersion obtenue, ces dernières étant particulièrement évidentes dans le cas des municipalités limitrophes. Les surestimations observées dans ces municipalités peuvent être en partie expliqués par le fait que les parcelles en bordure de l'État, appartenant parfois à des propriétés officiellement localisées dans les États voisins ne sont pas recensées dans les statistiques des municipalités du Tocantins. Néanmoins, bien que le nombre de clusters ait été empiriquement défini pour simplifier le traitement, la classification finale peut bénéficier d'une évaluation interne de différentes combinaisons de nombres de clusters afin d'obtenir des résultats de clusters optimaux d'un point de vue statistique (maximisant la somme des carrés intra-cluster) pouvant favoriser une meilleure séparabilité des classes.

L'inclusion d'autres étapes de traitement pourrait aussi contribuer à améliorer les résultats de classification au détriment de traitements plus lourds. En particulier, la radiométrie des objets pourrait être affectée par l'effet de bordure dû à la disparité des résolutions spatiales entre les jeux de données MODIS et Landsat. La méthodologie présentée pourrait ainsi bénéficier du « démixage » spectral et du rééchantillonnage spatial des pixels MODIS à la résolution spatiale Landsat (ex. Bisquert *et al.*, 2015; Gevaert et García-Haro, 2015; Knauer *et al.*, 2016). Des erreurs potentielles peuvent être également induites par les dates d'acquisition des scènes Landsat utilisées pour la segmentation. Par exemple, une parcelle qui présentait un aspect homogène à la période d'acquisition estivale des images Landsat sera délimitée par un seul objet mais peut être cultivée avec deux systèmes de culture différents pendant la saison culturale (i.e. l'information de NDVI dans l'objet résultera d'un mélange des informations temporelles de NDVI des deux systèmes). La méthodologie pourrait ainsi bénéficier d'une segmentation multi-temporelle d'images Landsat, mais au détriment de la simplicité des traitements et des données d'entrée.

Bilan : Classification des systèmes de culture à l'échelle régionale par la méthode « landscape-clustering »

- Ce chapitre présente une méthode de classification non supervisée simple et reproductible qui a permis de cartographier les principaux systèmes de culture du Tocantins en combinant deux étapes de réduction de la variabilité spectrale, une segmentation à l'échelle de la parcelle et une stratification en unités de paysage au préalable de la classification.
- L'inclusion d'une étape de stratification a significativement amélioré la précision de la classification, obtenant une précision globale pour quatre classes d'environ 93 % et une forte corrélation avec les estimations issues des statistiques agricoles, avec des coefficients de corrélation d'environ 0,9.
- La stratification régionale en unités de paysage semble être une contribution significative à la cartographie des systèmes de culture à l'échelle régionale, en particulier pour les systèmes de culture spatialement dispersés et sous-représentés par rapport à d'autres types de systèmes d'utilisation de terres.

Chapitre 5.
Cas d'application au Burkina Faso

5.1. ÉTAT DE L'ART DES APPROCHES DE TÉLÉDÉTECTION POUR LA CARTOGRAPHIE DU DOMAINE AGRICOLE AU BURKINA FASO 130
5.1.1. Exemples de cartographie à l'échelle globale et continentale 130
5.1.2. Exemples de cartographie à l'échelle régionale et nationale 132
5.1.3. Défis pour la télédétection 133
5.2. ÉVALUATION DE LA REPRODUCTIBILITÉ DE L'APPROCHE DÉVELOPPÉE 134
5.2.1. La taille du parcellaire et la variabilité intra-parcellaire 134
5.2.2. La persistante couverture nuageuse 139
5.2.3. La difficile séparabilité des classes 139
5.2.4. Adaptation des méthodes 139
5.3. RÉSULTATS 142
5.3.1. Les composantes principales 142
5.3.2. Les unités de paysage 145
5.3.3. Résultats de classification 145
5.4. DISCUSSION 150
5.4.1. Perspectives 154

Chapitre 5. Cas d'application au Burkina Faso

L'agriculture du Burkina Faso, représentative de la « petite agriculture » de l'Afrique de l'Ouest, est une agriculture principalement familiale caractérisée par un parcellaire de petite taille et un bas niveau d'intensification. Elle est très différente de l'agriculture entrepreneuriale caractéristique de la région du Tocantins sur laquelle on a réalisé le développement méthodologique.

Ainsi, dans ce dernier chapitre on va tester la reproductibilité des méthodes développées et présentées dans les chapitres précédents (chapitres 3 et 4) en appliquant la démarche sur le territoire du Burkina Faso, en réponse à la dernière question de recherche :

- Quelle est la reproductibilité de l'approche développée face au changement de type d'agriculture (de l'agriculture intensive à la petite agriculture familiale) ?

Dans un premier temps nous présentons des exemples de produits cartographiques du domaine cultivé du Burkina Faso établis à partir de données de télédétection, et les principales limites rencontrées pour aller plus en détail dans la description des classes agricoles, en lien avec les spécificités régionales. Dans un deuxième temps on évalue la reproductibilité de l'approche développée et on présente les adaptations réalisées. Finalement, les résultats obtenus sont présentés et discutés.

5.1. État de l'art des approches de télédétection pour la cartographie du domaine agricole au Burkina Faso

En parcourant la littérature en télédétection, on s'aperçoit qu'aucune étude n'a été publiée à ce jour sur la cartographie des systèmes agricoles (décrits par les espèces cultivées et les pratiques agricoles associées) du Burkina Faso à l'échelle nationale. Cependant, nous présentons ci-après quelques exemples non exhaustifs d'études récentes sur la cartographie du domaine agricole de l'échelle globale à l'échelle nationale.

5.1.1. Exemples de cartographie à l'échelle globale et continentale

Plusieurs exemples de produits globaux d'occupation des terres sont disponibles à ce jour. Ceux-ci dissocient généralement les surfaces cultivées et les surfaces de pâturages d'autres types de surface, et, selon le produit, dissocient l'agriculture irriguée et pluviale. Quelques exemples sont les produits "GlobCover" à 300 m de résolution spatiale (Arino *et al.*, 2007), "GlobeLand 30" à 30 m (Chen *et al.*, 2015) et "ESA CCI Land Cover" à 300 m

(Kirches *et al.*, 2014)[42]. Ces cartes sont produites à des résolutions spatiales allant de la haute à la moyenne résolution à partir de différentes données satellitaires (ex. MODIS, MERIS, SPOT VGT, Landsat TM/ETM+, HJ-1[43]) (Arino *et al.*, 2007; Chen *et al.*, 2015; Kirches *et al.*, 2014). Cependant, leur utilisation est souvent limitée pour des applications à l'échelle régionale en raison de leur mise à jour irrégulière et une précision qui est souvent remise en question, notamment sur l'Afrique de l'Ouest où ils tendent à négliger les spécificités régionales (Fritz, See et Rembold, 2010; Knauer *et al.*, 2017; Lambert, Waldner et Defourny, 2016; See *et al.*, 2015).

L'équipe de l'ESA CCI Land Cover a actualisé l'information sur l'occupation des terres du continent africain avec un « prototype » de carte à 20 m de résolution à partir d'une série annuelle avec une résolution temporelle de 10 jours de données Sentinel-2A (2015-2016)[44], mais la version finale n'est pas encore disponible, et la précision de la version prototype n'a pas été validée (Ramoino, Pera et Arino, 2018). Ce prototype de carte a été obtenu par classification supervisée à l'aide des classifieurs du maximum de vraisemblance et de forêts aléatoires (*Random Forests*) à partir d'une base de données d'entrainement issue de l'outil *open source* "Collect Earth" qui se base sur la photo-interprétation d'images issues de Google Earth, Bing Maps et Google Earth Engine (Bey *et al.*, 2016).

Xiong *et al.* (2017) ont également travaillé à l'échelle du continent, produisant une carte du domaine cultivé pour l'année 2015 à 30 m de résolution spatiale. La carte a été produite à partir de données Sentinel-2A (résolution temporelle de 10 jours), Landsat 8 (résolution temporelle de 16 jours) et des données d'élévation SRTM (*Shuttle Radar Topographic Mission*) sur la plateforme de traitement de données Google Earth Engine. La méthode de classification a utilisé des données d'apprentissage dérivées par photo-interprétation d'images de Google Earth et des produits globaux d'occupation des terres ("GlobCover" (Arino *et al.*, 2007), "GRIPC" (Salmon *et al.*, 2015), "GlobeLand 30" (Chen *et al.*, 2015)) pour entrainer les classifieurs supervisés de forêts aléatoires et SVM (*Support Vector Machine*) sur 7 zones agro-écologiques du continent (les classifications sont réalisées séparément par zone).

[42] La dernière mise à jour disponible de ces produits date de : Globcover (2006) ; GlobeLand 30 (2010) ; ESA CCI Land Cover (2015).

[43] Les capteurs « Medium Resolution Imaging Spectrometer » (MERIS, à bord du satellite Envisat), « Végétation » (VGT, à bord des satellites SPOT-4 et -5), « Thematic Mapper » (TM, à bord des satellites Landsat-4 et -5) et la constellation « Huan Jing-1 » (HJ-1), ne sont plus opérationnels en 2018.

[44] La version « prototype » est Consulté à l'adresse :http://2016africalandcover20m.esrin.esa.int/.

5.1.2. Exemples de cartographie à l'échelle régionale et nationale

À l'échelle de l'Afrique de l'Ouest, Gessner *et al.* (2015) proposent une approche « optimisée régionalement » pour la cartographie de l'occupation des terres à 250 m de résolution spatiale (testée sur l'année 2006, mais le résultat n'est pas accessible). Leur approche utilise différents types de données, optiques et radar, adaptées à la détection des différentes classes d'occupation des terres. Les surfaces cultivées et de pâturage sont classées par un arbre de décision et une classification supervisée en utilisant le classifieur de forêts aléatoires à partir d'une série annuelle MODIS d'indices de végétation (produits "MOD13Q1" et "MYD13Q1").

Lambert, Waldner et Defourny (2016) ont cartographié le domaine cultivé en 2014 sur une région d'environ 4 300 000 km^2 comprenant les « agrosystèmes » sahéliens et soudaniens de l'Afrique de l'Ouest y compris ceux du Burkina Faso (leur carte n'est pas accessible). La méthode utilisée est basée sur une classification supervisée avec un classifieur SVM (« *Suppot Vector Machines* ») à partir des données journalières PROBA-V à 100 m de résolution spatiale et des données d'entrainement dérivés des produits d'occupation des terres globaux "GlobeLand 30" (Chen *et al.*, 2015) et "ESA CCI Land Cover" (Kirches *et al.*, 2014).

L'étude la plus récente et complète à notre connaissance sur le Burkina à l'échelle nationale a été présentée par Knauer *et al.* (2017). Des cartes du domaine cultivé ont été produites à l'échelle nationale pour les années 2001 et 2014 pour analyser la trajectoire des surfaces cultivées pendant la période d'étude, notamment leur extension. L'agriculture pluviale, irriguée et les plantations ont été classées par une méthode de classification supervisée par le classifieur de forêts aléatoires à partir des données MODIS et Landsat et de données d'apprentissage générées par photo-interprétation à partir d'images de Google Earth. Un modèle de fusion de données, ESTARFM (*Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model*) (Knauer *et al.*, 2016), a été utilisé pour combiner les informations de plus de 500 scènes Landsat et 3000 scènes MODIS pour générer des séries annuelles à haute résolution temporelle (8 jours) et spatiale (30 m). La carte obtenue pour l'année 2014 est présentée dans la Figure 46, mais elle n'est pas accessible au téléchargement.

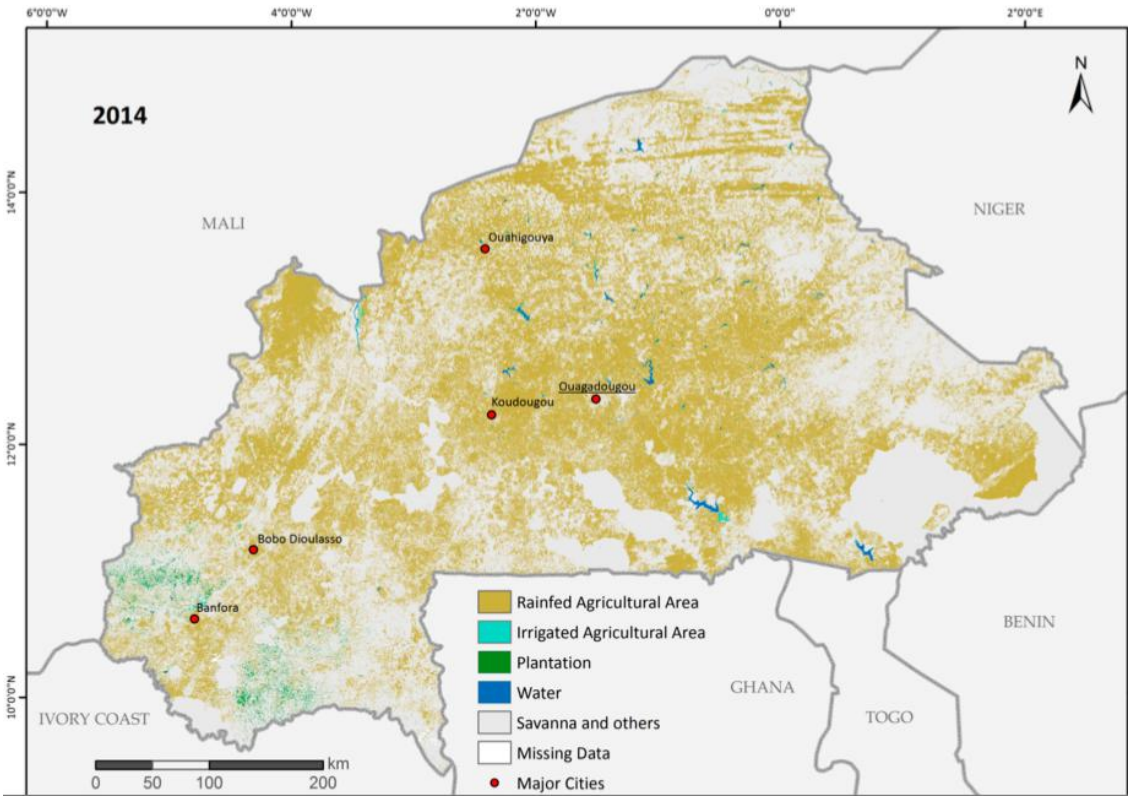


Figure 46. Carte du domaine cultivé en 2014 au Burkina Faso (source : Knauer *et al.*, 2017).

5.1.3. Défis pour la télédétection

Les études présentées révèlent plusieurs spécificités de l'Afrique de l'Ouest et du Burkina Faso qui s'avèrent contraignantes pour la télédétection. Celles-ci sont répertoriées par Bégué *et al.* (2016) et concernent principalement :

- le parcellaire agricole fragmenté, la taille des parcelles étant généralement inférieure à l'hectare ;
- la grande variabilité intra-parcellaire liée à l'association de plusieurs espèces et à la variabilité des conditions de croissance au sein d'une même parcelle (Malingreau, Bartholomé et Barisano, 1987), exacerbée par la présence des arbres dans les parcelles, et la grande variabilité inter-parcellaire liée à la diversification de la production et l'hétérogénéité de pratiques au sein des exploitations ;
- la persistante couverture nuageuse qui réduit la disponibilité de données optiques dégagées pendant la saison de pluies rendant difficile le suivi de la phénologie des cultures dominantes pluviales et limitant pratiquement l'information à la saison sèche lorsque le domaine cultivé est difficile à délimiter (les parcelles en sol nu,

sans limites claires se confondent avec les prairies clairsemées et les terres stériles) (Knauer *et al.*, 2017; Lambert *et al.*, 2016) ;

- le développement végétatif des cultures et de la végétation naturelle est synchronisé sur la courte saison de pluies ce qui rend difficile la séparabilité des surfaces cultivées des autres surfaces végétalisées notamment des savanes (Gessner *et al.*, 2015; Xiong *et al.*, 2017). Cette difficulté est exacerbée par les jachères qui sont mises en exploitation d'une année sur l'autre en fonction des nécessités avec différentes durées (allant de moins d'un an jusqu'à cinq ans) et présentent donc des réponses spectrales et temporelles très différentes ce qui les rend difficiles à distinguer des cultures et de la végétation naturelle (Xiong *et al.*, 2017).

Ces fortes contraintes à l'utilisation de la télédétection expliquent en partie le manque d'approches de cartographie des systèmes agricoles avec un niveau de finesse thématique suffisante pour distinguer les espèces cultivées et les pratiques agricoles à l'échelle régionale.

5.2. Évaluation de la reproductibilité de l'approche développée

Les défis pour la télédétection identifiés précédemment (section 5.1.1) anticipent des potentielles adaptations de l'approche développée sur le Tocantins (présentée dans les chapitres 3 et 4) aux spécificités des systèmes agricoles du Burkina. Les sous-sections suivantes illustrent les différences entre les deux régions d'étude en termes de besoin en imagerie et en méthodes compte tenu de ces défis, et les adaptations réalisées.

5.2.1. La taille du parcellaire et la variabilité intra-parcellaire

Comme présenté dans le chapitre introductif (section 1.3.1), les besoins en résolution spatiale diffèrent selon la région à cartographier compte tenu de la taille des parcelles agricoles. Les Figures 47 et 48 présentent différents rapports de dimension entre les images utilisées dans l'approche développée (Landsat et MODIS) et les parcelles agricoles des deux régions d'étude.

La Figure 47 présente une parcelle de taille moyenne de chaque région (~100 ha au Tocantins et ~0,6 ha au Burkina Faso) sur un extrait d'image MODIS NDVI et Landsat 8 à 250 m et 30 m de résolution spatiale sur 9 km^2 de surface agricole. La Figure 48 représente l'emprise des parcelles présentées dans la Figure 47 sous une grille représentant les dimensions des pixels MODIS et Landsat.

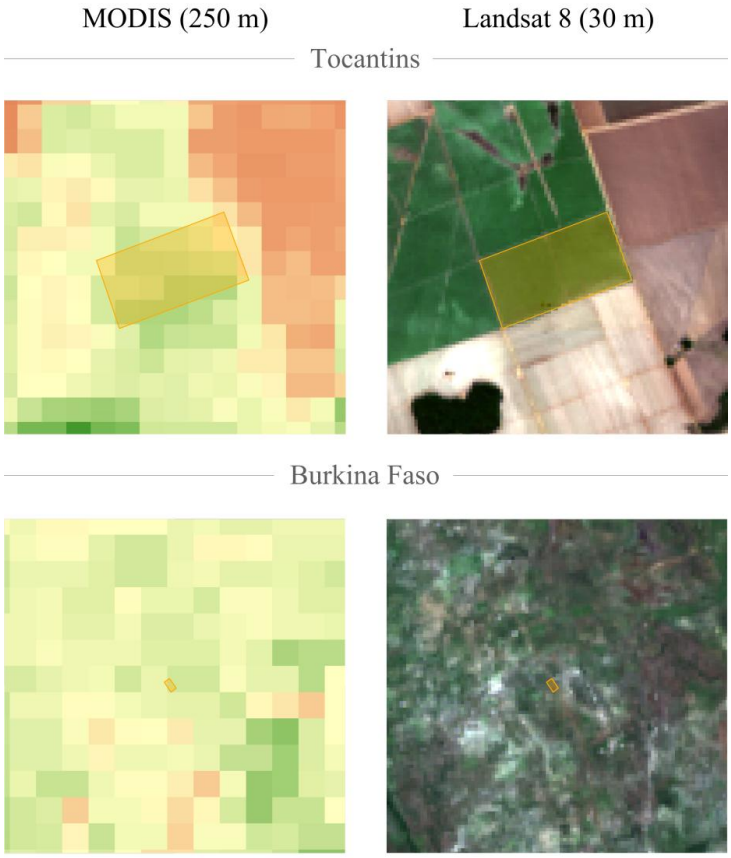


Figure 47. Représentation d'une parcelle de taille moyenne d'environ 100 ha au Tocantins et 0,6 ha au Burkina Faso, sur un extrait de 9 km^2 d'images MODIS NDVI et Landsat 8.

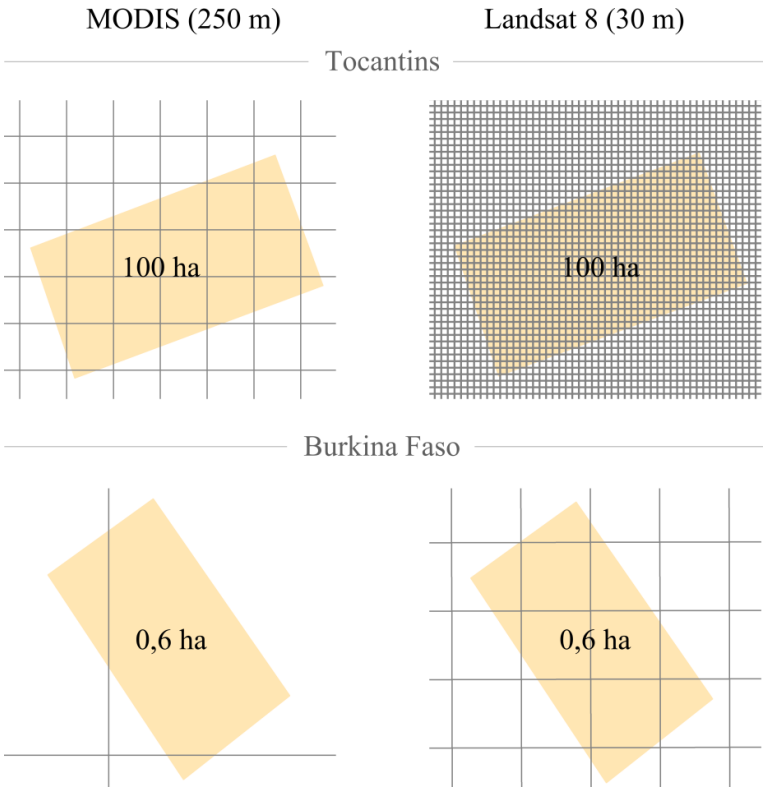


Figure 48. Représentation d'une parcelle de taille moyenne du Tocantins et du Burkina Faso sous une grille correspondant aux pixels MODIS NDVI et Landsat 8.

Comme observé dans ces deux figures (Figure 47 et 48), et selon la situation définie par Strahler, Woodcock et Smith (1986) (présentée dans la section 1.3.1), l'agriculture du Burkina Faso se trouve dans un « cas de basse résolution » la résolution spatiale des données MODIS étant supérieure à la taille moyenne du parcellaire burkinabé. En prenant comme exemple le seuil du « facteur d'échelle » proposé par Woodcock et Strahler (1987) par lequel on peut distinguer la limite d'un « cas de haute résolution » (une taille de pixels supérieure aux ¾ de la taille des unités spatiales d'intérêt), ceci se situerait à 67 m de résolution soit un pixel de 4 500 km^2 pour une parcelle moyenne burkinabé de 0,6 ha, or la taille d'un pixel MODIS est presque 14 fois plus élevée.

La résolution spatiale des données MODIS qui se situe par-dessous la limite du « cas de haute résolution » considérant la taille des parcelles agricoles du Tocantins était adaptée pour la classification des systèmes de culture, garantissant au moins un pixel avec une information spectrale « pure » à l'intérieur de la parcelle, celles-ci étant d'autant plus très homogènes. Les pixels MODIS au Burkina Faso contiennent des informations spectrales mixtes issues de la variabilité intra-parcellaire (ex. dans les cas de cultures mixtes et de présence des arbres dans les parcelles), variabilité inter-parcellaire et les autres types de surface entourant les parcelles. En outre, l'information temporelle mixte du comportement phénologique des différentes surfaces empêche de distinguer la phénologie des cultures, ce qui signifie que les données MODIS sont pauvrement adaptées à la cartographie des systèmes de culture au Burkina Faso.

Pour que la méthode de classification des systèmes de culture du Tocantins, présentée dans le chapitre 4, soit reproduite avec des données d'une résolution correspondante au Burkina Faso, les données MODIS à 250 m devraient être remplacées par de données à environ 20 m de résolution, et les données Landsat à 30 m par des données à une résolution de l'ordre de 2,5 m. Parmi les principaux satellites opérationnels à ce jour (cf. Tableau 2), les satellites Sentinel-2 (10 à 20 m) et SPOT 6/7 (bande panchromatique à 1,5 m et bandes multispectrales à 6 m) réunissent les caractéristiques nécessaires du point de vue de la résolution spatiale.

Les données Sentinel-2, avec une emprise d'environ 290 km × 290 km, sont mises à disposition avec un découpage en tuiles d'environ 100 km × 100 km. L'emprise de ces tuiles est considérablement plus petite que celle des données originelles (MODIS et Landsat) comme illustré dans la Figure 49. Les satellites SPOT 6/7, avec une fauchée de 60 km, présentent plusieurs modes d'acquisition, pouvant acquérir des images jusqu'à 600 km de long (et 60 km de large) en une seule passe. L'emprise d'une tuile d'une acquisition de 60 km × 60 km est représentée dans la Figure 49 à titre d'exemple.

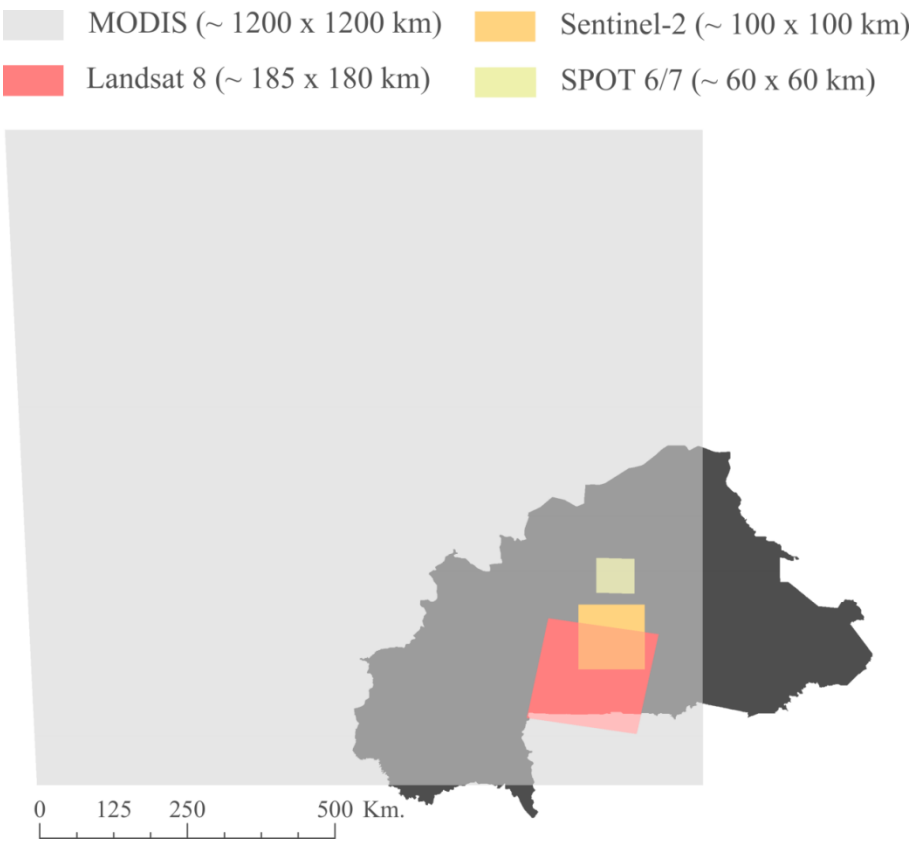


Figure 49. Emprise spatiale des tuiles MODIS, Landsat 8, Sentinel-2 et SPOT 6/7 sur la région du Burkina Faso.

En considérant ces nouvelles données, le nombre de tuiles nécessaires pour couvrir l'ensemble de la région augmente et les traitements deviennent plus lourds. Par exemple, 44 tuiles Sentinel-2 sont nécessaires pour obtenir une couverture de l'ensemble de la région d'étude. La fréquence d'acquisition de ces données étant de 5 jours, avec les deux satellites Sentinel-2A et -2B opérationnels depuis mars 2017, les séries temporelles annuelles se composent de 73 images ce qui fait un total de 3212 images à traiter. Chaque fichier contenant une image occupe environ 6 Go d'espace disque, ce qui génère des volumes de données conséquents qui nécessitent des moyens de stockage et des capacités de traitement importants.

La constellation de satellites commerciaux à haute résolution spatiale SPOT6/7 fournissent des images sur demande. Avec un prix compris entre 0,3 € et 1,2 € par km²[45], le prix d'une couverture nationale du Burkina Faso s'élèverait à un prix inabordable pour la plupart des utilisateurs (entre 100 k€ et 400 k€), et les volumes de données générés seraient

[45] Le prix au km² est le prix « acteur public » via le projet GEOSUD. La fourchette des prix est liée au nombre de programmations du satellite nécessaires pour obtenir une image dégagée. En saison sèche, l'acquisition d'une image dégagée est généralement obtenue en 1 ou 2 tentatives, en saison de pluies, le nombre de tentatives est plus élevé, de l'ordre de 3 à 5.

pratiquement ingérables. À titre d'exemple, une segmentation à l'échelle de la parcelle de la commune de Koumbia au Burkina Faso (1 553 km^2) a généré un fichier au format shapefile de 2,64 Go avec environ 900 000 segments (un extrait du résultat de la segmentation est présenté dans la Figure 50 ; Raffaele Gaetano, communication personnelle 2017).



Figure 50. Extrait de la segmentation réalisée sur une image SPOT 7 dans le cadre du projet JECAM (*Joint Experiment of Crop Assessment and Monitoring*) (Raffaele Gaetano, communication personnelle 2017).

Comme on peut l'observer dans la Figure 50, la segmentation est conditionnée par la grande variabilité intra-parcellaire. La taille des segments doit être suffisamment petite pour arriver à isoler les arbres au sein des parcelles afin d'obtenir des segments avec une information spectrale « homogène » des cultures à l'intérieur et faciliter ainsi les classifications postérieures, ce qui résulte en un nombre de segments très élevé. De plus, plusieurs segments sont parfois nécessaires pour délimiter des parcelles hétérogènes. En partant de cet exemple, nous avons extrapolé les résultats de la segmentation de la commune de Koumbia (nombre de segments et taille du fichier) à la superficie totale du Burkina (~274 000 km^2), et les avons comparés aux résultats obtenus dans le Tocantins sur une surface équivalente (Tableau 11). Cette estimation montre la grande quantité de segments et la grande taille du fichier qui seraient générés au Burkina par rapport au Tocantins.

Tableau 11. Ordres de grandeur des résultats de segmentation pour le Tocantins et le Burkina Faso sur une superficie équivalente de 240 000 km^2. Les valeurs ont été estimées à partir d'une segmentation réalisée sur le Tocantins (section 4.1.1.1) et une segmentation réalisée sur la commune de Koumbia.

	Nombre de segments	Taille du fichier shapefile (Go)
Tocantins	~185 000	0,450
Burkina Faso	~161 531 000	529

5.2.2. La persistante couverture nuageuse

Les images Sentinel-2 présentent une résolution spatiale adaptée à la taille des parcelles au Burkina. Cependant, la fréquence théorique des images Sentinel-2 (une image tous les 5 ou 10 jours) est fortement réduite pendant la saison de pluies en raison de la forte couverture nuageuse. Le pourcentage de couverture nuageuse moyen par tuile pendant la période culturale de 2016 (de juin à octobre) est de 43 %, et le nombre moyen d'images disponibles avec moins de 10 % de couverture nuageuse pendant la période est de 5 images par tuile[46]. Les acquisitions de 2016 ont été faites avec une répétitivité de 10 jours, lorsque seul le satellite Sentinel-2A était opérationnel. Avec la mise en orbite du satellite jumeau Sentinel-2B en mars 2017, la résolution temporelle est passée à 5 jours à l'équateur, qui devrait diminuer les lacunes d'information pendant la saison culturale.

5.2.3. La difficile séparabilité des classes

La synchronisation de la croissance de la végétation sur la courte période de pluies et la difficulté supplémentaire liée à la pratique courante des jachères suggère que le nombre de clusters initiaux utilisés pour le site du Tocantins devrait être augmenté afin de maximiser la séparabilité du domaine cultivé des autres surfaces végétalisées. En effet, un partitionnement spectral plus fin du jeu de données initial doit théoriquement faciliter l'isolement des classes spectrales qui présentent de faibles différences phénologiques.

5.2.4. Adaptation des méthodes

En ce qui concerne la méthode de stratification régionale en unités de paysage, nous considérons que les données MODIS restent adaptées pour la délimitation des unités de paysage, ces dernières se situant dans un « cas de haute résolution » par rapport à la résolution

[46] Les calculs ont été réalisés à partir de toutes les images disponibles pour les 30 tuiles dont 50 % ou plus de l'emprise recouvre la région d'étude (tuiles exclues du calcul : 30PXB, 30PZB, 30PVA, 31PBR, 30PUV, 31PCQ, 30PTU, 31PDP, 31PCN, 31PDN, 30PWS, 30PUR, 30PVR, 30PWR).

spatiale de MODIS. Cependant, les informations contenues dans un pixel MODIS sur le Burkina Faso étant plus synthétiques que sur le Tocantins, les limites des unités de paysage seront moins marquées par des éléments de la surface visibles sur les images au Tocantins (ex. les contours de parcelles).

Par rapport à la méthode de classification des systèmes de culture, celle-ci préciserait d'une substitution des données MODIS et les données Landsat originales par des données adaptées au changement d'échelle imposé par la petite taille du parcellaire burkinabé tels que Sentinel-2 et SPOT 6/7. Cependant, en raison des problèmes de nébulosité qui affectent les données Sentinel-2 pendant la saison culturale de 2016, il reste peu d'informations disponibles sur le développement végétatif des cultures, ce qui compromet la méthode classification des systèmes de culture pour l'année 2016. D'autre part, la couverture du territoire national avec des données SPOT 6/7 a un prix très élevé, voire prohibitif pour des applications agricoles, et dans le cas d'une délimitation du parcellaire burkinabé le résultat de segmentation serait difficilement exploitable dû à la puissance de calcul nécessaire.

Ainsi, dans le cadre de cette thèse nous avons choisi d'évaluer la reproductibilité de l'approche développée avec les séries temporelles d'images MODIS, compte tenu que le produit MOD13Q1 minimise les effets de la nébulosité avec le processus de synthèse sur 16 jours à partir de données journalières. Cependant, la résolution spatiale des données MODIS ne permet pas d'accéder à une information spectrale « pure » à l'intérieur des parcelles agricoles sans techniques de traitement supplémentaires telles que les méthodes de « démixage » spectral, la classification des systèmes de culture et des systèmes d'utilisation agricole des terres s'avère impossible. Néanmoins, la cartographie du domaine cultivé peut être envisagée en partant du principe que les pixels MODIS captant majoritairement les réponses spectrales mixtes des groupes de parcelles puissent être dissociés de ceux captant les réponses d'autres types de surface par la méthode de classification « landscape-clustering » présentée dans le chapitre 4.

La méthode adoptée dans ce chapitre se décompose ainsi en deux étapes : (i) une étape de délimitation des unités de paysage suivant la méthode présentée dans le chapitre 3 (section 3.1.1), et (ii) une étape de classification non supervisée des pixels de la série temporelle MODIS-NDVI du 2016 (présentée dans la section 2.2.1) suivant la méthode de « landscape-clustering » présentée dans le chapitre 4 (section 4.1.2.2). La Figure 51 montre les principales étapes méthodologiques adoptées dans ce dernier chapitre.

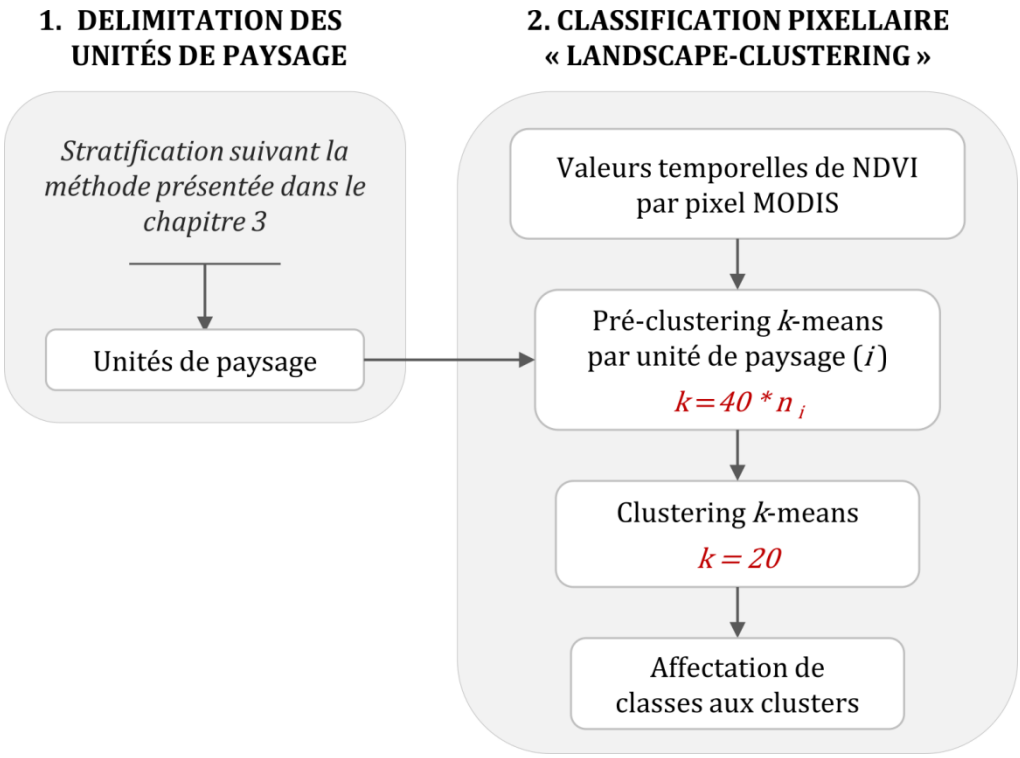


Figure 51. Représentation des étapes méthodologiques adoptées dans le chapitre 5.

Les adaptations réalisées portent uniquement sur l'étape de classification :

- La méthode de classification orientée-objet est remplacée par une classification par pixel. Un total de 5 037 865 pixels sont à classer, chacun des pixels contenant les informations temporelles des 23 images de la série annuelle de NDVI de 2016. L'étape de pré-clustering ne s'effectue donc plus sur un groupe d'objets mais sur un groupe de pixels au sein d'une unité de paysage (tous les pixels dont le centroïde est situé à l'intérieur des limites d'une unité de paysage sont regroupés) ;

- Une seule étape de réduction de la variabilité spatiale étant appliquée (stratification de la région en unités de paysage, mais pas de segmentation à l'échelle de la parcelle), et compte tenu de la difficulté de séparabilité des classes au Burkina Faso, nous augmentons de façon empirique le nombre de clusters dans l'étape de pré-clustering (de 20 à 40 clusters), et dans l'étape de clustering finale (de 10 à 20 clusters) pour augmenter la séparabilité des classes tout en restant sur un nombre final de clusters relativement facile à interpréter.

Les autres étapes et paramètres des méthodes originales, décrits en détail dans les chapitres 3 (section 3.1.1) et 4 (section 4.1.2.2), sont conservées. Les clusters finaux ont été interprétés par l'analyse de leurs profils temporels de NDVI et par photo-interprétation à l'aide des images à haute résolution spatiale Sentinel-2A et SPOT 7 (présentées dans la

section 2.2.1.3) et affectés aux classes « culture annuelle » et « autre type d'occupation des terres ».

Les points GPS collectés en 2016 de façon exhaustive dans les zones nord, centre, sud et ouest, couvertes par des acquisitions d'images à haute résolution spatiale (3663 points au total, 2357 points de référence pour la classe « culture annuelle » et 1306 points pour la classe « autre type d'occupation et utilisation des terres », cf. Figure 29), ont été croisés avec une grille de pixels MODIS et un seul point par pixel MODIS a été retenu pour l'évaluation de la classification pixellaire. Nous avons procédé de la façon suivante : pour les 1679 pixels qui se superposent spatialement aux points GPS nous avons calculé le nombre d'occurrences (nombre de points à l'intérieur de chaque pixel) par classe. En cas d'occurrence multiple, nous avons retenu la classe majoritaire. En cas de parité dans le nombre de points représentatifs de chaque classe, nous avons réalisé une sélection aléatoire. Au total, 1679 points ont été retenus, 966 pour la classe « culture annuelle » et 713 pour la classe « autre type d'occupation et utilisation des terres ».

La précision du résultat de classification est estimée en utilisant les indices de précision globale et le coefficient Kappa (présentés dans la section 2.3.5) calculés à partir des 1679 points de terrain retenus. De plus, la superficie cultivée (cultures annuelles) ainsi cartographiée à l'échelle nationale est comparée à la surface donnée par les statistiques agricoles de 2016 (présentées dans la section 2.2.3.2).

5.3. Résultats

5.3.1. Les composantes principales

La Figure 52 montre les vecteurs propres des trois premières composantes principales en fonction des images de la série annuelle de 2016 (représentées par leur code du jour julien).

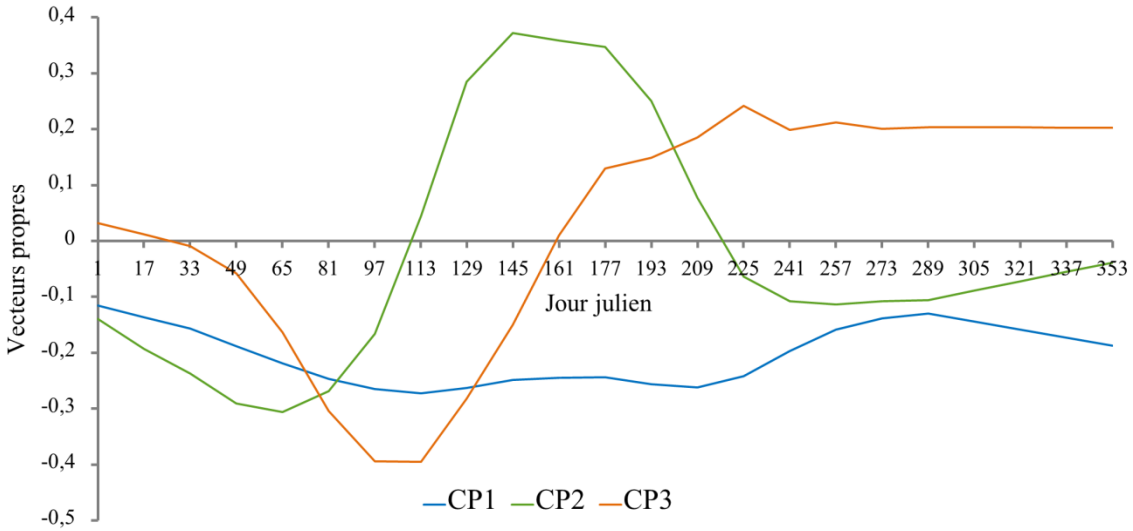


Figure 52. Magnitude des vecteurs propres (axe d'ordonnées) des trois premières composantes principales en fonction des images de la série annuelle de 2016 (représentées en abscisse par leur code du jour julien).

On observe dans la Figure 52, comme dans le résultat de la CP1 au Tocantins (Figure 32 CP1), une distribution temporelle relativement constante des vecteurs propres de la première CP sur toute la période annuelle avec un signe négatif, ce qui implique que cette CP n'a pas capté des variations temporelles particulières. Comme cela a été observé dans les résultats précédents au Tocantins (section 3.2.1), cette CP est fortement corrélée de façon négative avec le NDVI cumulé sur l'année, avec une corrélation linéaire globale de r = - 0,99. Le résultat de la corrélation par pixel est présenté dans la Figure 53. Ce résultat suggère que le principal facteur de variabilité dans la série temporelle capté par la première composante (représentant 84 % de la variance totale) est associé à la variabilité spatiale de valeurs de NDVI.

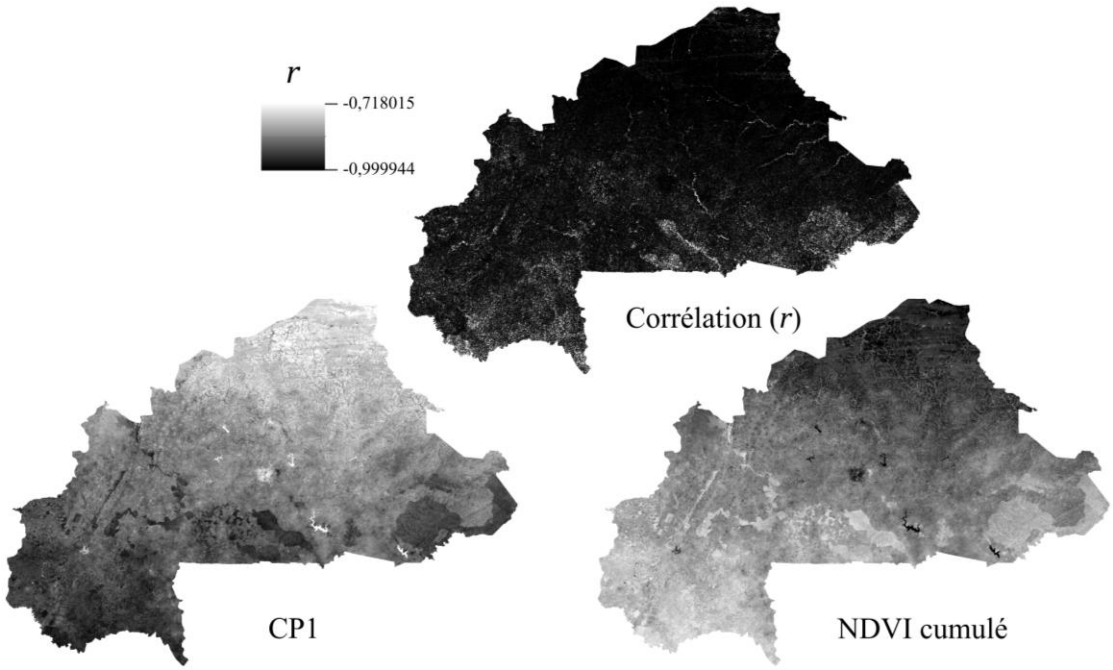


Figure 53. Corrélation (r) entre la CP1 et le NDVI cumulé sur l'année 2016.

Le profil des vecteurs propres de la deuxième CP (Figure 52 CP2) suggère que la CP2 a capté la variabilité saisonnière de la végétation associée au régime pluviométrique. Ceci est mis en évidence par les dates des pics maximaux et minimaux et les inversions de signe qui reflètent les phases du cycle pluviométrique annuel. En effet, comme dans le résultat de la CP2 au Tocantins (Figure 32 CP2), on observe que pendant la saison de pluies, ici entre fin avril et début mai (jour 113) et fin août (jour 225), les images sont positivement corrélées avec les vecteurs propres, et pendant la saison sèche le signe est inversé.

Les changements de signe autour des jours 113 et 225 seraient ainsi associés respectivement aux processus de reverdissement et sénescence de la végétation, le développement végétatif culminant entre les jours 129 et 193. Les vecteurs propres positifs ont une distribution gaussienne, le « maximum de croissance » étant ici moins marqué (représenté par une courbe en cloche) que celui du profil des vecteurs propres du CP2 au Tocantins (Figure 32 CP2). Ceci pourrait être expliqué par le décalage régional du début de la saison de pluies suivant le gradient nord-sud spécialement marqué au Burkina Faso.

Le profil des vecteurs propres de la CP3 (Figure 52 CP3), montre également un cycle annuel qui indique que cette CP a capté une variation temporelle de la végétation. La variation mise en évidence par ce troisième CP est plus précoce que la variation saisonnière dominante (captée par la CP2) et inversée par rapport à cette dernière, soulignant des événements avec une variation de NDVI importante autour des jours 97 et 113 (mois d'avril).

Globalement les résultats obtenus montrent que la CP1 capte majoritairement la variabilité spatiale des valeurs de NDVI et que les CP d'ordre supérieur captent la variabilité temporelle au cours de la série annuelle. Ces résultats sont cohérents avec les résultats obtenus pour la série annuelle 2013-2014 au Tocantins présentés dans la section 3.2.1, et des études précédentes présentées dans la section 3.1.1.1.

5.3.2. Les unités de paysage

La segmentation des 22 « variables phénologiques » ou CP d'ordre supérieur avec l'algorithme de segmentation multi-résolution (Baatz et Schäpe, 2000) (paramétré suivant la méthode décrite dans la section 3.1.1.2, cf. Tableau 7) a produit 80 unités de paysage (Figure 54), d'une superficie moyenne de 3 420 km^2, 49 km^2 pour la plus petite et 10 715 km^2 pour la plus large.

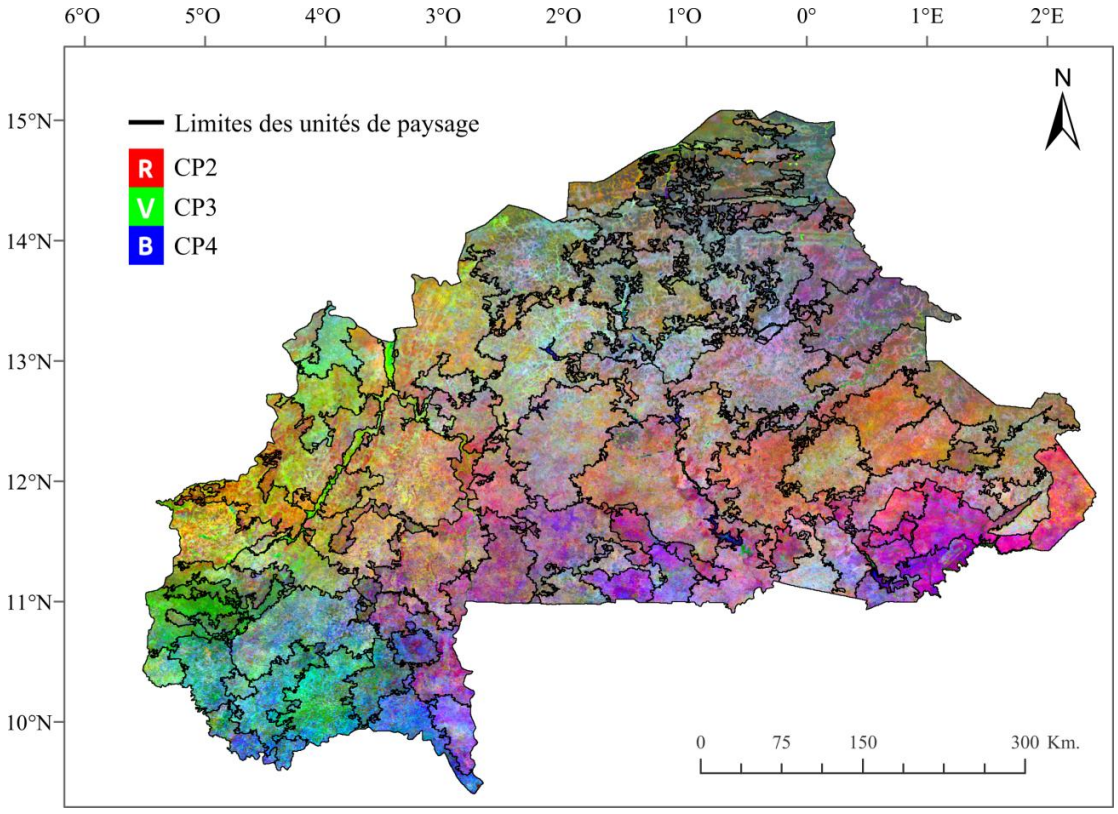


Figure 54. Résultat de la stratification régionale en unités de paysage obtenu suivant la méthode décrite dans le chapitre 3 à partir d'une série temporelle de données MOD13Q1 de l'année 2016.

5.3.3. Résultats de classification

L'analyse des profils temporels moyens de NDVI des vingt clusters finaux et l'analyse par photo-interprétation sur les images à haute résolution spatiale Sentinel-2A et SPOT6/7 nous ont permis de dissocier les clusters appartenant à la classe « culture annuelle » et les clusters « autre type d'occupation des terres ». Les profils temporels de NDVI des clusters

affectés à la classe « culture annuelle » sont présentés dans la Figure 55(a) et ceux des clusters affectés à la classe « autre type d'occupation des terres » sont présentés dans la Figure 55(b) et (c).

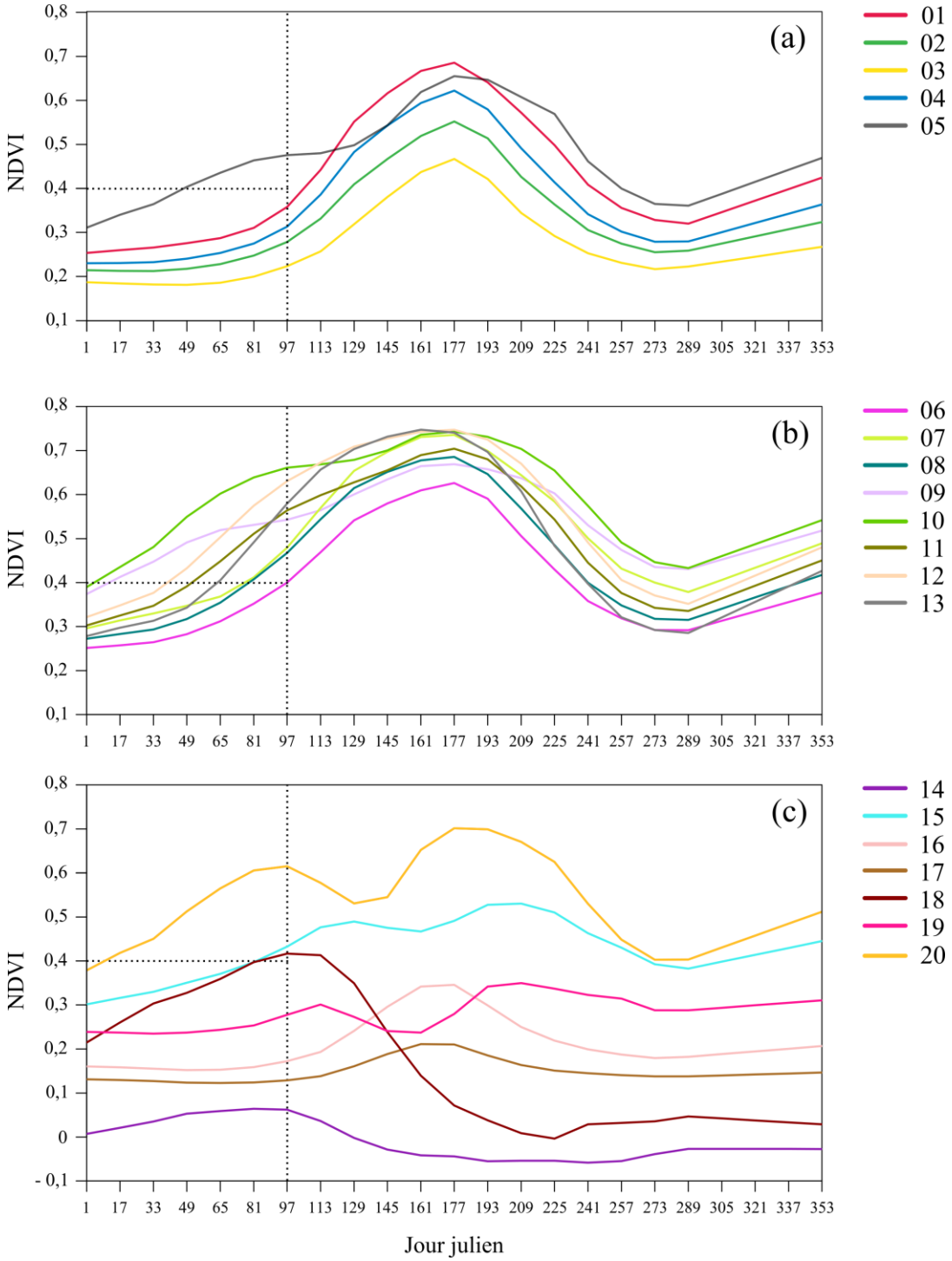


Figure 55. Profils temporels des vingt clusters finaux obtenus avec la méthode « landscape-clustering » adaptée au Burkina Faso. Les courbes continues représentent le NDVI moyen de tous les pixels appartenant à chaque cluster en fonction du jour julien. Les profils des clusters qui semblent avoir capté les cultures annuelles sont représentés dans la figure (a) et les profils des clusters qui semblent avoir capté d'autres types d'occupation et utilisation des terres sont représentés dans les figures (b) et (c).

La plupart des profils temporels de NDVI (Figure 55(a)(b)) montrent une forte augmentation des valeurs de NDVI jusqu'à atteindre un pic entre juin et juillet (du jour 161 au jour 193), puis une diminution des valeurs conduisant à des valeurs minimales au mois d'octobre (du jour 273 au jour 289). Ce résultat suggère que ces clusters ont capté le fort développement végétatif lié à la pluviométrie et reflète ainsi la phénologie saisonnière marquée qui est synchronisée sur la saison de pluies pour toutes les surfaces végétalisées.

Cependant, on peut observer une différence principale entre les profils temporels de NDVI des clusters affectés à la classe « culture annuelle » (Figure 55(a)) et ceux des clusters affectés à la classe « autre type d'occupation des terres » (Figure 55(b)) : les profils de la classe « culture annuelle » montrent un début de croissance de la végétation plus tardif que les profils de la classe « autre type d'occupation des terres ». Le profil du cluster k = 05 (Figure 55(a)) qui diffère de cette tendance a été néanmoins affecté à la classe « culture annuelle » après confirmation par photo-interprétation sur les images de haute résolution spatiale de la présence de parcelles agricoles, mais suggère un mélange de types d'occupation des terres.

À différence des profils temporels de NDVI obtenus avec la classification de 2014-2015 du Tocantins (Figure 43(iv)), les profils de la classe « autre type d'occupation des terres » (Figure 55(b)(c)) présentent aussi des oscillations des valeurs de NDVI. Outre que les profils qui reflètent une tendance positive pendant la saison de pluies (augmentation des valeurs), on retrouve des profils qui reflètent la tendance inverse. Ce résultat suggère que ces clusters ont capté les surfaces inondées pendant la saison de pluies, tels que les corps d'eau saisonniers et les bords des corps d'eau permanents, ce phénomène étant particulièrement net pour le profil du cluster k = 18 (Figure 55(c)).

La précision globale estimée à partir des 1679 points de terrain pour la classification finale à deux classes est de 64,8 % (intervalle de confiance à 95 % = 62,5-67,1 %) et le coefficient de Kappa est de 0,3. La Figure 56 présente la carte du domaine cultivé (cultures annuelles) obtenue pour l'année 2016 à 250 m de résolution spatiale.

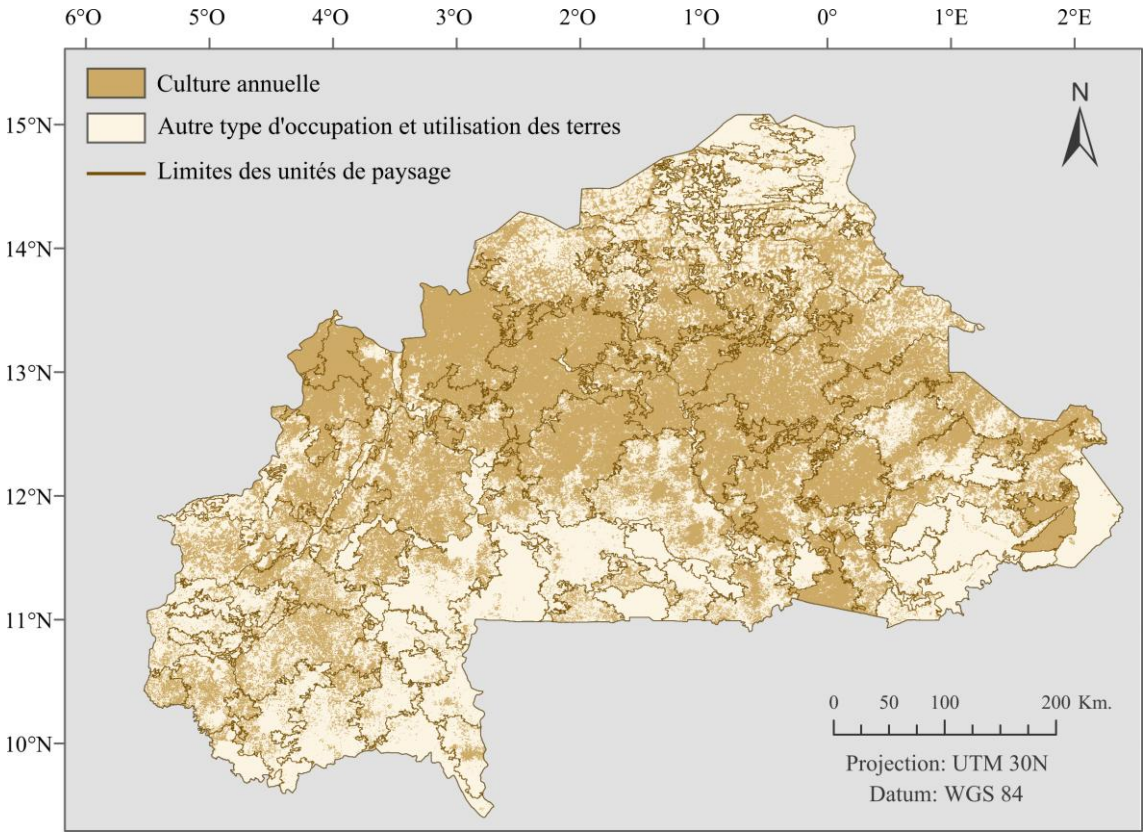


Figure 56. Carte du domaine cultivé (cultures annuelles) à 250 m de résolution spatiale pour la saison culturale de 2016 obtenue avec la méthode de classification non supervisée « landscape-clustering » et limites des unités de paysage (cf. Figure 54).

On observe des cultures annuelles réparties sur l'ensemble du pays avec une couverture dense dans une bande latitudinale dans la zone soudano-sahélienne (entre 12°N et 14°N). Les limites des unités de paysage mettent en évidence une différenciation « paysagère » de la répartition des surfaces cultivées avec des paysages allant d'une quasi-continuité des surfaces cultivées à une quasi-absence de cultures, en passant par des paysages avec différents niveaux d'intensité d'utilisation agricole des terres. L'absence de cultures annuelles est notable dans l'extrême nord aride du pays avec essentiellement une présence le long des bas-fonds mise en évidence par une distribution « ramifiée » des surfaces cultivées (Figure 57).

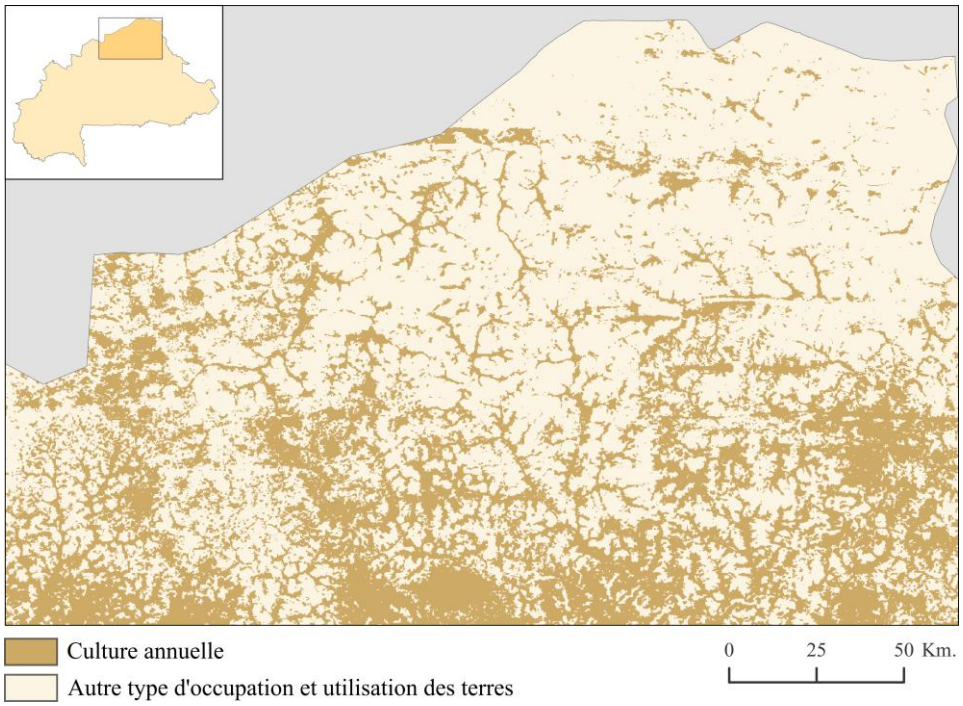


Figure 57. Extrait du résultat de classification du domaine cultivé de 2016 au nord du Burkina Faso.

L'absence de cultures est aussi observée au sein des parcs nationaux, réserves et forêts classées (Figure 58), même si certains de ces espaces subissent une exploitation forestière illégale et commencent à être mis en culture (Knauer *et al.*, 2017). Dans le sud-ouest du Burkina on observe une distribution relativement éparse des cultures annuelles. La superficie totale occupée par la classe « culture annuelle » est de 132 928 km^2 (soit environ 49 % de la superficie du pays).

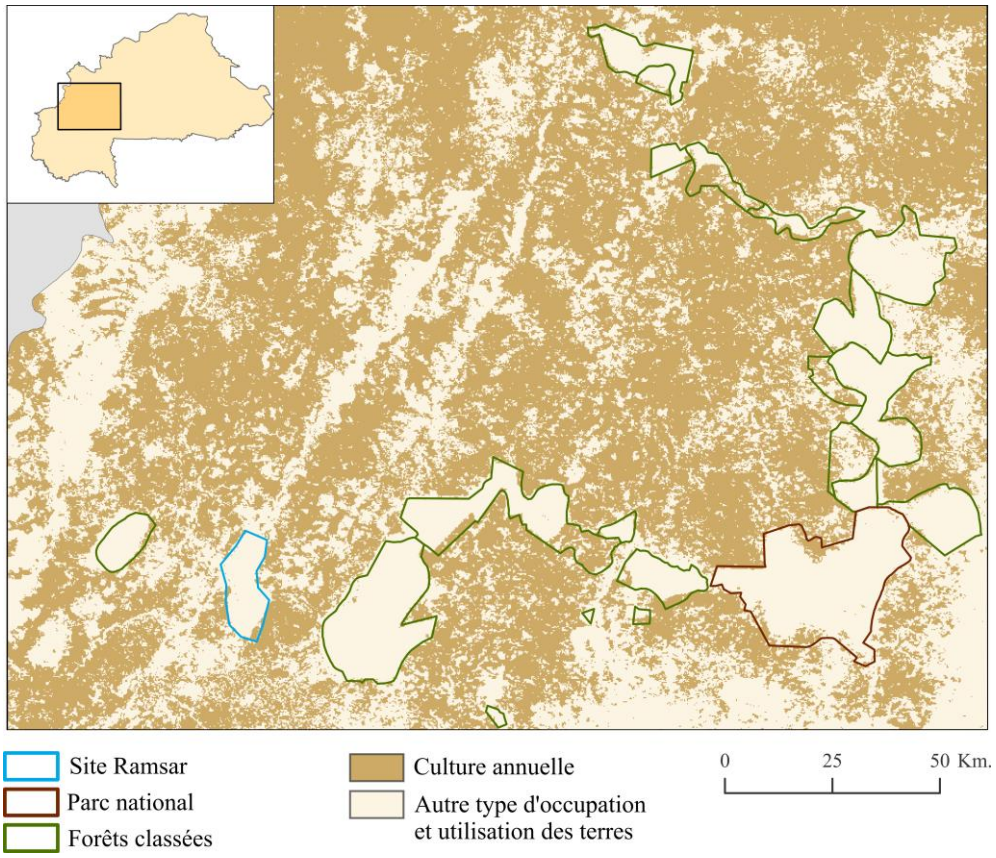


Figure 58. Extrait du résultat de classification du domaine cultivé de 2016 et limites des espaces protégés (source : limites des espaces protégés (IUCN/UNEP-WCMC, 2017)).

5.4. Discussion

Les résultats montrent que la méthode de classification non supervisée « landscape-clustering », adaptée aux spécificités de la « petite agriculture » burkinabé, est capable de distinguer le domaine cultivé (cultures annuelles) d'autres types d'occupation et utilisation des terres à l'échelle régionale avec une précision d'environ 65 %.

Le résultat du coefficient de Kappa de 0,3 est faible, ce qui indique un accord entre les points de référence et la classification « légèrement » supérieur à un accord avec une allocation aléatoire (« prévu par hasard »). Ce coefficient utilise la randomisation comme base de référence, ce qui est remis en question par plusieurs auteurs (Foody, 2002; Pontius et Millones, 2011; Strahler *et al.*, 2006). Pontius et Millones (2011) stipulent qu'une allocation aléatoire ne représente pas un niveau de référence raisonnable et proposent de calculer les erreurs d'allocation sur la base des points de référence. Le désaccord global entre la classification et les points de référence est de 35 % ce qui suggère une confusion modérée entre les deux classes évaluées.

L'évaluation de la précision des résultats de classification présente néanmoins des limites liées à la différence d'échelle entre les points de référence (collectés à l'échelle de la parcelle) et la résolution spatiale de la classification à évaluer (250 m), les points de référence retenus (un par pixel MODIS) pouvant ne pas être représentatifs du type d'occupation des terres majoritaire de la superficie occupée par les pixels classés (~ 6 ha par pixel).

Nous avons obtenu une estimation de la superficie totale occupée par les cultures annuelles en 2016 de 132 928 km², ce qui est beaucoup plus important que le chiffre dérivé des statistiques officielles qui est de 58 693 km^2 (DGESS/MAAH, 2017). Les résultats récents de l'étude réalisée par Knauer *et al.* (2017) donnent quant à eux une superficie totale[47] de 115 289 km^2 en 2014. En calculant un taux de croissance moyen annuel des surfaces cultivées de 4 215 km^2 à partir des estimations réalisées par Knauer *et al.* (2017) sur l'extension des surfaces agricoles sur une période de 13 ans (de 2001 à 2014), nous projetons une superficie totale d'environ 123 719 km^2 en 2016, soit une valeur très proche (moins de 8% de différence) de celle que nous avons obtenue par la méthode « landscape-clustering ». Le grand écart entre les estimations réalisées par télédétection et les estimations des statistiques agricoles pour l'année 2016 (DGESS/MAAH, 2017) s'explique soit par des fortes surestimations de la méthode de télédétection, soit par des fortes sous-estimations des surfaces cultivées par la méthode d'extrapolation des estimations collectées à partir d'enquêtes auprès des ménages échantillons (présentée dans la section 2.2.3.2).

Concernant les résultats obtenus par la méthode « landscape-clustering », des sources d'erreur ont été détectées à partir de l'analyse visuelle des images à haute résolution : (i) des surestimations des surfaces cultivées dans le nord aride du pays (ex. Figure 59(zone nord)) dues à la confusion des surfaces cultivées avec la végétation herbacée naturelle clairsemée, qui pourrait s'expliquer par la croissance synchronisée de la végétation avec la pluie, et (ii) des sous-estimations des surfaces cultivées dans le sud-ouest du pays (ex. Figure 59(zone ouest)) dues à la confusion entre savanes arborées et parcelles agro-forestières qui mixent cultures annuelles et forte densité d'arbres. Ainsi, on observe un gradient nord-sud de la qualité de la classification, avec une surestimation dans le nord du pays et une sous-estimation dans le sud (ex. Figure 59(zone centre) et Figure 59(zone sud)).

[47] Nous considérons ici les classes *« rainfed agricultural area »* et « *irrigated agricultural area* ».

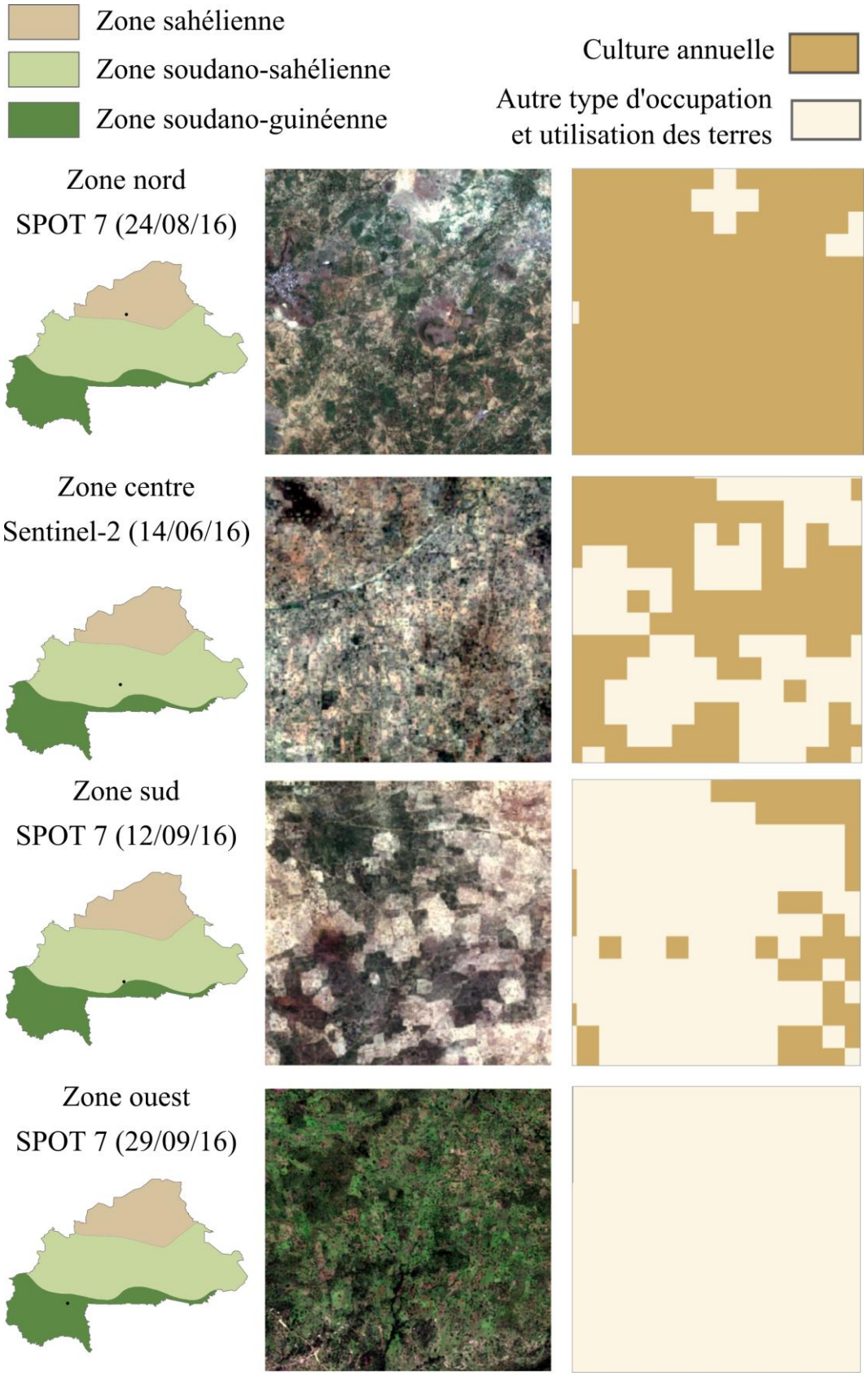


Figure 59 . Extraits de 9 km^2 des images à haute résolution Sentinel-2A et SPOT 7 pour illustrer les sources d'erreur de classification observées spatialement. La localisation des extraits et les limites des zones climatiques sont présentées à gauche.

Ces sources de confusion sont exacerbées par le mélange d'informations spectrales contenues dans les pixels MODIS, qui tend à homogénéiser la radiométrie des pixels. Ozdogan et Woodcock (2006) ont étudié la relation entre la résolution spatiale des images et les erreurs de classification des surfaces cultivées et montrent que lorsque la taille des pixels est grande par rapport à la taille des parcelles cultivées, la plupart des pixels ont des proportions de surface similaires et la variance entre pixels est faible, ce qui entraine des confusions entre classes. Notre hypothèse de départ était que les pixels MODIS captant majoritairement les réponses spectrales mixtes des groupes de parcelles pouvaient être dissociés de ceux captant les réponses d'autres types de surface par la méthode de classification « landscape-clustering ». Cette hypothèse n'est ainsi qu'en partie confirmée puisqu'elle n'est pas généralisable dans l'ensemble du Burkina i.e. la mosaïque complexe de surfaces, couverte par les pixels MODIS, ne permet pas, dans une grande partie des cas, de dissocier les groupes de parcelles d'autres types de surfaces par la méthode de classification « landscape-clustering ».

Malgré les erreurs de classification, la méthode non supervisée utilisée reste intéressante notamment car elle ne nécessite pas de données autres que les images satellitaires et par sa facilité de mise en œuvre qui la démarquent des méthodes utilisées dans les études présentées dans la section 5.1. Ces dernières sont en effet entièrement dépendantes de données externes (ex. images de Google Earth ou cartes précédentes) et reposent fortement sur des connaissances expertes pour l'obtention de données d'entrainement (ex. par photo-interprétation) pour des classifications supervisées (ex. Gessner *et al.*, 2015; Knauer *et al.*, 2017; Lambert *et al.*, 2016; Xiong *et al.*, 2017). Ces données externes d'entrainement, ou de validation sont une source d'erreur dans les classifications et fournissent des indicateurs de précision biaisés dû principalement au décalage temporel entre la période à cartographier et la période couverte par les données de référence et à leurs propres sources d'erreur. De plus, le manque d'informations multi-temporelles introduit des contraintes majeures pour la séparabilité de classes (telles que les surfaces cultivées et la végétation herbacée) par photo-interprétation (Knauer *et al.*, 2017).

Bien que la cartographie des systèmes de culture et des systèmes d'utilisation agricole des terres à l'échelle régionale n'ait pas pu être réalisée compte tenu de l'offre en imagerie disponible à ce jour, la carte du domaine cultivé obtenue représente un produit intéressant en soi. Ce type de produit cartographique est considéré fondamental dans les régions vulnérables à l'insécurité alimentaire car il aide à la prise des décisions sur la quantité de nourriture à stocker, à distribuer ou à exporter et contribue ainsi à mitiger les fluctuations de prix et les pénuries alimentaires (See *et al.*, 2015). Dans ce contexte, la stratification en unités

de paysage présente un potentiel complémentaire plausible, car il offre une information spatiale additionnelle : par croisement avec la carte du domaine cultivé, elle permet de dissocier des zones avec différents niveaux d'intensité d'utilisation agricole des terres à une échelle pertinente pour la prise de décision au niveau national.

5.4.1. Perspectives

À la fin de 2018, la première série annuelle de données Sentinel-2 sera disponible avec une résolution temporelle de 5 jours, alliant ainsi hautes résolutions spatiale et temporelle sur une longue période de temps. Cette haute fréquence d'acquisition devrait réduire les lacunes d'information pendant les saisons culturales au Burkina. En outre, les séries temporelles Sentinel-2 peuvent être complétées par d'autres données à haute résolution spatiale telles que des images Landsat 8, un produit harmonisé "*Harmonized Landsat 8 Sentinel-2* (HLS) MYD13Q1" étant mis en place à cette fin par la NASA. Le couplage avec des données radar, telles que celles issues des images Sentinel-1, offre aussi un potentiel important pour le suivi de l'agriculture, notamment dans les régions affectées par une forte nébulosité comme au Burkina, les données radar et optique apportant en plus des informations complémentaires sur les surfaces (Corgne, Hubert-Moy et Betbeder, 2016). Néanmoins, le traitement de ces données à haute résolution à l'échelle du pays va nécessiter des capacités de calcul et de stockage conséquentes et des méthodes de traitement adaptées.

Quelques pistes méthodologiques sont avancées. Nous avons observé, en cohérence avec les observations réalisées par Knauer *et al.* (2017) et Lambert *et al.* (2016), que les surfaces cultivées se dissocient du reste des surfaces végétalisées en début de saison de pluies lorsque la végétation naturelle commence sa croissance, alors que les surfaces agricoles pluviales sont encore en sol nu (période de semis). Ainsi, la méthode présentée pourrait être testée avec les données Sentinel-2 sur une période resserrée en début de saison culturale pendant laquelle l'écart spectral entre la végétation naturelle et les surfaces mises en culture est maximal, et la nébulosité est encore acceptable pour le traitement des images optiques.

Les systèmes de culture pluviaux ayant généralement la même phénologie (à la différence des systèmes au Tocantins qui présentaient des cycles culturaux distincts), ils sont difficilement dissociables sur la base d'une interprétation visuelle des profils temporels de NDVI à l'aide du calendrier cultural. Ainsi, pour augmenter la séparabilité des systèmes de culture, différentes stratégies pourraient être testées dont :

- une analyse des variables phénologiques simples telles que le cumul annuel, le maximum annuel et l'amplitude annuelle du NDVI ;

- l'inclusion d'une hiérarchie dans les étapes de clustering , par exemple, écarter dans un premier temps les clusters associés à des classes d'occupation des terres facilement identifiables telles que les surfaces en eau, forêts et surfaces bâties, et refaire ensuite un clustering uniquement sur les surfaces végétalisées pour maximiser la séparabilité des surfaces ayant un même comportement phénologique.

En se situant dans un « cas de haute résolution » (Strahler *et al.*, 1986) on pourrait envisager avec la méthode présentée et ses éventuelles améliorations une séparabilité de certains systèmes de culture, notamment les systèmes plus productifs des moins productifs tels que la culture de rente cotonnière riche en intrants et les cultures céréalières vivrières. Des indicateurs spatiaux (ex. taille et forme des parcelles) pourraient aussi servir à cette fin, les cultures de rente étant généralement associées à des parcelles de grande taille alors que les parcelles de petite taille sont caractéristiques des cultures vivrières. Cependant, pour détecter des informations sur les pratiques culturales au sein des parcelles, telles que la pratique répandue de la culture associée, il faudrait introduire des méthodes de « démixage » spectral et rééchantillonnage spatial (ex. Bisquert *et al.*, 2015; Gevaert et García-Haro, 2015; Knauer *et al.*, 2016) des images à une résolution décamétrique à une résolution métrique, ce qui est difficilement réalisable à l'échelle régionale, et pose un certain nombre de difficultés méthodologiques.

Bilan : Cas d'application au Burkina Faso

- Ce chapitre a permis de soulever les limites de reproductibilité des méthodes de cartographie des systèmes de culture et des systèmes d'utilisation des terres développées dans des régions caractérisées par la « petite agriculture familiale », compte tenu de la disponibilité d'imagerie satellitaire actuelle.
- Des spécificités régionales contraignantes pour la télédétection (parcellaire fragmenté et de petite taille, grande variabilité intra-parcellaire, forte couverture nuageuse pendant la saison culturale) et une disponibilité d'images limitée n'ont pas permis de cartographier les systèmes de culture et les systèmes d'utilisation des terres au Burkina Faso suivant les méthodes originales développées dans ce travail de thèse.
- La méthode de classification « landscape-clustering », adaptée aux spécificités régionales, a toutefois permis de produire une carte du domaine cultivé sur l'ensemble du pays pour la saison culturale de 2016, avec une précision globale d'environ 65 %.
- Cette méthode présente des avantages par rapport aux méthodes de classification supervisée couramment employées, notamment son indépendance à des données externes qui limitent fortement la production cartographique.

Conclusions et perspectives

CONTRIBUTIONS DE LA THÈSE 158
Apports méthodologiques 158
Pour la cartographie des systèmes d'utilisation agricole des terres à l'échelle régionale 158
Pour la cartographie des systèmes de culture à l'échelle régionale 160
Enseignements tirés du cas d'application au Burkina Faso 160
Apports thématiques 161
PERSPECTIVES 163
Perspectives de recherche 163
Pour la cartographie des systèmes d'utilisation agricole des terres 163
Pour la cartographie des systèmes de culture à l'échelle régionale 164
Sur l'échelle temporelle de la cartographie 165
Perspectives opérationnelles 166

Conclusions et perspectives

Les cartes de systèmes agricoles à l'échelle régionale fournissent des informations spatiales clés pour le suivi de la production agricole et des impacts environnementaux associés, et constituent donc un outil précieux pour une planification raisonnée de l'utilisation des terres. Il existe ainsi un intérêt grandissant pour la cartographie des systèmes agricoles de manière opérationnelle sur de grandes étendues afin de fournir des informations détaillées, régulières, objectives et spatialement explicites, qui sont manquantes dans la plupart des régions.

Cette thèse s'est attachée à développer et évaluer une nouvelle approche de télédétection, basée sur le traitement et l'analyse d'imagerie satellitaire optique, qui combine le potentiel conjoint des techniques de réduction de la variabilité spectrale et des méthodes de classification non supervisée pour la cartographie des systèmes agricoles à deux niveaux d'organisation spatiale : le paysage agricole et la parcelle. La principale originalité de ce travail de thèse réside dans la mise en place de méthodes simples et facilement reproductibles, pratiquement entièrement basées sur des données satellitaires, contribuant au développement d'outils génériques pour appuyer le suivi de l'agriculture à grande échelle.

Les méthodes ont été développées sur l'État du Tocantins au Brésil, une région où les activités agricoles s'organisent sur une longue saison culturale et principalement autour d'une production commerciale dans de grandes exploitations. Elles ont ensuite été testées au Burkina Faso, une région représentative de la « petite agriculture familiale » de l'Afrique de l'Ouest, caractérisée par un parcellaire de petite taille, avec une forte variabilité intra-parcellaire et par des activités agricoles synchronisées sur une courte saison de pluies. Ceci nous a permis d'évaluer la contribution des méthodes développées et de relever les limites et perspectives de reproductibilité dans des régions où les contraintes pour la télédétection sont majeures.

Contributions de la thèse

Apports méthodologiques

Pour la cartographie des systèmes d'utilisation agricole des terres à l'échelle régionale

Les méthodes de télédétection semi-automatiques développées pour la cartographie des systèmes d'utilisation agricole des terres à l'échelle régionale sont originales par rapport aux méthodes utilisées dans les approches traditionnelles, basées généralement sur

la délimitation d'unités cartographiques par des experts sur un choix subjectif de variables hétérogènes.

La série annuelle de NDVI du produit MOD13Q1 à 250 m de résolution spatiale s'est révélée une source majeure d'informations spectrales et temporelles liées à l'utilisation agricole des terres. La méthode innovante d'extraction des « variables phénologiques » a permis de dissocier de façon automatique la part de l'information liée à la variabilité temporelle des valeurs de NDVI, de celle liée à la variabilité spatiale dans la série annuelle de NDVI. Les images composantes principales retenues semblent pertinentes pour la délimitation des unités de paysage agricole car elles conservent le détail temporel des variations saisonnières et intra-saisonnières qui caractérisent les activités agricoles, tout en le dissociant du reste de l'information de la série annuelle.

La méthode de segmentation à l'aide de l'algorithme multi-résolution de Baatz et Schäpe (2000) a permis d'automatiser la délimitation des unités de paysage ayant une phénologie relativement homogène (basée sur le critère d'homogénéité de l'algorithme). Cette méthode permet de :

- minimiser l'intervention de l'opérateur : la délimitation des unités de paysage se fait automatiquement, à partir de l'ensemble de variables fournies ;
- réduire la quantité de données nécessaires : seules les informations synthétiques dans les valeurs numériques des « variables phénologiques » sont utilisées ;
- délimiter des unités de paysage relativement homogènes à différents niveaux organisationnels : en faisant varier le paramètre d'échelle de l'algorithme multi-résolution.

Ce paramètre d'échelle a été fixé empiriquement pour pouvoir caractériser l'organisation spatiale des systèmes agricoles à un niveau supérieur à la parcelle. Toutefois, nous tenons à souligner le caractère multiniveau des unités de paysage ; le paramètre d'échelle peut être ajusté selon l'objet d'étude, comme dans les approches en écologie des paysages où les unités de paysage sont représentées à plusieurs échelles afin de déterminer laquelle est la plus pertinente pour les processus écologiques étudiés.

Enfin, l'arbre de décision et l'analyse visuelle des profils temporels de NDVI nous ont permis de caractériser les unités de paysage, et de révéler l'organisation spatiale des différents types de systèmes d'élevage (décrits par leur niveau d'intensité) et de systèmes de culture, (décrits par les types et leurs modes de culture). Ces deux étapes méthodologiques simples permettent ainsi de transférer des informations locales, contenues dans une carte d'occupation

des terres à l'échelle de la parcelle et dans la série temporelle de NDVI, à l'échelle des unités de paysage.

Pour la cartographie des systèmes de culture à l'échelle régionale

Une des principales contraintes pour la cartographie des systèmes de culture à l'échelle régionale par télédétection est la forte variabilité spectrale des surfaces liée à la grande diversité des conditions environnementales et des pratiques agricoles au sein d'une région. Dans les études visant ce même objectif de cartographie, les méthodes de classification supervisées prédominent, mais elles nécessitent des campagnes de terrain consommatrices en ressources pour collecter des données d'entraînement.

Dans ce travail de thèse nous avons proposé une méthode originale de classification non supervisée, en suivant une approche orientée-objet, basée exclusivement sur des données satellitaires. Sur l'État du Tocantins, cela a permis de cartographier trois systèmes de culture, décrits par le type de culture et le mode de culture, avec une précision globale d'environ 93 %. Cette méthode de classification bénéficie de deux techniques de réduction de la variabilité spectrale (une segmentation à l'échelle de la parcelle et la stratification régionale en unités de paysage) et d'une étape d'hyperclustering qui maximise la séparabilité du jeu de données initial en un grand nombre de classes spectrales, garantissant que les classes d'intérêt soient dissociées des autres classes dans un nombre final de clusters réduit.

La méthode utilise exclusivement des données disponibles gratuitement à ce jour (MODIS et Landsat). Elle peut potentiellement être reproduite dans d'autres régions avec des parcelles de taille moyenne à grande (à partir d'environ 8 hectares), mais peut nécessiter un ajustement des paramètres de segmentation à l'échelle de la parcelle. Les résultats ont montré une contribution significative de la stratification régionale en unités de paysage dans l'amélioration de la performance de la classification. La classification non supervisée n'a pas besoin de l'intervention a priori de l'utilisateur sur la définition des classes **-** la séparabilité des informations contenues dans les images satellitaires étant objectivement basée sur la différence ou ressemblance spectrale entre classes - mais nécessite une connaissance du calendrier cultural de la région d'étude pour l'interprétation finale des résultats.

Enseignements tirés du cas d'application au Burkina Faso

Le cas d'application au Burkina Faso nous a permis de tester la reproductibilité des méthodes développées face à un changement d'échelle dicté par la taille du parcellaire des petites agricultures de la zone, et face à la disponibilité de l'imagerie satellitaire actuelle.

Ce travail a aussi contribué à définir des perspectives liées à l'offre en imagerie satellitaire future pour la cartographie des systèmes agricoles au Burkina Faso.

Les résultats ont montré que les limites de reproductibilité des méthodes au Burkina Faso résultent principalement des spécificités de la région (parcellaire fragmenté et de petite taille, grande variabilité intra-parcellaire et forte couverture nuageuse pendant une courte saison culturale), ainsi que de la disponibilité limitée en imagerie satellitaire (manque de produits qui combinent une haute résolution spatiale et temporelle, et prix élevé des images à très haute résolution spatiale). Ces limites ont nécessité l'adaptation des méthodes utilisées, notamment le remplacement de la classification orientée-objet par une classification pixellaire, l'augmentation du nombre de clusters, et la réduction du nombre de classes finales. Ne pouvant pas accéder à une information spectrale « pure » à l'intérieur des parcelles agricoles avec des données MODIS, la cartographie des systèmes de culture et des systèmes d'utilisation agricole des terres à l'échelle régionale n'a pas été possible, et seul le domaine cultivé a pu être cartographié, avec une précision globale d'environ 65 %. Les principales sources d'erreur de classification sont dues au mélange d'informations spectrales contenues dans les pixels MODIS, qui tend à homogénéiser la radiométrie des pixels, et à la similarité des propriétés spectrales et temporelles des cultures et de la végétation naturelle.

Globalement, les résultats ont montré que la série annuelle de MODIS-NDVI reste toujours adaptée à la méthode de délimitation des unités de paysage, l'extraction des « variables phénologiques » à l'aide d'une ACP et la stratification régionale ayant produit des résultats équivalents à ceux obtenus pour le Tocantins. En ce qui concerne la méthode de classification des systèmes de culture, la série annuelle de MODIS devrait être remplacée par une série annuelle à une résolution décamétrique afin d'accéder à une information spectrale plus « pure » à l'intérieur des parcelles et donc à l'information spectrale et temporelle réellement associée aux différents systèmes de culture. Les nouvelles données Sentinel-2 réunissent les caractéristiques nécessaires pour réaliser ce changement d'échelle, mais la première série annuelle complète (avec des acquisitions tous les 5 jours) ne devrait pas être disponible avant la fin de l'année 2018.

Apports thématiques

Au Brésil, les cartes des principaux systèmes d'utilisation agricole des terres (saison culturale 2013-2014) et des systèmes de culture (saison 2014-2015) du Tocantins apportent une information complémentaire, détaillée et spatialement explicite, de celle fournie par les statistiques agricoles de la région qui sont agrégées à l'échelle des municipalités. Cette information peut être utile aux acteurs locaux potentiellement concernés par la répartition

spatiale des activités agricoles dans la région (ex. investisseurs du secteur agricole, coopératives et syndicaux agricoles, producteurs, institutions gouvernementales, centres de recherche et développement) et peut être utilisée à la place des statistiques agricoles dans les travaux destinés à la planification de l'utilisation des terres à grande échelle (ex. les travaux présentés dans la section 1.2.5.1).

Ces cartes contribuent également à augmenter la visibilité internationale du front pionnier qui s'installe dans le biome du Cerrado dans le cadre du développement de la région de MATOPIBA. Elles fournissent des indicateurs d'extension des systèmes de production de soja et de pratiques culturales de double culture et monoculture à partir desquels des indicateurs d'intensification de pratiques peuvent être extraits. Ces informations pourraient potentiellement servir à la communauté scientifique pour évaluer le potentiel d'intensification des terres cultivées en vue de conserver les zones de Cerrado naturelles, ce qui est une préoccupation majeure actuellement comme le montrent les travaux récents de Spera (2017). Ces indicateurs peuvent à terme guider la planification de l'utilisation des terres dans la région et servir d'exemple pour les états brésiliens voisins.

Au Burkina Faso, la carte du domaine cultivé en 2016 apporte une information actualisée par rapport aux produits cartographiques existants (présentés dans la section 5.1). Cette carte fournit des indicateurs sur l'extension des cultures annuelles qui est une donnée importante pour la communauté scientifique et pour les acteurs du territoire qui n'ont pas accès aux statistiques agricoles à l'échelle sous-nationale. Cependant, cette carte est produite à 250 m de résolution spatiale et sa précision globale, estimée à 65 %, permet seulement une utilisation qualitative. Par exemple, elle peut être utilisée pour dériver des indicateurs spatiaux sur l'état de conservation des espaces protégés et aider à repérer les conflits émergents pour la ressource foncière toujours plus rare.

La carte du domaine cultivé, superposée à la stratification régionale en unités de paysage, peut également servir à produire des indicateurs régionaux sur l'intensité d'utilisation des terres par l'agriculture, permettant d'évaluer le degré de développement agricole et de saturation de l'espace dans les différentes unités spatiales. Ces indicateurs sont intéressants dans un contexte de prévision de l'installation des nouvelles zones agricoles et la planification de l'utilisation future des terres. Cette superposition des informations spatiales peut également aider à cibler des zones à prospecter avec des acquisitions d'images à haute résolution spatiale pour une étude plus fine des principaux systèmes agricoles sur place.

Perspectives

Plusieurs pistes de recherche et d'applications opérationnelles se dégagent de cette thèse. Notamment, la généricité des méthodes présentées mérite d'être évaluée sur d'autres régions et dans le temps pour identifier des améliorations potentielles dans une optique d'opérationnalisation des outils. Nous présentons ici quelques perspectives potentielles et les développements réalisés et envisagés pour mettre à disposition les outils dérivés de cette thèse et permettre de reproduire les méthodes développées.

Perspectives de recherche

Pour la cartographie des systèmes d'utilisation agricole des terres

Si l'on applique les mêmes traitements de délimitation des unités de paysage aux mêmes types de données, on peut s'attendre à obtenir des unités spatiales comparables d'une année sur l'autre pour une même région et entre régions (à quelques altérations près dans la forme et taille des unités en raison de l'évolution naturelle des surfaces et des différences régionales). Cette question n'a pas été abordée dans la thèse et mériterait d'être testée pour évaluer la robustesse de la méthode notamment face aux variations inter-annuelles des conditions climatiques.

La caractérisation des unités de paysage en termes de proportion des différents types d'occupation et utilisation agricole des terres, pourrait être complétée par d'autres informations. Par exemple, des variables complémentaires dérivées de données satellitaires (ex. altitude et pente issues des modèles numériques de terrain), de données externes environnementales (ex. types des sols, érodibilité des sols, données climatiques) ou de données socio-économiques (ex. accès aux axes routiers, incitations fiscales, infrastructures agricoles, différences sociales dans la communauté de producteurs) pourraient être croisées avec la stratification régionale des paysages pour révéler des facteurs sous-jacents à l'organisation spatiale des systèmes d'utilisation des terres au sein de la région. L'analyse de ces facteurs en complément de l'occupation et de l'utilisation agricole des terres au niveau du paysage permettrait une meilleure identification des voies d'intensification agricole, et ainsi de mieux orienter la gestion et la planification de l'utilisation des terres pour mitiger les impacts sur l'environnement (ex. Rizzo *et al.*, 2013; Rufino *et al.*, 2016).

Par ailleurs, des indicateurs paysagers pourraient être calculés à partir de la carte de systèmes de culture, établie rappelons-le à l'échelle de la parcelle, pour caractériser davantage l'organisation spatiale des systèmes de culture à l'intérieur des unités de paysage. Par exemple, des indices de composition (ex. indices de diversité : uniformité, richesse),

de configuration (ex. taille moyenne des parcelles, indices de complexité de la forme des parcelles) ou encore de connectivité (ex. indice de cohésion) peuvent fournir des informations supplémentaires sur les systèmes d'utilisation des terres au niveau du paysage et permettre notamment d'extraire des indicateurs sur leur « performance » et leur « durabilité » (ex. Colson, Bogaert et Ceulemans, 2011; Panigrahy, Manjunath et Ray, 2005; Plexida *et al.*, 2014; Ribeiro *et al.*, 2016).

Pour la cartographie des systèmes de culture à l'échelle régionale

La cartographie des systèmes agricoles à l'échelle de la parcelle a porté sur la détection de systèmes de cultures annuelles. Cependant, la détection des systèmes de culture pérennes et les systèmes d'élevage (au travers de la détection des pâturages) permettrait de représenter toute la diversité de systèmes agricoles et contribuerait à affiner notre démarche. Ces informations spatiales sont fondamentales pour le suivi et la planification de l'utilisation des terres dans de nombreuses régions du monde et mériteraient d'être traitées de façon approfondie.

Par exemple, au Tocantins, même si les systèmes de culture pérennes (ex. plantations d'eucalyptus ou hévéa, canne à sucre) présentent à ce jour une extension mineure dans la région par rapport à l'extension des systèmes de cultures annuelles dédiés à la production intensive de soja, leurs surfaces évoluent rapidement tirées par la demande. L'étude récente de de Souza *et al.*, (2017) a montré la forte extension des systèmes de canne à sucre dans la municipalité de Pedro Afonso au Tocantins entre 2008 et 2013 pour répondre à la demande en biocombustible et conclut sur l'importance du suivi de ces surfaces. Toujours au Brésil, et plus particulièrement dans la région du MATOPIBA, le suivi de l'état de dégradation des pâturages est une priorité nationale, les zones dégradées étant des zones privilégiées pour l'extension de la production de soja (Spera, 2017). La mise en culture ou la récupération des pâturages dégradés pour l'élevage représentent des stratégies de gestion décisives pour l'intensification de la production dans cette région car elles évitent l'ouverture de nouvelles zones et favorisent la préservation des espaces naturels menacés (Spera, 2017).

Les cultures pérennes et les pâturages présentent des comportements phénologiques qui peuvent être confondus à l'échelle régionale avec d'autres types de surfaces végétalisées. Il est ainsi nécessaire d'explorer des méthodes permettant de mieux séparer ces différentes surfaces. Nous présentons ci-dessous de façon synthétique quelques pistes à explorer qui nous semblent pertinentes pour maximiser la séparabilité de classes :

- tester l'influence de la longueur des séries temporelles d'images de NDVI (ex. tester des séries pluriannuelles et des séries resserrées sur des périodes intra-annuelles d'intérêt) ;
- tester l'apport de variables phénologiques telles que le cumul annuel, le maximum annuel et l'amplitude annuelle du NDVI ;
- tester l'apport de variables texturales calculées à partir d'images à haute résolution ;
- tester l'apport d'étapes de classification hiérarchiques.

En outre, compte tenu des limites de reproductibilité de la méthode de cartographie des systèmes de culture sur des régions caractérisées par des parcelles de petite taille avec la résolution spatiale des données MODIS et Landsat, la méthode mérite d'être adaptée aux nouvelles sources de données satellitaires. En particulier, les données de la constellation Sentinel-2 à 10 m de résolution spatiale et 5 jours de résolution temporelle sont des produits clés qui vont ouvrir de nouveaux axes de recherche pour la cartographie des systèmes de culture de ces régions et contribuer ainsi au suivi de la petite agriculture familiale.

Sur l'échelle temporelle de la cartographie

Les travaux présentés portent sur une cartographie des systèmes agricoles à un pas de temps annuel. Cette échelle temporelle permet de caractériser les systèmes agricoles par les cultures et les pratiques culturales adoptées pendant une saison culturale. Une mise à jour annuelle de ces informations spatiales est bien adaptée à l'échelle temporelle des politiques publiques et de la planification de l'utilisation des terres (à une échelle annuelle voire pluriannuelle). Cependant, d'autres échelles temporelles, pertinentes pour le suivi de la production en cours de saison (échelle intra-annuelle) et pour caractériser les trajectoires de l'utilisation des terres (échelle interannuelle), n'ont pas été traitées dans cette thèse.

Par exemple, des études récentes ont montré l'intérêt de techniques innovantes de télédétection pour la cartographie des types de culture en cours de saison, comme l'entraînement de classifications avec des données interannuelles (Zhong *et al.*, 2016), l'entraînement de réseaux de neurones à partir de variables biophysiques issues de données multi-capteur (Waldner *et al.*, 2015) ou l'utilisation du couplage des données radar et optiques et de variables texturales (Inglada *et al.*, 2016). Les cartes de types de culture en cours de saison issues de ces types de méthodes sont particulièrement intéressantes pour la prévision de la production dans les régions en situation d'insécurité alimentaire car elles permettent d'évaluer la diversification de la production avant la fin de la saison et peuvent être utilisées en entrée des modèles de prévision des rendements des cultures. La temporalité

de ces informations est essentielle pour les systèmes de surveillance tels que les systèmes d'alerte précoce (ex. FEWS NET) et contribuent à une planification efficace de l'utilisation des ressources agricoles. Il semblerait ainsi intéressant d'évaluer la capacité des méthodes simples basées sur des algorithmes non supervisées de clustering tells que *k*-means ou ISODATA pour une cartographie des systèmes de culture en cours de saison culturale.

Comme présenté dans le chapitre introductif (cf. Tableau 1) certaines pratiques culturales ne sont détectables que d'une année sur l'autre, telles que les rotations culturales ou les jachères. En ce qui concerne le suivi des trajectoires de l'utilisation agricole des terres pour évaluer des processus tels que l'intensification, l'extension des surfaces agricoles ou la mise en friche, il est aussi nécessaire de cartographier les systèmes agricoles sur plusieurs années. Différents travaux ont montré l'intérêt du suivi pluriannuel de l'agriculture à l'échelle de la parcelle à partir duquel les transitions peuvent être quantifiées avec exactitude (ex. Arvor *et al.*, 2012), mais l'analyse des transitions dans les paysages agricoles semble aussi pertinente et commence à être explorée dans la communauté de télédétection (ex. Eckert *et al.*, 2017). Ainsi, il pourrait être intéressant de reproduire les deux types de cartes proposées dans cette thèse sur plusieurs années et d'analyser les différents indicateurs qui pourraient être dérivés aux deux niveaux d'organisation spatiale.

Perspectives opérationnelles

La reproductibilité de la méthode « landscape-clustering » pour la cartographie des systèmes de culture sera prochainement testée sur l'État du Tocantins et sur l'État voisin du Mato Grosso pour la saison culturale 2015-2016. Des jeux de données de vérité terrain ont été collectés pour évaluer les résultats des classifications pour ces deux périodes et ainsi évaluer la robustesse de la méthode dans le temps et dans une autre région du Brésil. Sa robustesse pourrait être aussi évaluée en l'appliquant dans d'autres régions du monde caractérisées par des parcelles de taille moyenne à grande (considérant les mêmes données satellitaires en entrée).

Pour faciliter sa reproductibilité, un outil de stratification, fonctionnel et libre, a été développé et mis à disposition sous forme d'application en tant qu'extension de la librairie *open-source* Orfeo Toolbox (OTB). L'application "LandscapeStratification" du module "NDVITimeSeries" implémentée par Rémi Cresson (Irstea, UMR TETIS) est disponible sur : https://gitlab.irstea.fr/remi.cresson/NDVITimeSeries. Cette application utilise la segmentation large échelle par fusion de région (Lassalle *et al.*, 2015), incluant le critère d'homogénéité de Baatz et Schäpe (2000) implémentée dans OTB (remplaçant ainsi l'outil "Multiresolution Segmentation" implémenté dans le logiciel commercial

"eCognition® Developer"). Ainsi la totalité des méthodes présentées dans cette thèse peuvent être reproduites avec des librairies de traitement d'imagerie satellitaire sur des logiciels libres tels que Monteverdi, R ou QGIS.

Pour que les méthodes soient également indépendantes de données externes, autres que les données MODIS et Landsat, la méthode de cartographie des systèmes d'utilisation agricole des terres bénéficierait de la possibilité de pouvoir remplacer la carte d'occupation des terres (les données TerraClass dans le cas d'étude au Tocantins) par la carte de systèmes de culture produite suivant la méthode de « landscape-clustering ». Ainsi, les résultats de la stratification régionale en unités de paysage serviraient à la classification des systèmes de culture, et en retour, la carte des systèmes de culture servirait à caractériser les unités de paysage pour obtenir une carte de systèmes d'utilisation agricole des terres. La classification des unités de paysage à partir de la carte de systèmes de culture profiterait néanmoins d'un nouvel arbre de décision avec des règles de classification adaptées, et l'analyse visuelle des profils temporels de NDVI ne serait plus nécessaire car l'information serait directement disponible sur la carte de systèmes de culture. L'approche multiscalaire développée serait ainsi simplifiée et entièrement basée sur des images satellitaires.

Références bibliographiques

A

AfDB. (2011). *Amélioration des statistiques relatives à la sécurité alimentaire, à l'agriculture durable, et au développement rural Plan d'action pour l'Afrique 2011-2015*. African Development Bank. Consulté à l'adresse : https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/AfDB Agriculture_long_FR_web.pdf

Akoudjin, M., Kiéma, S., Sangare, M., César, J., Bouyer, J., et Kabore-Zoungrana, C. (2016). Influence des activités agricoles sur la végétation le long d'un gradient pluviométrique nord-sud du Burkina Faso. *VertigO-La Revue Électronique En Sciences de L'environnement*, *16*(1). http://doi.org/10.4000/vertigo.17015

Alexandratos, N., et Bruinsma, J. (2012). *World agriculture towards 2030/2050: the 2012 revision* (No. ESA/12-03). FAO, Rome.

Arino, O., Gross, D., Ranera, F., Leroy, M., Bicheron, P., Brockman, C., … Weber, J. L. (2007). GlobCover: ESA service for Global land cover from MERIS. In *International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 2412–2415). http://doi.org/10.1109/IGARSS.2007.4423328

Arvor, D., Jonathan, M., Meirelles, M. S. P., Dubreuil, V., et Durieux, L. (2011). Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *International Journal of Remote Sensing*, *32*(22), 7847–7871. http://doi.org/10.1080/01431161.2010.531783

Arvor, D., Meirelles, M., Dubreuil, V., Bégué, A., et Shimabukuro, Y. E. (2012). Analyzing the agricultural transition in Mato Grosso, Brazil, using satellite-derived indices. *Applied Geography*, *32*(2), 702–713. http://doi.org/10.1016/j.apgeog.2011.08.007

Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing*, *5*(2), 949–981. http://doi.org/10.3390/rs5020949

B

Baatz, M., et Schäpe, A. (2000). Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII*, *58*, 12–23.

BAfD. (2015). Développement territorial et inclusion spatiale, Perspectives économiques en Afrique 2015. Banque africaine de développement (BAfD), Organisation de coopération et de développement économiques (OCDE), Programme des Nations Unies pour le développement (PNUD).

Barona, E., Ramankutty, N., Hyman, G., et Coomes, O. T. (2010). The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environmental Research Letters*, *5*(2). http://doi.org/10.1088/1748-9326/5/2/024002

Barros de Sousa, P. A., Teixeira Borges, R. S., et Ribeiro Dias, R. (2012). *Atlas do Tocantins: subsídios ao planejamento da gestão territorial*. Palmas. Consulté à l'adresse : http://web.seplan.to.gov.br/Arquivos/download/ZEE/TO_AtlasTocantins2012/Atlas_do_Tocantins_2012.pdf

Bégué, A., Arvor, D., Bellon, B., Betbeder, J., de Abelleyra, D., Ferraz, R. P. D., … R. Verón, S. (2018). Remote Sensing and Cropping Practices: A Review. *Remote Sensing*, *10*(1), 99. http://doi.org/10.3390/rs10010099

Bégué, A., Arvor, D., Lelong, C., Vintrou, E., et Simoes, M. (2015). Agricultural Systems Studies Using Remote Sensing. In *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing* (pp. 113–130). CRC Press.

Bégué, A., Leroux, L., Lo Seen, D., Tonneau, J. P., et Morant, P. (2016). Observation spatiale pour l'agriculture en Afrique: potentiels et défis. (B. Charrier et C. Torre, Eds.). Notes techniques: AFD, 12. Paris: AFD-CIRAD.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., et Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*. http://doi.org/10.1016/j.isprsjprs.2003.10.002

Bey, A., Sánchez-Paus Díaz, A., Maniatis, D., Marchi, G., Mollicone, D., Ricci, S., … Miceli, G. (2016). Collect Earth: Land Use and Land Cover Assessment through Augmented Visual Interpretation. *Remote Sensing*, *8*(10).

Biancalani, R., Nachtergaele, F., Petri, M., et Bunning, S. (2013). Land degradation assessment in drylands: methodology and results. *LADA Project Report, FAO*.

Bisquert, M., Bégué, A., et Deshayes, M. (2015). Object-based delineation of homogeneous landscape units at regional scale based on MODIS time series. *International Journal of Applied Earth Observation and Geoinformation*, *37*, 72–82. http://doi.org/10.1016/j.jag.2014.10.004

Bisquert, M., Bégué, A., Deshayes, M., et Ducrot, D. (2017). Environmental evaluation of MODIS-derived land units. *GIScience et Remote Sensing*, *54*(1), 64–77. http://doi.org/10.1080/15481603.2016.1256861

Bisquert, M., Bordogna, G., Bégué, A., Candiani, G., Teisseire, M., et Poncelet, P. (2015). A simple fusion method for image time series based on the estimation of image temporal validity. *Remote Sensing*, *7*(1), 704–724. http://doi.org/10.3390/rs70100704

Bivand, R., Keitt, T., et Rowlingson, B. (2016). rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 1.2.7. Consulté à l'adresse : https://cran.r-project.org/package=rgdal

Bivand, R., et Lewin-Koh, N. (2016). maptools: Tools for Reading and Handling Spatial Objects. R package version 0.9.2. Disponible à l'adresse : https://cran.r-project.org/package=maptools

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., … Tiede, D. (2014). Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing : Official Publication of the International Society for Photogrammetry and Remote Sensing (ISPRS)*, *87*(100), 180–191. http://doi.org/10.1016/j.isprsjprs.2013.09.014

Blaschke, T., Johansen, K., et Tiede, D. (2011). Object-Based Image Analysis for Vegetation Mapping and Monitoring. *Advances in Environmental Remote Sensing: Sensor, Algorithms, and Applications*, (January), 241–271. http://doi.org/10.1201/b10599-13

Bolfe, É. L., Victória, D. D. C., Contini, E., Bayma-Silva, G., Spinelli-Araujo, L., et Gomes, D. (2016). Matopiba em crescimento agrícola Aspectos territoriais e socioeconômicos. *Revista de Política Agrícola*, *25*(4), 38–62.

Bonny, S. (2010). L'intensification écologique de l'agriculture : voies et défis. In *Innovation and Sustainable Development in Agriculture and Food Symposium* (pp. 1–11).

Borghi, E., Luchiari Junior, A., Avanzi, J. C., Bortolon, L., Oliveira Bortolon, E. S., Motta Campos, L. J., et Tavares Correa, L. V. (2015). *Estado da arte da agricultura e pecuária do estado do Tocantins*. *Embrapa Pesca e Aquicultura. Documentos 13*. Palmas. Consulté à l'adresse : http://www.agrisus.org.br/arquivos/Embrapa_Tocantins.pdf

Boryan, C., Yang, Z., Mueller, R., et Craig, M. (2011). Monitoring US agriculture: The US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International*, *26*(5), 341–358. http://doi.org/10.1080/10106049.2011.562309

Bouma, J. (1997). The land use systems approach to planning sustainable land management at several scales. *ITC Journal*, *1997*(3/4), 237–242.

Bourget, É., et Le Dû-Blayo, L. (2010). Définition d'unités paysagères par télédétection en Bretagne: méthodes et critiques. *Norois*, (3), 69–83.

Bridhikitti, A., et Overcamp, T. J. (2012). Estimation of Southeast Asian rice paddy areas with different ecosystems from moderate-resolution satellite imagery. *Agriculture, Ecosystems and Environment*, *146*(1), 113–120. http://doi.org/10.1016/j.agee.2011.10.016

Brossier, J. (1987). Système et système de production. *Cahiers Des Sciences Humaines*, *23*, 377-390.

Brown, J. C., Kastens, J. H., Coutinho, A. C., Victoria, D. de C., et Bishop, C. R. (2013). Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data. *Remote Sensing of Environment*, *130*, 39–50. http://doi.org/10.1016/j.rse.2012.11.009

Bruinsma, J. (2011). The resources outlook: by how much do land, water and crop yields need to increase by 2050? In P. Conforti (Ed.), *Looking ahead in world food and agriculture: perspectives to 2050.* (pp. 233–278). FAO, Rome.

Bunker, B. E., Tullis, J. A., Cothren, J. D., Casana, J., et Aly, M. H. (2016). Object-based Dimensionality Reduction in Land Surface Phenology Classification. *AIMS Geosciences*, *2*(4), 302–328.

Burger, P., Berton, S., Billaz, R., et Lebreton, A. (2011). Agroécologie, une transition vers des modes de vie et de développement viables. Paroles d'acteurs. *Groupe de Travail Désertification*, 1–93.

Busetto, L., et Ranghetti, L. (2016). MODIStsp: An R package for automatic preprocessing of MODIS Land Products time series. *Computers et Geosciences*, *97*, 40–48.

C

Cai, H., Zhang, S., Bu, K., Yang, J., et Chang, L. (2011). Integrating geographical data and phenological characteristics derived from MODIS data for improving land cover mapping. *Journal of Geographical Sciences*, *21*(4), 705–718.

Cano, E., Denux, J.-P., Bisquert, M., Hubert-Moy, L., et Chéret, V. (2017). Improved forest-cover mapping based on MODIS time series and landscape stratification. *International Journal of Remote Sensing*, *38*(7), 1865–1888. http://doi.org/10.1080/01431161.2017.1280635

Cheema, M. J. M., et Bastiaanssen, W. G. M. (2010). Land use and land cover classification in the irrigated Indus Basin using growth phenology information from satellite data to support water management analysis. *Agricultural Water Management*, *97*(10), 1541–1552. http://doi.org/10.1016/j.agwat.2010.05.009

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., … Mills, J. (2015). Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 7–27. http://doi.org/http://dx.doi.org/10.1016/j.isprsjprs.2014.09.002

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., et Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, *91*(3), 332–344. http://doi.org/10.1016/j.rse.2004.03.014

Cihlar, J. (2000). Land cover mapping of large areas from satellites: Status and research priorities. *International Journal of Remote Sensing*, *21*(6), 1093–1114. http://doi.org/10.1080/014311600210092

CILSS. (2016). *Les Paysages de l'Afrique de l'Ouest: Une Fenêtre sur un monde en pleine évolution*. Comité Inter-états de Lutte contre la Sécheresse dans le Sahel (CILSS), U.S. Geological Survey Earth Resources Observation and Science (USGS EROS) Center, United States Agency for International Development (USAID).

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Collin, Y. (2012). Rapport d'information n°504 fait au nom de la délégation sénatoriale à la prospective sur le défi alimentaire à l'horizon 2050. Sénat, Délégation à la prospective, le 18 Avril 2012.

Colson, F., Bogaert, J., et Ceulemans, R. (2011). Fragmentation in the Legal Amazon, Brazil: Can landscape metrics indicate agricultural policy differences? *Ecological Indicators*, *11*(5), 1467-1471. http://doi.org/10.1016/j.ecolind.2010.12.020

CONAB. (2018). Soja: série histórica de produtividade (safras 2000/01 a 2016/17). Disponible à l'adresse : http://www.conab.gov.br/

Conseil de l'Europe. (2000). Convention européenne du paysage. *Série Des Traités Européens - N° 176*. Florence, Italie.

Corgne, S., Hubert-Moy, L., et Betbeder, J. (2016). Monitoring of Agricultural Landscapes Using Remote Sensing Data. In N. Baghdadi et M. Zribi (Eds.), *Land Surface Remote Sensing in Agriculture and Forest* (pp. 221–247). ISTE Press Ltd, Elsevier Ltd. http://doi.org/10.1016/B978-1-78548-103-1.50006-6

D

Daughtry, C. S. T., Doraiswamy, P. C., Hunt Jr., E. R., Stern, A. J., McMurtrey III, J. E., et Prueger, J. H. (2006). Remote sensing of crop residue cover and soil tillage intensity. *Soil and Tillage Research*, *91*(1–2), 101–108. http://doi.org/10.1016/j.still.2005.11.013

de Almeida, C. A. (2016). *Paysage des systèmes de production agropastoraux de l'État du Rondônia-Amazonie brésilienne*. Thèse de doctorat. Université Montpellier.

de Almeida, C. A., Coutinho, A. C., Esquerdo, J. C. D. M., Adami, M., Venturieri, A., Diniz, C. G., … Gomes, A. R. (2016). High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amazonica*, *46*(3), 291–302. http://doi.org/10.1590/1809-4392201505504

de Almeida, C. A., Mourão, M., Dessay, N., Lacques, A.-E., Monteiro, A., Durieux, L., … Seyler, F. (2016). Typologies and Spatialization of Agricultural Production Systems in Rondônia, Brazil: Linking Land Use, Socioeconomics and Territorial Configuration. *Land*, *5*(2), 18. http://doi.org/10.3390/land5020018

de Almeida, T. I. R., Penatti, N. C., Ferreira, L. G., Arantes, A. E., et do Amaral, C. H. (2015). Principal component analysis applied to a time series of MODIS images: the spatio-temporal variability of the Pantanal wetland, Brazil. *Wetlands Ecology and Management*, *23*(4), 737-748. http://doi.org/10.1007/s11273-015-9416-4

de Bie, C. A. J. M. (2000). *Comparative performance analysis of agro-ecosystems.* Thèse de doctorat. Wageningen University, The Netherlands.

de Miranda, E. E., Magalhães, L. A., et de Carvalho, C. A. (2014). *Proposta de delimitação territorial do Matopiba. Nota Técnica 1.* Campinas, São Paulo. Consulté à l'adresse : https://www.embrapa.br/gite/publicacoes/NT1_DelimitacaoMatopiba.pdf

de Schutter, O. (2010). *Agroecology and the Right to Food. Report presented at the 16th Session of the United Nations Human Rights Council [A/HRC/16/49].* Consulté à l'adresse : http://www.srfood.org/images/stories/pdf/officialreports/20110308_a-hrc-16-49_agroecology_en.pdf

de Souza, C. H. W., Cervi, W. R., Brown, J. C., Rocha, J. V., et Lamparelli, R. A. C. (2017). Mapping and evaluating sugarcane expansion in Brazil's savanna using MODIS and intensity analysis: a case-study from the state of Tocantins. *Journal of Land Use Science*, *12*(6), 457-476. http://doi.org/10.1080/1747423X.2017.1404647

Deffontaines, J.-P. (1973). Analyse du paysage et étude régionale des systèmes de production agricole. *Economie Rurale*, *98*, 3–13.

Deffontaines, J.-P. (1997). Du paysage comme moyen de connaissance de l'activité agricole à l'activité agricole comme moyen de production du paysage. In *Thème et variations: nouvelles recherches rurales au sud* (pp. 305–322). ORSTOM, Paris. Consulté à l'adresse : http://www.documentation.ird.fr/hor/fdi:010012810

Deffontaines, J.-P., Thenail, C., et Baudry, J. (1995). Agricultural systems and landscape patterns: how can we build a relationship? *Landscape and Urban Planning*, *31*(1–3), 3–10. http://doi.org/10.1016/0169-2046(94)01031-3

Dembélé, Y. (2010). *Cartographie des zones socio-rurales : un outil d'aide à la planification pour la gestion de l'eau en agriculture. Fao Water.*

DGESS/MAAH. (2017). *Rapport général des résultats définitifs de la campagne agricole 2016/2017 et des perspectives de la situation alimentaire et nutritionnelle.* Consulté à l'adresse : http://www.sisa.bf/sisa/IMG/pdf/rapport_general_resultats_definitifs_campagne_agricole_2016-2017_def.pdf

Dheeravath, V., Thenkabail, P. S., Chandrakantha, G., Noojipady, P., Reddy, G. P. O., Biradar, C. M., … Velpuri, M. (2010). Irrigated areas of India derived using MODIS 500 m time series for the years 2001-2003. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 42–59. http://doi.org/10.1016/j.isprsjprs.2009.08.004

Di Gregorio, A., et Jansen, L. J. M. (2000). *Land Cover Classification System (LCCS): Classification Concepts and User Manual. FAO* (Vol. 53).

Dixon, J., Gulliver, A., et Gibbon, D. (2001). *Farming Systems and Poverty: Improving Farmers' Livelihoods in a Changing World. FAO et World Bank, Rome, Italy et Washington DC, USA.*

Driessen, P. M., et Konijn, N. T. (1992). *Land-use systems analysis*. Wageningen Agricultural University Wageningen, The Netherlands.

E

Eastman, J. R., et Fulk, M. (1993). Long sequence time series evaluation using standardized principal components. *Photogrammetric Engineering and Remote Sensing*, *59*(6), 991–996.

Eckert, S., Kiteme, B., Njuguna, E., et Zaehringer, J. G. (2017). Agricultural expansion and intensification in the foothills of Mount Kenya: A landscape perspective. *Remote Sensing*, *9*(8). http://doi.org/10.3390/rs9080784

Even, M.-A., et Vert, J. (2011). Centre d'études et de prospective du Ministère de l'agriculture, de l'alimentation, de la pêche, de la ruralité et de l'aménagement du territoire. *Analyse N° 28, Juin 2011*.

F

FAO. (s.d.-a). Agro-MAPS (Database). Disponible à l'adresse : http://ref.data.fao.org/database?entryId=7c7aaeb7-7de5-45b2-9024-fbb4315dc771

FAO. (s.d.-b). AQUASTAT (Database). Disponible à l'adresse : http://www.fao.org/nr/water/aquastat/data/query/index.html?lang=en%0Ahttp://www.fao.org/nr/aquastat

FAO. (1976). *A framework for land evaluation. FAO soils bulletin n.32*.

FAO. (2011). CA Adoption Worldwide. Disponible à l'adresse : http://www.fao.org/ag/ca/6c.html

FAO. (2017a). FAOSTAT (Database). Disponible à l'adresse : http://www.fao.org/faostat/en/#data

FAO. (2017b). Global Agro-Environmental Stratification (GAES). FAO GeoNetwork. Consulté à l'adresse : www.fao. org/geonetwork/

FAO. (2017c). *The future of food and agriculture – Trends and challenges*. Rome.

FAO-UN. (1994). National administrative boundaries of Africa. Disponible à l'adresse : http://www.fao.org/geonetwork/srv/en/main.home?uuid=f7e7adb0-88fd-11da-a88f-000d939bc5d8

FEG Consulting and Save the Children. (2008). *The Practitioner's Guide to the Household Economy Approach*. Johannesburg.

FEWS NET/USAID. (2010). *Zones et profils de moyens d'existence au Burkina Faso*. Famine Early Warning Systems Network (FEWS NET), United States Agency for International Development (USAID).

Fischer, G., Hizsnyik, E., et Wiberg, D. (2011). *Scarcity and abundance of land resources: competing uses and the shrinking land resource base. SOLAW Background Thematic Report - TR02.*

Fischer, G., Nachtergaele, F., Prieler, S., van Velthuizen, H. T., Verelst, L., et Wiberg, D. (2012). *Global Agro-ecological Zones (GAEZ v3.0) - Model Documentation. IIASA, Laxenburg, Austria and FAO, Rome, Italy. iv*. http://doi.org/10.1029/97GB03657

Fischer, G., Velthuizen, H. Van, Shah, M., et Nachtergaele, F. (2002). *Global Agro-ecological Assessment for Agriculture in the 21st Century: Methodology and Results. Analysis* (Vol. RR-02-02). Consulté à l'adresse : http://www.iiasa.ac.at/Admin/PUB/Documents/RR-02-002.pdf

Fisher, P. F., Comber, A. J., et Wadsworth, R. (2005). Land use and land cover: contradiction or complement. *Re-Presenting GIS*, 85–98.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*(1), 185–201. http://doi.org/10.1016/S0034-4257(01)00295-4

Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, *21*(3), 768–769. Consulté à l'adresse : http://ci.nii.ac.jp/naid/10009668881/

Fornaro, A. C. (2012). *Logística e agronegócio globalizado no Estado do Tocantins: um estudo sobre a expansão das fronteiras agrícolas modernas no território brasileiro*. Universidade Estadual de Campinas.

Franklin, S. E., et Wulder, M. A. (2002). Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. *Progress in Physical Geography*, *26*(2), 173–205.

Fresco, L. O. (1984). *Comparing Anglophone and Francophone approaches to farming systems research and extension.* Networking Paper no.1. Farming Systems Support Project. University of Florida and US Agency for International Development.

Fresco, L. O., et Westphal, E. (1988). A hierarchical classification of farm systems. *Experimental Agriculture*, *24*(4), 399–419.

Friedrich, T., Derpsch, R., et Kassam, A. (2012). Overview of the global spread of conservation agriculture. *Field Actions Science Reports*, (Special Issue 6), 1–7. http://doi.org/10.1201/9781315365800-4

Fritz, S., See, L., Mccallum, I., You, L., Bun, A., Moltchanova, E., … Obersteiner, M. (2015). Mapping global cropland and field size. *Global Change Biology*, *21*(5), 1980–1992. http://doi.org/10.1111/gcb.12838

Fritz, S., See, L., et Rembold, F. (2010). Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *International Journal of Remote Sensing*, *31*(9), 2237–2256. http://doi.org/10.1080/01431160902946598

G

Galford, G. L., Mustard, J. F., Melillo, J., Gendrin, A., Cerri, C. C., et Cerri, C. E. P. (2008). Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment*, *112*(2), 576–587. http://doi.org/10.1016/j.rse.2007.05.017

Galochet, M., Godard, V., et Hotyat, M. (2001). Unités paysagères et biodiversité des îlots boisés : De l'image satellitale à l'analyse de terrain - Version française de: Land Units and the Biodiversity of Forest Islets: From Satellite Images to Ground Analysis. In L. Ryszkowski (Ed.), *Landscape ecology in agroecosystems management* (pp. 317–330). CRC Press.

Garrity, D., Dixon, J., et Boffa, J.-M. (2012). *Understanding African farming systems. Science and policy implications. Australian International Food Security Centre*.

Geng, S., Hess, C. E., et Auburn, J. (1990). Sustainable Agricultural Systems: Concepts and Definitions. *Journal of Agronomy and Crop Science*, *165*(2–3), 73–85. http://doi.org/10.1111/j.1439-037X.1990.tb00837.x

Gessner, U., Machwitz, M., Esch, T., Tillack, A., Naeimi, V., Kuenzer, C., et Dech, S. (2015). Multi-sensor mapping of West African land cover using MODIS, ASAR and TanDEM-X/TerraSAR-X data. *Remote Sensing of Environment*, *164*, 282–297. http://doi.org/10.1016/j.rse.2015.03.029

Gevaert, C. M., et García-Haro, F. J. (2015). A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sensing of Environment*, *156*, 34–44. http://doi.org/10.1016/j.rse.2014.09.012

Gibson, C. C., Ostrom, E., et Ahn, T. K. (2000). The concept of scale and the human dimensions of global change: A survey. *Ecological Economics*, *32*(2), 217–239. http://doi.org/10.1016/S0921-8009(99)00092-0

Girard, M.-C., et Girard, C.-M. (2010). *Traitement des données de télédétection - Environnement et ressources naturelles* (2nd ed.). Dunod.

Girard, N. (2004). La région: une notion géographique? *Ethnologie Française*, *37*(2), 107–112. http://doi.org/10.3917/ethn.041.0107.

Gonin, A., et Tallet, B. (2012). Quel avenir pour l'élevage dans le bassin cotonnier de l'Ouest du Burkina Faso ? Dynamiques agro-pastorales et recompositions territoriales. *Autrepart*, *60*(1), 95–110. http://doi.org/10.3917/autr.060.0095

Grillo, J., et Holt, J. (2009). *Application of the livelihood zone maps and profiles for food security analysis and early warning - guidance for famine early warning systems network (FEWS NET) representatives and partners*. Consulté à l'adresse : http://www.fews.net/sites/default/files/documents/reports/Guidance_Application_of_Livelihood_Zone_Maps_and_Profiles_en.pdf

Guan, X., Huang, C., Liu, G., Meng, X., et Liu, Q. (2016). Mapping rice cropping systems in Vietnam using an NDVI-based time-series similarity measurement based on DTW distance. *Remote Sensing*, *8*(1), 19.

Guenot, A., et Huchet-Bourdon, M. (2014). Rôle du coton sur la filière maïs au Burkina Faso. *Économie Rurale, Agricultures, Alimentations, Territoires*, *341*, 107–119. Consulté à l'adresse : http://economierurale.revues.org/4353

Gumma, M. K., Thenkabail, P. S., Teluguntla, P., Rao, M. N., Mohammed, I. A., et Whitbread, A. M. (2016). Mapping rice-fallow cropland areas for short-season grain legumes intensification in South Asia using MODIS 250 m time-series data. *International Journal of Digital Earth*, *9*(10), 981–1003. http://doi.org/10.1080/17538947.2016.1168489

Gupta, R. P., Tiwari, R. K., Saini, V., et Srivastava, N. (2013). A Simplified Approach for Interpreting Principal Component Images. *Advances in Remote Sensing*, *2*(2), 111–119. http://doi.org/10.4236/ars.2013.22015

Gurgel, H. C., et Ferreira, N. J. (2003). Annual and interannual variability of NDVI in Brazil and its connections with climate. *International Journal of Remote Sensing*, *24*, 3595–3609. http://doi.org/10.1080/0143116021000053788

H

Hadria, R., Duchemin, B., Baup, F., Le Toan, T., Bouvet, A., Dedieu, G., et Le Page, M. (2009). Combined use of optical and radar satellite data for the detection of tillage and irrigation operations: Case study in Central Morocco. *Agricultural Water Management*, *96*(7), 1120-1127. http://doi.org/10.1016/j.agwat.2009.02.010

Hagolle, O., Huc, M., Villa Pascual, D., et Dedieu, G. (2015). A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENuS and Sentinel-2 Images. *Remote Sensing*, *7*(3), 2668–2691.

Hall-Beyer, M. (2003). Comparison of Single-Year and Multiyear NDVI Time Series Principal Components in Cold Temperate Biomes. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(11 PART I), 2568–2574. http://doi.org/10.1109/TGRS.2003.817274

Hartigan, J. A., et Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 100–108.

Hay, G. J., et Castilla, G. (2008). Geographic object-based image analysis (GEOBIA): A new name for a new discipline. *Lecture Notes in Geoinformation and Cartography*, 75–89. http://doi.org/10.1007/978-3-540-77058-9_4

Hazeu, G. W., Metzger, M. J., Mücher, C. A., Perez-Soba, M., Renetzeder, C., et Andersen, E. (2011). European environmental stratifications and typologies: An overview. *Agriculture, Ecosystems and Environment*, *142*(1–2), 29–39. http://doi.org/10.1016/j.agee.2010.01.009

Hijmans, R. J. (2016). raster: Geographic Data Analysis and Modeling. R package version 2.5.8. Disponible à l'adresse : https://cran.r-project.org/package=raster

Hirano, C., da Silva Chagas, C., Saraiva do Amaral, F. C., de Sousa Neto, N. C., Rendeiro Pereira, N., et Soares Pinheiro da Cunha, M. (1989). *Aptidão agrícola das terras do Estado de Tocantins. Boletim de pesquisa n°40*. Rio de Janeiro: EMBRAPA-SNLCS. Consulté à l'adresse : http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/327947

Hirosawa, Y., Marsh, S. E., et Kliman, D. H. (1996). Application of standardized principal component analysis of land-cover characterization using multitemporal AVHRR data. *Remote Sensing of Environment*, *58*(3), 267–281. http://doi.org/10.1016/S0034-4257(96)00068-5

HLPE. (2017). *Nutrition and food systems. A report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security*. Rome, Italy.

Huete, A., Didan, K., Miura, T., Rodriguez, E. ., Gao, X., et Ferreira, L. . (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1), 195–213. http://doi.org/10.1016/S0034-4257(02)00096-2

I

IBGE. (2004). Mapa de Biomas do Brasil (escala 1:5 000 000). Instituto Brasileiro de Geografia e Estatística e Ministério do Meio Ambiente. Disponible à l'adresse : ftp://ftp.ibge.gov.br/Cartas_e_Mapas/Mapas_Murais/

IBGE. (2014). Instituto Brasileiro de Geografia e Estatística, Produção Agrícola Municipal (PAM). Disponible à l'adresse : http://www.ibge.gov.br/home/estatistica/economia/pam/2014/

IBGE. (2015). Instituto Brasileiro de Geografia e Estatística, Produção Agrícola Municipal (PAM). Disponible à l'adresse : http://www.ibge.gov.br/home/estatistica/economia/pam/2015/

IBGE. (2017). Base cartográfica contínua do Brasil (escala 1:250 000), versão 2017. Diretoria de geociências coordenação de cartografia do Instituto Brasileiro de Geografia e Estatística. Disponible à l'adresse : ftp://geoftp.ibge.gov.br/cartas_e_mapas/bases_cartograficas_continuas/bc250/versao2017/shapefile/

Inglada, J., Vincent, A., Arias, M., et Marais-Sicre, C. (2016). Improved early crop type identification by joint use of high temporal resolution sar and optical image time series. *Remote Sensing*, *8*(5). http://doi.org/10.3390/rs8050362

INPE. (s.d.-a). Projeto TerraClass Amazônia : Mapeamento do Uso e Cobertura da Terra na Amazônia Legal Brasileira. Consulté à l'adresse : http://www.inpe.br/cra/projetos_pesquisas/dados_terraclass.php

INPE. (s.d.-b). Projeto TerraClass Cerrado : Mapeamento do Uso e Cobertura Vegetal do Cerrado. Consulté à l'adresse : http://www.dpi.inpe.br/tccerrado/

IUCN/UNEP-WCMC. (2017). The World Database on Protected Areas (WDPA). Version octobre 2017. Disponible à l'adresse : www.protectedplanet.net

J

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666. http://doi.org/10.1016/j.patrec.2009.09.011

Jenkerson, C. (2013). User guide: Earth resources observation and science (EROS) center science processing architecture (ESPA) on demand interface.

Jouve, P. (2006). La dimension spatiale des systèmes de culture: Comparaison entre agriculture tempérée et agriculture tropicale. *Cahiers Agricultures*, *15*(3), 255–260.

K

Karlson, M., Ostwald, M., Reese, H., Bazié, H. R., et Tankoano, B. (2016). Assessing the potential of multi-seasonal WorldView-2 imagery for mapping West African agroforestry tree species. *International Journal of Applied Earth Observation and Geoinformation*, *50*, 80–88. http://doi.org/10.1016/j.jag.2016.03.004

Kirches, G., Brockmann, C., Boettcher, M., Peters, M., Bontemps, S., Lamarche, C., … Defourny, P. (2014). Land Cover CCI-Product User Guide-Version 2. *ESA Public Document CCI-LC-PUG*. Louvain-la-Neuve, Belgium: UCL, ESA.

Klink, C. A., et Machado, R. B. (2005). Conservation of the Brazilian Cerrado. *Conservation Biology*. http://doi.org/10.1111/j.1523-1739.2005.00702.x

Knauer, K., Gessner, U., Fensholt, R., Forkuor, G., et Kuenzer, C. (2017). Monitoring agricultural expansion in Burkina Faso over 14 years with 30 m resolution time series: The role of population growth and implications for the environment. *Remote Sensing*, *9*(2). http://doi.org/10.3390/rs9020132

Knauer, K., Gessner, U., Fensholt, R., et Kuenzer, C. (2016). An ESTARFM fusion framework for the generation of large-scale time series in cloud-prone and heterogeneous landscapes. *Remote Sensing*, *8*(5). http://doi.org/10.3390/rs8050425

L

Labatut, V., et Cherifi, H. (2012). Accuracy measures for the comparison of classifiers. In *The 5th International Conference on Information Technology*. Amman, Jordanie.

LADA. (2008). *Mapping Land Use Systems at global and regional scales for Land Degradation Assessment Analysis.* Nachtergaele F. et Petri M. LADA Technical report n.8, version 1.1. Rome, Italy.

Lam, N. S. N. (2008). Fractals and Scale in Environmental Assessment and Monitoring. In *Scale and Geographic Inquiry: Nature, Society, and Method* (pp. 23–40). http://doi.org/10.1002/9780470999141.ch2

Lambert, M. J., Waldner, F., et Defourny, P. (2016). Cropland mapping over Sahelian and Sudanian agrosystems: A Knowledge-based approach using PROBA-V time series at 100-m. *Remote Sensing*, *8*(3). http://doi.org/10.3390/rs8030232

Lambin, E. F., Geist, H. J., et Lepers, E. (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, *28*(1), 205–241. http://doi.org/10.1146/annurev.energy.28.050302.105459

Lambin, E. F., Geist, H. J., et Rindfuss, R. R. (2006). Introduction: Local Processes with Global Impacts. *Land-Use and Land-Cover Change, Local Processes and Global Impact*, (Turner), 2-8. http://doi.org/10.1007/3-540-32202-7_1

Lassalle, P., Inglada, J., Michel, J., Grizonnet, M., et Malik, J. (2015). A Scalable Tile-Based Framework for Region-Merging Segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(10), 5473–5485. http://doi.org/10.1109/TGRS.2015.2422848

Leenhardt, D., Angevin, F., Biarnès, A., Colbach, N., et Mignolet, C. (2010). Describing and locating cropping systems on a regional scale. A review. *Agronomy for Sustainable Development*. http://doi.org/10.1051/agro/2009002

Lelong, C., Alexandre, C., et Dupuy, S. (2014). Discrimination of tropical agroforestry systems in very high resolution satellite imagery using object-based hierarchical classification: A case-study in Cameroon. *South-Eastern European Journal of Earth Observation and Geomatics*.

Leutner, B., et Horning, N. (2017). RStoolbox: tools for remote sensing data analysis. R Package version 0.1.8.

Ligges, U., Short, T., et Kienzle, P. (2015). signal: Signal processing. R package version 0.7.6. Disponible à l'adresse : https://cran.r-project.org/package=signal

Lima, A. A. C., Oliveira, F. N. S., et de Aquino, A. R. L. (2000). *Solos e aptidão agrícola das terras do Estado do Tocantins. Embrapa Agroindústria Tropical. Documentos, 31*. Fortaleza.

Lloyd, S. P. (1982). Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. http://doi.org/10.1109/TIT.1982.1056489

Lobell, D. B., et Field, C. B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters*, *2*(1), 7pp. http://doi.org/10.1088/1748-9326/2/1/014002

Loveland, T. R., et DeFries, R. S. (2004). Observing and monitoring land use and land cover change. *Ecosystems And Land Use Change*, *153*, 231–246.

M

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297).

Malingreau, J. P., Bartholomé, E., et Barisano, E. (1987). Surveillance de la production agricole en Afrique de l'Ouest. Nécessité d'une intégration de différentes plates-formes satellitaires. In *Proceedings of the Symposium SPOT 1, utilisation des images, bilans, résultats* (pp. 353–370). Paris.

Massey, R., Sankey, T. T., Congalton, R. G., Yadav, K., Thenkabail, P. S., Ozdogan, M., et Sánchez Meador, A. J. (2017). MODIS phenology-derived, multi-year distribution of conterminous U.S. crop types. *Remote Sensing of Environment*, *198*, 490–503. http://doi.org/10.1016/j.rse.2017.06.033

Mathys, E., et Gardner, A. (2009). *USAID Office of Food for Peace Burkina Faso Food Security Country Framework FY 2010 – FY 2014*. Washington, D.C: Food and Nutrition Technical Assistance II Project (FANTA-2), Academy for Educational Development (AED).

Matson, P. A., Parton, W. J., Power, A. G., et Swift, M. J. (1997). Agricultural Intensification and Ecosystem Properties. *Science*, *277*(5325), 504–509. http://doi.org/10.1126/science.277.5325.504

Maus, V., Câmara, G., Cartaxo, R., Sanchez, A., Ramos, F. M., et De Queiroz, G. R. (2016). A Time-Weighted Dynamic Time Warping Method for Land-Use and Land-Cover Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(8), 3729–3739. http://doi.org/10.1109/JSTARS.2016.2517118

McNairn, H., Champagne, C., Shang, J., Holmstrom, D., et Reichert, G. (2009). Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(5), 434–449. http://doi.org/10.1016/j.isprsjprs.2008.07.006

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157.

Mingwei, Z., Qingbo, Z., Zhongxin, C., Jia, L., Yong, Z., et Chongfa, C. (2008). Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data. *International Journal of Applied Earth Observation and Geoinformation*, *10*(4), 476–485. http://doi.org/10.1016/j.jag.2007.11.002

Ministère de l'éducation nationale et ministère de l'enseignement supérieur et de la recherche. Arrêté du 28 novembre 1980 relatif à la terminologie de la télédétection arospatiale. Journal Officiel numéro complémentaire du 11 décembre 1980, p. 10848 (1980).

Monfreda, C., Ramankutty, N., et Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, *22*(1). http://doi.org/10.1029/2007GB002947

Mücher, S., De Simone, L., Kramer, H., de Wit, A., Roupioz, L., Hazeu, G., … Cormont, A. (2016). A new Global Agro-Environmental Stratification (GAES). *Wageningen Environmental Research. Report 2761*, 70. http://doi.org/10.18174/400815

Mücher, S., De Simone, L., Kramer, H., de Wit, A., Roupioz, L., Hazeu, G., … Cormont, A. (2016). A new Global Agro-Environmental Stratification (GAES). http://doi.org/10.18174/400815

Müller, C., Bondeau, A., Popp, A., Waha, K., et Fader, M. (2010). *Climate Change Impacts on Agricultural Yields*. Washington D.C. © World Bank: License: CC BY 3.0 IGO.

N

Nachtergaele, F., et Petri, M. (2013). *Mapping Land Use Systems at Global and Regional Scales for Land Degradation Assessment Analysis. Lada*.

NASA LP DAAC. (s.d.-a). MOD09A1 data product. NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center.

NASA LP DAAC. (s.d.-b). MOD13Q1 version 5 data product. NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center. Disponible à l'adresse : https://lpdaac.usgs.gov/

Nir, D. (1987). Regional geography considered from the systems' approach. *Geoforum*, *18*(2), 187-202. http://doi.org/10.1016/0016-7185(87)90005-4

O

O'Neill, R. V, DeAngelis, D. L., Waide, J. B., et Allen, T. F. H. (1986). *Hierarchical Concept of Ecosystems. Monographs in Population Biology Series. Vol.23*. Consulté à l'adresse : http://search.barnesandnoble.com/booksearch/isbninquiry.asp?r=1etean=0691084378

OCHA ROWCA. (2015). Burkina Faso: Settlements. United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Regional Office for West and Central Africa (ROWCA). Disponible à l'adresse : https://data.humdata.org/dataset/settlements

OCHA ROWCA. (2017). Second Administrative Level Boundaries Project (SALB), Institut Géographique du Burkina Faso (IGB). United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Regional Office for West and Central Africa (ROWCA). Disponible à l'adresse : https://data.humdata.org/dataset/burkina-faso-administrative-boundaries

OSS. (2015). *Burkina Faso : Atlas des cartes d'occupation du sol. Projet amélioration de la résilience des populations sahéliennes aux mutations environnementales - REPSAHEL*. Observatoire du Sahara et du Sahel (OSS).

Ozdogan, M., et Woodcock, C. E. (2006). Resolution dependent errors in remote sensing of cultivated areas. *Remote Sensing of Environment*, *103*(2), 203–217. http://doi.org/10.1016/j.rse.2006.04.004

P

Pacheco, A. M., McNairn, H., et Merzouki, A. (2010). Evaluating TerraSAR-X for the identification of tillage occurrence over an agricultural area in Canada. *Remote Sensing for Agriculture, Ecosystems, and Hydrology Xii*, *7824*, 1–7. http://doi.org/78240p\r10.1117/12.868218

Pacheco, A., et McNairn, H. (2010). Evaluating multispectral remote sensing and spectral unmixing analysis for crop residue mapping. *Remote Sensing of Environment*, *114*(10), 2219–2228. http://doi.org/10.1016/j.rse.2010.04.024

Panigrahy, S., Manjunath, K. R., et Ray, S. S. (2005). Deriving cropping system performance indices using remote sensing data and GIS. *International Journal of Remote Sensing*, *26*(12), 2595-2606. http://doi.org/10.1080/01431160500114698

Panigrahy, S., Ray, S. S., Manjunath, K. R., Pandey, P. S., Sharma, S. K., Sood, A., … Parihar, J. S. (2011). A Spatial Database of Cropping System and its Characteristics to Aid Climate Change Impact Assessment Studies. *Journal of the Indian Society of Remote Sensing*, *39*(3), 355–364. http://doi.org/10.1007/s12524-011-0093-3

Parmentier, S. (2014). *Scaling-up agroecological approaches: what, why and how*. *Oxfam-Solidarity, Brussels*.

Pearson, K. (1901). Principal components analysis. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *6*(2), 559.

Pedroso da Silva, A. R., et Almeida, M. G. (2007). O agronegócio e o Estado do Tocantins: o atual estágio de consolidação. *Caminhos de Geografia*, *8*(21).

Plet, F. (2003). La geographie rurale française : quelques jalons. *Sociétés Contemporaines*, *49–50*(1), 85–106. http://doi.org/10.3917/soco.049.0085

Plexida, S. G., Sfougaris, A. I., Ispikoudis, I. P., et Papanastasis, V. P. (2014). Selecting landscape metrics as indicators of spatial heterogeneity-Acomparison among Greek landscapes. *International Journal of Applied Earth Observation and Geoinformation*, *26*(1), 26–35. http://doi.org/10.1016/j.jag.2013.05.001

Plourde, J. D., Pijanowski, B. C., et Pekin, B. K. (2013). Evidence for increased monoculture cropping in the Central United States. *Agriculture, Ecosystems and Environment*, *165*, 50–59. http://doi.org/10.1016/j.agee.2012.11.011

Pontius, R. G., et Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, *32*(15), 4407-4429. http://doi.org/10.1080/01431161.2011.552923

Q

Qiu, B., Zhong, M., Tang, Z., et Wang, C. (2014). A new methodology to map double-cropping croplands based oncontinuous wavelet transform. *International Journal of Applied Earth Observation and Geoinformation*, *26*(1), 97–104. http://doi.org/10.1016/j.jag.2013.05.016

R

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramoino, F., Pera, F., et Arino, O. (2018). *"S2 prototype LC map at 20m of Africa 2016" Users Feedback Compendium [6th February 2018]*. Rome, Italy. Consulté à l'adresse : http://due.esrin.esa.int/files/S2_prototype_LC_map_at_20m_of_Africa_2016-Users_Feedback_Compendium-6-Feb-2018.pdf

Ribeiro, P. F., Santos, J. L., Santana, J., Reino, L., Leitão, P. J., Beja, P., et Moreira, F. (2016). Landscape makers and landscape takers: links between farming systems and landscape patterns along an intensification gradient. *Landscape Ecology*, *31*(4), 791–803. http://doi.org/10.1007/s10980-015-0287-0

Richards, J. a, et Jia, X. (2006). *Remote Sensing Digital Image Analysis: An Introduction*. *New York*. http://doi.org/10.1007/978-3-642-30062-2

Rizzo, D., Marraccini, E., Lardon, S., Rapey, H., Debolini, M., Benoît, M., et Thenail, C. (2013). Farming systems designing landscapes: Land management units at the interface between agronomy and geography. *Geografisk Tidsskrift*, *113*(2), 71–86. http://doi.org/10.1080/00167223.2013.849391

Robbez-Masson, J. M., Foltête, J. C., Cabello, L., et Flitti, M. (1999). Prise en compte du contexte spatial dans l'instrumentation de la notion de paysage. Application à une segmentation géographique assistée. *Revue Internationale de Géomatique*, *9*, 173–195.

Rodrigues de Melo, K. D., Polastrine, A., Amaral, R. M., Grando Lopes, A., Sayão Lobato, C. L. D., et Rosanova, C. (2010). Caracterização da bovinocultura leiteira nos municípios de Taguatinga, Combinado e Lavandeira, região sudeste do Estado do Tocantins, Brasil. In *XX CONGRESSO BRASILEIRO DE ZOOTECNIA* (p. 3). Palmas, Tocantins.

Rosenfield, G. H., et Fitzpatrick-Lins, K. (1986). A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing*, *52*(2), 223–227.

Rouse, J. W., Haas, R. H., Schell, J. A., et Deering, D. W. (1973). Monitoring Vegetation Systems in the Great Okains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, *1*, 325–333. http://doi.org/10/citeulike-article-id:12009708

Rudorff, B., Risso, J., Aguiar, D., Gonçalves, F., Salgado, M., Perrut, J., … Virtuoso, M. (2015). *Análise Geoespacial da Dinâmica das Culturas Anuais no Bioma Cerrado: 2000 a 2014*.

Rufino, M. C., Atzberger, C., Baldi, G., Butterbach-Bahl, K., Rosenstock, T. S., et Stern, D. (2016). Targeting landscapes to identify mitigation options in smallholder agriculture. In *Methods for Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture* (pp. 15–36). http://doi.org/10.1007/978-3-319-29794-1_2

S

Sahajpal, R., Zhang, X., Izaurralde, R. C., Gelfand, I., et Hurtt, G. C. (2014). Identifying representative crop rotation patterns and grassland loss in the US Western Corn Belt. *Computers and Electronics in Agriculture*, *108*, 173–182. http://doi.org/10.1016/j.compag.2014.08.005

Salmon, J. M., Friedl, M. A., Frolking, S., Wisser, D., et Douglas, E. M. (2015). Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *International Journal of Applied Earth Observation and Geoinformation*, *38*, 321–334. http://doi.org/10.1016/j.jag.2015.01.014

Sauer, S., et Leite, S. P. (2012). Agrarian structure, foreign investment in land, and land prices in Brazil. *Journal of Peasant Studies*, *39*(3–4), 873–898. http://doi.org/10.1080/03066150.2012.686492

Savitzky, A., et Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*(8), 1627–1639. http://doi.org/10.1021/ac60214a047

Sayre, N. F. (2005). Ecological and geographical scale: parallels and potential for integration. *Progress in Human Geography*, *29*(3), 276–290. http://doi.org/10.1191/0309132505ph546oa

Sebillotte, M. (1982). Les systèmes de culture. Réflexion sur l'intérêt et l'emploi de cette notion à partir de l'expérience acquise en région de grande culture. In *Séminaire du département d'agronomie de l'INRA* (pp. 63–80). INRA, Vichy.

Sebillotte, M. (1990). *Système de culture, un concept opératoire pour les agronomes*. (L. Combe et D. Picard, Eds.). INRA, Paris, 165-196.

See, L., Fritz, S., You, L., Ramankutty, N., Herrero, M., Justice, C., … Obersteiner, M. (2015). Improved global cropland data as an essential ingredient for food security. *Global Food Security*. http://doi.org/10.1016/j.gfs.2014.10.004

SEPLAN. (2012). Base de Dados Geográficos do Tocantins - Atualização 2012. Palmas: Secretaria do Planejamento (SEPLAN), Superintendência de Planejamento e Gestão Central de Políticas Públicas, Diretoria de Zoneamento Ecológico-Econômico. Disponible à l'adresse : http://seplan.to.gov.br/zoneamento/bases-vetoriais/base-de-dados-geograficos-do-tocantins-atualizacao-2012/

Siebert, S., Hoogeveen, J., et Frenken, K. (2006). Irrigation in Africa, Europe and Latin America - Update of the Digital Global Map of Irrigation Areas to Version 4. *Frankfurt Hydrology Paper*, (5), 135p.

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., et Scanlon, B. (2015). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, *19*(3), 1521–1545. http://doi.org/10.5194/hess-19-1521-2015

Silva, L. A. G. C. (2007). *Biomas presentes no Estado do Tocantins.* Consultoria Legislativa da Câmara dos Deputados.

Skakun, S., Kussul, N., Shelestov, A. Y., Lavreniuk, M., et Kussul, O. (2016). Efficiency Assessment of Multitemporal C-Band Radarsat-2 Intensity and Landsat-8 Surface Reflectance Satellite Imagery for Crop Classification in Ukraine. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(8), 3712–3719. http://doi.org/10.1109/JSTARS.2015.2454297

Slonim, N., Aharoni, E., et Crammer, K. (2013). Hartigan's k-means versus Lloyd's k-means - Is it time for a change? In *IJCAI International Joint Conference on Artificial Intelligence* (pp. 1677–1684).

Solano, R., Didan, K., Jacobson, A., et Huete, A. (2010). *MODIS Vegetation Index User's Guide (MOD13 Series).*

Soudani, K., le Maire, G., Dufrêne, E., François, C., Delpierre, N., Ulrich, E., et Cecchini, S. (2008). Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sensing of Environment*, *112*(5), 2643–2655. http://doi.org/10.1016/j.rse.2007.12.004

Spangler, K. R., Lynch, A. H., et Spera, S. A. (2017). Precipitation Drivers of Cropping Frequency in the Brazilian Cerrado: Evidence and Implications for Decision-Making. *Weather, Climate, and Society*, *9*(2), 201–213. http://doi.org/10.1175/WCAS-D-16-0024.1

Spera, S. (2017). Agricultural Intensification Can Preserve the Brazilian Cerrado: Applying Lessons From Mato Grosso and Goiás to Brazil's Last Agricultural Frontier. *Tropical Conservation Science*, *10*, 194008291772066. http://doi.org/10.1177/1940082917720662

Stomph, T. J., Fresco, L. O., et van Keulen, H. (1994). Land use system evaluation: Concepts and methodology. *Agricultural Systems*, *44*(3), 243–255. http://doi.org/10.1016/0308-521X(94)90222-2

Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., … Woodcock, C. E. (2006). *Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps*. *Scientific and Technical Research series* (Vol. EUR 22156).

Strahler, A. H., Woodcock, C. E., et Smith, J. A. (1986). On the nature of models in remote sensing. *Remote Sensing of Environment*, *20*(2), 121–139. http://doi.org/10.1016/0034-4257(86)90018-0

T

Thiombiano, A., et Kampmann, D. (2010). *Atlas de la Biodiversité de L'Afrique de L'Ouest, Tome II: Burkina Faso*. (A. Thiombiano et D. Kampmann, Eds.). Ouagadougou et Frankfurt/Main: Goethe-Universität Frankfurt am Main, Geowissenschaften/Geographie.

Trimble Germany GmbH. (2014). eCognition Developer© 9.0 User Guide. Munich, Germany.

U

United Nations. Department of Economic and Social Affairs. Population Division. (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables* (No. ESA/P/WP/248).

USGS. (s.d.). Landsat Level-2 Science Products. U.S. Geological Survey Earth Resources Observation and Science Center (USGS/EROS), Science Processing Architecture (ESPA) On Demand Interface. Disponible à l'adresse : https://espa.cr.usgs.gov/

USGS. (2016). *Landsat 8 (L8) Data Users Handbook*. Sioux Falls, South Dakota: U.S. Geological Survey (USGS) Landsat Project Science Office at the Earth Resources Observation and Science (EROS) Center. Consulté à l'adresse : https://landsat.usgs.gov/sites/default/files/documents/Landsat8DataUsersHandbook.pdf

V

Van Wart, J., van Bussel, L. G. J., Wolf, J., Licker, R., Grassini, P., Nelson, A., … Cassman, K. G. (2013). Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Research*, *143*, 44–55. http://doi.org/10.1016/j.fcr.2012.11.023

Veldkamp, A., Kok, K., De Koning, G. H. J., Schoorl, J. M., Sonneveld, M. P. W., et Verburg, P. H. (2001). Multi-scale system approaches in agronomic research at the landscape level. *Soil and Tillage Research*, *58*(3–4), 129–140. http://doi.org/10.1016/S0167-1987(00)00163-X

Vermote, E. F., El Saleous, N. Z., et Justice, C. O. (2002). Atmospheric correction of MODIS data in the visible to middle infrared: First results. *Remote Sensing of Environment*, *83*(1–2), 97-111. http://doi.org/10.1016/S0034-4257(02)00089-5

Vintrou, E., Desbrosse, A., Bégué, A., Traoré, S., Baron, C., et Lo Seen, D. (2012). Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. *International Journal of Applied Earth Observation and Geoinformation*, *14*(1), 83–93. http://doi.org/10.1016/j.jag.2011.06.010

Vintrou, E., Soumaré, M., Bernard, S., Bégué, A., Baron, C., et Lo Seen, D. (2012). Mapping Fragmented Agricultural Systems in the Sudano-Sahelian Environments of Africa Using Random Forest and Ensemble Metrics of Coarse Resolution MODIS Imagery. *Photogrammetric Engineering et Remote Sensing*, *78*(8), 839–848. http://doi.org/10.14358/PERS.78.8.839

Viña, A., Peters, A. J., et Ji, L. (2003). Use of multispectral ikonos imagery for discriminating between conventional and conservation agricultural tillage practices. *Photogrammetric Engineering and Remote Sensing*.

W

Waldner, F., Lambert, M. J., Li, W., Weiss, M., Demarez, V., Morin, D., … Defourny, P. (2015). Land cover and crop type classification along the season based on biophysical variables retrieved from multi-sensor high-resolution time series. *Remote Sensing*, *7*(8), 10400–10424. http://doi.org/10.3390/rs70810400

Wang, T., Kou, X., Xiong, Y., Mou, P., Wu, J., et Ge, J. (2010). Temporal and spatial patterns of NDVI and their relationship to precipitation in the Loess Plateau of China. *International Journal of Remote Sensing*, *31*(7), 1943–1958. http://doi.org/10.1080/01431160902929263

Wardlow, B. D., et Egbert, S. L. (2008). Large-area crop mapping using time-series MODIS 250m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, *112*(3), 1096–1116. http://doi.org/10.1016/j.rse.2007.07.019

Warrens, M. J. (2015). Properties of the quantity disagreement and the allocation disagreement. *International Journal of Remote Sensing*, *36*(5), 1439–1446. http://doi.org/10.1080/01431161.2015.1011794

Wiens, J. A. (1989). Spatial Scaling in Ecology. *Functional Ecology*, *3*(4), 385. http://doi.org/10.2307/2389612

Willer, H., et Lernoud, J. (2017). *The World of Organic Agriculture. Statistics and Emerging Trends 2017*. Research Institute of Organic Agriculture (FiBL), Frick, and IFOAM - Organics International. Bonn.

Woodcock, C. E., et Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*(3), 311–332. http://doi.org/10.1016/0034-4257(87)90015-0

World Bank. (2010). *World Development Report 2010 : Development and Climate Change*. Washington D.C. © World Bank: License: CC BY 3.0 IGO.

World Bank. (2017). World Development Indicators (Agriculture, value added, % of GDP). Disponible à l'adresse : https://data.worldbank.org/

World Bank. (2018). Population estimates and projections (Population growth, annual %). Disponible à l'adresse : http://databank.worldbank.org/

X

Xiong, J., Thenkabail, P. S., Gumma, M. K., Teluguntla, P., Poehnelt, J., Congalton, R. G., … Thau, D. (2017). Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *126*, 225–244.

Xiong, J., Thenkabail, P. S., Tilton, J. C., Gumma, M. K., Teluguntla, P., Oliphant, A., … Gorelick, N. (2017). Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on google earth engine. *Remote Sensing*, *9*(10). http://doi.org/10.3390/rs9101065

Y

You, L., et Wood, S. (2006). An entropy approach to spatial disaggregation of agricultural production. *Agricultural Systems*, *90*(1–3), 329–347. http://doi.org/10.1016/j.agsy.2006.01.008

You, L., Wood, S., Wood-Sichra, U., et Wu, W. (2014). Generating global crop distribution maps: From census to grid. *Agricultural Systems*, *127*, 53–60. http://doi.org/10.1016/j.agsy.2014.01.002

Young, S. S., et Anyamba, A. (1999). Comparison of NOAA/NASA PAL and NOAA GVI data for vegetation change studies over China. *Photogrammetric Engineering and Remote Sensing*, *65*, 679–696.

Z

Zhong, L., Yu, L., Li, X., Hu, L., et Gong, P. (2016). Rapid corn and soybean mapping in US Corn Belt and neighboring areas. *Scientific Reports*, *6*. http://doi.org/10.1038/srep36240

Zonneveld, I. S. (1989). The land unit - A fundamental concept in landscape ecology, and its applications. *Landscape Ecology*, *3*(2), 67–86. http://doi.org/10.1007/BF00131171

Annexe - Publications

Bellón, B., Bégué, A., Lo Seen, D., de Almeida, C.A., et Simões, M. (2017). "A Remote Sensing Approach for Regional-Scale Mapping of Agricultural Land-Use Systems Based on NDVI Time Series". *Remote Sensing*, 9(6), 600. http://doi.org/10.3390/rs9060600

Bellón, B., Bégué, A., Lo Seen, D., Lebourgeois, V., Evangelista, B. A., Simões, M., et Demonte Ferraz, R.P. (2018). "Improved regional-scale Brazilian cropping systems' mapping based on a semi-automatic object-based clustering approach". *International Journal of Applied Earth Observations and Geoinformation*, 68C, 127-138. http://doi.org/10.1016/j.jag.2018.01.019

remote sensing

MDPI

Article

A Remote Sensing Approach for Regional-Scale Mapping of Agricultural Land-Use Systems Based on NDVI Time Series

Beatriz Bellón [1,*], Agnès Bégué [1], Danny Lo Seen [1], Claudio Aparecido de Almeida [2] and Margareth Simões [3,4]

1 Cirad, UMR TETIS (Land, Environment, Remote Sensing and Spatial Information), Maison de la Télédétection, Rue Jean-François Breton, 34090 Montpellier, France; agnes.begue@teledetection.fr (A.B.); loseen@teledetection.fr (D.L.S.)
2 National Institute for Space Research (INPE), Av. dos Astronautas, 1758, São José dos Campos, SP 12227-010, Brazil; claudio.almeida@inpe.br
3 Embrapa Solos, Rua Jardim Botânico, 1024, Rio de Janeiro, RJ 22460-000, Brazil; margareth.simoes@embrapa.br
4 Department of Computer Engineering, Rio de Janeiro State University (UERJ/FEN/DESC/PPGMA), Rua São Francisco Xavier, 524, 5031 D, Maracanã, Rio de Janeiro, RJ 20550-900, Brazil
* Correspondence: beatriz.bellon@teledetection.fr; Tel.: +33-467-558-645

Academic Editors: Raul Zurita-Milla, James Campbell and Prasad S. Thenkabail
Received: 2 May 2017; Accepted: 10 June 2017; Published: 13 June 2017

Abstract: In response to the need for generic remote sensing tools to support large-scale agricultural monitoring, we present a new approach for regional-scale mapping of agricultural land-use systems (ALUS) based on object-based Normalized Difference Vegetation Index (NDVI) time series analysis. The approach consists of two main steps. First, to obtain relatively homogeneous land units in terms of phenological patterns, a principal component analysis (PCA) is applied to an annual MODIS NDVI time series, and an automatic segmentation is performed on the resulting high-order principal component images. Second, the resulting land units are classified into the crop agriculture domain or the livestock domain based on their land-cover characteristics. The crop agriculture domain land units are further classified into different cropping systems based on the correspondence of their NDVI temporal profiles with the phenological patterns associated with the cropping systems of the study area. A map of the main ALUS of the Brazilian state of Tocantins was produced for the 2013–2014 growing season with the new approach, and a significant coherence was observed between the spatial distribution of the cropping systems in the final ALUS map and in a reference map extracted from the official agricultural statistics of the Brazilian Institute of Geography and Statistics (IBGE). This study shows the potential of remote sensing techniques to provide valuable baseline spatial information for supporting agricultural monitoring and for large-scale land-use systems analysis.

Keywords: geographic object-based image analysis (GEOBIA); Moderate Resolution Imaging Spectroradiometer (MODIS); principal components analysis (PCA); cropping systems; Stratification

1. Introduction

Ensuring food security through sustainable agricultural development sets a double challenge to agricultural systems: Food production must substantially increase to satisfy the demand of a continuously growing population while improving the stewardship of natural resources and minimizing environmental impacts [1]. This transition towards sustainability is further challenged by the consequences of climate change and an increased competition for land, generating an urgent need

for repetitive spatial information to help monitor the dynamics of the agricultural land-use systems at the regional and global scales.

Land use is defined as "the sequence of operations carried out with the purpose to obtain goods and services from the land" [2]. Land-use systems (LUS) can thus be considered as coupled human–environment systems that can be characterized by two main aspects: land resources and management practices [3]. LUS mapping originally consisted in the delineation of relatively homogeneous areas of land (referred to as land units), which are directly linked to a specific type of land use [4,5]. Regional-scale agricultural land-use systems (ALUS) maps typically represent the major agricultural zones for a given region, each agricultural zone (land unit) being directly linked to a particular agricultural system. The land units of a regional-scale ALUS map are naturally composed of multiple agricultural fields which are spatially "concentrated" and representative of a particular agricultural system, but also contain a mosaic of other land-cover types, e.g., patches of natural vegetation among the agricultural fields, due to their extent. The plotting of the land units' boundaries is, however, challenged by limited access to spatially explicit and detailed land-use information over large areas, which may account for the general lack of LUS maps.

The few existing large-area LUS mapping approaches, such as the one developed by FAO [2], are usually based on a subjective choice of socio-economic and environmental variables. These are derived from heterogeneous sources which are empirically categorized and combined to map the different land-use systems. The resulting LUS maps are subject to error propagation due to the disparity of spatial resolution, production date, and quality of the original data. In addition, access to the data is not always guaranteed, which limits the reproducibility of the LUS maps in time and across regions. As a result, the description and location of the different agricultural land-use systems remain highly unclear for most world regions.

Earth observation satellite systems can significantly contribute to LUS mapping since they provide timely and detailed land-use information over large areas due to their synoptic coverage and high revisiting frequency. In particular, remote sensing-derived time series of vegetation indices allow monitoring of the phenology (seasonal vegetation variation) and the intra-seasonal variations of the agricultural land cover (i.e., cropland, pastures, and rangelands), from which the agricultural land use (including some management practices) can be determined [6–9]. Furthermore, the geographic object-based image analysis (GEOBIA) [10] segmentation techniques, which automatically delineate homogeneous objects at multiple scales, seem particularly adapted for the extraction of land units. Bisquert et al. [11] introduced an innovative approach based on GEOBIA of spectral and textural variables derived from coarse-resolution vegetation index images that produced a spatial segmentation of the French territory in broad-scale land units. The resulting land units proved to be consistent in terms of land cover [12], but remained unclassified, and the variables used are not descriptive enough for effective delineation of agricultural land-use systems since they do not account for the intra-seasonal variations that characterize agricultural land-use practices.

The aim of this work is to provide a simple and reproducible approach for regional-scale mapping of ALUS based on GEOBIA and vegetation index time series analysis. The presented approach involves the extraction of temporal variations of the vegetation cover in relation to agricultural land use from a vegetation index time series, by performing a principal components analysis and selecting the principal component images that capture the seasonal and intra-seasonal variability of the vegetation index time series. Segmentation and classification of the selected principal components images follow, finally producing an ALUS map. The approach was applied to an annual Moderate Resolution Imaging Spectroradiometer Normalized Difference Vegetation Index (MODIS NDVI) dataset and tested on a new agricultural expansion region of Brazil. The spatial distribution of the cropping systems in the final ALUS map was evaluated with the annual crop estimates of the Municipal Agricultural Production database (PAM) of the Brazilian Institute of Geography and Statistics (IBGE).

2. Materials

2.1. Study Area

The study area corresponds to the Brazilian State Tocantins, a region that spans 50°W to 45°W and 5°S to 13°S, covering 277,621 km^2 between the Amazon and the Cerrado biomes (Figure 1). The geographical location between two biomes provides this region with a rich diversity of environmental conditions, which have led to a diversification of human activities and thus a variety of agricultural landscapes [13].



Figure 1. The study area corresponds to the Tocantins state (TO, red-striped), located in the center-north of Brazil, which is part of the MATOPIBA region (color-coded orange). This region is the latest agricultural frontier of Brazil, composed of parts of the states of Maranhão (MA), Piauí (PI), Bahia (BA), and the full extent of Tocantins.

While livestock production on extensive rangelands still dominates most of the land outside of the protected area's limits, this region has experienced rapid agricultural expansion within the last decade. This trend is mainly characterized by an expansion of the soybean-cultivated area, which is mostly produced for export (with an increase of 804,800 hectares in the last 15 years [14]). Sequential cropping systems (mainly soybean-based double-cropping systems, where soybean is preceded or succeeded by a cereal) are widespread, since the cultural season benefits from a 7-month rainy season (from October to April) and agriculture is mostly mechanized, with field sizes around 100 hectares.

2.2. TerraClass Land-Cover Map

Among the efforts to produce datasets which will effectively help monitor the recent land-use transitions in Brazil, the National Institute for Space Research (INPE) (in partnership with Brazilian Agricultural Research Corporation (EMBRAPA)) has contributed to the TerraClass project [15]. TerraClass products include land-cover maps of the Amazon biome for every two years since 2004, together with a first version of the Cerrado biome that was recently released for the year 2013 [16,17].

For this study, we combined the results of the latest TerraClass products (TerraClass Amazônia 2014 and TerraClass Cerrado 2013) to generate a land-cover map at 30 m spatial resolution covering the

full extent of the Tocantins state. The combined product is used to extract an agricultural land-cover map, composed of the land-cover classes' pasture and rangelands, and annual agriculture (Figure 2).

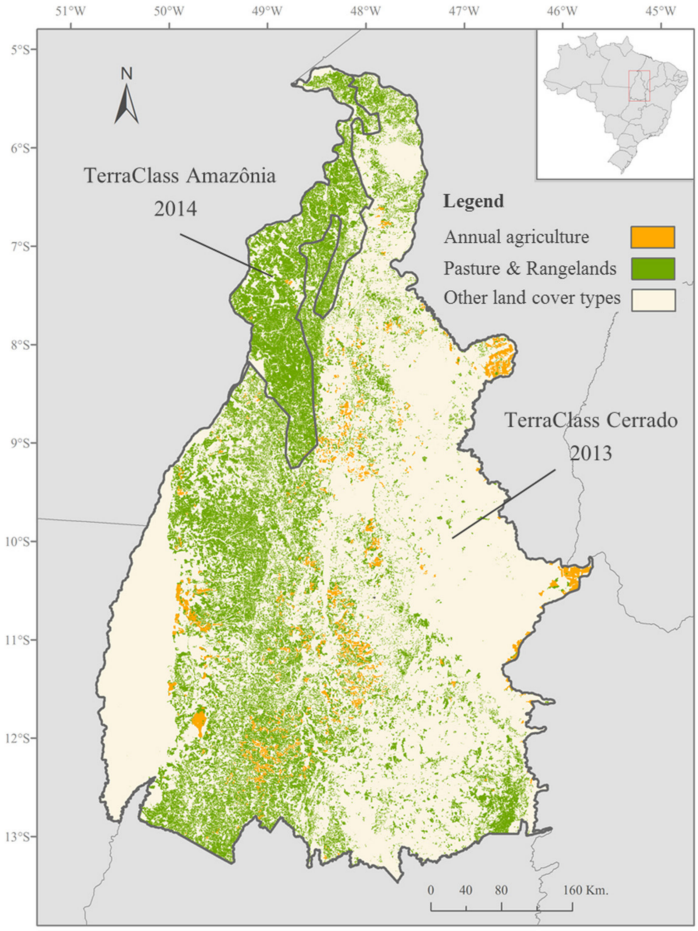


Figure 2. Agricultural land-cover map extracted from the combination of the TerraClass Amazônia 2014 map (covering the North East of the State of Tocantins) and the TerraClass Cerrado 2013 map (covering the rest of the State of Tocantins) [16,17].

2.3. Agricultural Statistics

The official agricultural statistics at the Brazilian municipality administrative level were acquired from the Municipal Agricultural Production database (PAM) of the Brazilian Institute of Geography and Statistics (IBGE) [18]. This database provides annual estimates of planted and harvested area, average yield, average income, and production of the different crops since 1973.

For our study, we collected the harvested area estimates of the 2014 annual crops of the municipalities of Tocantins. The collected data were processed and used as the reference for the spatial distribution of the main cropping systems of Tocantins, against which our results were evaluated.

2.4. MODIS NDVI Time Series

Although the NDVI is known to be affected by background, aerosol effects, and saturation in high biomass regions [19], it is highly sensitive to small increases in the amount of photosynthetic vegetation [20] and is therefore well adapted to capture the vegetation dynamics of the agricultural land-use systems.

A time series of 23 16-day composited NDVI images from October (start of the growing season) 2013 to October 2014 (of the MOD13Q1 v5 product at a 250 m spatial resolution) is used in this

study. The dataset was retrieved from the NASA Land Processes Distributed Active Archive Center (LP DAAC) through the processing R package MODIStsp Version 1.3.2 [21]. Two MOD13Q1 tiles (h13v09 and h13v10) are needed for covering the entire state of Tocantins.

3. Methods

The methods applied in this study follow a typical GEOBIA approach [10] through automatic image partition into meaningful image objects (the land units), and subsequent classification of the resulting objects by an assessment of their characteristics to generate new geographic information (the agricultural land-use systems (ALUS) map). The methodological framework is presented in Figure 3 and described in the following three sub-sections: Pre-processing of the NDVI time series (Section 3.1); Extraction of the land units (Section 3.2); Classification of the land units (Section 3.3). A final sub-section presents the unsupervised evaluation approach that was used to evaluate the cropping systems' classification results of the ALUS map (see Section 3.4. Unsupervised Evaluation of the Classification Results).

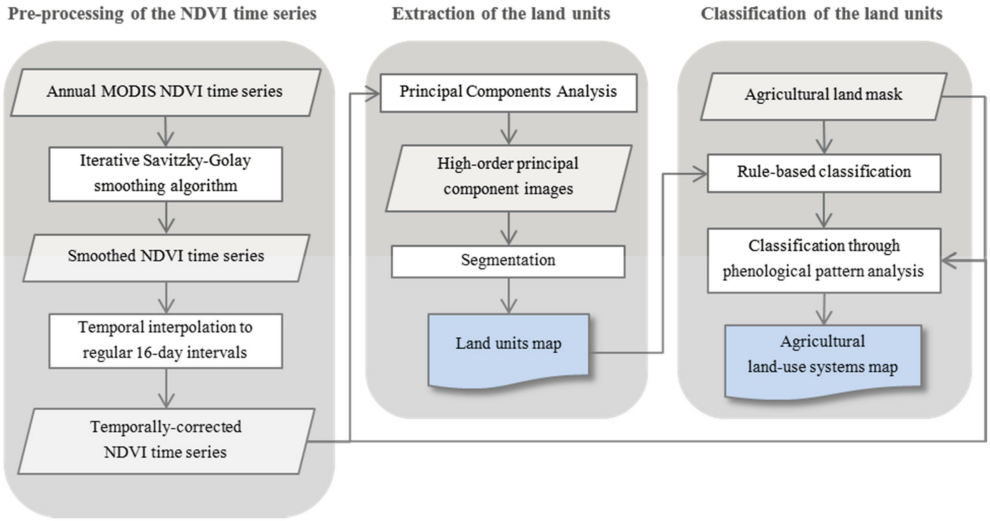


Figure 3. Representation of the methodological framework used in this study.

3.1. Pre-Processing of the NDVI Time Series

The compositing process of the MOD13Q1 product uses the Constrained View angle-Maximum Value Composite (CV-MVC) algorithm, which involves selecting (pixel-wise) the observation with the highest NDVI value and the smallest view angle out of all the observations made during a 16-day cycle [22]. However, remaining effects such as cloudiness or bi-directional effects are still present in the dataset.

In order to generate a noise-reduced time series, an adapted version of the function-fitting smoothing method introduced by Chen et al. [23] was applied. This noise-reduction algorithm is based on a linear interpolation of cloudy pixels, followed by the use of the Savitzky–Golay (SG) filter through an iteration process until the upper envelope of the original NDVI profile is best approached [23,24].

The cloud flag data supplied with the MOD13Q1 product (Pixel Reliability layer) showed systematic overestimations of clouds in the cropland areas, which are particularly evident during the harvest stages in the crop calendar. Consequently, only the pixels followed by an unnatural NDVI increase (greater than 0.4 between two composited images) were considered as cloudy dates and linearly interpolated. In order to reduce the noise of other residual effects such as consecutive cloudiness, the iterative SG filter was applied. After several tests, the parameters d (degree of the smoothing polynomial) and m (half-width of the smoothing window) of the SG filter (for the long-term change curve and for the fitting iteration) were set to $d = 2$, $m = 2$, and $d = 6$, $m = 4$, respectively. This

smoothing method effectively improved the quality of the time series, reducing signal noise while preserving the detail of the NDVI temporal profiles (Figure 4).

To avoid the temporal discontinuities generated by the compositing process (which affect the shape of the NDVI profiles), the values of the smoothed time series were linearly interpolated. This was done using their actual acquisition dates provided in the "Composite Day of Year (CDOY)" layer of the MOD13Q1 product, to place them at 16-day regular intervals.

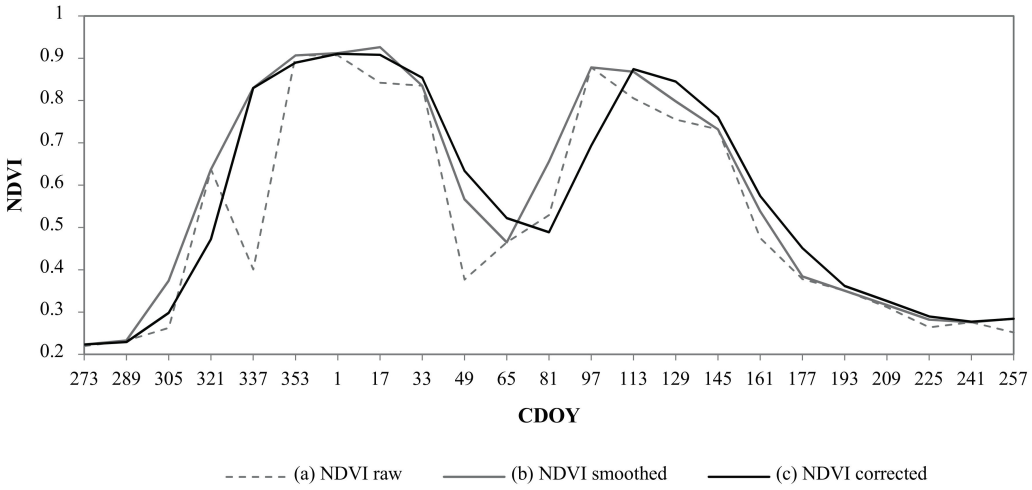


Figure 4. Sample Normalized Difference Vegetation Index (NDVI) time series profile of (**a**) the raw data; (**b**) the filtered and smoothed data; and (**c**) the final temporally-corrected data.

3.2. Extraction of the Land Units

The extraction of the land units is developed in two steps which are described in the following sub-sections: Principal Components Analysis (Section 3.2.1) and Delineation of the Land Units (Section 3.2.2).

3.2.1. Principal Components Analysis

A principal components analysis (PCA) transformation was applied to the corrected NDVI time series. Originally formulated by Pearson (1901), the PCA technique consists of a linear orthogonal transformation (or rotation) of the original variables into a set of uncorrelated variables known as the principal components (PC). When applied to remote sensing time series, PCA computes new values for each pixel of the original time series' images to produce a set of PC images (also known as eigenchannels). The resulting PC images are uncorrelated with one another and ordered according to the amount of variance retained within the total variance of the original time series dataset (from most to least). Low-order PC images (PC1, PC2) concentrate most of the variance and therefore inevitably capture the information redundancy among the original images, while the higher-order PC images capture less redundant information (associated with specific localized change events).

PCA has proven to be particularly effective (when applied to NDVI time series) at identifying seasonal changes of the land cover and revealing their relationship to different factors—mainly climate variability and human activities factors [25–30]. The type of information captured by each PC image will vary according to the characteristics of the input NDVI time series (temporal resolution and length of the time series, spatial resolution and extent) and the PCA method used (either unstandardized if the PC images are calculated using the covariance matrix, or standardized if the correlation matrix is used instead). However, some common outcomes are observed when the area under analysis is large enough to include different types of land cover: The first PC image captures the major element of variability in the NDVI time series (which for large areas corresponds to the spatial variability of the NDVI values, and is correlated to the integrated NDVI over all seasons), whilst the other PC images

capture the seasonal and intra-seasonal variability of NDVI, each related to a particular factor or a combination of factors [25–30].

We expect the phenological patterns of different agricultural land-use systems (configured by seasonal development of the vegetation and intra-seasonal variability and short-term fluctuations induced by the agricultural practices), to be detectable in these latter PC images. Therefore, after performing a PCA on the corrected NDVI time series of 23 images, we retained the last 22 PC images (the high-order PC images) out of the 23 resulting PC images for further analysis, which contain 27% of the total variance. The first PC image was deliberately discarded after verification that the temporal variability of NDVI was not significantly captured, and that (as expected) the 73% of the total variance captured by the first PC image is mostly related to the spatial variability of the NDVI values, as shown in other studies (e.g., [27,28]).

The PCA was performed by means of the rasterPCA tool in the processing R package RStoolbox Version 0.1.8 [31] and was based on the covariance matrix (unstandardized PCA). Unstandardized PCA is preferred since sudden changes in the magnitude of the NDVI related to land use (such as a rapid NDVI decrease due to crop harvest) will contribute significantly to the development of the new component images. Conversely, normalizing the input images through standardized PCA would weaken the contribution of these sudden change events by forcing all the original input images to have equal weight in the derivation of the PC images.

3.2.2. Delineation of the Land Units

The segmentation of the high-order PC images was performed using the iterative region-growing algorithm based on the Baatz and Schäpe criterion, implemented in the Multiresolution Segmentation tool of the object-orientated image analysis software eCognition® Developer 9.0 [32]. This algorithm starts with each pixel of the image forming an image object. At each iteration, neighboring objects are merged based on their similar characteristics described by a homogeneity criterion, until a user-defined threshold (known as the scale parameter) is reached. The multiresolution segmentation algorithm ensures a regular growth of the image objects over the whole scene, and so the resulting objects have a similar size, which is determined by the scale parameter threshold. In this way, a low scale parameter will lead to many objects of small size and, conversely, a high scale parameter will allow the growth of few large objects. The influence of the radiometry versus the shape in the definition of the homogeneity criterion is determined by the weight assigned to the color and shape parameters [33].

The objective of the segmentation being the delineation of land units with relatively homogeneous phenological patterns, the shape of the objects in the definition of the homogeneity criterion becomes irrelevant and, therefore, we assign the total weight to the color by setting the color parameter to 1 and the shape parameter to 0. In practice, setting the color parameter to 1 means that the relative homogeneity of the resulting image objects (land units) will be entirely based on the pixel values of the PC images and their implicit information related to the temporal variability of the vegetation cover. Besides, each of the 22 PC images is made to contribute equally to the segmentation by setting the weight of each input PC image to 1.

A scale (according to Benz et al. [34]) is defined as the level of aggregation and abstraction at which an object can be clearly described. Since our geographic objects of interest (the land units) do not relate to an inherent scale but can rather be described at different scales of analysis, we consider that there is no optimal segmentation scale and that different scale parameters should be tested in each particular case to approach adequately sized, thematically meaningful land units. We assume that the multiresolution segmentation algorithm based on homogeneity criterion delineates relatively homogeneous objects regardless of the scale parameter used. However, some simple guiding principles were followed in this study to choose among different scale parameters aiming at meaningful land units (in terms of relatively homogeneous phenological patterns): Objects should be larger than the minimum mapping unit of an agricultural land-use system (i.e., the field) and there should be a sensible

trade-off between the number of objects for further analysis and their degree of heterogeneity. After testing several segmentation levels, the scale parameter threshold was empirically set to 850 (unitless).

3.3. Classification of the Land Units

The classification of the land units is developed in two steps which are described in the following sub-sections: Rule-based Classification (Section 3.3.1) and Classification through Phenological Pattern Analysis (Section 3.3.2).

3.3.1. Rule-Based Classification

The agricultural land-cover map extracted from the TerraClass product was used for the classification of the land units belonging to the agricultural domain (either to the crop agriculture domain or to the livestock domain) into four major ALUS: Dominant crop agriculture system, coexistence (crop-livestock) system, semi-intensive livestock system, and intensive livestock system. This preliminary classification of the ALUS was performed following the decision tree ruleset developed by Almeida et al. [35] for the localization of agricultural production systems in the Brazilian State Rondônia at the municipal level, presented in Figure 5. This ruleset seems well adapted for the classification of the ALUS in the Tocantins state, owing to its closeness to the Rondônia state (in terms of land-cover and land-use systems' characteristics).

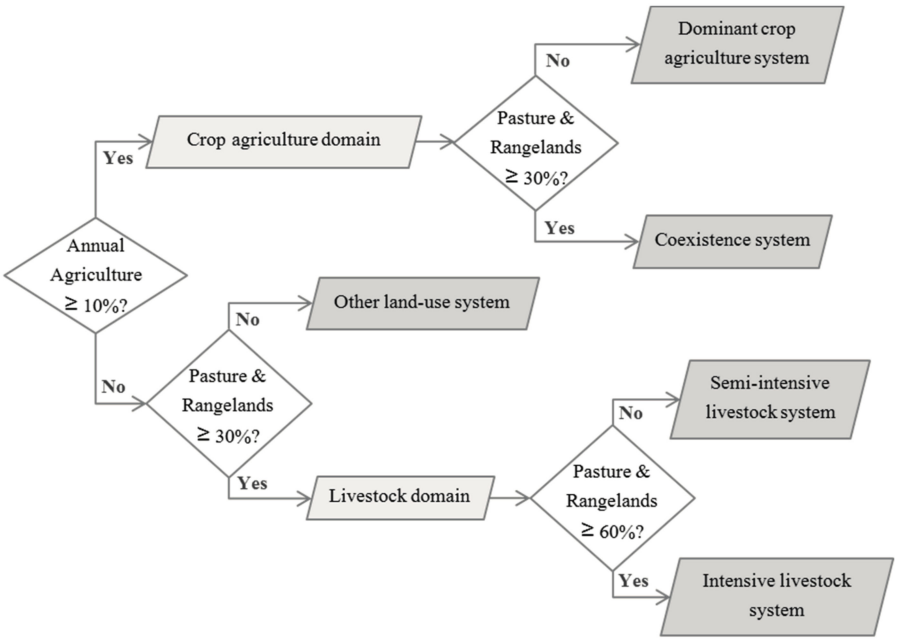


Figure 5. Decision tree ruleset used for the classification of the land units into four major agricultural land-use systems adapted from Almeida et al. [35].

3.3.2. Classification through Phenological Pattern Analysis

The land units belonging to the crop agriculture domain were further classified into different cropping systems based on the analysis of their phenological patterns. The mean temporal NDVI profile of the cropland cover inside each land unit was extracted from the MODIS NDVI time series using the TerraClass annual agriculture class as a mask. The general NDVI variations over the growing season were analyzed through visual interpretation of the shape of the resulting mean temporal NDVI profiles, which revealed the principal crop development stages of the cropland cover inside each land unit. The NDVI profiles with visually equivalent shapes were grouped together and were

finally classified into different cropping systems by correspondence with the crop calendar of the main cropping systems of Tocantins.

3.4. Unsupervised Evaluation of the Classification Results

Since there are no available thematically-equivalent maps with which to validate the accuracy of the classification of the ALUS, an unsupervised evaluation was carried out with a reference map extracted from the official agricultural statistics (area estimates of the annual crops in the municipalities of Tocantins for the year 2014). The evaluation was only carried out for the cropping systems' results since no official statistics were available on the pasture and rangelands with which to evaluate the livestock systems' classification results.

In order to produce a synthetic reference map out of the detailed agricultural statistics dataset that could be compared to the cropping systems in the ALUS map, we first calculated the proportion of area occupied by each crop type in the municipalities with more than 2000 hectares of cultivated area. The predominant crop types (accounting for most of the cultivated area) were differentiated from minor crops and were used to characterize the main cropping system related to each municipality.

The ALUS map was finally evaluated based on the spatial agreement between the classified cropping systems and the cropping systems in the reference map.

4. Results

4.1. Land Units

The segmentation of the high-order principal component images produced a total of 90 land units (Figure 6), with an average size of 3090 km^2, 44 km^2 for the smallest one, and 15,069 km^2 for the largest one.

4.2. Agricultural Land-Use Systems

The decision tree ruleset based on the agricultural land-cover map extracted from the TerraClass product identified a total of 44 land units belonging to the agricultural domain (114,114 km^2) among the 90 automatically delineated land units. Out of the 44 land units, 21 were classed as dominant crop agriculture systems (11,193 km^2), 15 as semi-intensive livestock systems (82,576 km^2), 7 as intensive livestock systems (18,329 km^2) and 1 was classed as a coexistence system (2016 km^2) (the latter corresponding to an area where both a crop agriculture system and a livestock system are present).

Three distinct phenological patterns were identified from the analysis of the shape of the temporal NDVI profiles of the 22 land units belonging to the crop agriculture domain. The plots of the 22 temporal NDVI profiles are presented in Figure 7 in three groups of phenological patterns.

In the first group a major single crop cycle prevails, with a peak NDVI period around February (from day 33 to day 65) (Figure 7). Since only minor increases of NDVI appear before and after the main growing cycle, we can consider that there is no secondary meaningful development cycle of the vegetation cover, and that, therefore, the land units belonging to the first group are mainly related to a single cropping system. Since the soybean crop is the dominant crop of Tocantins, the first group of land units seems distinctive of the soybean single cropping system.

The second and the third groups of land units both show a characteristic double NDVI peak representing separate crop cycles (Figure 7). Although the first cycle of both groups has a common peak NDVI period around January (from day 1 to day 33), these two groups are easily dissociable by their temporal NDVI profile shapes, for their second cycle appears with more than a month difference between the two groups. The phenological pattern of the second group is distinctive of the soybean–maize double cropping system, which is characterized by a summer soybean crop during the rainy season followed by a winter maize crop. This phenological pattern exhibits low NDVI values around Mars (from day 65 to day 81) as a result of the soybean harvest, after which the winter maize is

usually immediately sown and subsequently harvested during June and July (shown by the low NDVI values after day 161).

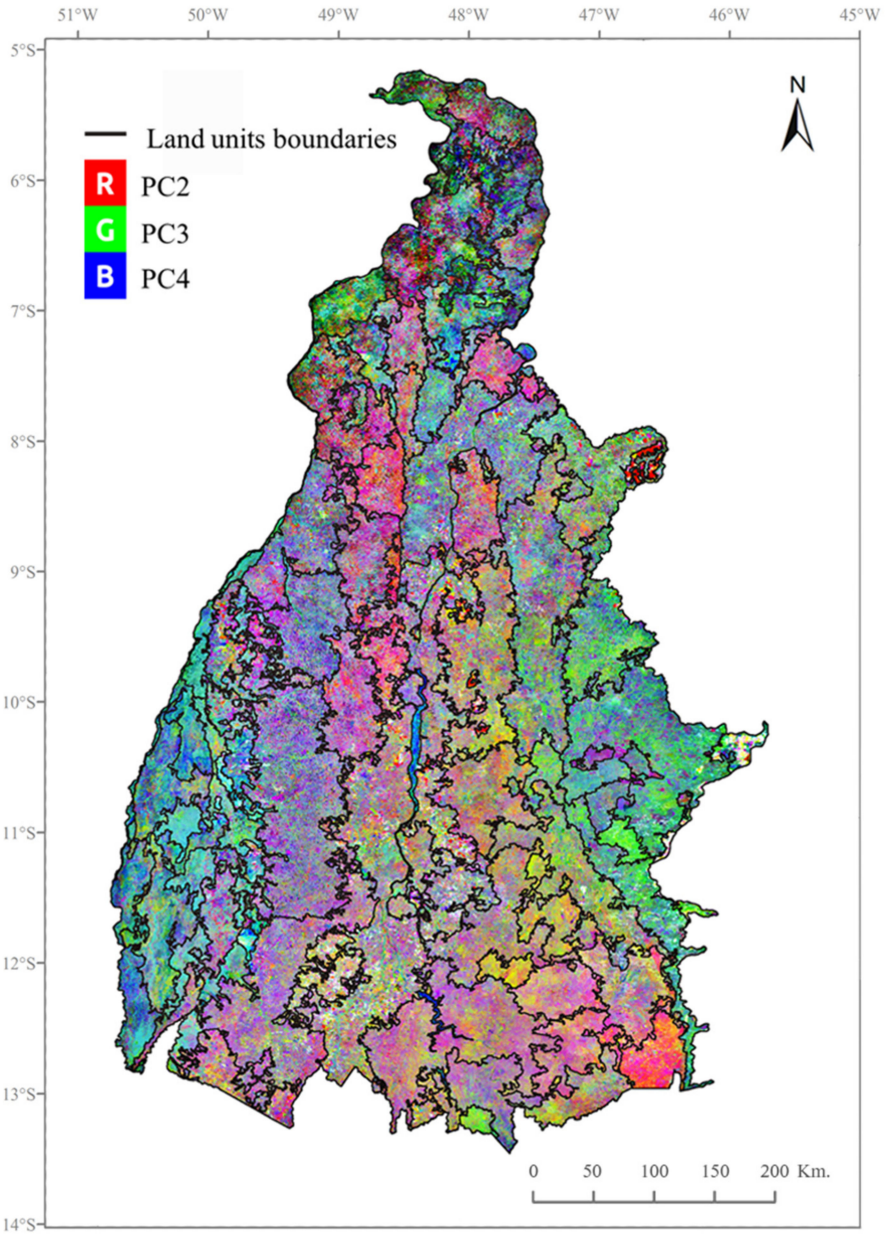


Figure 6. Land units' boundaries over a color composition of the first three principal component (PC) images (RGB PC2, PC3, PC4) of the 22 PC images used as the segmentation variables.

The phenological pattern of the third group is distinctive of the rice–soybean double cropping system, which is the characteristic cropping system of the lowlands located in the southwest of Tocantins in the Formoso river basin. This particular cropping system is characterized by a flood-irrigated rice crop during the rainy season followed by a sub-irrigated soybean crop from May to September (mostly for seed production). The low NDVI values around May (from day 113 to day 145), followed by an increase until the second peak NDVI period is reached around July (from day 193 to day 225—shown in Figure 7), evidences the late second soybean crop cycle, which is unique to the rice–soybean double cropping system.

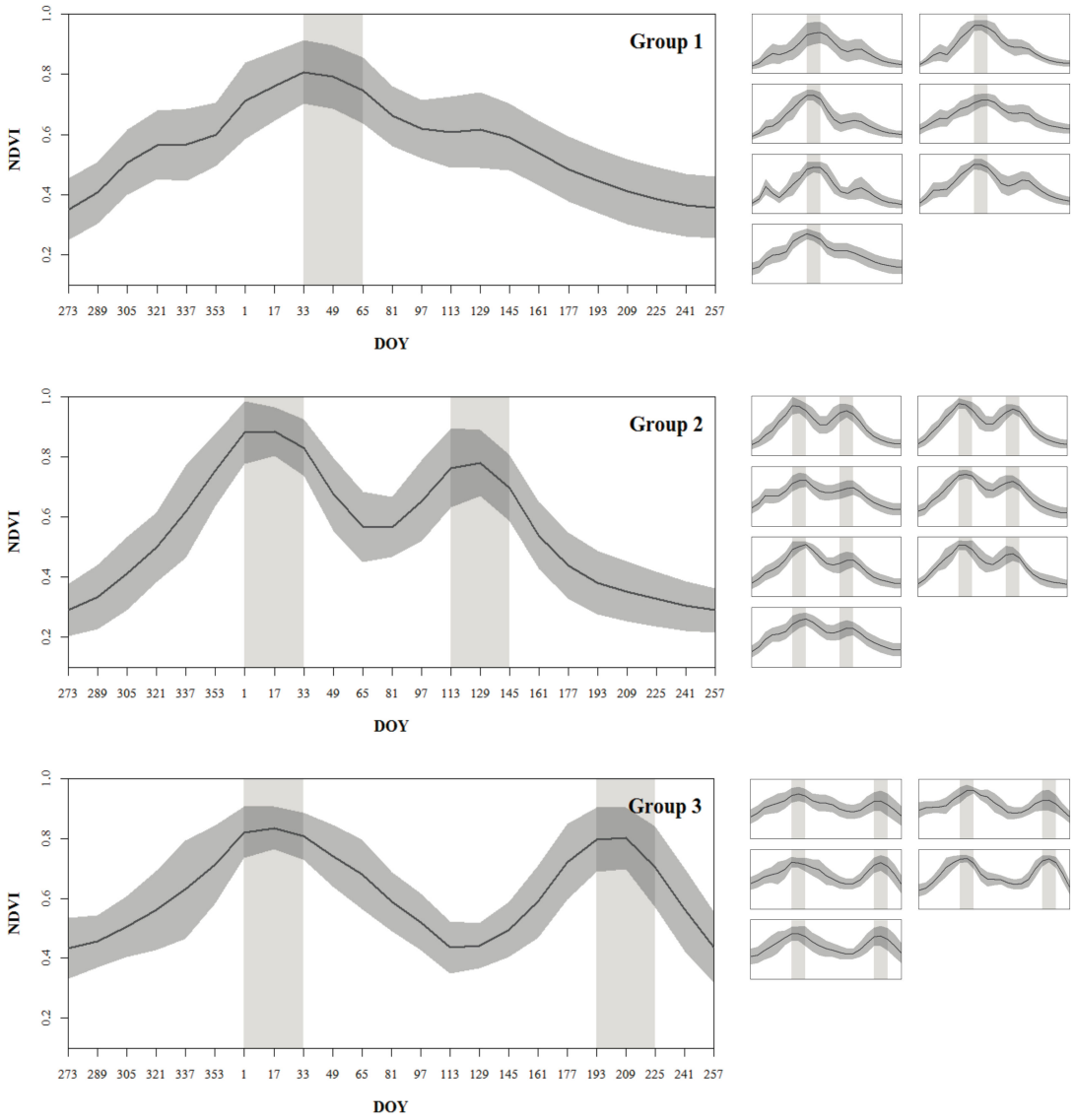


Figure 7. The temporal NDVI profiles' plots of the 22 crop agriculture domain land units. The plots show the mean (solid black curve) and the standard deviation (shaded dark grey curve) of the NDVI pixel values inside the annual agriculture TerraClass class for each land unit against the day of the year (DOY). A representative land-unit profile for each of the three phenological pattern groups is shown in detail at the left and the profiles of the other land units belonging to each group are represented assembled at their right side. The "monthly" (32-day) periods (at which the NDVI profiles peak) are represented by shaded light grey vertical bands, and are consistent between the plots of each group.

The crop agriculture domain land units were therefore sub-classified into the three main cropping systems of Tocantins according to the correspondence of their phenological patterns with the characteristic cropping pattern of each cropping system. The map of the main ALUS of Tocantins with the final classification is presented in Figure 8.

As shown in Figure 8, the livestock systems occupy a larger area (about eight times larger) than the cropping systems, and are mainly distributed along a longitudinal transect located in the center-west of the state. The cropping systems are concentrated in specific locations that are sparsely distributed over most of the state—except for the northern third of the state, where the livestock systems are predominant. The livestock systems and the cropping systems are mostly separated, except for the only coexistence area located in the southern region of the livestock-dominant longitudinal transect.

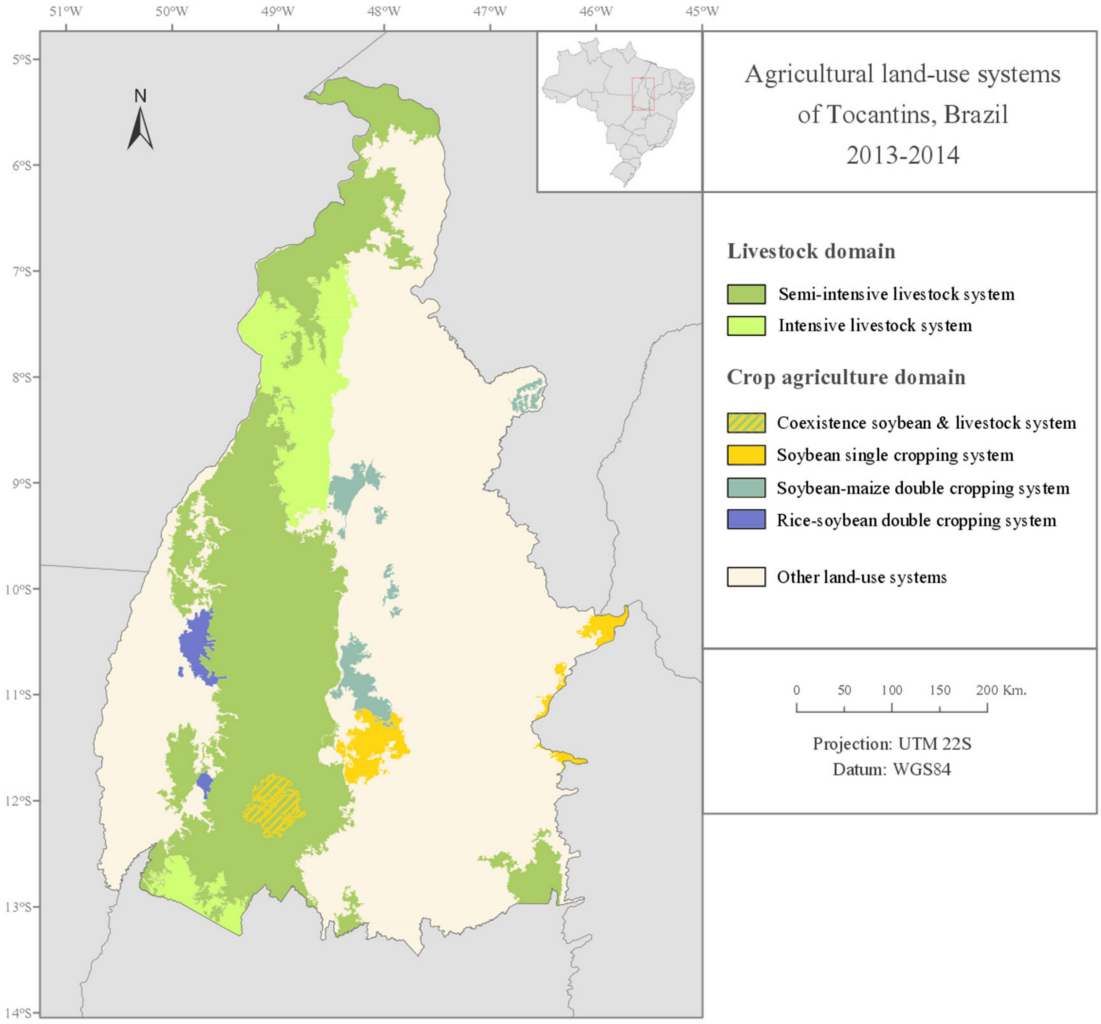


Figure 8. Map of the main agricultural land-use systems of the Tocantins state in the 2013–2014 growing season.

4.3. Unsupervised Evaluation of the Classification Results

The agricultural statistics at the municipal level revealed three predominant crops accounting for more than 94% of the total cultivated area: Soybean (75%), rice (11.1%), and winter maize (8.6%). All of the municipalities with more than 2000 hectares of cultivated area present a high proportion of the soybean crop, with 29–100% of their cultivated area dedicated to its production (which confirms the dominance of this crop type in Tocantins). However, three distinct classes of municipalities were observed. A first class (composed of 15 municipalities)—where the cultivated area is almost exclusively covered by soybean—was related to the soybean single cropping system. A second class (composed of 7 municipalities)—where the proportion of winter maize is considerable—was related to the soybean–maize double cropping system. Finally, a third class (composed of 3 municipalities)—where the proportion of rice is nearly equivalent to the proportion of soybean—was related to the rice–soybean double cropping system. The average proportion of the different crop types for each class of municipalities is presented in Figure 9.

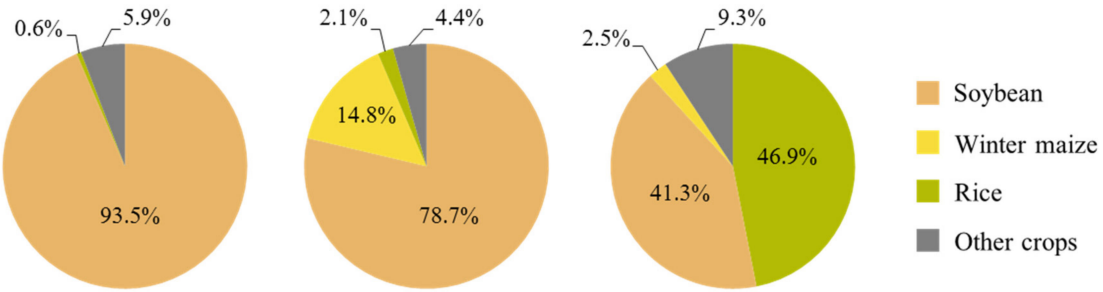
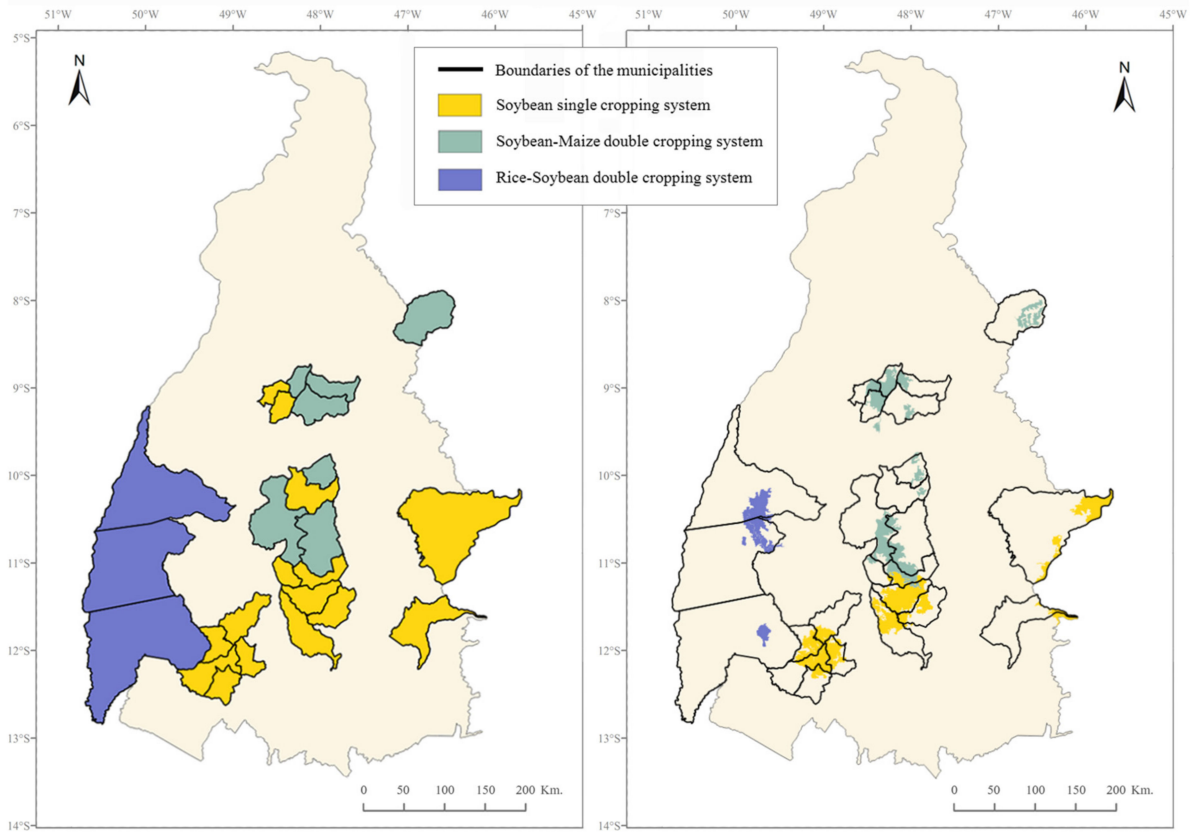


Figure 9. The average proportion of the crops of the municipalities belonging to the first class (**left**), the second class (**center**), and the third class (**right**), related to the soybean single cropping system, to the soybean–maize double cropping system, and to the rice–soybean double cropping system, respectively.

The resulting classification of the municipalities is represented on a map together with the classification of the cropping systems in the ALUS remote sensing-based map (Figure 10).



(**a**) (**b**)

Figure 10. The spatial distribution of the three main cropping systems (**a**) in the reference map at the municipal level, extracted from the agricultural statistics, and (**b**) in the agricultural land-use systems remote sensing-based map. The boundaries of the municipalities are included in both maps to facilitate their comparison.

Although the reference map and the ALUS map are produced at different scales, an apparent coherence between both maps can be recognized for the spatial distribution of the soybean single cropping system and the rice–soybean double cropping system (Figure 10). The spatial distribution of the soybean–maize double cropping system shows general coherence between both maps but seems

overestimated locally for some regions, covering some municipalities which are classified as being dominated by the soybean single cropping system in the reference map.

5. Discussion

The study area was partitioned into relatively homogeneous land units that were linked to a specific type of agricultural land use, resulting in a regional level agricultural land-use systems' map. Two livestock systems could be differentiated by the proportion of pasture and rangelands in the land units, and the three main cropping systems of the study area could be differentiated through the proportion of annual cropland in the land units and through their specific phenological patterns (Figure 7); this allowed not only to characterize the land units in terms of the crop type but also in terms of the cropping pattern.

The results showed a general agreement between the location of the three main cropping systems in the final agricultural land-use systems map and the reference map, except for some localized overestimations of the presence of the soybean–maize double cropping system over some of the reference soybean single cropping system municipalities. This overestimation might be due to the non-inclusion of the non-commercial summer grass cover crops such as millet, sorghum, or brachiaria in the agricultural statistics and, therefore, in the reference map. Cover crops are sometimes sown after the winter soybean harvest instead of maize, as part of the widely adopted no-tillage system in the study area. These cover crops might contribute to the second peak of NDVI that is apparent in the NDVI temporal profiles of the land units, which partially intersect municipalities where only the soybean surface area has been officially censused.

The reference map that was used for the evaluation of the classification results shows that land-use systems' maps can be produced from the agricultural statistics. However, the thematic resolution of this type of map is limited by the data available in the statistics, thus not accounting for certain systems (e.g., the livestock systems). Furthermore, the spatial accuracy is greatly limited by the data aggregation level (e.g., the municipal level)—the administrative boundaries not being representative of the actual spatial distribution of the land-use systems. The presented remote sensing approach produces a land-use systems' map which overcomes these limitations, by automatically delineating boundaries which are representative of the actual physical differences between zones related to the temporal variations of the vegetation cover.

While the ability of principal components analysis to uncover significant intra-annual change events out of vegetation index time series has been widely proven (e.g., [25–30]), studies mainly focus on the low-order principal component (PC) images—generally the first three—which capture most of the variance. We decided to include the higher-order PC images in our analysis, since they are highly sensitive to subtle changes in the data and therefore capture change events which are localized both spatially and in time, such as the intra-seasonal variations and rapid changes induced by the agricultural practices. Overall, the results confirm that the high-order PC images do capture the information related to the temporal variability of the vegetation cover, which was the essential portion of information of the NDVI time series (represented by 27% of the total variance in the time series) needed for the segmentation of the study site in land units with relatively homogeneous phenological patterns. Furthermore, the phenological patterns shown in Figure 7 reflected the intra-seasonal variations which were linked to cropping practices such as the approximate planting and harvesting dates of different crops.

The automatic extraction of relatively homogeneous land units in terms of phenological patterns, through a principal components analysis and an automatic segmentation, is a simple and innovative approach that can be reproduced in other regions. This approach overcomes the potential error propagation of the expert-based land units' border delineation and the subjective choice and use of variables in the traditional land-use systems' mapping approaches. Furthermore, the presented approach is scale-independent and therefore allows multi-scale analysis, which might help improve the understanding of the land-use systems by characterizing them at different organizational levels.

In addition, the decision tree ruleset used for the classification of four major agricultural land-use systems (presented in the Figure 5—adapted from Almeida, et al. [35]), is conceived with simple threshold rules based on the percentage cover of agricultural land, which can be easily adapted or directly applied, to other agricultural regions. This classification is, however, dependent on a land-cover map. Even if land-cover maps are now produced and available for many regions, their accuracy will condition the agricultural land-use systems' classification results, and they may be rarely updated. Consequently, the rule-based classification step would largely benefit from the inclusion of an automatic remote sensing-based classification method to produce timely and accurate regional-scale agricultural land-cover maps at the field level. Some studies have shown that partitioning the study region into sub-regions improves the accuracy of regional-scale land-cover classifications [36,37]. Hence, the presented land units' extraction approach could therefore also be used as a preliminary step for automatic remote sensing-based classifications, producing regional-scale land-cover and land-use maps at the field level.

6. Conclusions

Remote sensing-derived time series of NDVI represent an important source of spectral and temporal information related to agricultural land use and, when combined with GEOBIA techniques in the present study, allowed the discrimination of different land-use systems (in particular, cropping systems) at a regional scale. Remote sensing-based approaches, therefore, present a significant potential for agricultural land-use system mapping over large areas and benefit from the spatial and temporal continuity of the satellite data and technological advances such as the new high temporal and spatial resolution Sentinel-2 satellite systems of the Copernicus program.

The presented approach can be potentially reproduced in other regions with minimal adaptation to specific context, thus contributing to the development of generic and simple tools for the location and characterization of the agricultural land-use systems across regions. This baseline-detailed spatial information is highly valuable for further large-scale agronomic and environmental assessments. If reproduced periodically, this type of approach can help with long-term analysis of the dynamics of the major agricultural land-use systems, which can help monitor large-scale land-use change processes such as cropland expansion or intensification (being particularly useful in changing regions such as the MATOPIBA region, for effective land-use planning leading to sustainable use of the resources).

Acknowledgments: Beatriz Bellón received a fellowship from CIRAD (French Agricultural Research Centre for International Development) and from the FP7 SIGMA (Stimulating Innovation for Global Monitoring of Agriculture) project. This work was also supported by the French Space Agency (CNES) through the funding of the TOSCA AGRIZONE project and by the CAPES/COFECUB program through the funding of the GeoABC Project (Methodologies and technological innovation for satellite monitoring of low carbon agriculture in support to Brazil's ABC Plan), project No. 845/15. The authors would like to gratefully acknowledge the members of the NTSA (Thematic Unit on Agricultural Systems) of EMBRAPA (Brazilian Agricultural Research Corporation) Fisheries and Aquaculture for their valuable contribution to this work.

Author Contributions: Beatriz Bellón, Agnès Bégué and Danny Lo Seen conceived and designed the experiments; Beatriz Bellón processed the data and performed the proposed methods; Beatriz Bellón, Agnès Bégué and Danny Lo Seen analyzed the results; Beatriz Bellón wrote the paper; and Agnès Bégué, Danny Lo Seen, Claudio Aparecido De Almeida and Margareth Simões supervised the research and contributed to the editing and review of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. FAO. How to feed the world in 2050. In Proceedings of the Expert Meeting on How to Feed the World in 2050, Rome, Italy, 24–26 June 2009.
2. Nachtergaele, F.; Petri, M. *Mapping Land Use Systems at Global and Regional Scales for Land Degradation Assessment Analysis*; FAO: Rome, Italy, 2013.

3. Bégué, A.; Arvor, D.; Lelong, C.; Vintrou, E.; Simoes, M. Agricultural systems studies using remote sensing. In *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing*; CRC Press: Boca Raton, FL, USA, 2015; pp. 113–130.
4. Driessen, P.M.; Konijn, N.T. *Land-Use Systems Analysis*; Wageningen Agricultural University: Wageningen, The Netherlands, 1992.
5. FAO. *Guidelines for Land-Use Planning*; FAO: Rome, Italy, 1993.
6. Arvor, D.; Jonathan, M.; Meirelles, M.S.P.; Dubreuil, V.; Durieux, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *Int. J. Remote Sens.* **2011**, *32*, 7847–7871. [CrossRef]
7. Cheema, M.J.M.; Bastiaanssen, W.G.M. Land use and land cover classification in the irrigated Indus Basin using growth phenology information from satellite data to support water management analysis. *Agr. Water Manag.* **2010**, *97*, 1541–1552. [CrossRef]
8. Kiptala, J.K.; Mohamed, Y.; Mul, M.L.; Cheema, M.J.M.; Van Der Zaag, P. Land use and land cover classification using phenological variability from MODIS vegetation in the Upper Pangani River Basin, Eastern Africa. *Phys. Chem. Earth* **2013**, *66*, 112–122. [CrossRef]
9. Wardlow, B.D.; Egbert, S.L. Large-area crop mapping using time-series MODIS 250m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116. [CrossRef]
10. Hay, G.J.; Castilla, G. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In *Object-Based Image Analysis*; Blaschke, T., Lang, S., Hay, G.J., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 75–89.
11. Bisquert, M.; Bégué, A.; Deshayes, M. Object-based delineation of homogeneous landscape units at regional scale based on MODIS time series. *Int. J. Appl. Earth Obs.* **2015**, *37*, 72–82. [CrossRef]
12. Bisquert, M.; Bégué, A.; Deshayes, M.; Ducrot, D. Environmental evaluation of MODIS-derived land units. *GIsci. Remote Sens.* **2016**, *54*, 64–77. [CrossRef]
13. Silva, L.A.G.C. *Biomas Presentes no Estado do Tocantins*; Consultoria Legislativa da Câmara dos Deputados: Brasília, Brazil, 2007.
14. CONAB. Soja: Série Histórica de Produtividade (Safras 2000/01 a 2015/16). Available online: http://www.conab.gov.br/ (accessed on 15 April 2017).
15. Almeida, C.A.; Coutinho, A.C.; Esquerdo, J.C.D.M.; Adami, M.; Venturieri, A.; Diniz, C.G.; Dessay, N.; Durieux, L.; Gomes, A.R. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. *Acta Amaz.* **2016**, *46*, 291–302. [CrossRef]
16. INPE. Projeto Terraclass Cerrado: Mapeamento do uso e Cobertura Vegetal do Cerrado. Available online: http://www.dpi.inpe.br/tccerrado/ (accessed on 22 March 2017).
17. INPE. Projeto Terraclass Amazônia: Mapeamento do uso e Cobertura da Terra na Amazônia Legal Brasileira. Available online: http://www.inpe.br/cra/projetos_pesquisas/dados_terraclass.php (accessed on 22 March 2017).
18. IBGE, Instituto Brasileiro de Geografia e Estatística. Produção Agrícola Municipal (PAM) 2014. Available online: http://www.ibge.gov.br/home/estatistica/economia/pam/2014/ (accessed on 15 April 2017).
19. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
20. Soudani, K.; le Maire, G.; Dufrêne, E.; François, C.; Delpierre, N.; Ulrich, E.; Cecchini, S. Evaluation of the onset of green-up in temperate deciduous broadleaf forests derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sens. Environ.* **2008**, *112*, 2643–2655. [CrossRef]
21. Busetto, L.; Ranghetti, L. MODIStsp: An R Package for Preprocessing of MODIS Time Series. R Package Version 1.3.2. 2017. Available online: https://cran.rstudio.com/web/packages/MODIStsp/index.html (accessed on 15 April 2017).
22. Solano, R.; Didan, K.; Jacobson, A.; Huete, A. *MODIS Vegetation Index User's Guide (MOD13 Series)*; Vegetation Index and Phenology Lab, University of Arizona: Tucson, AZ, USA, 2010.
23. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. [CrossRef]
24. Savitzky, A.; Golay, M.J.E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]

25. Eastman, J.R.; Fulk, M. Long sequence time series evaluation using standardized principal components. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 991–996.
26. Gurgel, H.C.; Ferreira, N.J. Annual and interannual variability of NDVI in Brazil and its connections with climate. *Int. J. Remote Sens.* **2003**, *24*, 3595–3609. [CrossRef]
27. Hall-Beyer, M. Comparison of single-year and multiyear NDVI time series principal components in cold temperate biomes. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2568–2574. [CrossRef]
28. Hirosawa, Y.; Marsh, S.E.; Kliman, D.H. Application of standardized principal component analysis of land-cover characterization using multitemporal AVHRR data. *Remote Sens. Environ.* **1996**, *58*, 267–281. [CrossRef]
29. Wang, T.; Kou, X.; Xiong, Y.; Mou, P.; Wu, J.; Ge, J. Temporal and spatial patterns of NDVI and their relationship to precipitation in the Loess Plateau of China. *Int. J. Remote Sens.* **2010**, *31*, 1943–1958. [CrossRef]
30. Young, S.S.; Anyamba, A. Comparison of NOAA/NASA PAL and NOAA GVI data for vegetation change studies over China. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 679–696.
31. Leutner, B.; Horning, N. Rstoolbox: Tools for Remote Sensing Data Analysis. R Package Version 0.1.8. 2017. Available online: https://cran.r-project.org/web/packages/RStoolbox/index.html (accessed on 15 April 2017).
32. Trimble. *Ecognition Developer© 9.0 User Guide*; Trimble Germany GmbH: Munich, Germany, 2014.
33. Baatz, M.; Schäpe, A. Multiresolution segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. In *Angewandte Geographische Informations-Verarbeitung, XII*; Wichmann Verlag: Karlsruhe, Germany, 2000; pp. 12–23.
34. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm.* **2004**, *58*, 239–258. [CrossRef]
35. Almeida, C.A.; Mourão, M.; Dessay, N.; Lacques, A.-E.; Monteiro, A.; Durieux, L.; Venturieri, A.; Seyler, F. Typologies and spatialization of agricultural production systems in Rondônia, Brazil: Linking land use, socioeconomics and territorial configuration. *Land* **2016**, *5*, 18. [CrossRef]
36. Cai, H.; Zhang, S.; Bu, K.; Yang, J.; Chang, L. Integrating geographical data and phenological characteristics derived from MODIS data for improving land cover mapping. *J. Geogr. Sci.* **2011**, *21*, 705–718. [CrossRef]
37. Cano, E.; Denux, J.-P.; Bisquert, M.; Hubert-Moy, L.; Chéret, V. Improved forest-cover mapping based on MODIS time series and landscape stratification. *Int. J. Remote Sens.* **2017**, *38*, 1865–1888. [CrossRef]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Int J Appl Earth Obs Geoinformation 68 (2018) 127–138

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Improved regional-scale Brazilian cropping systems' mapping based on a semi-automatic object-based clustering approach

Beatriz Bellón[a,b,*], Agnès Bégué[a,b], Danny Lo Seen[a,b], Valentine Lebourgeois[a,b], Balbino Antônio Evangelista[c], Margareth Simões[d,e], Rodrigo Peçanha Demonte Ferraz[d]

[a] CIRAD, UMR TETIS, F-34398 Montpellier, France
[b] TETIS, Univ Montpellier, AgroParisTech, CIRAD, CNRS, IRSTEA, Montpellier, France
[c] Embrapa Pesca e Aquicultura, Prolongamento da Avenida NS 10, cruzamento com a Avenida LO 18, sentido Norte, loteamento Água Fria, 77008-900 Palmas, TO, Brazil
[d] Embrapa Solos, Rua Jardim Botânico, 1024, 22460−000, Rio de Janeiro, RJ, Brazil
[e] Rio de Janeiro State University (UERJ), Rua São Francisco Xavier, 524, Maracanã, 20550-900, Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Keywords:
Cropping systems
Geographic object-based image analysis (GEOBIA)
Stratification
Clustering
Moderate resolution imaging spectroradiometer (MODIS)
Landsat

ABSTRACT

Cropping systems' maps at fine scale over large areas provide key information for further agricultural production and environmental impact assessments, and thus represent a valuable tool for effective land-use planning. There is, therefore, a growing interest in mapping cropping systems in an operational manner over large areas, and remote sensing approaches based on vegetation index time series analysis have proven to be an efficient tool. However, supervised pixel-based approaches are commonly adopted, requiring resource consuming field campaigns to gather training data. In this paper, we present a new object-based unsupervised classification approach tested on an annual MODIS 16-day composite Normalized Difference Vegetation Index time series and a Landsat 8 mosaic of the State of Tocantins, Brazil, for the 2014–2015 growing season. Two variants of the approach are compared: an hyperclustering approach, and a landscape-clustering approach involving a previous stratification of the study area into landscape units on which the clustering is then performed. The main cropping systems of Tocantins, characterized by the crop types and cropping patterns, were efficiently mapped with the landscape-clustering approach. Results show that stratification prior to clustering significantly improves the classification accuracies for underrepresented and sparsely distributed cropping systems. This study illustrates the potential of unsupervised classification for large area cropping systems' mapping and contributes to the development of generic tools for supporting large-scale agricultural monitoring across regions.

1. Introduction

Agricultural systems are constantly evolving to meet the increasing global demand for food production (FAO, 2009). Both the cropland expansion and the intensification of management practices contribute to meet such demand, increasing the concern on the environmental impact associated with these land-use changes (Tilman et al., 2011). In view of these large-scale dynamics, there is a growing need for regularly updated spatial information on the cropping systems at regional scales.

Cropping systems result of a combination of human and environmental factors, which naturally vary through space. A given cropping system thus presents local specificities, leading to a high spatial variability over large areas. In addition, when considering large areas, the cropping systems are embedded in a complex landscape mosaic composed of a high diversity of land use and land cover (LULC) types. Therefore, large-area mapping of cropping systems and their management practices, remains a challenging task, and as a result, their description and location remain rather unclear for most regions (Leenhardt et al., 2010).

Recent studies have however shown the usefulness of earth observation (EO) data to successfully map cropping systems, and some management practices over large areas (Atzberger, 2013; Bégué et al., 2015). In the last decade, various of these studies have successfully used vegetation index (VI) time series, mostly MODIS-derived, to map, at a regional scale, intensification practices such as the adoption of sequential cropping (e.g. Arvor et al., 2011; Brown et al., 2013; Cheema and Bastiaanssen, 2010; Guan et al., 2016; Gumma et al., 2016; Mingwei et al., 2008; Wardlow and Egbert, 2008)

Most of these studies, however, rely entirely on reference data, usually ground observations, to train supervised classifications (e.g.

* Corresponding author.
E-mail addresses: beatriz.bellon@cirad.fr (B. Bellón), agnes.begue@cirad.fr (A. Bégué), danny.lo_seen@cirad.fr (D. Lo Seen), valentine.lebourgeois@cirad.fr (V. Lebourgeois), balbino.evangelista@embrapa.br (B.A. Evangelista), margareth.simoes@embrapa.br (M. Simões), rodrigo.demonte@embrapa.br (R.P. Demonte Ferraz).

https://doi.org/10.1016/j.jag.2018.01.019
Received 6 November 2017; Received in revised form 29 January 2018; Accepted 30 January 2018
0303-2434/ © 2018 Elsevier B.V. All rights reserved.

Arvor et al., 2011; Brown et al., 2013; Mingwei et al., 2008). Gathering reference data twice or more during a growing season to train classifications of sequential cropping systems represents a time and resource consuming task. This is especially the case for large areas since it is difficult to extensively cover the study area to guarantee representativeness of the intra-class spectral variability linked to the high spatial variability (Franklin and Wulder, 2002).

Contrastingly, unsupervised classifications, being solely based on the VI time series' information, are more operational without affecting the accuracy of the resulting maps, as shown by Cheema and Bastiaanssen (2010), Wardlow and Egbert (2008), Gumma et al. (2016). Hyperclustering techniques are commonly used, producing an initial large number of clusters that are then reduced by successive processing steps. This technique ensures that under-represented classes of interest are correctly isolated, and has the advantage of reducing the impact of the user-defined parameters (e.g. the number of clusters) on the results (Cihlar, 2000). However, the classification refinement techniques used in these studies, such as merging or splitting clusters, are highly expert knowledge-driven and often rely on external data (e.g. cropland masks, reference ground data), which limits their reproducibility.

In order to deal with the high intra-class spectral variability without depending on external data, cropping systemś mapping over large areas would largely benefit from spectral variability reduction steps prior to classification. Geographic object-based image analysis (GEOBIA) techniques are widely used to reduce spectral variability by grouping individual pixels into thematically meaningful objects that summarize their spectral properties (Blaschke et al., 2014; Bunker et al., 2016; Hay and Castilla 2008). Another spectral variability reduction strategy, referred to as stratification, consists in partitioning the study regions into "biophysical-meaningful" sub-regions prior to classification (Cai et al., 2011; Cano et al., 2017; Franklin and Wulder, 2002; Vintrou et al., 2012; Wardlow and Egbert, 2008; Xiong et al., 2017). Traditionally, external maps are used (e.g. ecoregion, agro-ecological zoning maps), but some novel approaches have started to explore the potential of EO data-based landscape stratifications for large area land-cover classifications (e.g. Cano et al., 2017; Vintrou et al., 2012).

Based on this context, we hypothesize that the use of stratification prior to unsupervised classification improves the accuracy of the regional-scale cropping systems' classifications at the field-level by reducing the spectral variability of the EO data. To test our hypothesis, we propose a new semi-automatic GEOBIA clustering approach, and compare the classification accuracies obtained with two variants of the approach: an hyperclustering approach, which includes a field-level segmentation step prior to classification, and a landscape-clustering approach, which combines two spectral variability reduction steps prior to classification, a field-level segmentation and a landscape stratification.

This work contributes to the development of simple and generic tools to improve large-area cropping systems' mapping with new classification methods fully based on EO data. The approach benefits from the high spatial resolution of a Landsat mosaic and the high temporal resolution of a MODIS VI time series, and is tested on the Tocantins state for the 2014–2015 growing season. The classification results are evaluated with ground-truth data and the annual crop estimates of the Brazilian Institute of Geography and Statistics (IBGE) (IBGE, 2014).

2. Materials

2.1. Study region

The Tocantins state, located in the centre-north of Brazil (Fig. 1a), is an agricultural region of strong anthropic influence which has undergone fast land-use changes over the last decades. Since its creation in 1988, Tocantins experiences a growing population and a fast-evolving economy based on a developing agricultural sector (Pedroso da Silva and de Almeida, 2007; Fornaro, 2012). Apart from the extensive development of the livestock systems which has been induced by the spreading of the transport network since the 1970s, this region undergoes, since the last decade, an expansion of the cropland mainly to the detriment of areas of degraded pasture, but also of some native vegetation areas (Barona et al., 2010; Silva, 2007; Fornaro, 2012). This trend is greatly driven by the global demand for Brazilian soybean, which is responsible for the expansion of the soybean cultivated areas in the latest agricultural frontier of Brazil, the MATOPIBA region, named after the states of Maranhão, Tocantins, Piauí and Bahia (Bolfe et al., 2016; Fornaro, 2012).

As a result, three soybean-based cropping systems, characterized by highly mechanized management practices and large field sizes, of around 100 ha, dominate the agricultural landscape of Tocantins. One of the cropping systems is a single cropping system, characterized by a single soybean crop cycle, and the other two cropping systems are characterized by a double-cropping pattern with two distinct crop cycles in the same growing season. This form of intensification is commonly adopted in this region since it allows increasing yields while reducing the need for expansion of the cropland areas (Spangler et al., 2017).

In one of the double cropping systems, henceforth referred to as the 'soybean-cereal double cropping system', a soybean crop is grown in the summer season, followed, in succession, by a cereal crop (mostly maize) during the end of summer to fall season. The other double-cropping system henceforth referred to as the 'rice-soybean double cropping system', is localized in the lowlands of the Formoso River basin and concerns a rainy season rice crop (grown under flood irrigation) succeeded by a late soybean crop.

2.2. Ground-truth data collection

Field surveys were conducted in the study area during October 2015, at the end of the 2014–2015 growing season, just before the rainy summer period and beginning of the next season. The sampling strategy aimed at collecting multiple observations for the three major cropping systems' classes and other LULC types over a wide variety of zones.

In particular, different agricultural zones, scattered throughout the study area were targeted and surveyed thoroughly, in order to capture the regional variability of the cropping systems. The selection of zones to survey was assisted by the NTSA (Thematic Unit on Agricultural Systems) of EMBRAPA (Brazilian Agricultural Research Corporation) Fisheries and Aquaculture, and was mainly focused on identifying zones where the crop fields are spatially "concentrated".

The survey equipment included a Trimble Yuma2 rugged tablet with an integrated GPS and Quantum GIS software, which allowed real-time navigation and direct recording of the waypoints in a spatial database. The tablet was equipped with the multispectral Landsat 8 OLI 2014–2015 mosaic (presented in Section 2.4.2), overlaid by the field-level segmentation (presented in Section 3.1.1). The sampling consisted in associating the observed class to its corresponding object in the segmentation layer. The commonly adopted no-tillage system allowed us to identify the cropping systems by the crop residues that cover the fields after the harvest of the crops.

A total of 900 GPS waypoints belonging to 35 different municipalities were registered (presented in Fig. 2), representing the three major cropping systems belonging to the annual cropland and other LULC types.

2.3. Area-based agricultural statistics

Area-based estimates of the soybean, second crop maize and rice crops were acquired from the Municipal Agricultural Production database (PAM) of the IBGE (2014). The 2015 harvested area estimates of the 139 municipalities of Tocantins were used to evaluate our classification results. The 'second crop maize' area estimates were used to evaluate the classification of the soybean-cereal double cropping

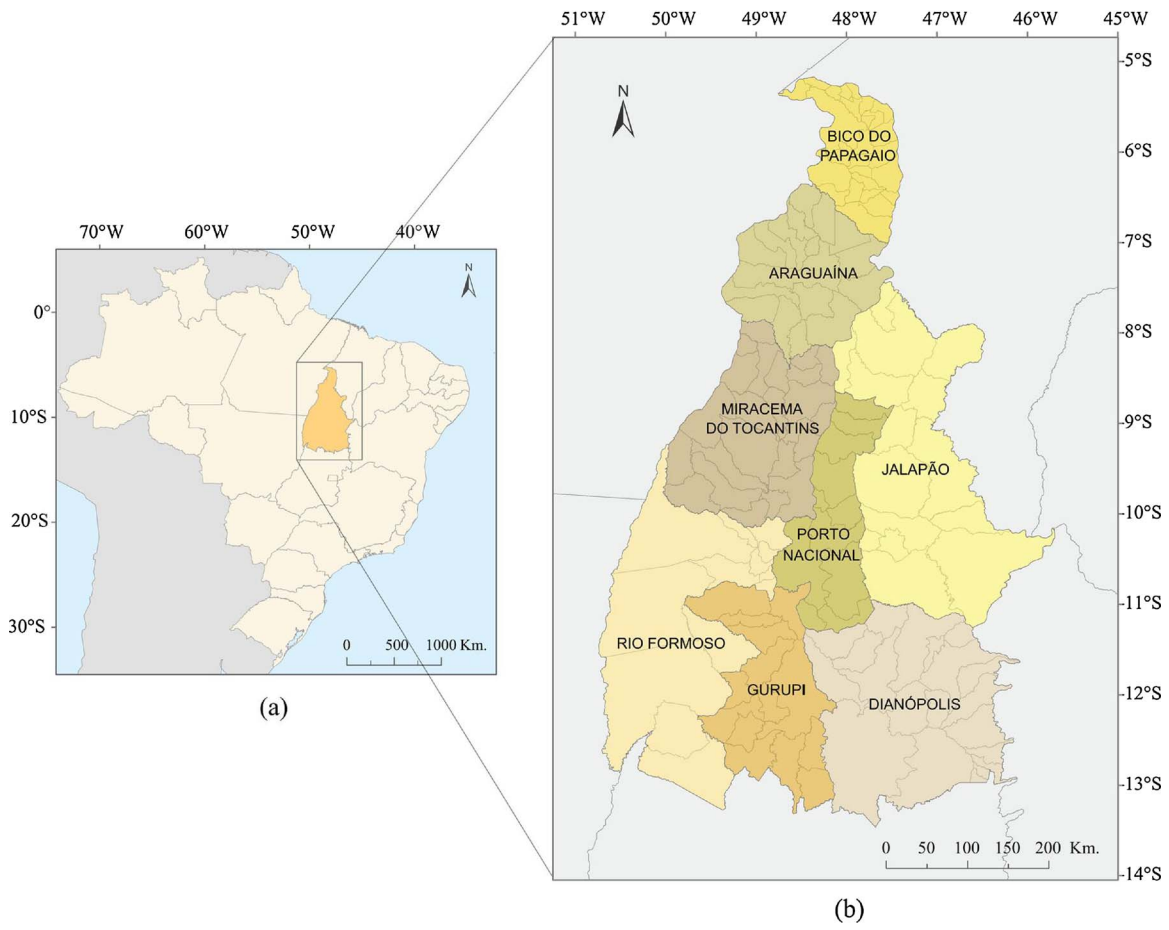


Fig. 1. (a) Location of Tocantins state, Brazil, (b) Administrative Micro-regions that compose the Tocantins state and limits of municipalities (grey lines).

system since no statistics are available on the other cereals which are alternatively grown as cover or fodder crops after the soybean crop harvest (e.g. sorghum or millet).

2.4. Earth observation data acquisition and pre-processing

2.4.1. MODIS NDVI time series

An annual time series of the MODIS VIs product (MOD13Q1 collection 5) (Solano et al., 2010) was acquired from the NASA Land processes distributed active archive centre through the processing R package MODIStsp (Busetto and Ranghetti, 2016). Two tiles (h13v09, h13v10) of the 16-day composited Normalized Difference Vegetation Index (NDVI) band were mosaicked pairwise for each of the composite periods to generate a total of 23 NDVI mosaic images, from the beginning of the growing season in October 2014 to October 2015, covering the study region at a 250 m spatial resolution.

Although the compositing process of the MOD13Q1 product already reduces noise effects by selecting the observation with the highest NDVI value and the smallest view angle among the 16 days' acquisitions (Solano et al., 2010), the dataset was further pre-processed to reduce remaining effects. The dataset was pre-processed in two steps (presented in detail in the '3.1. Pre-processing of the NDVI time series' section in Bellón et al. (2017)): First, a noise-reduction algorithm based on the Savitzky-Golay function-fitting smoothing method introduced by Chen et al. (2004) was applied to correct pixel values mainly affected by clouds and other atmospheric effects and improve the overall quality of the NDVI time series. Finally, the smoothed time series was linearly interpolated using the acquisition dates for each pixel retained during the compositing process, which are provided in the 'Composite Day of Year (CDOY)' band of the MOD13Q1 product, to reconstruct the NDVI time series at regular 16-day intervals, and therefore avoid temporal discontinuities in the time series.

2.4.2. Landsat 8 OLI mosaic

Nineteen Landsat 8 Operational Land Imager (OLI) scenes were mosaicked to generate a seamless image covering the whole state of Tocantins at 30 m spatial resolution. The surface reflectance Landsat 8 OLI scenes were retrieved from the United States Geological Survey (USGS) Landsat archive via the ESPA Ordering interface (Jenkerson, 2013; Zanter, 2016). In order to obtain a cloud-free mosaic, we selected scenes which were cloud-free inside the Tocantins state's boundaries, acquired during July 2014 and July 2015. Table 1 presents the paths and rows numbers of the Landsat scenes and their corresponding acquisition dates.

3. Methods

The presented approach was conceived as a GEOBIA classification approach. The high resolution Landsat mosaic was segmented into thematically meaningful field level objects which were subsequently classified through clustering. Two clustering approaches were tested and compared: the hyperclustering approach and the landscape-clustering approach. Both approaches employed the k-means classifier and conducted clustering in two steps: a pre-clustering step partitioned the objects into a large number of initial sub-clusters, and a final-clustering step partitioned the resulting sub-clusters into the desired number of final clusters. The landscape-clustering approach, unlike the hyperclustering approach, uses landscape units, which are obtained though stratification of the study region with a landscape-level segmentation, for the pre-clustering step. The methodological framework

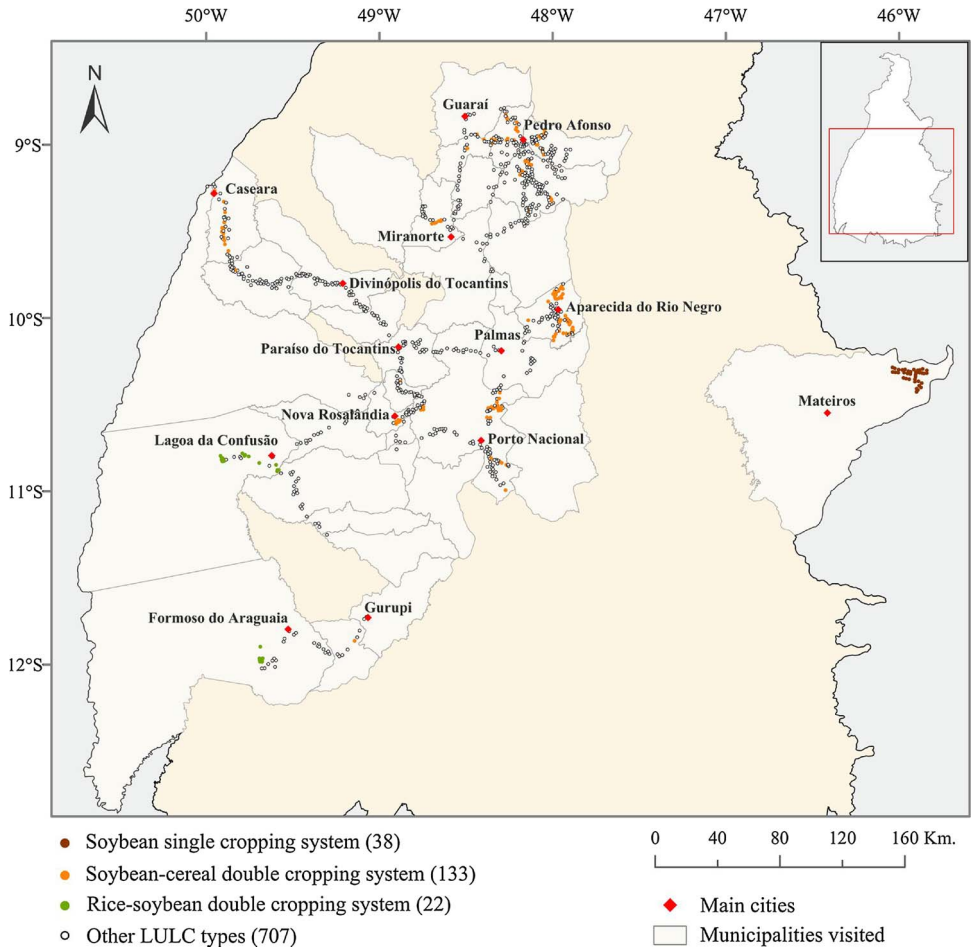


Fig. 2. Location and number of waypoints registered for each of the four classes in the spatial database during the ground-truth data collection field surveys in October 2015.

used in this study is presented in Fig. 3 and described in the following sub-sections.

3.1. Segmentation at the field and landscape levels

3.1.1. Field-level segmentation

A field-level segmentation was performed on the Landsat mosaic to delineate the boundaries of the crop fields and other landscape features. The segmentation was carried out using the 'Multiresolution Segmentation' algorithm proposed by Baatz and Schäpe in the eCognition® Developer 9.0 software (Trimble Germany GmbH, 2014; Baatz and Schäpe, 2000).

The segmentation algorithm was applied to five spectral bands of the Landsat mosaic, which were equally weighted, and the segmentation parameters were empirically set after sequential testing of different parameter combinations and visual evaluation of the different segmentation results (Table 2 presents the final segmentation parameters). The objective of the segmentation being the delineation of the overall rectangular-shaped crop fields, we heavily weighted the 'shape parameter' compared to the 'colour' (radiometry), and the 'compactness' compared to the 'smoothness', resulting in compact, regular-shaped objects, which allowed an effective delineation of the agricultural fields over different parts of the study area (Fig. 4 presents subset images of the segmentation results). A total of 187,877 objects at a spatial resolution of 30 m were obtained over the entire study area.

3.1.2. Landscape-level segmentation

The study region was partitioned into landscape-level units following the stratification method proposed by Bellón et al. (2017) based on the phenological information of a MODIS VI time series. This method consists in an automatic segmentation of the study area into units of land which present certain homogeneity in their physiognomy and land use that differs from adjacent units, while being

Table 1
Acquisition details of the Landsat 8 OLI scenes retrieved from the USGS archive to compose the cloud-free mosaic of Tocantins.

Acquisition date	Path	Row	Sun elevation (degrees)	Sun azimuth (degrees)	Scene cloud cover (%)[a]
17/07/2015	223	64	50.52	46.03	0.06
		65	49.35	45.00	0.05
		66	48.18	44.04	0.00
		67	46.98	43.15	0.00
		68	45.78	42.33	0.00
		69	44.56	41.56	0.00
07/07/2014	222	64	49.89	44.06	0.04
		65	48.70	43.10	0.00
		66	47.50	42.21	0.00
		67	46.28	41.38	0.00
		68	45.05	40.61	0.00
		69	43.81	39.90	0.21
19/07/2015	221	66	48.39	44.46	0.00
		67	47.20	43.56	0.00
		68	46.00	42.72	0.00
		69	44.79	41.94	0.44
25/07/2014	220	67	48.08	44.87	0.68
		68	46.90	43.97	0.00
		69	45.70	43.14	0.00

[a] The cloud covered areas in the scenes correspond to areas outside of the Tocantins state's boundaries.

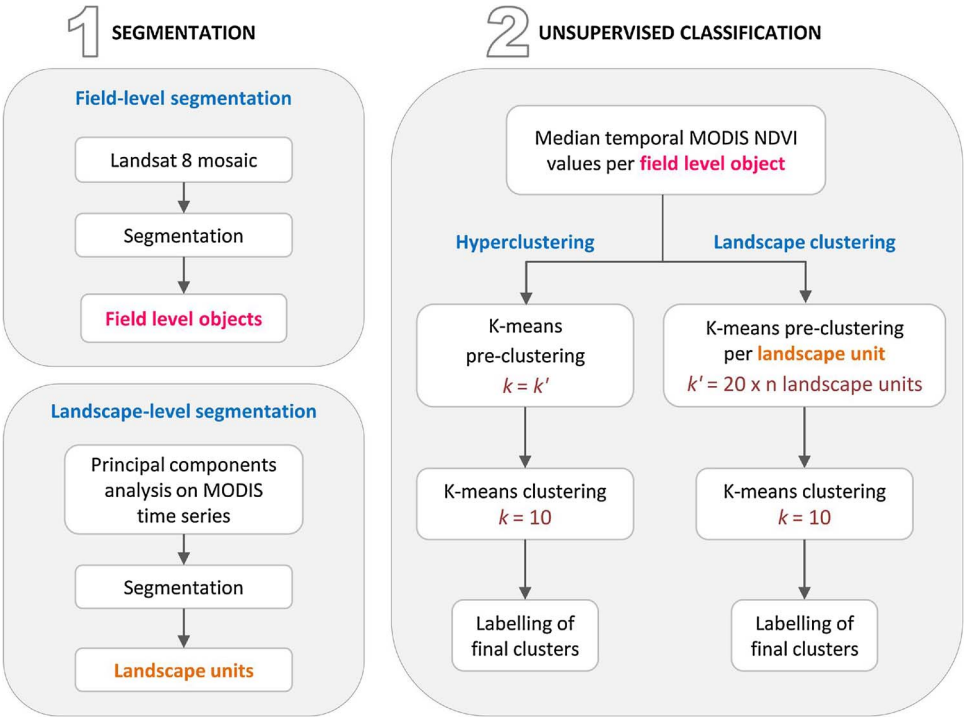


Fig. 3. Methodological framework.

heterogeneous in their structure, since they are extensive enough to contain a mosaic of patches with different land cover types, and thus are henceforth referred to as landscape units (see Bellón et al. (2017) for further details).

This landscape stratification method was originally conceived to produce an agricultural land-use systems' map at the landscape level by classifying the large-scale land units. In this study, the landscape stratification is otherwise used as a spectral variability reduction strategy, to assist regional-scale cropping systems' mapping at the field-level.

The landscape-level segmentation is performed using the 'Multiresolution Segmentation' algorithm previously presented (see Section 3.1.1), and the segmentation parameters were set according to the original method presented in Bellón et al. (2017) (shown in Table 2). A total of 120 landscape units were extracted following this method (Fig. 5).

3.2. Unsupervised classification

The two variants of the clustering approach were based on the temporal NDVI information contained in the MODIS time series. A common dataset was generated by extracting the median NDVI value per Landsat field-level object for each of the 23 dates in the time series. Due to the difference in spatial resolution between the MODIS pixels at 250 m and the Landsat objects at 30 m, only the values belonging to the pixels whose centroids were located inside of an object were used in the calculation of the median NDVI value for that given object. The median value was chosen as the measure of central tendency since it is less sensitive to skewness in the distribution of values than the mean, and therefore is less affected by the influence of the mixed pixels (whose radiometry results from different cover types) situated on the object's boundary.

Both variants of the clustering approach were based on the k-means algorithm to classify the dataset of field-level objects containing the median NDVI value for each of the 23 dates in the MODIS time series. Originally formulated by MacQueen (1967), the k-means algorithm aims at partitioning the data into '*k* groups' such that the within-cluster sum of squares is minimized. This widely used partitional clustering algorithm has a number of variations. For this study, the clustering was performed by means of the algorithm proposed by Hartigan and Wong (1979), implemented in the 'kmeans' tool in the processing R package 'stats' (R Core Team, 2016).

3.2.1. Hyperclustering approach

In the hyperclustering approach, a pre-clustering was performed on the dataset of field-level objects. The number of sub-clusters *k* was set to 2400 to match the number of sub-clusters *k'* in the landscape-clustering approach presented in the next section (Section 3.2.2).

The mean NDVI temporal profile per sub-cluster was then calculated, resulting in a total of 2400 mean NDVI temporal profiles which were finally partitioned into ten final clusters. The number of final

Table 2
Field-level and landscape-level segmentation parameters (see Bellón et al. (2017) for further details on the landscape-level segmentation).

Segmentation level	Input data	Spatial resolution	Bands (equally weighted)	Scale	Colour	Shape
Field level	Landsat 8 OLI mosaic	30 m	Blue (B2) Green (B3) Red (B4) NIR (B5) SWIR 1 (B6)	110	0.2	0.8 (compactness = 1; smoothness = 0)
Landscape level	Principal component (PC) images computed from the MODIS NDVI time series	250 m	22 PC images	850	1	0

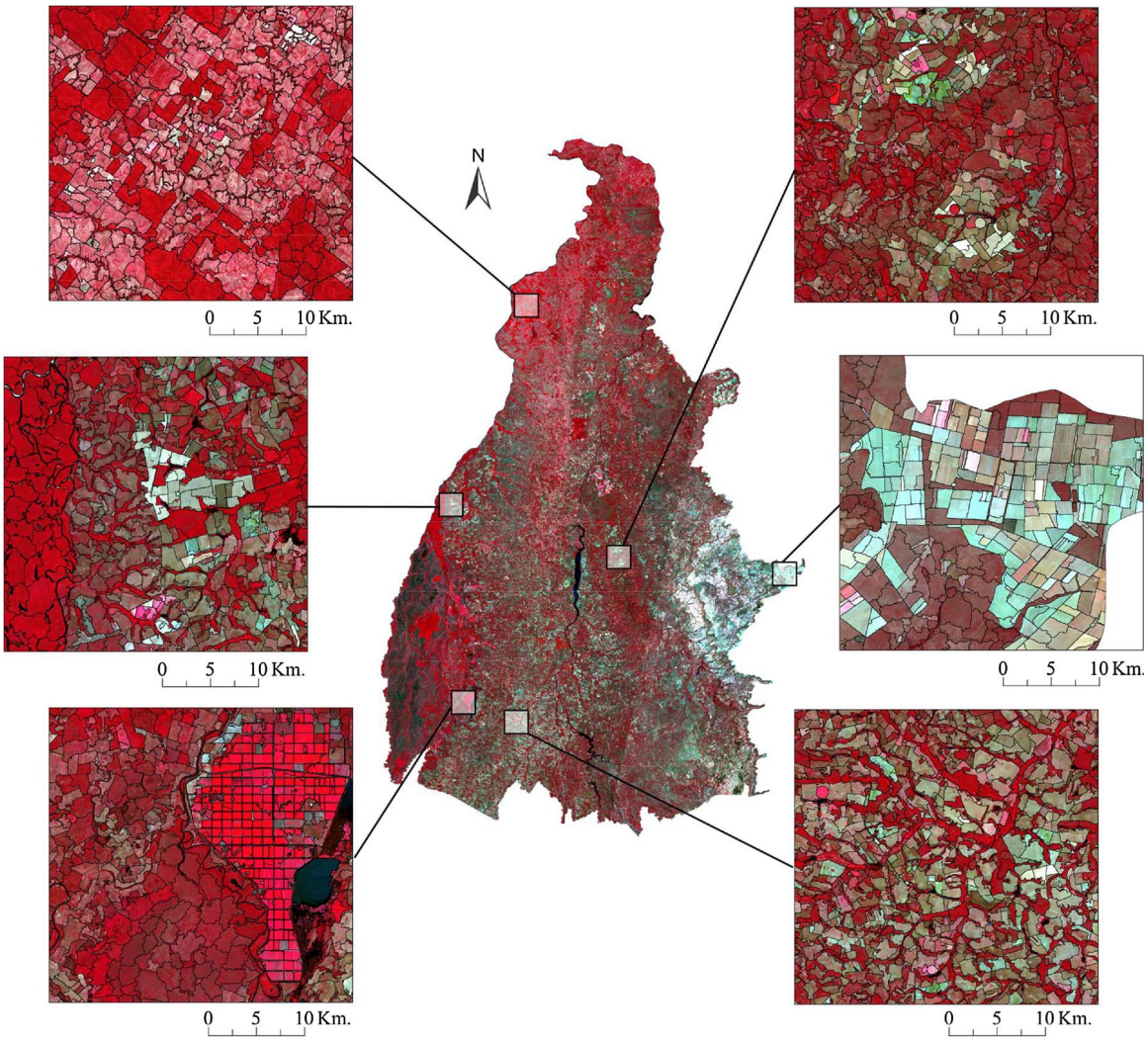


Fig. 4. Image subsets of the Landsat field-level segmentation results (field-level objects) over different landscapes. The background image corresponds to the false-colour Landsat mosaic.

clusters was empirically set to 10, expecting 10 clusters to adequately discriminate the four targeted classes while still being easily interpretable.

Since random seeding was used for initial clustering in the k-means algorithm, ten replicates were conducted and the clustering result with the highest degree of difference between clusters and similarity within clusters was finally retained. This was justified by the highest ratios of the between-clusters sum of squares to the total sum of squares (R^2) and of the between-clusters sum of squares to the total within-clusters of squares ($R^2/(1 - R^2)$).

3.2.2. Landscape-clustering approach

In the landscape-clustering approach, the landscape units were first spatially overlaid by the field-level objects, and all of the objects whose centroid was located inside of the boundaries of a given landscape unit were grouped, resulting in one group of field-level objects per landscape unit. A pre-clustering was then performed on each group of objects, making as much partitions as number of landscape units (n = 120). The exact number of spectral classes per landscape unit being unknown, the number of sub-clusters per group of objects was empirically set to 20, which represents 5 times more clusters than final targeted classes, summing up to a total of 2400 k' sub-clusters ($k' = 20$ sub-clusters × 120 landscape units) for the whole study region.

The mean NDVI temporal profile was then calculated for each of the sub-clusters, resulting in a total of 2400 mean NDVI temporal profiles. A final clustering was performed on the k' mean NDVI temporal profiles to obtain 10 final clusters.

Ten replicates of the final clustering on the mean NDVI temporal profiles were evaluated and the clustering result with the highest R^2 and $R^2/(1 - R^2)$ ratios was retained.

3.2.3. Labelling of final clusters

The mean NDVI temporal profiles of the ten final clusters of each approach were calculated and plotted. In general, the NDVI temporal profiles of cropping systems present fast NDVI increase and decrease trends related to the different stages of crop growth and harvest periods and thus of the crop cycles. This type of pattern allowed us to identify and label the final clusters through visual analysis of their NDVI variations over the growing season.

The clusters whose NDVI profile's shapes followed the cropping calendar of the main cropping systems of Tocantins were labelled with their corresponding cropping system class, while the rest were labelled as the 'other LULC types' class.

3.3. Accuracy assessment of the classification results

The accuracy of both the landscape clustering and the hyperclustering results was estimated and compared by using the ground-truth data and the agricultural statistics as a reference.

Six accuracy measures were calculated using the 900 reference

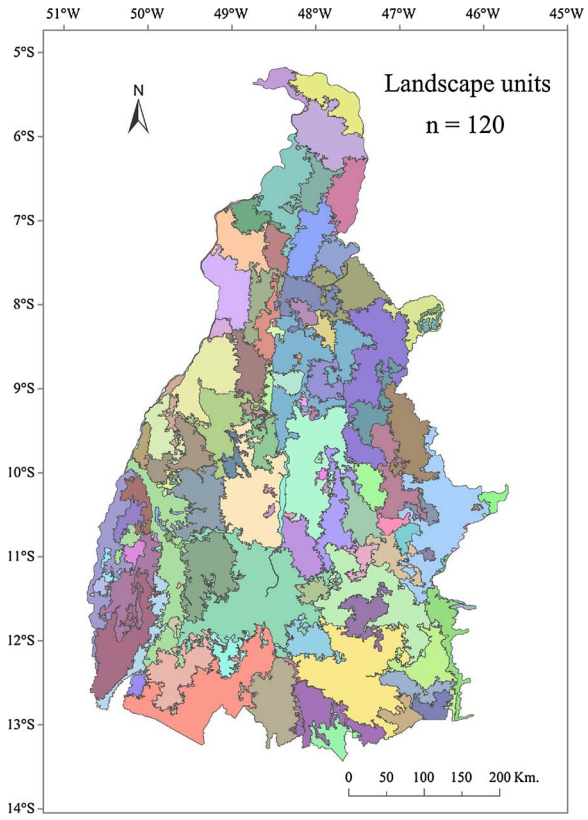


Fig. 5. The landscape-level segmentation result.

objects in the ground-truth database: the overall accuracy with its associated confidence interval (Rosenfield and Fitzpatrick-Lins, 1986), the Cohen's Kappa coefficient (Cohen, 1960), the quantity disagreement and allocation disagreement (Pontius and Millones, 2011), the McNemar's Chi-squared test (McNemar, 1947) and the *F*-measure (Labatut and Cherifi, 2012).

The overall accuracy corresponds to the total number of the correctly classified objects divided by the total number of objects being evaluated and is expressed as a percentage. The Kappa coefficient measures the difference between the observed agreement and the probability of agreement occurring by chance and ranges between 0 (no more agreement than expected by chance) and 1 (perfect agreement). The Kappa coefficient is often used as an overall measure of accuracy, but its use is being questioned by numerous articles (Strahler et al., 2006; Foody, 2002; Pontius and Millones, 2011). The quantity disagreement and allocation disagreement have therefore been included as complementary measures for the accuracy assessment. These two measures decompose the overall disagreement (which is the complement of the overall accuracy) into two types of disagreement and thus provide additional insight on the accuracy of the classification results.

In our case quantity disagreement expresses the amount of difference between the reference objects and the classified objects that is due to the less than perfect match in the proportions of the classes, and allocation disagreement expresses the amount of difference between the reference objects and the classified objects that is due to the less than optimal match in the spatial allocation of the classes given the proportions of the classes in the reference and classified objects. These two measures are expressed as a percentage.

The McNemar's Chi-squared test was used to evaluate if there is a statistically significant difference between the two classification results' accuracies. Finally, the class-specific accuracy was evaluated using the *F*-measure, which corresponds to the weighted harmonic mean of the producer's and user's accuracy and ranges from 0 to 1, being the lowest and the highest accuracy, respectively.

The crop area-based estimates in the agricultural statistics at the municipality level provided by the IBGE were finally used to evaluate the accuracy of the best classification result according to the ground-truth data accuracy assessment tests. A correlation coefficient was

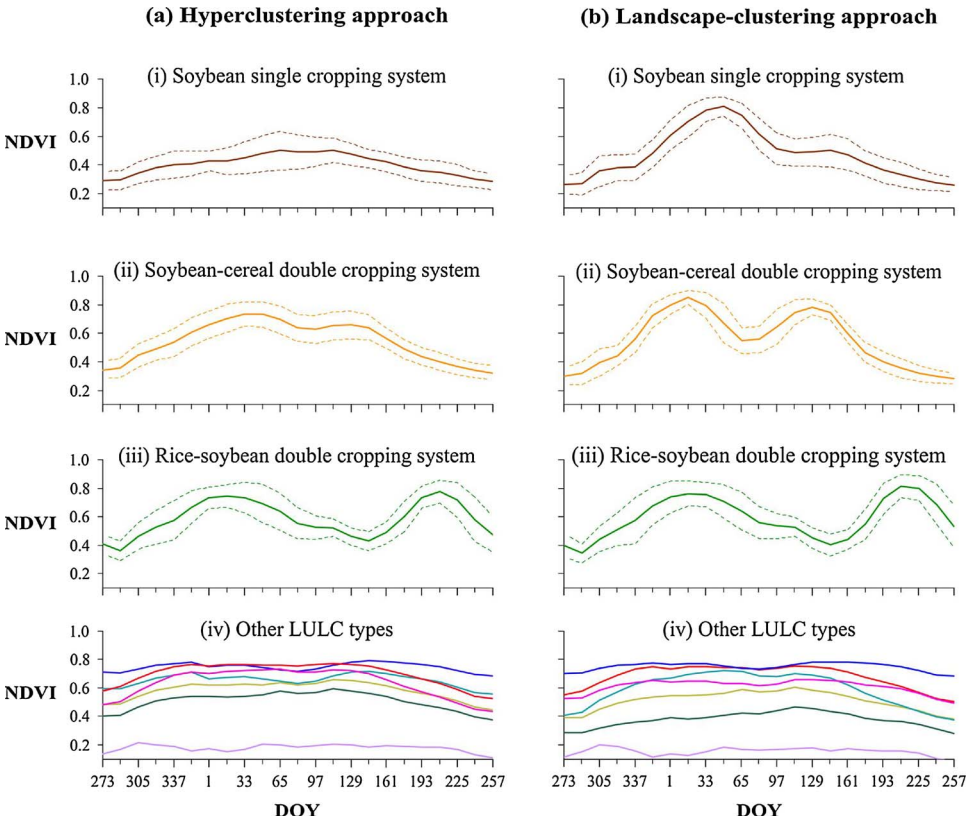


Fig. 6. The temporal NDVI profiles' plots of the 10 final clusters obtained with (a) the hyperclustering approach, and (b) the landscape-clustering approach, and their corresponding labels. The continuous line curves represent the mean NDVI of all of the objects belonging to each cluster against the day of the year (DOY). The mean NDVI profiles of the clusters which seem to have captured the three main cropping systems of Tocantins are represented in separate plots (i), (ii) and (iii) with their standard deviation (discontinuous line curves), and the mean NDVI profiles of the clusters which seem to have captured other LULC types are represented in the same plot (iv).

computed between the estimated crop area and the reference IBGE estimates at the municipality level.

4. Results

4.1. Clustering results

Since 120 landscape units were obtained, the number of pre-clusters for the whole study area was 2400 for both the landscape-clustering approach (20 pre-clusters per landscape unit) and the hyperclustering approach. The plots of the mean NDVI temporal profiles of the final ten clusters are presented in Fig. 6 for the hyperclustering (Fig. 6(a)), and the landscape-clustering approach (Fig. 6(b)).

Three distinct cropping patterns related to the main cropping systems of Tocantins can be identified among the NDVI temporal profiles of the final clustering results. These are particularly noticeable in three clusters' profiles obtained with the landscape-clustering approach (Fig. 6(b)(i–iii)), and in one of the clusters' profiles obtained with the hyperclustering approach (Fig. 6(a)(iii)), since they present fast NDVI increase and decrease trends. The rest of the cluster's profiles present remarkably smoother shapes.

Nevertheless, two clusters' profiles obtained with the hyperclustering approach which present slight NDVI oscillations seem to have captured the temporal variations related to the cropping systems and thus can be labelled accordingly, even if the general flatness of their profiles suggests mixture of classes in both these clusters (Fig. 6(a)(i)(ii)).

The first cropping pattern which is clearly recognizable in the temporal NDVI profile plot in Fig. 6(b)(i) exhibits a major single crop cycle that peaks around February, suggesting that the corresponding cluster has grouped the objects belonging to the soybean single cropping system. The cropping pattern in the temporal NDVI profile plot in Fig. 6(b)(ii) seems distinctive of the soybean-cereal double cropping system since the NDVI values are low around Mars, reflecting the soybean harvest, after which the NDVI increases as a result of the second cereal's growth until it peaks around May. Finally, the cropping pattern in the temporal NDVI profile plot in Fig. 6(b)(iii) seems distinctive of the rice-soybean double cropping system, with a long crop cycle during the rainy season corresponding to the rice crop cycle, followed by a shorter, late crop cycle from May to September. This late second soybean crop cycle is evidenced by the NDVI peak period around July.

4.2. Accuracy assessment results

Table 3 presents the overall accuracy, the Cohen's Kappa coefficient and the quantity and allocation disagreement of the classification results in four classes (the three main cropping systems and the 'other LULC types' class), derived from the hyperclustering and the landscape-clustering approaches.

The overall accuracy is higher for the landscape-clustering classification than for the hyperclustering classification i.e. the landscape-clustering approach correctly classified a larger number of objects than the hyperclustering approach, which suggests that the landscape-clustering approach performed better than the hyperclustering approach in the classification of the four classes. Accordingly, the overall disagreement (composed of the quantity disagreement and the allocation disagreement) was lower for the landscape-clustering classification (6.8%) than for the hyperclustering classification (22.1%). Quantity disagreement was higher than allocation disagreement for both classification results, suggesting that the classification errors were related in a higher degree to a quantification error than to a location error.

The Kappa coefficient result for the landscape-clustering approach is also higher than that of the hyperclustering approach i.e. the difference between the observed agreement between the classified objects and the reference objects and the probability of agreement occurring by chance was higher for the landscape-clustering classification than for the hyperclustering classification, suggesting that the landscape-clustering classification's reliability is higher.

Furthermore, the results of the McNemar's test (presented in Table 4), showed a statistically significant improvement of the landscape-clustering classification results over the hyperclustering results (p-value $< .0001$).

Class-specific accuracy is presented in Table 5 through the F-measure results of the landscape-clustering and the hyperclustering approaches. The 'soybean single cropping system' class obtained the lowest F-measure value for both classifications (F-measure < 0.7), and contrastingly, the 'rice-soybean double cropping system' class was correctly classified by both approaches, obtaining the highest F-measure (F-measure $= 1$). Both the 'soybean-cereal double cropping system' and the 'other LULC types' classes obtained high F-measure values (F-measure > 0.8) for the landscape-clustering approach but only the 'other LULC types' class did for the hyperclustering approach, the 'soybean-cereal double cropping system' class presenting the greatest difference between both approaches. Except for the 'rice-soybean double cropping system' class which obtained the same F-measure for both approaches, the rest of the classes obtained lower scores with the hyperclustering approach, which indicates a better performance of the landscape-clustering approach.

The soybean crop area-based estimates in the municipality-level agricultural statistics provided by the IBGE are compared to the landscape-clustering approach's soybean area results in Fig. 7. The total area of soybean (considering all three cropping systems) estimated by the landscape-clustering approach accounts for 73% of the total area reported by the IBGE, obtaining a strong positive correlation (r) of 0.87 ($r^2 = 0.75$) at the municipality level.

Likewise, the second crop maize IBGE area estimates show a strong correlation with the landscape-clustering 'soybean-cereal double cropping system' classification results, with a correlation coefficient at the municipality level of 0.87 ($r^2 = 0.76$). Finally, the rice IBGE crop area estimates also show a strong correlation with the landscape clustering 'rice-soybean double cropping system', with a correlation coefficient of 0.99 ($r^2 = 0.97$).

4.3. Cropping systems' map obtained with the landscape-clustering approach

Fig. 8 presents the map derived from the classification results of the main cropping systems of Tocantins at the field level obtained with the landscape-clustering approach. The accuracy of the map corresponds to the accuracy of the landscape-clustering classification results presented in Section 4.2.

As shown in the map (Fig. 8), the rice-soybean double cropping system is exclusively localized in the southwest of Tocantins, while the

Table 3
The overall classification accuracy and its associated confidence interval (CI), the overall Kappa coefficient and the quantity and allocation disagreement results for 900 classified objects from each clustering approach.

Clustering approach	Overall accuracy (%)	95% CI (%)	Overall Kappa	Quantity disagreement (%)	Allocation disagreement (%)
Landscape-clustering	93.2	91.4–94.8	0.81	4.5	2.3
Hyperclustering	77.9	75.0–80.6	0.52	15.3	6.8

Table 4
Contingency table summarizing the outcomes of the two clustering approaches (number of objects correctly classified or misclassified) and McNemar's Chi-squared test (χ^2) results.

	Landscape-clustering approach (correctly classified)	Landscape clustering approach (misclassified)	
Hyperclustering approach (correctly classified)	684	17	$\chi^2 = 110.72$, with degree of freedom = 1 (*p*-value < .0001)
Hyperclustering approach (misclassified)	155	44	

Table 5
Class *F*-measure results.

Class	Landscape-clustering approach	Hyperclustering approach
Soybean single cropping system	0.65	0.46
Soybean-cereal double cropping system	0.81	0.59
Rice-soybean double cropping system	1.00	1.00
Other LULC types	0.96	0.85

soybean single cropping system and the soybean-cereal double cropping system are locally concentrated in the south-east border of Tocantins and in the north-east border of the state, respectively, and present a scattered distribution over the rest of state.

5. Discussion

Globally, the results demonstrated that the new object-based clustering approach, based exclusively on EO data, can be used to produce field-level maps of cropping systems, described by the crop type and cropping pattern, at a regional scale.

The accuracy assessment results showed that both the hyperclustering and the landscape-clustering variants of the approach produced good classifications, scoring overall accuracies of more than 75%, even if the landscape-clustering approach clearly outperformed the hyperclustering approach. This is particularly evident in the results for the two cropping systems which are spatially dispersed, the single soybean single cropping system and the soybean cereal double cropping system, the landscape-clustering approach obtaining an *F*-measure of around 0.2 higher than the hyperclustering approach for these two classes. The Kappa coefficient results further suggest a high reliability of the landscape-clustering classification, which is somewhat lower for the hyperclustering classification.

The significant improvement in the classification accuracy when using a stratification prior to clustering is in accordance with the results of Cano et al. (2017). This result suggests that partitioning the study region into sub-regions, in this case through landscape-level segmentation on the phenological information of a MODIS time series, is an effective spectral variability reduction strategy for improving LULC classification accuracy. This variability reduction strategy seems to particularly benefit the LULC types which are spatially dispersed and under-represented compared to more extensive LULC types since it maximizes the possibility of them being correctly isolated with initial clustering per landscape unit. On the contrary, the hyperclustering approach without stratification, even with a high number of initial clusters (in our case 2400), did not avoid the mixture of classes, thus making its final clusters hard to label and resulting in lower classification accuracies.

The soybean single cropping system obtained the lower accuracies with both approaches (*F*-measures of 0.65 and 0.46 for the landscape-clustering and the hyperclustering approaches, respectively). This was mostly due to its confusion with pasture areas with strong vegetative development during the rainy season, and thus presenting a high NDVI peak over the same period of the soybean crop's maximum growth stage. Contrastingly, all of the ground-truth objects for the rice-soybean double cropping system were correctly classified with both approaches. This excellent result may be related to the singular phenological pattern of this cropping system, with two distinct and temporary distant crop

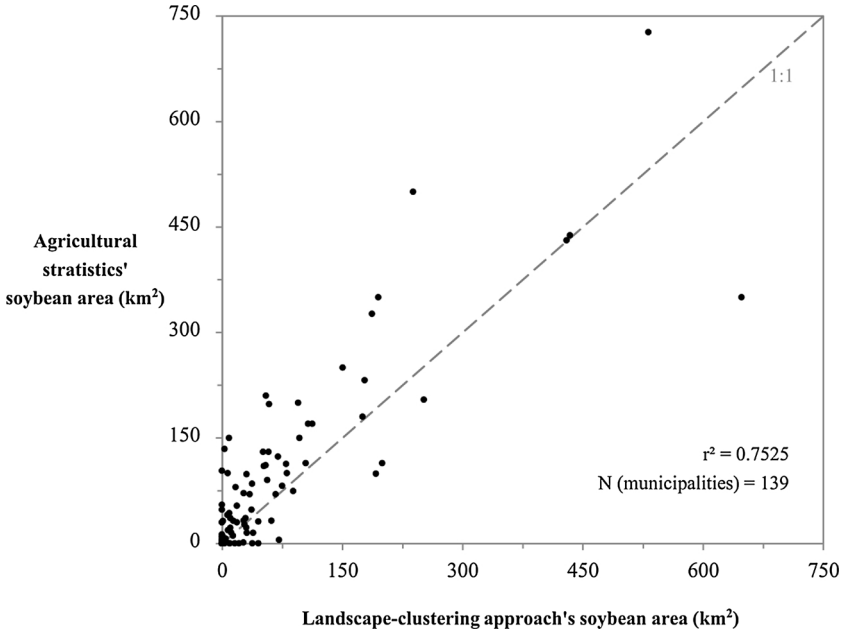


Fig. 7. Comparison of Tocantins' municipal-level soybean area estimates in square kilometres between the remote sensing landscape-clustering results and the IBGE agricultural statistics for 2015.

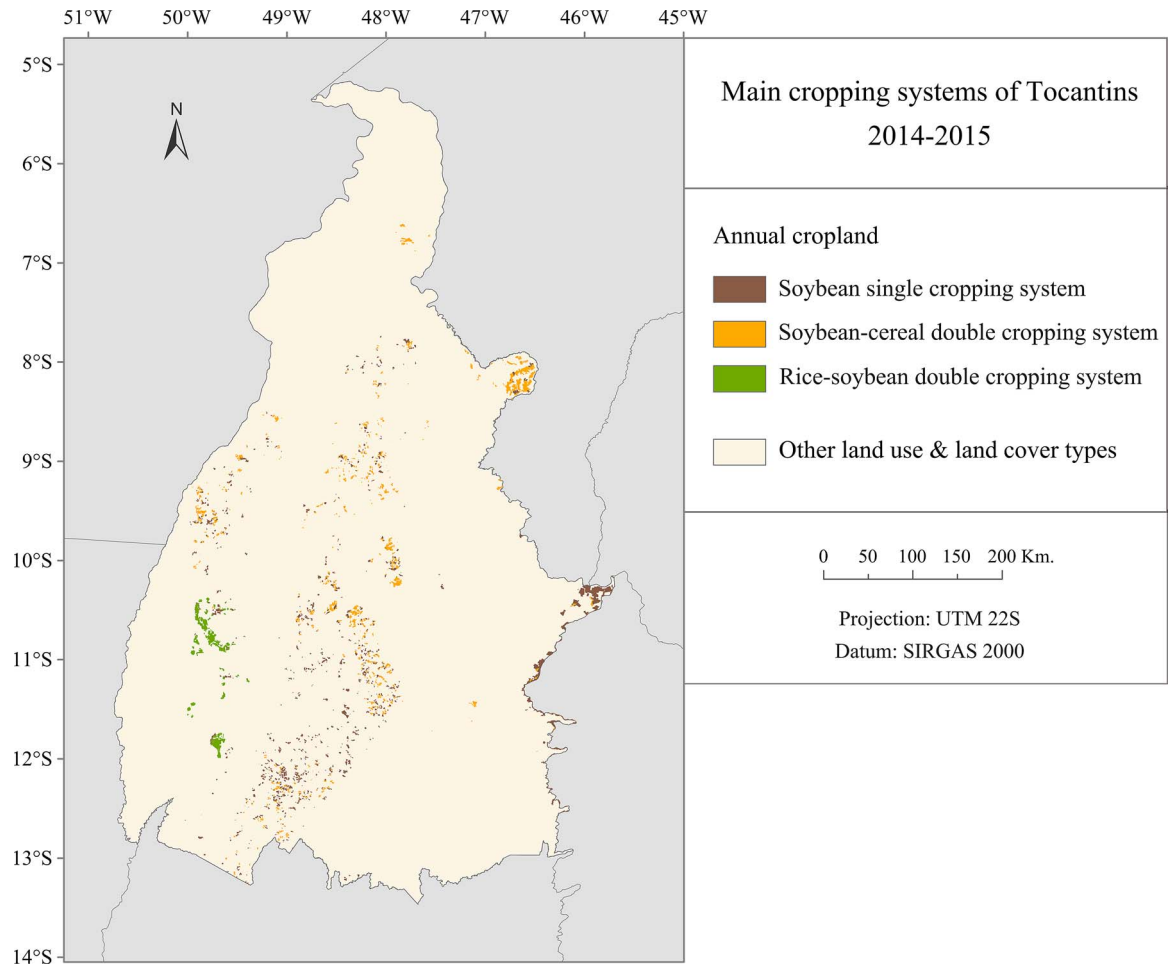


Fig. 8. Map of the main cropping systems of Tocantins at the field level in the 2014–2015 growing season obtained with the landscape-clustering approach.

cycles, which does not resemble the phenological pattern of any other LULC class. Moreover, the landscape-clustering approach did not show an improvement in the classification of this particular cropping system, since all of the fields are localized in the same micro-region (the Rio Formoso micro-region) and so limited spatial variability is expected. The good location of this cropping system might partly explain the lower location disagreement in comparison to the quantity disagreement between the classifications of both approaches and the reference dataset.

The combination of spectral variability reduction steps (field-level and landscape-level segmentations) and a two-step unsupervised classification, appears to be an effective way of reducing the datasets spectral variability with no need of complex expert knowledge-driven processing steps and external data. Covering a territory such as the Tocantins state (surface area of 277,621 km^2) with a coarse spatial resolution dataset such as MODIS at 250 m represents an initial dataset of around 10 million pixels (per MOD13Q1 composite). Following the presented methodology, this high spatial dimensionality was successively reduced to 187,877 field-level objects, then to 2400 pre-clusters, and finally to 10 clusters. Other than the user-defined segmentation parameters and the number of clusters, the processing is fully automatic, and the low number of final clusters allows labelling through visual analysis, yet requiring basic knowledge of the crop calendar of the cropping systems in order to relate the final clusters to the cropping systems of interest.

The comparison of the landscape-clustering results to the annual soybean, rice, and second crop maize area estimates of the IBGE showed strong positive correlations at the municipality level of around $r = 0.9$. The slight underestimation of the soybean area in the clustering results is due to the misclassification of cropland fields with areas of grassland or pasture. These underestimations together with some overestimations are responsible for the dispersion obtained, which in the case of bordering municipalities is particularly evident since crop fields which are actually located inside the mapped state's boundary are sometimes censused on the neighbouring state's statistics. Nonetheless, although the number of clusters was empirically set to simplify the processing, the final classification may benefit from internal evaluation of different combinations of cluster numbers in order to obtain optimal clustering results from a statistical point of view (maximizing the variance between clusters and minimizing the variance within clusters) which may result in a better separability of the misclassified classes.

The inclusion of other processing steps may help improve the classification results at the expense of more computationally-intensive and time-consuming processing. In particular, radiometry of the field-level objects might still be affected by the border effect due to the disparity of spatial resolutions between the MODIS and the Landsat datasets. The presented methodology might thus benefit from spectral unmixing and spatial resampling of the MODIS pixels to the Landsat spatial resolution (e.g. Bisquert et al., 2015; Gevaert and García-Haro, 2015; Knauer et al., 2016). Potential errors also might be induced by the acquisition dates of the Landsat scenes used for the field-level segmentation. For example, some fields might have been divided into two different cropping systems during the growing season but had a homogeneous aspect at the Landsat summer acquisition date. The methodology might thus benefit from multi-temporal segmentation but at the expense of increased complexity and data requirements.

The reproducibility of the presented landscape-clustering approach is currently being tested in the neighbouring Brazilian state of Mato

Grosso for the 2015–2016 growing season and could be reproduced in other states of the MATOPIBA region to provide soybean-based cropping systems' maps over the whole interstate region. This approach could also be regularly reproduced to provide timely insight on the evolution of the cropping systems in such dynamic agricultural areas, allowing the analysis of land-use trajectories like the ones presented by Spera et al. (2014) and Dubreuil et al. (2012).

Reproducibility of the presented approach in other Brazilian regions or other world regions with large crop field sizes may require adaptation of the scale parameter of the 'Multiresolution segmentation' algorithm (Baatz and Schäpe, 2000) to obtain equivalently scaled objects in adequacy to the agricultural landscapes of the targeted region: field-level objects adapted to the region's fields and landscape units extensive enough to contain multiple field-level objects. The choice of the weight given to the other segmentation parameters may be guided by the indications provided (ex. heavily weighting the 'shape parameter' and the 'compactness' to obtain the field-level objects). The use of coarse spatial resolution data however prevents the approach from being reproducible in smallholder agriculture regions with fields smaller than a MODIS pixel. In such regions, higher spatial resolution datasets are needed to access the spectrally-pure information of the cropping systems.

6. Conclusions

The presented object-based clustering approach, being exclusively based on freely available MODIS and Landsat data, can be potentially reproduced to map cropping systems of other Brazilian regions or other world regions with large field sizes with minimal adaptation. The simplicity of the methods, which neither need external data nor computationally intensive processing demarcates from the data-intensive or highly-expert driven remote sensing approaches which are commonly used when addressing cropping systems' mapping over large areas. This makes it particularly adapted for cropping systems' classifications when field data is hard to obtain and external data is lacking.

The novel landscape-clustering approach successfully mapped the main cropping systems of the Brazilian state of Tocantins at the field level by combining two spectral variability reduction steps, a field-level segmentation and a landscape stratification, and an unsupervised classification. The inclusion of a landscape stratification step prior to unsupervised classification significantly improved the classification accuracy, from 78% overall accuracy to 93%, which confirms our initial hypothesis. The classification results of the landscape-clustering approach were strongly correlated to the official crop area-based estimates, with correlation coefficients of around 0.9. Landscape stratification seems to be a significant contribution towards large area cropping systems' mapping, particularly in regions where cropping systems are spatially spread and under-represented compared to other LULC types.

The produced map can be used as fine-scale complementary information to the agricultural statistics estimates which are generally aggregated at administrative levels, and could eventually be used as an input to crop yield forecasting models. If reproduced yearly, this map could also contribute to the monitoring of cropping systems' dynamics such as cropping frequency change and their implications for the environment.

Conflicts of interest

None.

Acknowledgements

This work was supported by the CIRAD (French Agricultural Research Centre for Development); the FP7 SIGMA (Stimulating Innovation for Global Monitoring of Agriculture) project under grant agreement no.603719; the French Space Agency (CNES) through the funding of the TOSCA AGRIZONE project; and the CAPES/COFECUB program through the funding of the GeoABC Project (Methodologies and technological innovation for satellite monitoring of low carbon agriculture in support of Brazil's ABC Plan), project No. 845/15.

References

Arvor, Damien, Jonathan, Milton, Penello Meirelles, Margareth Simões, Dubreuil, Vincent, Durieux, Laurent, 2011. Classification of MODIS EVI time series for crop mapping in the state of mato grosso, Brazil. Int. J. Remote Sens. 32 (22), 7847–7871. http://dx.doi.org/10.1080/01431161.2010.531783.

Atzberger, Clement, 2013. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. Remote Sens. 5 (2), 949–981. http://dx.doi.org/10.3390/rs5020949.

Bégué, Agnès, Damien, Arvor, Camille, Lelong, Elodie, Vintrou, Margareth, Simoes, 2015. Agricultural systems studies using remote sensing. Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. CRC Press, pp. 113–130.

Baatz, Martin, Schäpe, Arno, 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. Angewandte Geographische Informationsverarbeitung XII 58. Wichmann-Verlag, Heidelberg, pp. 12–23.

Barona, Elizabeth, Ramankutty, Navin, Hyman, Glenn, Coomes, Oliver T., 2010. The role of pasture and soybean in deforestation of the Brazilian Amazon. Environ. Res. Lett. 5 (2). http://dx.doi.org/10.1088/1748-9326/5/2/024002.

Bellón, Beatriz, Agnès Bégué, Danny Lo Seen, Claudio Aparecido de Almeida, and Margareth Simões 2017. 'A Remote Sensing Approach for Regional-Scale Mapping of Agricultural Land-Use Systems Based on NDVI Time Series'. Remote Sensing 9 (6). Multidisciplinary Digital Publishing Institute: 600.

Bisquert, Mar, Bordogna, Gloria, Bégué, Agnès, Candiani, Gabriele, Teisseire, Maguelonne, Poncelet, Pascal, 2015. A simple fusion method for image time series based on the estimation of image temporal validity. Remote Sens. 7 (1), 704–724. http://dx.doi.org/10.3390/rs70100704.

Blaschke, Thomas, Geoffrey Hay, J. Hay, Maggi, Kelly, Lang, Stefan, Hofmann, Peter, Addink, Elisabeth, Feitosa, Queiroz, et al., 2014. Geographic object-based image analysis–towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191. http://dx.doi.org/10.1016/j.isprsjprs.2013.09.014. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS).

Bolfe, Édson L., De C. Victória, Daniel, Contini, Elisio, Bayma-Silva, Gustavo, Spinelli-Araujo, Luciana, Gomes, Daniel, 2016. Matopiba Em Crescimento Agrícola Aspectos Territoriais E. Revista de Política Agrícola 25 (4), 38–62.

Brown, J., Christopher, Jude H., Camargo Coutinho, Alexandre, de Castro Victoria, Daniel, Bishop, Christopher R., 2013. Classifying multiyear agricultural land use data from mato grosso using time-series MODIS vegetation index data. Remote Sens. Environ. 130, 39–50. http://dx.doi.org/10.1016/j.rse.2012.11.009.

Bunker, Brian E., Tullis, Jason A., Cothren, Jackson D., Casana, Jesse, Aly, Mohamed H., 2016. Object-based dimensionality reduction in land surface phenology classification. AIMS Geosci. 2 (4), 302–328 Aims Press.

Busetto, Lorenzo, Ranghetti, Luigi, 2016. MODIStsp: an R package for automatic preprocessing of MODIS land products time series. Comput. Geosci. 97, 40–48 Elsevier.

Cai, Hongyan, Zhang, Shuwen, Bu, Kun, Yang, Jiuchun, Chang, Liping, 2011. Integrating geographical data and phenological characteristics derived from MODIS data for improving land cover mapping. J. Geog. Sci. 21 (4), 705–718 Springer.

Cano, Emmanuelle, Denux, Jean-philippe, Bisquert, Mar, Hubert-, Laurence, Chéret, Véronique, Cano, Emmanuelle, Denux, Jean-philippe, Bisquert, Mar, Hubert, Laurence, 2017. Improved forest-cover mapping based on MODIS time series and landscape stratification. Int. J. Remote Sens. 38 (7), 1865–1888. http://dx.doi.org/10.1080/01431161.2017.1280635. Taylor & Francis.

Cheema, M.J.M., Bastiaanssen, W.G.M., 2010. Land use and land cover classification in the irrigated indus basin using growth phenology information from satellite data to support water management analysis. Agric. Water Manage. 97 (10), 1541–1552. http://dx.doi.org/10.1016/j.agwat.2010.05.009. Elsevier B.V.

Chen, Jin, Jönsson, Per, Tamura, Masayuki, Gu, Zhihui, Matsushita, Bunkei, Eklundh, Lars, 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sens. Environ. 91 (3), 332–344. http://dx.doi.org/10.1016/j.rse.2004.03.014.

Cihlar, J., 2000. Land cover mapping of large areas from satellites: status and research priorities. Int. J. Remote Sens. 21 (6), 1093–1114. http://dx.doi.org/10.1080/014311600210092.

Cohen, Jacob, 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37–46 Sage Publications Sage CA: Thousand Oaks, CA.

Dubreuil, Vincent, Damien, Arvor, Nathan, Debortoli, 2012. Monitoring the pioneer frontier and agricultural intensification in mato grosso using SPOT vegetation images. Revue Française de Photogrammétrie et de Télédétection 200, 2–10.

FAO. 2009. 'How to Feed the World in 2050'. Insights from an Expert Meeting at FAO 2050 (1): 1–35. 10.1111/j.1728-4457.2009.00312.x.

Foody, Giles M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80 (1), 185–201. http://dx.doi.org/10.1016/S0034-4257(01)00295-4.

Fornaro, Alexandre Caselli, 2012. Logística E Agronegócio Globalizado No Estado Do Tocantins: Um Estudo Sobre a Expansão Das Fronteiras Agrícolas Modernas No Território Brasileiro.

Franklin, S.E., Wulder, M.A., 2002. Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. Prog. Phys. Geogr. 26 (2), 173–205 Sage Publications Sage CA: Thousand Oaks, CA.

Gevaert, Caroline M., García-Haro, F. Javier, 2015. A comparison of STARFM and an unmixing-based algorithm for landsat and MODIS data fusion. Remote Sens. Environ.

156, 34–44. http://dx.doi.org/10.1016/j.rse.2014.09.012.
Guan, Xudong, Huang, Chong, Liu, Gaohuan, Meng, Xuelian, Liu, Qingsheng, 2016. Mapping rice cropping systems in vietnam using an NDVI-based time-series similarity measurement based on DTW distance. Remote Sens. 8 (1). http://dx.doi.org/10.3390/rs8010019.
Gumma, Murali Krishna, Thenkabail, Prasad S., Teluguntla, Pardharsadhi, Rao, Mahesh N., Mohammed, Irshad A., Whitbread, Anthony M., 2016. Mapping rice-fallow cropland areas for short-season grain legumes intensification in south asia using MODIS 250 M time-series data. Int. J. Digital Earth 9 (10), 981–1003. http://dx.doi.org/10.1080/17538947.2016.1168489.
Hartigan, John A., Wong, Manchek A., 1979. Algorithm AS 136: a K-Means clustering algorithm. J. R. Stat. Soc. Ser. C (Appl. Stat.) 28 (1), 100–108 JSTOR.
Hay, Geoffrey J., Castilla, Guillermo, 2008. Geographic object-based image analysis (GEOBIA): a new name for a new discipline. Lecture Notes in Geoinformation and Cartography. pp. 75–89. http://dx.doi.org/10.1007/978-3-540-77058-9_4.
IBGE, 2014. Instituto Brasileiro de Geografia E Estatística, Produção Agrícola Municipal (PAM). http://www.ibge.gov.br/home/estatistica/economia/pam/2014/.
Calli, Jenkerson, 2013. User Guide: Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA) on Demand Interface.
Knauer, Kim, Gessner, Ursula, Fensholt, Rasmus, Kuenzer, Claudia, 2016. An ESTARFM fusion framework for the generation of large-Scale time series in cloud-prone and heterogeneous landscapes. Remote Sens. 8 (5). http://dx.doi.org/10.3390/rs8050425.
Labatut, Vincent, Cherifi, Hocine, 2012. Accuracy measures for the comparison of classifiers. In: The 5th International Conference on Information Technology. Amman, Jordanie.
Leenhardt, Delphine, Angevin, Frédérique, Biarnès, Anne, Colbach, Nathalie, Mignolet, Catherine, 2010. Describing and locating cropping systems on a regional scale. A review. Agron. Sustain. Dev. 30 (1), 131–138. http://dx.doi.org/10.1051/agro/2009002.
MacQueen, James, 1967. Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1, 281–297.
McNemar, Quinn., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12 (2), 153–157 Springer.
Mingwei, Zhang, Qingbo, Zhou, Zhongxin, Chen, Jia, Liu, Yong, Zhou, Chongfa, Cai, 2008. Crop discrimination in northern China with double cropping systems using fourier analysis of time-series MODIS data. Int. J. Appl. Earth Obs. Geoinf. 10 (4), 476–485. http://dx.doi.org/10.1016/j.jag.2007.11.002.
Pedroso da Silva, Antenor Roberto, de Almeida, Maria Geralda, 2007. Agronegócio E O Estado Do Tocantins: O Atual Estágio de Consolidação. Caminhos de Geografia 8 (21), s1.
Pontius, Robert Gilmore, Millones, Marco, 2011. Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32 (15), 4407–4429. http://dx.doi.org/10.1080/01431161.2011.552923.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Rosenfield, George H., Fitzpatrick-Lins, Katherine, 1986. A coefficient of agreement as a measure of thematic classification accuracy. Photogramm. Eng. Remote Sens. 52 (2), 223–227.
Silva, Luís Antônio Guerra Conceição, 2007. Biomas presentes No Estado Do tocantins. Consultoria Legislativa da Câmara dos Deputados.
Solano, Ramon, Kamel Didan, Andree Jacobson, and Alfredo Huete. 2010. 'MODIS Vegetation Index User's Guide (MOD13 Series)'.
Spangler, Keith R., Lynch, Amanda H., Spera, Stephanie A., 2017. Precipitation drivers of cropping frequency in the Brazilian Cerrado: evidence and implications for decision-making. Weather Clim. Soc. 9 (2), 201–213. http://dx.doi.org/10.1175/WCAS-D-16-0024.1.
Spera, Stephanie A., Avery, S. Cohn, VanWey, Leah K., Mustard, Jack F., Bernardo, Bernardo F., Risso, Joel, Adami, Marcos, 2014. Recent cropping frequency, expansion, and abandonment in Mato Grosso, Brazil had selective land characteristics. Environ. Res. Lett. 9 (6) IOP Publishing: 64010.
Strahler, Alan H., Luigi Boschetti, Giles M. Foody, Mark A. Friedl, Matthew C. Hansen, Martin Herold, Philippe Mayaux, Jefrey T. Morisette, Stephen V. Stehman, and Curtis E. Woodcock. 2006. 'Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Land Cover Maps'. Scientific and Technical Research Series. Vol. EUR 22156. 10.1080/01431160512331326521.
Tilman, David, Balzer, Christian, Hill, Jason, Befort, Belinda L., 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. U. S. A. 108 (50), 20260–20264. http://dx.doi.org/10.1073/pnas.1116437108.
Trimble Germany GmbH, 2014. eCognition Developer© 9.0 User Guide. Munich, Germany.
Vintrou, Elodie, Desbrosse, Annie, Bégué, Agnès, Traoré, Sibiry, Baron, Christian, Lo Seen, Danny, 2012. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. Int. J. Appl. Earth Obs. Geoinf. 14 (1), 83–93. http://dx.doi.org/10.1016/j.jag.2011.06.010. Elsevier B.V.
Wardlow, Brian D., Egbert, Stephen L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: an assessment for the U.S. central great plains. Remote Sens. Environ. 112 (3), 1096–1116. http://dx.doi.org/10.1016/j.rse.2007.07.019.
Xiong, Jun, Thenkabail, Prasad S., Gumma, Murali K., Teluguntla, Pardhasaradhi, Poehnelt, Justin, Congalton, Russell G., Yadav, Kamini, Thau, David, 2017. Automated cropland mapping of continental africa using google earth engine cloud computing. ISPRS J. Photogramm. Remote Sens. 126, 225–244 Elsevier.
Zanter, K., 2016. Landsat 8 (L8) Data Users Handbook. LSDS-1574 Version 2.0.