



HAL
open science

Exploitation des séries temporelles d'images satellite à haute résolution spatiale par fusion d'informations multi-sources pour le suivi des opérations culturales. Application à la détection des coupes de canne à sucre à La Réunion

M. El Hajj

► **To cite this version:**

M. El Hajj. Exploitation des séries temporelles d'images satellite à haute résolution spatiale par fusion d'informations multi-sources pour le suivi des opérations culturales. Application à la détection des coupes de canne à sucre à La Réunion. Sciences de l'environnement. Doctorat de l'Institut des Sciences et Industries du Vivant et de l'Environnement, AgroParisTech, 2008. Français. NNT : . tel-02591842

HAL Id: tel-02591842

<https://hal.inrae.fr/tel-02591842>

Submitted on 15 May 2020

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

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

N° / / / / / / / / / / / / / / /

THÈSE

pour obtenir le grade de

Docteur

de

**l'Institut des Sciences et Industries du Vivant et de l'Environnement
(Agro Paris Tech)**

*présentée et soutenue publiquement le 19 décembre 2008
par*

Mahmoud EL HAJJ

Titre de la thèse :

**Exploitation des séries temporelles d'images satellite à haute résolution
spatiale par fusion d'informations multi-sources
pour le suivi des opérations culturales**

Application à la détection des coupes de canne à sucre à La Réunion

Devant le jury :

Mme Laurence HUBERT-MOY , Professeur à l'Université de Rennes 2	Rapporteur
M. Olivier COLOT , Professeur à l'Université des Sciences et Technologies de Lille...	Rapporteur
Mme Véronique BELLON-MAUREL , Professeur à SupAgro Montpellier.....	Examineur
M. Sébastien GARRIGUES , Ingénieur CNES.....	Examineur
M. Jean-Stéphane BAILLY , Enseignant chercheur Engref – AgroParisTech.....	Examineur
Mme Agnès BÉGUÉ , Chercheur CIRAD, HDR.....	Directrice de thèse
M. Florent MARAUX , Chercheur CIRAD.....	Invité

Table des matières

Introduction.....	11
1. Chapitre 1 : Contexte méthodologique.....	15
1.1. Introduction.....	16
1.2. Séries temporelles d'images satellite et suivi des opérations culturales	17
1.3. Détection de changement	20
1.3.1. Méthodes bi-temporelles	20
1.3.2. Méthodes d'analyse de profils temporels	21
1.3.3. Discussion	21
1.4. Fusion d'informations	23
1.4.1. Fusion de données de nature identique.....	23
1.4.2. Fusion de données de nature hétérogène.....	23
1.4.3. Cadres de fusion	24
1.4.4. Discussion	26
1.5. Besoin méthodologique.....	28
2. Chapitre 2 : Application agricole et données	29
2.1. Introduction.....	30
2.2. L'application agricole	31
2.2.1. La canne à sucre à La Réunion.....	31
2.2.2. Télédétection et canne à sucre.....	36
2.3. Les données satellite et agronomiques	40
2.3.1. Séries temporelles d'images SPOT	40
2.3.2. Sites d'étude et base de données	43
2.4. Conclusion	48
3. Chapitre 3 : Pré-traitement des images dans les séries temporelles.....	49
3.1. Introduction.....	50
3.2. Correction géométrique.....	51
3.3. Correction radiométrique	53
3.3.1. Calcul de la réflectance au sommet de l'atmosphère	55
3.3.2. Normalisation radiométrique relative.....	57
3.3.3. Correction atmosphérique	63
3.3.4. Résultats et discussion.....	67
3.4. Conclusion	77

4. Chapitre 4 : Système d'aide à la détection des coupes de canne à sucre.....	79
4.1. Introduction.....	80
4.2. Sources d'information.....	81
4.2.1. Séries temporelles	81
4.2.2. Modèle de culture.....	86
4.2.3. Connaissances expertes.....	92
4.3. Formalisme	94
4.3.1. Ensembles flous et termes linguistiques.....	95
4.3.2. Inférence à base de règles.....	96
4.4. Structure et composants	98
4.4.1. Etage d'entrée.....	98
4.4.2. Sortie	105
4.4.3. Base de règles.....	106
4.4.4. Moteur d'inférence.....	114
4.5. Prise de décision	116
4.5.1. Système à base de règles expertes.....	116
4.5.2. Système à base de règles induites.....	118
4.6. Conclusion	119
5. Chapitre 5 : Système : résultats et discussion	121
5.1. Introduction.....	122
5.2. Système à base de règles expertes.....	123
5.2.1. Performances.....	123
5.2.2. Analyse des erreurs	130
5.2.3. Contribution des ensembles flous.....	131
5.2.4. Analyse de la robustesse par rapport au nombre d'images disponibles	132
5.3. Système à base de règles induites par apprentissage.....	135
5.3.1. Performances.....	135
5.3.2. Analyse de la robustesse par rapport au nombre d'images disponibles	139
5.3.3. Impact des partitions	140
5.3.4. Contribution des différentes sources d'informations.....	142
5.3.5. Validation pluriannuelle.....	147
5.4. Comparaison	148
5.5. Conclusion	151
6. Conclusions et perspectives.....	153
6.1. Contributions de la thèse.....	153
6.2. Perspectives	156
Liste des tableaux.....	161
Liste des figures.....	163
Références.....	167
Annexes des chapitres.....	177
Annexes publications	233

Remerciements

Ce travail a été réalisé au sein de l'unité mixte de recherche TETIS (Territoires, Environnement, Télédétection et Information Spatiale), à la Maison de la Télédétection à Montpellier. Il a été co-financé par une bourse Région Languedoc-Roussillon / Cemagref et un budget de fonctionnement du CIRAD. Je tiens donc à remercier dans un premier temps ces organismes financeurs, hébergeurs et collaborateurs.

Trois années de travail acharné ne peuvent se terminer sans une vive émotion et une légère sensation de « c'est passé trop vite ». C'est pourquoi, j'espère pouvoir remercier suffisamment toutes les personnes qui auront marqué cette période de ma vie que je n'oublierais pas de si tôt.

Parmi les acteurs incontournables de cette thèse, je tiens à remercier Agnès Bégué, l'initiatrice du projet et la directrice de thèse, pour m'avoir accordé autant de confiance tout au long de ces trois années. Elle a marqué le projet par la pertinence de ces remarques et de sa direction. Elle a été un soutien certain dans cette thèse et également pour mes autres projets personnels et professionnels, merci à toi. Je suis également très reconnaissant envers Serge Guillaume pour son appui et sa participation active en tant qu'encadrant. Il aura su me guider dans mes questionnements et mes orientations.

Je remercie d'avance tous les membres du jury de thèse, en particulier les rapporteurs Mme Laurence Hubert-Moy et M. Olivier Colot qui ont eu la gentillesse d'accepter ce rôle.

Qu'aurait été cette thèse sans les quatre comités de pilotage qui ont à chaque fois apporté un dynamisme certain à ma thèse. Je remercie donc Olivier Hagolle (CNES, Cesbio), Gérard Dedieu (CNES, Cesbio), Pierre Langelier (CIRAD) et Michel Deshayes (Cemagref).

J'aimerais également exprimer ma gratitude à d'autres personnes ayant beaucoup apporté à ma thèse : Hélène De Boissezon pour les images Kalideos (une vraie mine d'or pour une thèse en télédétection), Nicolas Baghdadi pour son soutien à chaque fois que j'en ai eu besoin et Ghaleb Faour qui m'a initié au monde de la télédétection au sein du CNRS Liban.

Merci également à Aurore Bury et à Jean François Martiné qui ont organisé ma mission à La Réunion et qui l'ont rendue agréable et très convivial. Ils m'ont aussi donné d'autres coups de pouce de grandes valeurs durant la thèse.

Il ne faut pas non plus oublier les personnes que j'ai rencontrées à Montpellier, en commençant par mes amis de bureau : Thierry Tormos, Valentine Lebourgeois et Nicolas Moyroud qui ont supporté avec beaucoup de joie bien sûr, mes états d'âmes de thésard et de futur chercheur. Vianney Houlès m'aura assuré son soutien et m'aura initié à une science tout aussi complexe que le domaine de la télédétection, à savoir les jeux de société.

Je remercie également les collègues du CIRAD avec lesquels je m'échappais du travail quotidien pendant la pause du midi, à savoir Valérie Soti, Guéric Lemaire, Pascal Degenne, Danny Lo seen, Stéphane Dupuy, Camille Lelong, Nathalie Jean, Baptiste Mougel et aussi Christian Baron avec qui j'ai quand même parlé « travail », il m'a ainsi apporté des conseils en modélisation.

Je voudrais remercier affectueusement ma famille libanaise et française qui m'ont assuré leur support et m'ont beaucoup encouragé dans cette thèse. Merci notamment à mon frère Maher et à mes sœurs Mirna, Mayssam et Mahy pour leur soutien continu. J'ajouterai également une « spéciale dédicace » à mes parents, Mohamed et Maha, qui ont fait de moi l'homme que je suis ; c'est grâce à leur gentillesse, leurs sacrifices de tous les instants et leur amour que j'ai pu m'accomplir ainsi, merci à vous.

Enfin, un grand merci plein de reconnaissance et d'amour pour ma femme Hayate. Je la remercie pour son soutien constant, son aide et ses encouragements.

Résumé

Les séries temporelles d'images satellite acquises à haute résolution spatiale sont une source d'information importante pour le suivi des changements sur des grandes surfaces. Ces données sont particulièrement intéressantes pour les applications agricoles car elles permettent d'appréhender à l'échelle parcellaire les changements d'état de surface induits par les pratiques culturales. Cependant, le traitement des séries temporelles est souvent limité par l'irrégularité des acquisitions et par la nécessité d'une intervention experte récurrente. Cette thèse présente une méthodologie innovante qui répond à cette problématique. L'application agricole considérée est le suivi des coupes de canne à sucre à La Réunion à partir d'images SPOT (Kalideos ISLE-Réunion[©]). Afin d'assurer la comparabilité des images, nous avons tout d'abord développé une méthode de normalisation radiométrique relative basée sur une technique de sélection automatique de cibles invariantes. Ensuite, nous avons conçu et développé un système qui exploite, en temps réel, le contenu informatif des séries temporelles en s'appuyant sur des simulations faites à partir d'un modèle de culture et sur des connaissances expertes. Le formalisme de fusion d'informations utilisé est basé sur la modélisation linguistique et sur la logique floue. Il permet de manipuler des données imprécises, incertaines et de nature hétérogène. Il permet également de construire des règles de décision interprétables qui reproduisent, en partie, le raisonnement humain. Deux méthodes de construction des règles sont proposées : la première repose sur des règles définies par l'expert et la deuxième sur une induction automatique des règles par apprentissage. Les performances du système ont été évaluées sur différents sites et sur différentes années. Les résultats obtenus sont satisfaisants : la précision globale atteint 98,8% et le pourcentage de bonne détection des coupes atteint 96,1%. L'analyse de la contribution des différentes sources d'informations a montré, entre autres, que le modèle de culture apporte 6,4% de précision supplémentaire dans la détection des coupes. L'apport de l'expertise a été difficile à évaluer car elle intervient dans différentes parties du système. La méthodologie présentée dans cette thèse est très prometteuse. Elle est applicable à d'autres cultures et transférable à d'autres applications telle que la cartographie dynamique de l'occupation du sol.

Mots-clés : Série temporelle, modèle de culture, connaissances expertes, système d'inférence floue, canne à sucre, détection de changement, fusion d'informations, opérations culturales.

Abstract

Time series of satellite images acquired at high spatial resolution constitute an important source of information for monitoring changes over large areas. These data are particularly interesting for agricultural applications. They can be used to track field scale changes of land surface induced by farming practices. However, operational processing of the time series is often limited by the irregularity of acquisitions and the recurring need for expert intervention. This thesis presents an innovative methodology that addresses this issue. The studied application concerns the sugarcane harvest detection on Reunion Island using SPOT images (Kalideos ISLE-Réunion[©]). To ensure the comparability of the images, we first developed a method for relative radiometric normalization based on an automatic technique of invariant targets selection. Then, we designed and developed a real-time decision system that processes the time series by integrating information from crop growth modeling and expert knowledge. The information fusion formalism we used is based on linguistic modeling and fuzzy logic. This formalism copes with the imprecision, the uncertainty and the heterogeneous nature of the data. It also allows building interpretable decision rules that implement human reasoning. Two methods of rule generation are proposed: the first is manual and relies on expert knowledge, whereas the second is automatic and requires learning data. System performances were evaluated on different study sites and different years. Results are satisfactory: the overall accuracy reaches 98.8 % and the percentage of good harvest detections is 96.1%. The analysis of the contribution of the different sources showed, among other things, that the crop model provides up to 6.4 % of additional precision of harvest detection. The contribution of expert knowledge was difficult to assess because this source is involved in different parts of the system. The methodology presented in this thesis is very promising. It can be applied to other crops and transferred to other agricultural applications like the dynamic mapping of land cover.

Keywords: Time series, crop model, expert knowledge, fuzzy inference system, sugarcane, change detection, information fusion, agricultural practices.

Introduction

Les séries temporelles d'images satellite sont une source d'information importante pour le suivi de la surface terrestre à différentes échelles. Les applications sont multiples ; elles vont de la surveillance et gestion des ressources naturelles au suivi des dynamiques d'occupation du sol, de l'évaluation des risques à la cartographie des dommages, de l'évaluation de l'expansion urbaine à la mise à jour des cartes topographiques. Nombre de ces applications sont liées à des questions environnementales et ont un rôle stratégique pour les planificateurs, les décideurs et les gestionnaires.

Aujourd'hui, des séries d'images acquises à hautes résolutions spatiale et temporelle sont disponibles. Ces séries permettent d'élaborer un suivi « fin » d'une grande diversité de structures spatio-temporelles, et sont particulièrement intéressantes pour les applications agricoles à l'échelle de la parcelle. Le développement de méthodes efficaces pour l'analyse de ces séries est l'une des questions de recherche, à la fois pertinentes et ouvertes, à laquelle la communauté de télédétection doit faire face. L'importance et l'actualité de cette problématique sont directement liées à l'augmentation de la quantité d'images multi-temporelles acquises par les nombreux satellites qui sont en orbite autour de notre planète.

La composante temporelle, rajoutée aux dimensions spectrale et spatiale, est une précieuse source d'information qui, si elle est judicieusement exploitée, permet de révéler des motifs complexes et importants. Toutefois, l'utilisation de cette composante augmente la complexité habituellement associée au traitement des images mono-dates de télédétection. Le pré-traitement et l'analyse des séries temporelles d'images nécessitent des développements méthodologiques particuliers qui prennent en compte la « comparabilité » des images d'une date à l'autre et les contraintes temporelles liées au processus ou au phénomène observé. Ces contraintes temporelles peuvent être formalisées par la connaissance du processus étudié

(expertise) et par les conditions environnementales qui conditionnent le processus (modélisation).

L'objectif de cette thèse est le développement d'une méthodologie d'exploitation des séries temporelles d'images satellite à haute résolution spatiale, en coopération avec d'autres sources, pour le suivi des opérations culturales. Cette coopération doit permettre de suppléer à l'éventuelle absence de données dans les séries d'images (*e.g.* liée à la présence des nuages ou aux conflits de programmation satellite) par l'intégration de contraintes temporelles et environnementales. L'application choisie est le suivi des coupes de canne à sucre à La Réunion, qui constitue un exercice de détection de changement particulièrement difficile en raison de la forte variabilité spatio-temporelle des parcelles de canne.

Pour répondre à cette question, la thèse aborde les axes de recherche suivants :

- la normalisation radiométrique de séries temporelles d'images à haute résolution, en proposant une méthode de détection automatique des cibles invariantes ;
- la détection des changements d'état de surface des parcelles agricoles, en développant un système d'aide à la décision, basé sur la logique floue et sur la modélisation linguistique, qui fusionne les informations issues des images avec des simulations d'un modèle de culture et des connaissances expertes. Ce système fournit des décisions auxquelles sont associées des valeurs de confiance.

Le mémoire s'articule suivant cinq chapitres. Le premier chapitre introduit le contexte de la thèse en proposant une revue rapide des méthodes de détection de changement développées pour l'analyse des séries temporelles, ainsi que des méthodes de fusion d'informations multi-sources.

Le deuxième chapitre est consacré à la présentation de l'application agricole et à la description des séries temporelles d'images SPOT¹ sur La Réunion utilisées dans cette thèse.

Le troisième chapitre aborde les besoins de prétraitement d'images dans les séries temporelles. Nous discutons dans ce chapitre des différentes phases du pré-traitement à réaliser avant l'extraction de l'information, et présentons la méthode automatique de normalisation radiométrique que nous avons développée. Nous décrivons également une

¹ Satellite Pour l'Observation de la Terre.

méthode de correction atmosphérique (développée par le CNES²) qui a servi de référence pour la validation de notre méthode.

Le quatrième chapitre décrit le système d'aide à la décision que nous proposons pour détecter les coupes de canne à sucre à La Réunion. Ce système innovant exploite les informations issues des séries temporelles d'images SPOT en intégrant les simulations d'un modèle de canne à sucre et des connaissances expertes. Son formalisme permet de gérer l'incertitude, l'imprécision et la nature hétérogène des données. Il permet également de construire des règles de décision linguistiques faciles à interpréter. Deux méthodes de construction de la base de règles sont proposées : la première repose sur la définition de règles par l'expert, et la deuxième sur un apprentissage automatique de ces règles.

Le dernier chapitre présente les résultats du système, avec les deux versions de base de règles : règles expertes et règles induites par apprentissage. Les performances du système sont analysées, ainsi que sa robustesse par rapport au nombre d'images disponibles. Le caractère opérationnel du système est également étudié.

Enfin, nous concluons la thèse par un bilan des travaux réalisés et par une discussion sur les perspectives, notamment sur la possibilité de transfert du système d'aide à la décision vers d'autres cultures.

² Centre National d'Etudes Spatiales.

Chapitre 1 :

Contexte méthodologique

Sommaire

1.1. Introduction	16
1.2. Séries temporelles d'images satellite et suivi des opérations culturelles.....	17
1.3. Détection de changement	20
1.3.1. Méthodes bi-temporelles.....	20
1.3.2. Méthodes d'analyse de profils temporels.....	21
1.3.3. Discussion.....	21
1.4. Fusion d'informations	23
1.4.1. Fusion de données de nature identique	23
1.4.2. Fusion de données de nature hétérogène.....	23
1.4.3. Cadres de fusion.....	24
1.4.4. Discussion.....	26
1.5. Besoin méthodologique.....	28

1.1. Introduction

Dans ce chapitre, nous introduisons le contexte méthodologique de la thèse à savoir le suivi des opérations culturelles à partir des séries temporelles d'images satellite par fusion d'informations multi-sources.

Etant donné que le suivi des opérations culturelles est un exercice de détection de changement, nous faisons, dans un premier temps, un rapide tour d'horizon des différentes méthodes de détection de changement développées pour l'analyse des séries temporelles d'images. Nous divisons ces méthodes en deux grandes familles : (1) la famille des méthodes bi-temporelles qui s'intéressent à la détection du changement par analyse de couples d'images ; et (2) celle des méthodes d'analyse des profils temporels qui s'appuient sur des comparaisons de profils d'indicateurs issus de la série tout entière.

Dans un deuxième temps, nous abordons le thème de fusion d'informations multi-sources vers lequel les travaux menés en télédétection se sont orientés ces dernières années afin d'améliorer les performances du traitement d'images. Nous brossons un bref état de l'art des différentes études réalisées en télédétection fusionnant, pour certaines, des données de nature identique, et pour d'autres, des données de nature hétérogène. Nous présentons également les différents cadres de fusion sur lesquels se basent les différentes méthodes.

Enfin, nous identifions les questions méthodologiques particulières pour lesquelles les méthodes de suivi des opérations culturelles à partir des séries temporelles d'images satellite devraient répondre ; ces questions sont au cœur de l'objectif général de cette thèse.

1.2. Séries temporelles d'images satellite et suivi des opérations culturelles

Durant les dernières décennies, de nombreuses études ont porté sur l'analyse et l'exploitation des séries temporelles d'images à basse ou moyenne résolution spatiale pour le suivi de la végétation (Boles *et al.*, 2004; McCloy and Lucht, 2004; Maignan *et al.*, 2008). Grâce à leur haute résolution temporelle, ces séries, acquises par des capteurs tels que NOAA/AVHRR, SPOT/VEGETATION, TERRA/MODIS et autres (Tableau 1), ont été utilisées pour surveiller et prédire des structures globales à des larges échelles (*e.g.* bassin de production). D'autres études se sont intéressées à l'analyse des séries temporelles d'images acquises avec une répétitivité beaucoup moins importante mais avec une haute résolution spatiale adaptée à l'échelle de la parcelle agricole ; citons à titre d'exemple les séries d'images acquises avec les capteurs MSS *e.g.* (Lyon *et al.*, 1998), TM *e.g.* (Collins and Woodcock, 1996) et ETM + *e.g.* (Andréfouët *et al.*, 2001) embarqués sur les satellites du programme américain Landsat.

De nos jours, compte tenu des progrès techniques et de la possibilité de programmation d'images sur des sites précis, des séries temporelles à haute résolution spatiale et temporelle sont désormais acquises par des satellites comme SPOT, FORMOSAT 2 et RapidEye (Tableau 1). Ces nouvelles séries permettent d'élaborer un suivi plus fin d'une grande diversité de nouvelles structures spatio-temporelles. Elles sont particulièrement intéressantes pour caractériser les pratiques culturelles, un sujet qui représente aujourd'hui l'un des enjeux majeurs de la recherche en environnement (Hubert-Moy, 2004).

Le suivi des opérations culturelles permet d'améliorer et d'adapter les modes de gestion des terres agricoles afin d'augmenter leur productivité. Il permet également d'évaluer l'impact de ces opérations sur l'environnement (*e.g.* la sédimentation par l'érosion des sols, l'enrichissement des plans d'eau par une sur-fertilisation des cultures, la contamination de l'eau et du sol par les pesticides, etc.).

Jusqu'à présent, peu d'études ont porté sur ce sujet à cause du manque de données adaptées. La majorité des travaux réalisés se basent sur des séries d'images à basse ou moyenne résolution spatiale. (Young and Wang, 2001) ont utilisé des séries d'images NOAA/AVHRR pour suivre les changements d'occupation du sol en Chine sur une période de 10 ans ; ils ont remarqué une baisse de productivité dans les régions forestières et une

hausse dans les régions agricoles. (Hubert-Moy *et al.*, 2005) ont utilisé des séries d'images MODIS pour identifier, à une échelle régionale, les zones récurrentes de sol nu ; ce suivi aide à changer les pratiques de gestion des terres. (Sakamoto *et al.*, 2005) ont également utilisé les données MODIS pour estimer la répartition spatio-temporelle de la phénologie du riz, du nombre de récoltes par an, et des modes de cultures dans le delta du Mékong. . (Xiao *et al.*, 2006) se sont servis du même type de données pour faire une cartographie du riz et identifier la période initiale d'inondation et de transplantation dans les champs. Par cette étude, ils ont signalé la présence de certaines sources d'erreur liées aux capteurs optiques, tels que la contamination des nuages, les effets topographiques, et les limites de la résolution à la fois spatiale et temporelle.

L'utilisation des séries à haute résolution spatiale pour le suivi des opérations culturales sur des grandes surfaces a été limitée, d'une part, par la sensibilité de leur traitement à l'irrégularité d'acquisition (à cause de la présence des nuages, du coût des images et/ou de la programmation satellite) et d'autre part, par l'augmentation de la complexité des scènes étudiées (due à la haute résolution spatiale) (Defries and Belward, 2000). Une classification multi-saisonnière a généralement été adoptée (Wardlow and Egbert, 2008). Cette classification produit une précision plus élevée que la classification mono-date pour la cartographie de l'occupation du sol (Wardlow and Egbert, 2003) et de type de culture (Price *et al.*, 1997). L'acquisition des images sans nuages au moment optimal au cours de la saison de croissance a également été un problème en raison de la faible répétitivité (Defries and Belward, 2000).

Satellite	Capteur	Résolution spatiale (m)			Fauchée (km)	Répétitivité (nb de jours)	Synthèse (jours)
		Pan	MS	TIR			
NOAA	AVHRR	-	1100	1100	3000	1	10
Terra/Aqua	MODIS	-	250, 500, 1000	1000	2330	2	16, 32
Landsat	MSS	-	30 ou 79	120 ou 79	185	16 ou 18	-
	TM	-	30	120	185	16	-
	ETM+	15	30	60	180	16	-
SPOT	VEGETATION		1000		2200	1	10
	HRVIR	10	10, 20	-	60	26 ou 4 avec dépointage	-
	HRG	5 ou 2,5	10	-	60 ou 120	26 ou 2,5 avec dépointage	-
FORMOSAT 2		2	8	-	24	1	-
RapidEye		-	5	-	77	1	-

Tableau 1. Caractéristiques des principaux capteurs fournissant des séries temporelles d'images.

Par conséquent, afin de pallier ces limitations et d'exploiter le contenu informatif des séries temporelles acquises à haute résolution spatiale, le développement de nouvelles méthodes est nécessaire.

1.3. Détection de changement

De nombreuses méthodes de détection de changement à partir des séries temporelles d'images satellite ont été proposées dans la littérature. Nous retrouvons des revues bibliographiques exhaustives sur ces méthodes dans (Singh, 1989; Mas, 2000; Inglada, 2001; Coppin *et al.*, 2004). Ces méthodes peuvent être regroupées en deux grandes familles (Coppin *et al.*, 2004) : les méthodes bi-temporelles et les méthodes d'analyse des profils temporels (ou encore dites méthodes multi-temporelles).

1.3.1. Méthodes bi-temporelles

Cette famille de méthodes s'intéresse à la détection des changements d'une zone géographique à partir d'images acquises à deux dates différentes. Ces méthodes reposent sur des algorithmes variés de traitements d'images, tels que la différence d'images (Lyon *et al.*, 1998), le ratio (Jensen, 1981), les transformations linéaires (Crist and Cicone, 1984; Nielsen *et al.*, 1998; Moisan *et al.*, 1999), l'analyse par vecteur de changement (Malila, 1980), la classification conjointe (Muchoney and Haack, 1994), la comparaison des classifications (Hubert-Moy, 2004), les réseaux de neurones (Woodcock *et al.*, 2001), ou encore la régression d'images (Ridd and Liu, 1998). Les applications de ces méthodes ont porté principalement sur des images acquises à haute résolution spatiale (*e.g.* Landsat, SPOT).

L'analyse détaillée des méthodes bi-temporelles réalisée par (Coppin *et al.*, 2004) a montré que le choix de l'algorithme de détection de changement était dans presque tous les cas pragmatique plutôt que scientifique, et que ce choix a plus été conduit par l'application elle-même que par les questions principales de détection de changement. (Hubert-Moy, 2004) a rapporté également qu'il n'existe pas de méthode de détection du changement qui soit réellement générique, ou au moins facilement transposable d'une application à une autre. Le type de méthode utilisée, d'après elle, dépend largement de la zone d'étude de changement à détecter, et de sa résolution spatiale et temporelle.

1.3.2. Méthodes d'analyse de profils temporels

L'analyse des profils temporels consiste à comparer des signatures temporelles de différents indicateurs (*e.g.* indices de végétation), extraites à partir de séries d'images acquises régulièrement durant plusieurs saisons ou années successives. La fréquence temporelle élevée d'acquisition de données permet de détecter les changements subtils de la zone étudiée et de dégager les tendances : quand le profil temporel d'un ou plusieurs indicateurs d'un pixel (ou d'un objet) particulier s'écarte de la normale (ou de la moyenne), un événement de changement saisonnier ou un processus interannuel est détecté.

Plusieurs méthodes d'analyse de profils temporels se basent sur des techniques de traitement du signal, telles que les transformées d'ondelettes (Sakamoto *et al.*, 2005; Blackburn and Ferwerda, 2008) et de Fourier (Azzali and Menenti, 2000; Westra and De Wulf, 2007). D'autres méthodes utilisent des techniques de différence d'image (Ingram and Dawson, 2005), de vecteur de changement multi-temporel (Bayarjargal *et al.*, 2006), des modèles de régression (Fraser *et al.*, 2005) et d'analyse en composantes principales (Young and Wang, 2001).

La plupart des méthodes développées ont été appliquées à des images acquises à basse ou moyenne résolution spatiale et à haute répétitivité temporelle (*e.g.* AVHRR, SeaWIFS, VÉGÉTATION, MODIS) pour suivre des structures globales. Etant donnée la résolution spatiale de l'imagerie utilisée, et compte tenu de la couverture de larges surfaces, la validation de ces méthodes avec des ensembles de données indépendants reste un défi important (Coppin *et al.*, 2004).

1.3.3. Discussion

Les méthodes de détection de changement à partir de l'imagerie multi-temporelle sont nombreuses et variées. Le choix de ces méthodes dépend principalement de l'application traitée, de la zone d'étude et des données satellite disponibles. En ce qui concerne le suivi opérationnel des opérations culturales à partir des séries à haute résolution spatiale, la méthode choisie doit répondre à deux critères essentiels :

(1) la robustesse par rapport à l'irrégularité des acquisitions due aux conditions atmosphériques (nuages, brumes, etc.), au coût des images et/ou aux conflits de programmation satellite ;

(2) l'automatisation de l'exploitation du contenu informatif complexe des séries d'images.

Malgré le large éventail de méthodes proposées jusqu'à présent, aucune ne semble satisfaisante à la fois ces deux critères. Pour y aboutir, l'intégration d'informations provenant d'autres sources est nécessaire (ce point sera discuté dans la section suivante).

Par ailleurs, la question des dépendances temporelles entre les observations a été clairement négligée jusqu'à présent (Bontemps *et al.*, 2008). Seules quelques études ont exploité ces dépendances en se basant sur des automates temporisés (Largouët, 2000), cependant les finalités de ces travaux ne sont pas appropriées pour des prises de décision en temps réel. Afin de tirer profit de l'information temporelle, il est donc nécessaire de développer des méthodes qui prennent cette information en considération par l'intégration, dans les procédures de traitement de données, d'algorithmes capables d'exploiter les relations entre les images de la série. La prise en compte de ces relations passe par l'incorporation des connaissances expertes sur les milieux étudiés. La solution de ce problème méthodologique complexe peut se traduire par une augmentation de la précision fournie par la méthode de détection de changement.

Indépendamment du grand intérêt du processus de changement, l'état de l'art a souligné l'importance de la phase de prétraitement lors de la comparaison multi-temporelle des données de télédétection (Chen *et al.*, 2005). Une superposition précise des images (*i.e.* recalage) est un pré-requis primordial, notamment lors de l'utilisation des méthodes ponctuelles (à l'échelle du pixel). La correction radiométrique est également une condition indispensable pour rendre les images comparables (Song *et al.*, 2001; Du *et al.*, 2002; Coppin *et al.*, 2004).

1.4. Fusion d'informations

Ces dernières années, les travaux menés en télédétection en général et en détection du changement en particulier se sont orientés vers des approches qui exploitent conjointement différentes sources d'informations. Le but principal de la fusion des informations multi-sources est de profiter de la complémentarité des sources pour améliorer les performances. Les différentes sources fusionnées peuvent être de nature identique (*i.e.* images) ou hétérogènes (*e.g.* images et connaissances expertes).

1.4.1. Fusion de données de nature identique

Plusieurs travaux ont porté sur la fusion de données provenant de sources de nature identique : des images issues de plusieurs capteurs avec des résolutions (spatiale, spectrale et temporelle) identiques ou différentes, des images issues de plusieurs canaux d'un même capteur, ou des images dérivées de données de télédétection (*e.g.* classifications, transformées d'images, images d'indices de changement). (Le Hégarat-Masclé and Seltz, 2004) ont développé une méthode de détection du changement qui fusionne différents indices calculés à partir des images multi-temporelles afin d'évaluer les dommages causés aux forêts. Cette méthode a ensuite été utilisée dans différentes applications (*e.g.* suivi de l'exploitation forestière dans des forêts de pins et des forêts mixtes, suivi des parcelles dans des zones agricoles) avec des images SPOT HRV (Le Hégarat-Masclé *et al.*, 2006). (Hyde *et al.*, 2006) ont fusionné des informations structurales issues des données LiDAR, radar (SAR/InSAR) et optiques (Landsat ETM+, Quickbird) afin d'améliorer la cartographie de la structure forestière. Ils ont montré que l'intégration des données ETM+ améliore considérablement les estimations issues des données LiDAR et que les contributions des données radar et des données Quickbird sont marginales.

1.4.2. Fusion de données de nature hétérogène

Les données de télédétection peuvent aussi être fusionnées avec des informations hétérogènes provenant d'autres sources. Elles ont été intégrées avec des connaissances expertes *e.g.* (Lobell *et al.*, 2003), des données auxiliaires (*i.e.* informations cartographiques) (Lucas *et al.*, 2007) et des modèles agronomiques ou d'écosystèmes (Yan *et al.*, 2007). L'objectif de l'intégration diffère suivant les applications. (Lobell *et al.*, 2003) ont combiné des connaissances sur la phénologie des cultures avec des images multi-temporelles acquises par Landsat TM/ETM+ pour estimer la rotation des cultures régionales. (Lucas *et al.*, 2007)

ont évalué l'utilisation des séries temporelles d'images Landsat ETM+ acquises sur un cycle annuel pour la cartographie des habitats semi-naturels et de l'occupation des terres agricoles, par l'intégration de cartes topographiques, de données numériques d'élévation, de l'orthophotographie numérique et d'autres données auxiliaires. (Yan *et al.*, 2007) ont couplé des données issues des images NOAA/AVHRR avec un modèle d'écosystème pour estimer le potentiel de séquestration du carbone dans les sols agricoles et pour évaluer la durabilité de son absorption sous différents modes de gestion. (Mercier *et al.*, 2005) ont proposé une méthode pour l'estimation des états de surface à une échelle fine à partir des observations à basse résolution spatiale (provenant des images SPOT-VEGETATION), des observations terrain et des observations antérieures à haute résolution spatiale (provenant des images SPOT-HRVIR). Leur méthode est basée sur un réseau de neurones qui prend en considération des connaissances *a priori* en intégrant le potentiel du changement temporel pour l'apprentissage. (Metternicht, 2001) a cartographié la distribution de la salinité avec une approche qui intègre la classification multi-temporelle des images Landsat TM, des propriétés physiques et chimiques du sol et des attributs de relief.

L'intégration des données de télédétection avec des modèles agronomiques (*e.g.* par assimilation, forçage, etc.) a également été utilisée pour estimer la croissance des cultures et pour améliorer les prévisions du rendement (Delecolle *et al.*, 1992; Moulin *et al.*, 1998; Guerif and Duke, 2000; Launay and Guerif, 2005). Cette intégration a porté sur l'utilisation directe dans les modèles d'une ou plusieurs variables estimées à partir des données de télédétection, sur la mise à jour d'une variable d'état du modèle (*e.g.* LAI³), sa ré-initialisation, ou sa re-calibration pour obtenir des simulations en accord avec les données de télédétection.

1.4.3. Cadres de fusion

Différentes méthodes de fusion d'informations multi-sources ont été développées en se basant sur des cadres différents : théorie des probabilités, théorie des fonctions de croyance, théorie des possibilités, etc.

Les premières approches développées en télédétection ont été basées sur un concept probabiliste de type bayésien, *e.g.* (Swain, 1985). Toutefois, bien que l'intérêt de ce concept soit de reposer sur de solides bases mathématiques, ses limites concernant la combinaison d'informations multi-sources ont été rapidement mises en évidence. En effet, ce concept ne

³ Indice de surface foliaire (*Leaf Area Index*).

permet pas la manipulation de données symboliques ; les informations issues de diverses sources ne peuvent être combinées que si elles obéissent à une même loi de distribution, et les notions d'importance et de concordance des sources les unes par rapport aux autres ne sont pas bien intégrées dans les méthodes existantes.

Un autre point très important que le concept probabiliste ne gère pas est celui de l'imperfection des informations multi-sources. Ces informations sont dans la plupart des cas incertaines, imprécises, incomplètes, voire erronées. Les erreurs sont situées à toutes les étapes de la chaîne de traitement des images : du phénomène observé, du capteur, des phases de pré-traitement (corrections géométriques et radiométriques, normalisation, etc.) et de traitement. Malgré le manque d'exactitude de ces données on est toujours amené à prendre des décisions, et pour que cette décision soit rationnelle et cohérente il faut prendre en compte ces imperfections.

De nouveaux travaux se sont alors tournés vers la théorie des fonctions de croyance et la théorie des possibilités (*i.e.* ensembles flous). La théorie des fonctions de croyance regroupe l'ensemble des outils méthodologiques élaborés à partir des travaux de A.P. Dempster (Dempster, 1968) et de G. Shafer (Shafer, 1976), le plus souvent fédérés par des appellations différentes : *Dempster-Shafer Theory*, *Theory of Evidence*, etc. Cette théorie fournit un cadre pour traiter simultanément l'incertitude et l'imprécision des données à fusionner, ce que la théorie des probabilités ne permet pas. Elle présente aussi l'avantage de pouvoir manipuler des données qualitatives et de modéliser explicitement l'ignorance, qui peut être affectée *a priori* à des propositions inconnues. Cette ignorance peut, par la suite, évoluer avec l'apport de nouvelles connaissances. La fusion des données peut être faite avec la règle de la combinaison de Dempster. La notion de masse utilisée dans cette règle permet de définir l'incertitude, l'imprécision, et les conflits entre les informations multi-sources.

Dans un premier travail, (Lee *et al.*, 1986) ont présenté la différence de philosophie entre l'approche probabiliste et la théorie des fonctions de croyance, avant de les comparer sur la production d'une carte thématique par la fusion de quatre canaux Landsat. Des méthodes basées sur la théorie de fonctions de croyance ont ensuite été développées pour de nombreuses applications, et notamment pour la caractérisation de l'occupation du sol à partir des images multi-temporelles (Shimoda *et al.*, 1991), pour la fusion d'indices dérivés des images multi-temporelles afin d'améliorer la détection des changements (Le Hégarat-Masclé and Seltz, 2004), et pour la combinaison des classifications (Foucher *et al.*, 2003).

Le problème le plus délicat dans l'application de la théorie des fonctions de croyance est celui de l'attribution des masses aux sources ou aux sous-ensembles de sources. Si ces masses sont relatives à des informations abrégées, elles peuvent être directement définies par des experts (Baldwin, 1991). A un niveau intermédiaire, les masses sont calculées à partir des caractéristiques de l'image que l'on intègre dans un modèle géométrique simple (Van Cleynenbreugel *et al.*, 1991).

Une autre théorie héritant de la théorie des fonctions de croyance est la théorie des possibilités. Cette théorie, basée sur des ensembles flous (Dubois and Prade, 1979; Dubois and Prade, 1985), est une branche importante de la logique floue (Zadeh, 1965) qui suggère que l'inclusion d'un élément dans un ensemble est une question de détermination de son degré d'appartenance à cet ensemble. Si la théorie des ensembles flous assouplit le cadre de la théorie classique des ensembles pour pouvoir traiter l'imprécision, la théorie des possibilités propose un cadre permettant de traiter les incertitudes difficiles, voire impossibles à traiter par la théorie des probabilités. En effet, raisonner en termes de probabilités suppose de pouvoir estimer la probabilité de chaque événement. Pour cela, il faut en avoir une bonne connaissance. Si cette connaissance n'est pas disponible, alors on raisonne en termes de possibilité (et de son dual, la nécessité).

La théorie des possibilités et les ensembles flous ont été utilisés pour traiter différentes applications dans le domaine de la télédétection. (Andrefouet, 2000) a décrit l'étude d'un système écologique en utilisant des données de télédétection et des outils d'analyse d'images provenant de la théorie des possibilités. (Waser *et al.*, 2008) ont développé une méthode de détection de changement des surfaces forestières basée sur une classification floue qui utilise à la fois des modèles de régression logistique et des images aéroportées. (Solaiman, 1999) a fusionné des données issues de différents capteurs (ERS-L et JERS-1) avec des informations contextuelles pour faire une classification de l'occupation du sol. (Fauvel *et al.*, 2005) ont mis en place des techniques floues pour fusionner les résultats de plusieurs classifieurs, en tenant compte de la performance de chaque classifieur pour chaque classe.

1.4.4. Discussion

La fusion d'informations permet d'améliorer les performances du traitement d'images de télédétection en intégrant des données de natures identique ou hétérogène. Dans la littérature, cette fusion a porté principalement sur (1) des images de sources différentes qui sont issues de capteurs différents, de bandes spectrales ou temporelles différentes, *e.g.* (Simone *et al.*, 2002),

(2) sur des transformées d'images ou des indices de changement, *e.g.* (Le Hegarat-Masclé *et al.*, 2006), (3) et sur des images avec des données exogènes, *e.g.* (Desachy *et al.*, 1996).

Afin de répondre aux besoins de « robustesse » et d'« automatisation » signalés dans la section 1.3.3 pour le suivi des opérations culturales, il nous semble pertinent d'intégrer des informations provenant de la modélisation agronomique et des connaissances expertes.

Jusqu'ici, le bénéfice de l'intégration des données de télédétection avec les modèles agronomiques était l'amélioration des performances des modèles (Delecolle *et al.*, 1992; Guerif and Duke, 2000; Launay and Guerif, 2005). Peu d'attention a été accordée à l'utilisation des « sorties » des modèles (*i.e.* simulations) en tant que source d'information pour l'analyse des données de télédétection. Basés sur des variables climatiques et des paramètres biophysiques du sol, ces modèles peuvent fournir des informations indépendantes des séries temporelles, qui permettent de compenser leurs lacunes. Ainsi, le bénéfice de l'intégration du modèle et des séries temporelles que nous proposons, est d'améliorer les performances de l'exploitation des images.

L'autre source substantielle d'information que nous intégrons, est la connaissance experte sur l'application étudiée. Cette source permet d'automatiser l'analyse des données issues des images tout en prenant en considération les relations temporelles entre les observations ; par ailleurs, elle fournit des informations et des contraintes très utiles en cas de manque d'acquisition.

En pratique, la fusion d'informations multi-sources n'est pas un problème simple. Les sources apportent des informations souvent imprécises, incertaines et parfois contradictoires. Il s'agit alors de tenir compte du degré de fiabilité que l'on accorde aux différentes sources, et de choisir un cadre de fusion qui soit adapté à la nature de ces sources et qui permette de gérer l'incertitude de leurs informations, les imprécisions et les conflits. Il est également judicieux de fournir des résultats de fusion qui soient associés à des facteurs de confiance.

1.5. Besoin méthodologique

L'amélioration du suivi des opérations culturelles à partir des séries temporelles acquises à haute résolution spatiale est intimement liée au développement de méthodes automatiques, fiables et robustes de détection de changement. D'après l'état de l'art présenté dans les sections précédentes, il s'avère que ces méthodes doivent :

- répondre aux besoins de prétraitement d'images dans les séries pour les rendre superposables et comparables ;
- exploiter l'information temporelle en considérant les relations entre les images ;
- réduire les interventions humaines en automatisant le plus possible les phases de traitement ;
- intégrer des informations multi-sources indépendantes des séries temporelles pour compenser le manque de données lié à la présence des nuages, au coût des images et à la programmation satellite;
- gérer l'incertitude, l'imprécision et les conflits des informations multi-sources ;
- fournir une cartographie dynamique associée à des facteurs de confiance.

L'objectif général de cette thèse est de développer des méthodes qui répondent à ces besoins. A terme, ces méthodes pourront être appliquées aux produits des systèmes spatiaux conçus pour le suivi temporel, *e.g.* Formosat-2, Sentinel-2 (ESA-UE), RapidEye (DLR), etc.

Chapitre 2 :

Application agricole et données

Sommaire

2.1. Introduction	30
2.2. L'application agricole	31
2.2.1. La canne à sucre à La Réunion	31
2.2.1.1. La Réunion	31
2.2.1.2. La filière canne à sucre.....	31
2.2.1.3. Le besoin de suivi des coupes.....	34
2.2.2. Télédétection et canne à sucre.....	36
2.2.2.1. Le suivi des coupes à partir de l'imagerie multi-temporelle.....	37
2.2.2.2. Besoin méthodologique.....	38
2.3. Les données satellite et agronomiques	40
2.3.1. Séries temporelles d'images SPOT	40
2.3.2. Sites d'étude et base de données	43
2.3.2.1. Exploitations de canne à sucre.....	43
2.3.2.2. Parcellaire graphique.....	46
2.3.2.3. Données agronomiques	46
2.3.2.4. Données météorologiques	47
2.4. Conclusion.....	48

2.1. Introduction

Dans ce chapitre, nous présentons l'application agricole à laquelle nous nous intéressons dans cette thèse et qui servira de base pour la description de la méthodologie de suivi des opérations culturales proposée. Il s'agit du suivi des coupes de canne à sucre à La Réunion à partir des séries temporelles d'images SPOT⁴. Le choix de cette application est motivé, d'une part, par son importance par rapport à l'industrie de la canne à sucre à travers le monde, et d'autre part, par le défi méthodologique lié à la forte variabilité spatio-temporelle de cette culture.

Après un exposé général sur la canne à sucre à La Réunion, nous mettons en évidence le besoin de suivi des coupes de parcelles par la filière. Nous faisons un état de l'art sur les quelques travaux réalisés en télédétection qui touchent à la canne à sucre. Nous décrivons ensuite les principales méthodes de suivi de coupes rapportées dans la littérature, nous discutons leurs limites et dégageons les besoins méthodologiques nécessaires pour mieux répondre à l'application.

Dans la deuxième partie du chapitre, nous présentons les séries temporelles d'images SPOT utilisées, ainsi que les sites d'étude sélectionnés pour la mise au point et l'évaluation de notre méthode de suivi des coupes. Nous décrivons également les données agronomiques et météorologiques disponibles sur ces sites.

⁴ Satellites Pour l'Observation de la Terre.

2.2. L'application agricole

2.2.1. La canne à sucre à La Réunion

2.2.1.1. La Réunion

La Réunion est un département et une région français d'outre mer (DROM) situé dans l'Océan Indien (Figure 1). Il s'agit d'une île d'origine volcanique située dans l'archipel des Mascareignes, 800 km à l'Est de Madagascar (21°7' à 19°40' Sud, 55° 13' à 61°13' Est). Sa superficie de 2 512 km² occupe une forme ovoïde, compacte et fortement accidentée. Les régions côtières en pentes douces laissent place à une zone fortement montagneuse constituée de trois cirques naturels au centre de l'île. Le climat à La Réunion est tropical humide. L'année y est divisée en deux saisons : une saison pluvieuse et chaude de novembre à avril, et une saison « froide » et sèche de mai à octobre.

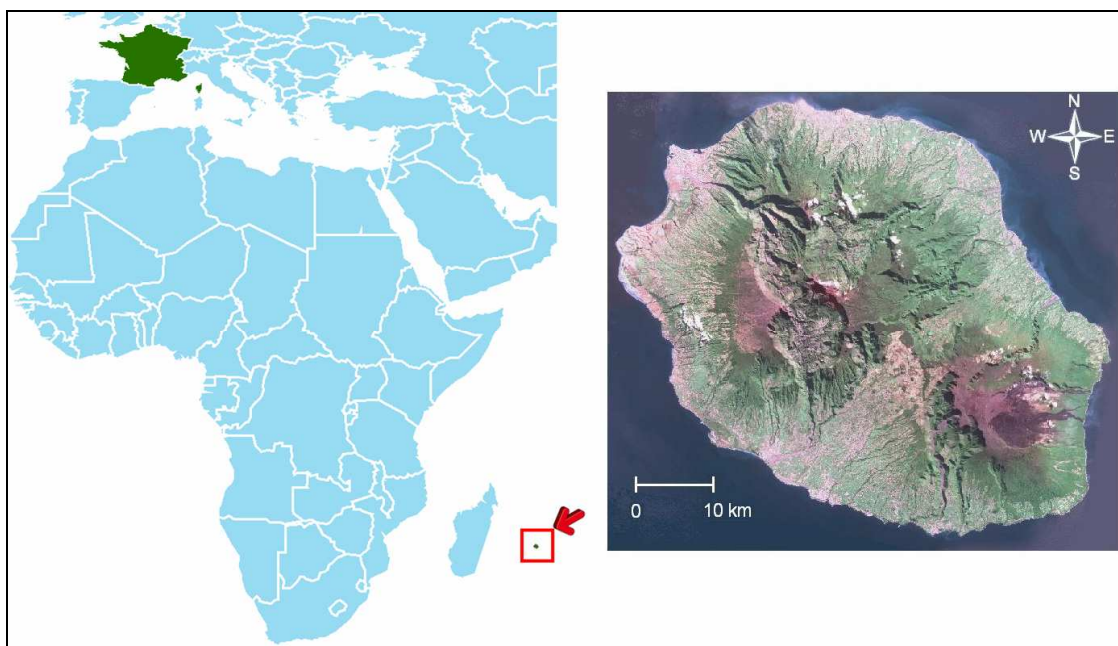


Figure 1. Localisation de l'île de La Réunion.

2.2.1.2. La filière canne à sucre

La canne à sucre (Figure 2) représente la principale culture de La Réunion. Elle est cultivée tout le long de la côte, sur une surface de 25 400 ha (Source: Agreste Réunion⁵, 2006). Deux millions de tonnes de tiges sont récoltées chaque année par 4 000 planteurs pour

⁵ <http://agreste.agriculture.gouv.fr>.

produire environ 200 000 tonnes de sucre ; La Réunion est ainsi le premier producteur européen de sucre de canne.

Mises à part quelques grandes plantations appartenant à une poignée de propriétaires terriens, la canne à sucre à La Réunion est essentiellement cultivée par de petits planteurs, gérant des exploitations d'une surface moyenne de 6,48 ha (Figure 3). Dans le Nord-Est, la partie humide de l'île, la culture de la canne à sucre est pluviale, alors qu'elle est irriguée dans le Sud-Ouest plus sec.

Les plantations de canne à sucre à la Réunion ont lieu généralement entre les mois de novembre et de mars, *i.e.* pendant la période chaude et pluvieuse. La longueur moyenne du cycle total est entre 7 et 8 ans, soit une canne vierge suivie de 6 à 7 repousses (voir le cycle de la canne à La Réunion dans l'Annexe 1). L'âge des cannes à la récolte varie entre 10 et 12 mois pour les repousses et il est aux alentours de 18 mois pour les cannes vierges. La campagne de récolte se déroule généralement de juillet (saison sèche) à décembre (avant le retour des pluies).



Figure 2. Photo de canne à sucre à La Réunion (© V. Lebourgeois).

Un mois avant le début de la campagne, les deux usines sucrières de l'île définissent le quota journalier que pourra fournir chaque agriculteur (tonnage par nombre de jours de campagne), en se basant sur les estimations individuelles de production. Le prix de la tonne de canne livrée dépend de sa richesse en sucre et de la richesse moyenne des livraisons de la même période.

Pivot de l'agriculture, la canne à sucre occupe plus que 60% des surfaces agricoles disponibles à la Réunion, et génère 15 000 emplois, directs et indirects. Sur cette île, dont la démographie est croissante, le défi est de taille : augmenter la production agricole tout en développant l'urbanisme. A la fois social, économique et culture propre, la canne est le ciment d'une réflexion collective pour le développement durable de l'île (CIRAD, 2008).

Une présentation générale sur la culture de la canne à sucre dans le monde est fournie dans l'Annexe 2.

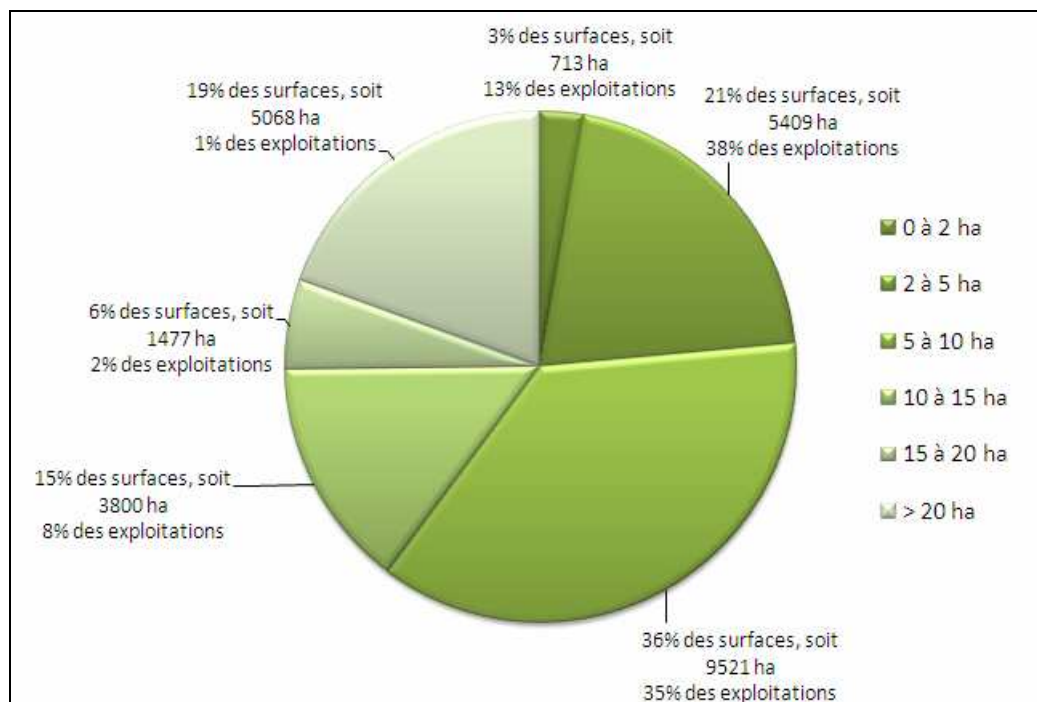


Figure 3. Classification des exploitations cannières selon leur surface en canne (Source : DAF Réunion, Février 2006).

2.2.1.3. Le besoin de suivi des coupes

L'un des principaux besoins exprimés par l'industrie de la canne à sucre à La Réunion et à travers le monde, est d'avoir une information continue sur l'avancement de la récolte tout au long de la saison, et ceci pour accroître l'efficacité à la fois dans l'usine et au champ. En effet, la récolte de la canne est une étape cruciale et demande une grande organisation logistique. Sa durée, généralement de 4 à 6 mois, est fonction des conditions climatiques qui déterminent la richesse en sucre, et du dimensionnement des usines qui fonctionnent nuit et jour à flux tendu. La coupe de la canne peut s'effectuer dans certaines régions du monde en continu toute l'année.

La récolte se fait encore beaucoup à la main dans les pays où la main d'œuvre reste bon marché, mais la mécanisation tend à se développer (Figure 4). Après la coupe, la canne doit être broyée dans les plus brefs délais, de 24h à 48h selon le mode de récolte, car on observe des processus chimiques entraînant l'apparition de sucres non cristallisables. Compte tenu de cette contrainte et de la quantité de biomasse à traiter (entre 60 et 120 tonne/ha), l'organisation de la récolte, des transports et du traitement de la canne doit être soigneusement planifiée sur les quelques mois que dure la récolte. A La Réunion, la filière souhaite avoir une information en temps réel sur les surfaces coupées durant la campagne avec une erreur inférieure à 5%.



Figure 4. Coupe mécanique d'une parcelle de canne à la Réunion (© M. El Hajj).

Un outil prometteur permettant de faire un suivi dynamique de la récolte de canne à sucre sur de grandes surfaces est l'analyse et l'exploitation des séries temporelles d'images satellites acquises à haute résolution spatiale. Nous nous intéressons dans cette thèse à cette application agricole qui représente un exercice difficile de suivi des opérations culturales. Cette application s'inscrit dans la continuité du projet SUCRETTE (Bégué *et al.*, 2005) mené par le CIRAD entre 2002 et 2004.

2.2.2. Télédétection et canne à sucre

Malgré la place importante occupée par la canne à sucre dans le monde, avec une contribution de 23% à la masse totale produite par l'agriculture mondiale (Source : FAO⁶, 2007), peu d'études basées sur la télédétection ont été menées sur cette culture. Ceci est dû à la très forte variabilité spatio-temporelle de la canne, à la longueur de la saison de récolte (Bégué *et al.*, 2008), et à la difficulté de réaliser des mesures *in situ* sur de grandes surfaces (Simoes *et al.*, 2005).

Les recherches réalisées pour la canne à sucre dans le domaine de la télédétection ont été largement inspirées de ce qui a été fait pour d'autres graminées (Abdel-Rahman and Ahmed, 2008). Des approches directes ont été élaborées pour relier la biomasse de la canne à sucre à des mesures multi-spectrales acquises par des radiomètres (Simoes *et al.*, 2005), par des capteurs aéroportés (Schmidt *et al.*, 2000), par des satellites à haute résolution spatiale (Krishna Rao *et al.*, 2002; Ueno *et al.*, 2005; Almeida *et al.*, 2006), et des satellites à basse et moyenne résolution spatiale (Schmidt *et al.*, 2000; Bastidas-Obando and Carbonell-Gonzalez, 2007). D'autres approches indirectes qui combinent des données multi-spectrales avec des modèles agrométéorologiques et/ou écophysiologicals de production ont également été explorées (Rudorff and Batista, 1990; Bastiaanssen and Ali, 2003; Bappel *et al.*, 2005).

Ces études ont montré l'apport de la télédétection en général et de l'imagerie satellite en particulier pour différentes applications liées à la canne. (Almeida *et al.*, 2006) ont proposé une méthode de prévision de rendement de canne à sucre en utilisant des données ASTER et Landsat ETM+. (Fortes and Dematte, 2006) ont mis au point une méthodologie pour la discrimination des variétés de canne en utilisant les images Landsat ETM+. D'autres chercheurs ont étudié la contribution des images multi-temporelles optiques pour classer les parcelles de canne à sucre, *e.g.* (Xavier *et al.*, 2006), pour identifier leurs caractéristiques, *e.g.* (Gers, 2003; Bappel *et al.*, 2005), et pour avoir des informations sur le progrès de la récolte durant la saison de coupe, *e.g.* (Bégué *et al.*, 2004; Lebourgeois *et al.*, 2007).

⁶ Organisation des Nations Unies pour l'alimentation et l'agriculture (www.fao.org).

2.2.2.1. Le suivi des coupes à partir de l'imagerie multi-temporelle

L'utilisation de l'imagerie multi-temporelle pour le suivi des coupes de canne à sucre est récente et liée à l'évolution des solutions satellitaires (*i.e.* progrès techniques au niveau des résolutions spatiales des capteurs, et possibilité de programmer des images sur des sites précis). En 2001, (Gers and Schmidt, 2001) ont utilisé quatre images SPOT-4 (une peu avant la campagne de récolte et 3 durant la campagne) pour le suivi des coupes dans une région d'Afrique du sud. La précision de détection des coupes qu'ils ont obtenue varie entre 62% et 85% ; ils ont rapporté des limitations liées à la résolution spatiale des images qui, parfois, n'est pas forcément adaptée pour de petites parcelles. (Bégué *et al.*, 2004) ont montré l'apport des séries temporelles d'images SPOT-4&5 (avec 4 dates) pour la détection et l'estimation des surfaces de canne coupées en Guadeloupe, avec des résultats plutôt encourageants : 90% de pixels bien classés. (Lebourgeois *et al.*, 2007) ont également utilisé des données SPOT (3 images acquises durant la campagne de récolte) pour faire une cartographie des coupes à la Réunion, avec une précision globale entre 93,8% et 96,5%, et une précision de détection des coupes qui atteint 89,5%. Toutefois le nombre de parcelles impliquées dans cette étude était faible et les résultats ont été évalués sur la surface dégagée sur toutes les images (échantillonnage spatial dû à la nébulosité).

Les principales méthodes rapportées dans la littérature, pour faire le suivi des coupes se basent soit sur une classification multi-spectrale soit sur une différence bi-temporelle d'images dans le Moyen Infra-Rouge⁷.

- La méthode de classification multi-spectrale consiste à faire une classification supervisée (Gers and Schmidt, 2001) ou non-supervisée (Lebourgeois *et al.*, 2007), de chaque image satellite selon les principaux états de surface d'une parcelle de canne (*e.g.* « canne verte », « résidus de récolte », « sol nu »), et à identifier ensuite les parcelles coupées en regroupant certaines classes (*e.g.* regroupement des classes « résidus de récolte » et « sol nu »). Pour réduire quelques confusions, certains auteurs (Bégué *et al.*, 2005) ont procédé à une comparaison de la classification obtenue avec celles des dates antérieures, en se basant sur les connaissances expertes sur le cycle de la canne.

Les deux limitations majeures de la méthode de classification multi-spectrale sont :

⁷ Les caractéristiques spectrales de la canne seront discutées plus en détails dans le chapitre 4.

- La subjectivité et le temps considérable à investir dans la phase de photo-interprétation qui est nécessaire (i) pour réduire le nombre de classes (cas d'une classification non-supervisée) ou pour faire l'apprentissage (cas d'une classification supervisée), et (ii) pour détecter les coupes en comparant les classifications;
- La confusion entre les différents états de parcelle lorsque le délai entre deux images dégagées est important (supérieur à 2 mois) ; étant donné que la campagne de récolte s'étend sur plusieurs mois (6 environ), il est difficile dans certains cas de distinguer une canne sur pied (non récoltée) d'une repousse qui a été récoltée au moins deux mois auparavant.
- La méthode de différence bi-temporelle dans le Moyen Infra-Rouge (Bappel, 2005) consiste à créer pour chaque couple d'images consécutives une image de différence, et à définir un seuil qui permet de discriminer les parcelles coupées des autres parcelles. Cette méthode a également montré des limitations importantes liées à la définition du seuil de changement, et à la grande sensibilité à la nébulosité ; en effet, avec ce type de méthode (*i.e.* bi-temporelle), les zones couvertes par les nuages dans les différentes images sont accumulées et seule l'intersection des parties dégagées dans les deux images est exploitable.

2.2.2.2. Besoin méthodologique

Afin de mieux répondre à l'application de suivi des coupes de canne à sucre à partir des séries temporelles d'images satellites, il est nécessaire d'élaborer et de développer une nouvelle méthode qui :

- intègre les connaissances expertes sur le système de culture de la canne à sucre dans le processus du traitement. Cette intégration permet d'automatiser la détection des coupes, et de réduire l'intervention humaine et le temps de traitement ;
- soit robuste au manque de données lié aux conditions atmosphériques ou à la fréquence d'acquisition. Pour cela le recours à des données provenant d'autres sources indépendantes des images est nécessaire ;
- exploite l'information temporelle dans la série afin d'améliorer les performances ;

- fournisse des cartes de coupes associées à des facteurs de confiance pour assurer la crédibilité des produits auprès de la filière.

2.3. Les données satellite et agronomiques

2.3.1. Séries temporelles d'images SPOT

Les données satellite utilisées dans cette thèse sont constituées d'une quarantaine d'images SPOT provenant de la base de données (BD) Kalideos ISLE-Réunion^{®8}. Cette BD, mise en place par le CNES⁹ (DeBoissezon and Sand, 2006; CNES, 2007), regroupe des images acquises par les satellites SPOT-1, 2, 3, 4 et 5 sur l'île de la Réunion entre 1986 et 2008. Son objectif est de promouvoir auprès de scientifiques de divers domaines, l'utilisation de la télédétection, tout en leur offrant la possibilité d'accéder, de façon aisée, à des données spatiales adaptées à leurs problématiques.

Les séries temporelles d'images que nous avons choisies à partir de la BD Kalideos ISLE-Réunion[®] sont acquises par les satellites SPOT-4 et SPOT-5. Une dizaine d'images environ dans chacune des années 2003, 2004, 2006 et 2007 ont été sélectionnées¹⁰ (Figure 5). Les caractéristiques de ces images sont recensées dans le Tableau 2 et les *quick-looks* (*i.e.* imagettes) figurent dans l'Annexe 3.

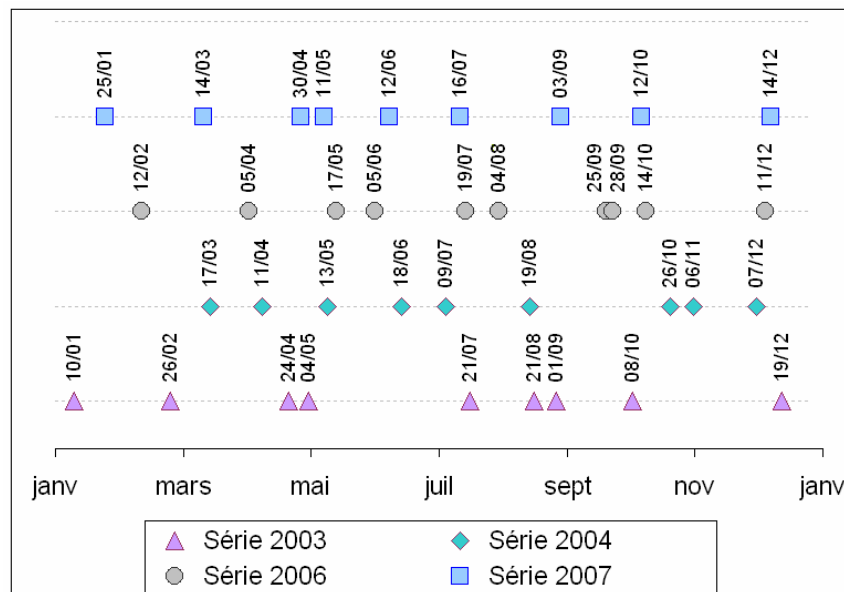


Figure 5. Dates d'acquisition des images SPOT dans les séries temporelles de 2003, 2004, 2006 et 2007.

⁸ Information Satellitaire pour L'Environnement - Réunion.

⁹ Centre National d'Etudes Spatiales.

¹⁰ Le choix des années a été fait en fonction de la disponibilité des données terrain sur les sites d'étude présentés dans la section 2.3.2.

Date	Satellite	Capteur	Mode	Angle d'incidence ¹¹ (en °)	Élévation solaire (en °)	Angle de phase ¹² (en °)
10/01/2003	SPOT 5	HRG 2	J	-04,65	64,10	21,28
26/02/2003	SPOT 5	HRG 1	J	-11,94	58,54	22,07
24/04/2003	SPOT 5	HRG 1	J	-04,38	48,02	40,23
04/05/2003	SPOT 5	HRG 1	J	10,90	46,80	47,99
21/07/2003	SPOT 5	HRG 1	J	10,58	41,20	53,13
21/08/2003	SPOT 5	HRG 1	J	18,17	48,90	51,00
01/09/2003	SPOT 5	HRG 1	J	-04,42	50,63	37,31
08/10/2003	SPOT 5	HRG 1	J	-25,95	60,40	19,75
19/12/2003	SPOT 5	HRG 1	J	-02,90	67,20	19,90
17/03/2004	SPOT 5	HRG 2	J	-19,10	54,20	25,24
11/04/2004	SPOT 5	HRG 1	J	17,96	52,45	48,42
13/05/2004	SPOT 5	HRG 1	J	-11,80	42,90	43,86
18/06/2004	SPOT 5	HRG 2	J	03,25	39,10	51,95
09/07/2004	SPOT 5	HRG 1	J	-04,73	38,83	49,70
19/08/2004	SPOT 5	HRG 1	J	17,96	48,50	51,24
26/10/2004	SPOT 5	HRG 2	J	03,30	67,90	24,94
06/11/2004	SPOT 5	HRG 1	J	-19,16	66,63	9,07
07/12/2004	SPOT 5	HRG 1	J	-12,28	66,65	11,19
12/02/2006	SPOT 5	HRG 2	J	-25,74	57,11	12,60
05/04/2006	SPOT 5	HRG 2	J	-25,87	49,58	31,00
17/05/2006	SPOT 4	HRVIR 2	XI	10,38	43,59	51,10
05/06/2006	SPOT 5	HRG 2	J	24,74	41,04	59,68
19/07/2006	SPOT 4	HRVIR 2	XI	-12,04	39,20	46,54
04/08/2006	SPOT 4	HRVIR 2	XI	-28,79	40,88	41,16
25/09/2006	SPOT 4	HRVIR 2	XI	-25,80	56,53	22,91
28/09/2006	SPOT 5	HRG 2	J	03,34	59,77	32,50
14/10/2006	SPOT 5	HRG 1	J	-11,86	62,88	19,26
11/12/2006	SPOT 4	HRVIR 2	XI	11,10	69,10	32,00
25/01/2007	SPOT 5	HRG 1	J	-24,79	65,07	4,98
14/03/2007	SPOT 5	HRG 1	J	19,15	54,21	51,72
30/04/2007	SPOT 4	HRVIR 1	XI	25,10	48,42	55,96
11/05/2007	SPOT 4	HRVIR 2	XI	02,90	44,33	46,93
12/06/2007	SPOT 4	HRVIR 1	XI	-26,20	37,35	46,64
16/07/2007	SPOT 5	HRG 1	J	10,63	39,58	54,93
03/09/2007	SPOT 4	HRVIR 1	XI	-19,60	49,54	32,09
12/10/2007	SPOT 4	HRVIR 2	XI	-25,80	62,80	13,38
14/12/2007	SPOT 5	HRG 2	J	03,29	66,01	27,28

Tableau 2. Caractéristiques des images SPOT issues de la base de données Kalideos ISLE-Réunion et utilisées dans cette thèse. (XI : mode multi-spectral de SPOT 4 ; J : mode multi-spectral de SPOT 5).

¹¹ Angle entre la normale à l'ellipsoïde référencé qui passe par le centre de la scène et la direction de visée du satellite vers le même point.

¹² Angle entre les projections sur le plan horizontal des directions (soleil, centre de la scène) et (satellite, centre de la scène).

Les instruments HRVIR 1&2 et HRG 1&2 des satellites SPOT-4 et SPOT-5 respectivement, permettent d'acquérir le rayonnement dans quatre bandes spectrales : Vert, Rouge, Proche Infra-Rouge (PIR) et Moyen Infra-Rouge (MIR). Ces quatre bandes correspondent respectivement aux intervalles de longueurs d'ondes allant de 0,50 à 0,59 μm , de 0,61 à 0,68 μm , de 0,78 à 0,89 μm et de 1,58 à 1,75 μm . Les images acquises par ces instruments sont caractérisées par une haute résolution spatiale : un pixel dans une image SPOT-4 représente un carré de taille 20 x 20 m² dans toutes les bandes spectrales, et dans une image SPOT-5, il représente un carré de 10 x 10 m² dans le Vert, Rouge et PIR, et 20 x 20 m² dans le MIR.

La plupart des produits multi-spectraux SPOT-4 et SPOT-5 disponibles dans la BD Kalideos ISLE-Réunion[©], pour les acquisitions à partir de l'année 2000, sont actuellement fournis avec trois niveaux de correction (niveau N3) : image ortho-rectifiée, image de réflectances au sommet de l'atmosphère (*Top Of Atmosphere*), et image de réflectances au sol (*Top Of Canopy*). Les phases de pré-traitement des séries temporelles d'images SPOT seront discutées plus en détail dans le chapitre suivant.

2.3.2. Sites d'étude et base de données

2.3.2.1. Exploitations de canne à sucre

Trois exploitations de canne à sucre ont été choisies pour la mise au point et l'évaluation de la méthodologie de suivi des coupes (Figure 6). Il s'agit des exploitations de M. Barau (au Nord), M. Caruel (au Nord) et M. Isautier (au Sud).

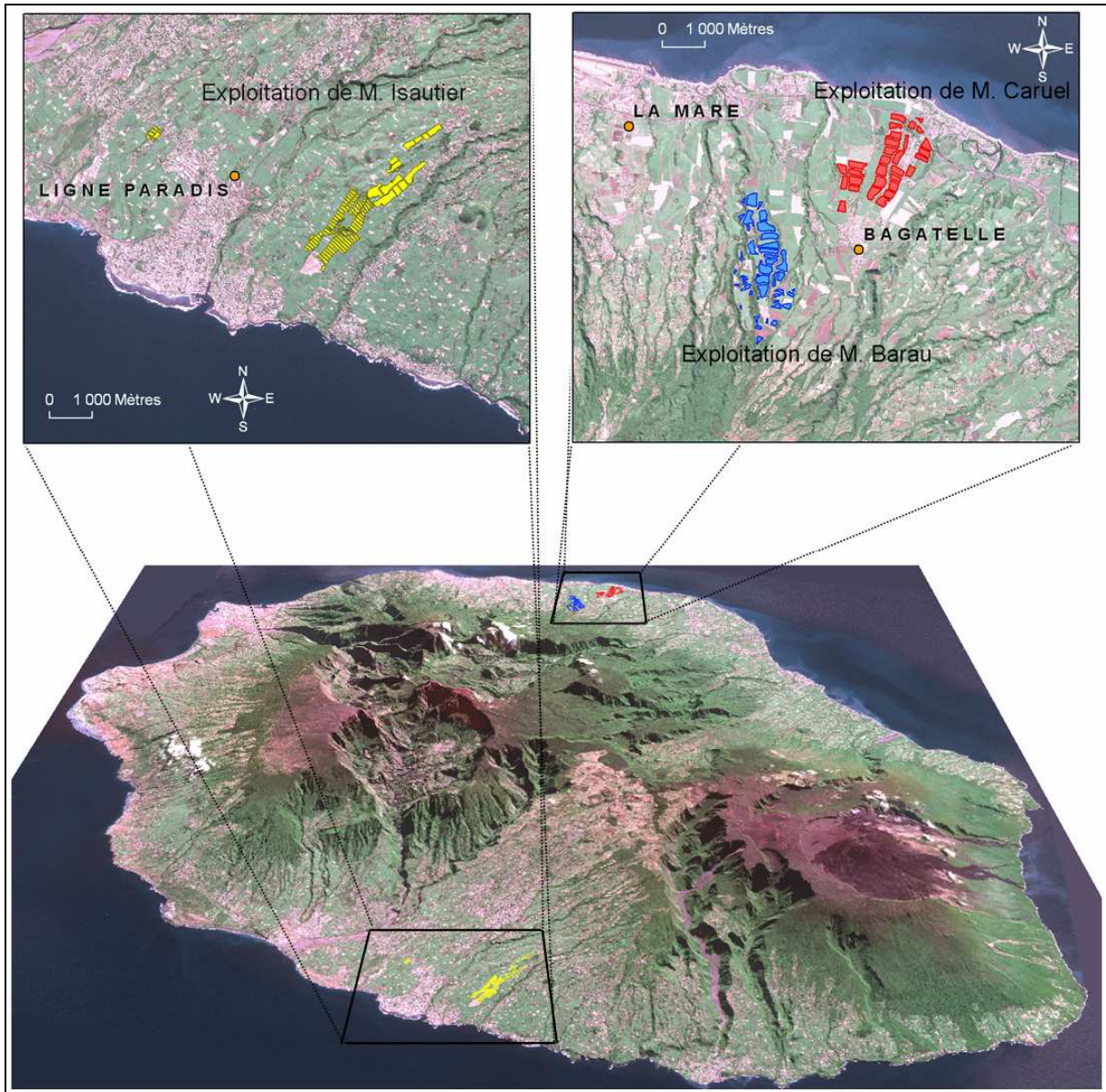


Figure 6. Localisation des exploitations de M.Barau, M. Caruel et M.Isautier sur l'île de la Réunion, et des stations météorologiques.

Critères du choix

Le choix des exploitations a été fait en fonction de plusieurs critères :

- La disponibilité de la vérité terrain (archive) : les exploitations des agriculteurs qui sont en relation avec le CIRAD ont été privilégiées afin d'utiliser les bases de données terrain, nécessaires pour la validation de la méthodologie.
- La localisation géographique en fonction des différentes zones climatiques : le climat de la Réunion est caractérisé par une grande variabilité pluviométrique entre la zone Est-Nord Est et la zone Ouest-Sud Ouest de l'île, du fait des hauts reliefs. Les précipitations tendent à se concentrer sur la zone directement exposée à l'humidité océane apportée par les vents du secteur Est. On parle ainsi de zone « au vent ». A l'inverse, la zone « sous le vent », qui bénéficie de l'abri du relief, est beaucoup plus sèche. Cette variabilité climatique Est-Ouest entraîne une variabilité dans le comportement de la canne à sucre au cours de son cycle annuel. Deux exploitations, celles de M. Barau et M. Caruel, ont été choisies dans la zone « au vent », et la troisième exploitation, celle de M. Isautier, a été choisie dans la zone « sous le vent ». La Figure 7 montre, entre autres, les différences entre les précipitations mensuelles enregistrées en 2007 dans les stations météorologiques près des exploitations de M. Caruel et de M. Isautier.
- Le système cultural (*i.e.* la prise en compte du facteur « irrigation ») : situées dans la zone humide de l'île, les exploitations de M. Barau et M. Caruel ne sont pas irriguées, tandis que celle de M. Isautier l'est en partie.

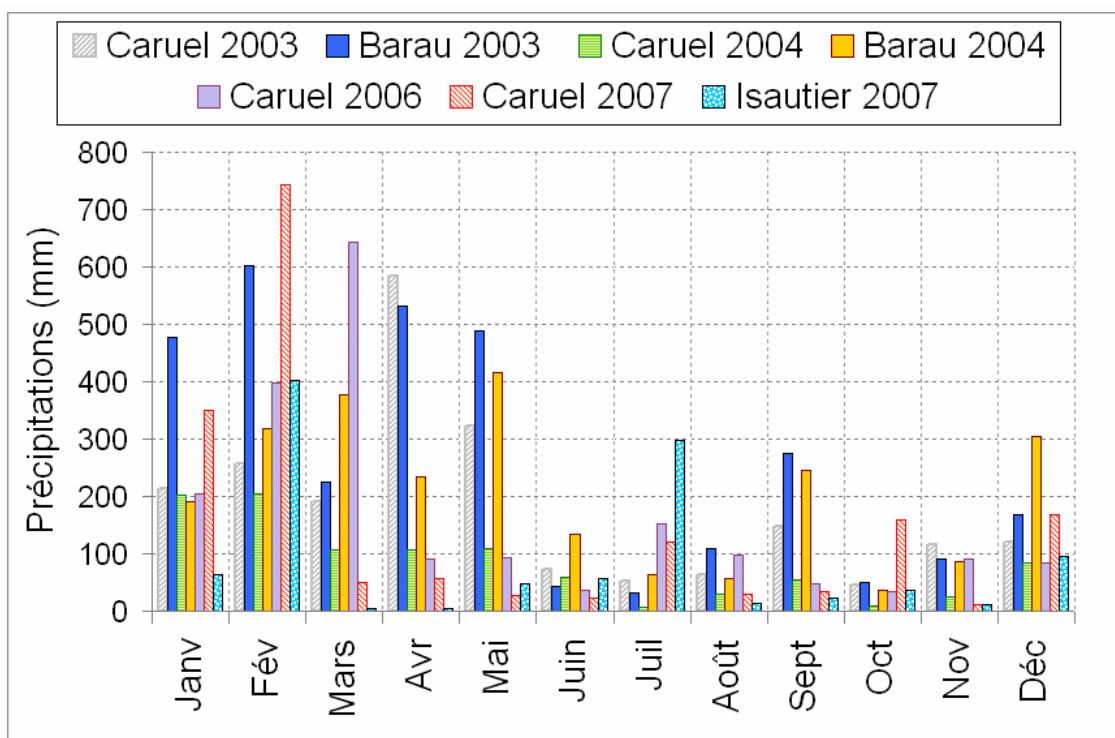


Figure 7. Précipitations mensuelles enregistrées en 2003, 2004, 2006 et 2007 dans les stations météorologiques situées à côté des exploitations de M. Caruel (Nord), M. Barau (Nord) et de M. Isautier (Sud).

Caractéristiques

L'exploitation de M. Caruel se trouve à une altitude moyenne de 107 m et comporte 30 parcelles. La surface moyenne des parcelles est de 5,4 ha et la variété principale de canne plantée sur cette exploitation est la R579.

La deuxième exploitation, celle de M. Barau, se situe un peu dans les hauteurs de l'île. Elle est à une altitude moyenne de 346 m, et elle dispose de 44 parcelles avec une surface moyenne de 3,5 ha. Les variétés principales plantées sur cette exploitation sont à surfaces égales la R570 et la R579.

Quant à l'exploitation de M. Isautier, elle est à une altitude moyenne de 259 m et est composée de 111 parcelles. La surface moyenne des parcelles est de 1,68 ha et la variété plantée est principalement la R570. Les surfaces minimum et maximum des parcelles des trois exploitations, ainsi que la gamme d'altitudes à laquelle elles sont situées sont données dans le Tableau 3.

	Nombre de parcelles	Surface min. de parcelle (ha)	Surface max. de parcelle (ha)	Altitude min (m)	Altitude max (m)
Caruel	30	0,58	8,55	44,13	204,01
Barau	44	0,30	10,30	161,11	545,39
Isautier	111	0,08	10,38	105,17	540,48

Tableau 3. Caractéristiques des trois exploitations de canne à sucre à La Réunion.

2.3.2.2. Parcelle graphique

Un parcellaire définissant les limites de chaque parcelle de canne à sucre à La Réunion était disponible. Ce parcellaire a été fourni par la DAF à l'échelle de l'îlot en utilisant les orthophotos IGN 2003 dont la résolution spatiale est de 50 cm. Il a ensuite été affiné par le CIRAD (Lebourgeois *et al.*, 2007) à l'échelle de la parcelle agronomique, en utilisant une image SPOT à 2,5 m.

2.3.2.3. Données agronomiques

Les bases de données agronomiques des trois exploitations étaient accessibles pour certaines années. Dans ces bases de données, sont renseignées les dates de coupe et de plantation ainsi que la surface de chaque parcelle. Dans certaines bases (M. Barau et M. Caruel), le rendement des parcelles est également renseigné.

Les données disponibles de M. Barau sont celles correspondant aux années 2003 et 2004. Celles de M. Caruel correspondent aux années de 2003, 2004, 2006 et 2007. Seules les données sur l'année de 2007 étaient disponibles pour l'exploitation de M. Isautier.

En utilisant les dates de récolte rapportées par les trois exploitants, nous avons construit une base de vérité terrain « coupes » durant les campagnes de récolte. Cette base de données indique, pour chaque campagne, l'état de chaque parcelle (si elle a été coupée ou non) entre deux images satellites dégagées. Nous signalons que les dates de coupe rapportées par les exploitants sont approximatives : elles correspondent aux dates situées au milieu des périodes de coupe des parcelles, périodes qui peuvent durer plusieurs semaines pour des raisons logistiques.

2.3.2.4. Données météorologiques

Des données climatiques journalières enregistrées dans les stations météorologiques les plus proches de chacune des trois exploitations de canne à sucre (Figure 6) ont été recueillies pour la période d'acquisition satellite. Ces données sont les précipitations journalières (mm), le potentiel d'évapo-transpiration (mm), le rayonnement global (J/m^2), et les valeurs minimale, maximale et moyenne de température ($^{\circ}C$). Nous présentons dans la Figure 7 les précipitations mensuelles enregistrées sur chacun des sites.

2.4. Conclusion

Les séries temporelles d'images SPOT acquises sur La Réunion et disponibles pour cette thèse, constituent un jeu de données pertinent pour le développement de méthodes de suivi des opérations culturales à partir des images à haute résolution spatiale. L'intérêt de ce jeu de données réside à la fois dans son volume (des séries temporelles sur plusieurs années), mais aussi dans le fait qu'il est acquis avec un pas de temps irrégulier, et présente une nébulosité saisonnière tout à fait caractéristique de la plupart des surfaces terrestres agricoles.

L'application agricole à laquelle nous nous intéressons, le suivi des coupes de canne à sucre, répond à l'un des principaux besoins exprimés par la filière à La Réunion et à travers le monde. En effet, en raison de la complexité du système de culture de la canne à sucre qui induit une forte variabilité spatio-temporelle des parcelles, l'utilisation des séries temporelles à haute résolution spatiale apparaît comme un outil approprié pour suivre les coupes sur de grandes surfaces. Toutefois, les méthodes développées jusqu'à présent pour suivre cette opération culturale restent limitées et souffrent de subjectivité, du manque d'automatisation, ainsi que d'une grande sensibilité à la présence des nuages et au rythme irrégulier d'acquisition d'image.

Etant donnée la disponibilité des données satellites, des données agronomiques et également des données météorologiques sur certains sites à La Réunion, nous proposons de développer un système d'aide à la détection des coupes de canne à sucre. La démarche de conception de ce système se veut générique même si l'ensemble est illustré par l'application.

Avant d'aborder la conception de ce système, nous nous intéresserons au pré-traitement des séries temporelles d'images satellites.

Chapitre 3 :

Pré-traitement des images dans les séries temporelles

Sommaire

3.1. Introduction	50
3.2. Correction géométrique	51
3.3. Correction radiométrique.....	53
3.3.1. Calcul de la réflectance au sommet de l'atmosphère	55
3.3.2. Normalisation radiométrique relative	57
3.3.2.1. Image de référence	57
3.3.2.2. Sélection automatique des cibles invariantes	57
3.3.2.3. Calculs et inversion des régressions	60
3.3.3. Correction atmosphérique.....	63
3.3.3.1. Données atmosphériques	63
3.3.3.2. Calcul de la réflectance au sol	65
3.3.4. Résultats et discussion	67
3.3.4.1. Validation de la sélection automatique des cibles invariantes.....	68
3.3.4.2. Comparaison entre normalisation relative et correction atmosphérique.....	71
3.3.4.3. Impact sur les propriétés spectrales de la canne à sucre	74
3.4. Conclusion.....	77

3.1. Introduction

L'utilisation quantitative des séries temporelles d'images satellites nécessite des phases de pré-traitement géométrique et radiométrique pour rendre les images superposables et comparables.

Au début de cette thèse, les images SPOT disponibles dans la BD Kalideos ISLE-Réunion[©] étaient au niveau 1A¹³ ; ainsi les corrections géométriques et radiométriques ont été réalisées au sein du CIRAD. Nos travaux de recherche ont porté principalement sur le développement d'une méthode automatique de normalisation radiométrique relative des images.

Au cours de la thèse, le CNES a lancé une campagne de pré-traitement des données de Kalideos ISLE-Réunion[©]. Ainsi, un deuxième jeu de données images a été fourni avec trois niveaux de pré-traitement pour chaque date : image ortho-rectifiée, image de réflectance au sommet de l'atmosphère et image de réflectance au sol. En utilisant ce nouveau jeu de données, nous avons comparé la méthode de normalisation relative que nous avons développée à celle de correction atmosphérique utilisée par le CNES pour produire les images de réflectance au sol ; le but de cette comparaison est d'étudier si la normalisation relative permet d'« éviter » la correction atmosphérique.

Dans ce chapitre, nous décrivons chacune des phases de pré-traitement des séries temporelles d'images SPOT. Nous consacrons la majeure partie à la correction radiométrique, notamment à la normalisation, et à la comparaison de la méthode de normalisation relative avec celle de correction atmosphérique.

Les principaux résultats de ce chapitre ont fait l'objet :

- d'un article « *Relative radiometric normalization and atmospheric correction of a SPOT 5 time series* » (El Hajj *et al.*, 2008c) publié dans la revue *Sensors* (Facteur d'impact = 1,573 ; 2007) ;
- de deux communications dans les colloques internationaux ISPRS (Houlès *et al.*, 2006) avec un poster et SPIE (El Hajj *et al.*, 2007b) avec une présentation orale.

Ces publications sont fournies dans l'Annexe P.1.

¹³ Les données SPOT de niveau 1A ont subi une correction radiométrique des distorsions dues aux écarts de sensibilité entre les détecteurs élémentaires de l'instrument de prise de vue.

3.2. Correction géométrique

Les images acquises par les satellites SPOT sont, comme toutes les données de télédétection, sujettes à des distorsions géométriques. Ces distorsions sont dues notamment à des erreurs de positionnement du satellite sur son orbite (mouvement du système de balayage, mouvement de la plate-forme, etc.), au fait que la Terre tourne autour de son axe durant l'enregistrement de l'image, et aux effets du relief et de la courbure de la Terre. Ces distorsions sont amplifiées par la visée oblique des capteurs ; le dépointage latéral des instruments HRVIR 1&2 et HRG 1&2 des satellites SPOT-4&5 varie entre $\pm 27^\circ$ (Figure 8). Il est donc nécessaire avant de réaliser n'importe quel processus de traitement d'images, de passer par une phase de correction géométrique pour :

- (1) compenser les distorsions géométriques affectant les images, de telle sorte que la représentation de chaque image soit aussi proche que possible de la réalité ;
- (2) rendre toutes les images dans la série superposables, en assortissant à chaque pixel sa véritable position en coordonnées au sol.

Certaines distorsions, comme les effets du relief, de rotation de la Terre et de visée oblique, sont systématiques ou prévisibles. Il est donc possible de corriger ces effets par une modélisation précise du mouvement de la plate-forme et du capteur, et par une relation géométrique entre la plate-forme et la Terre en utilisant un MNT (Modèle Numérique de Terrain).

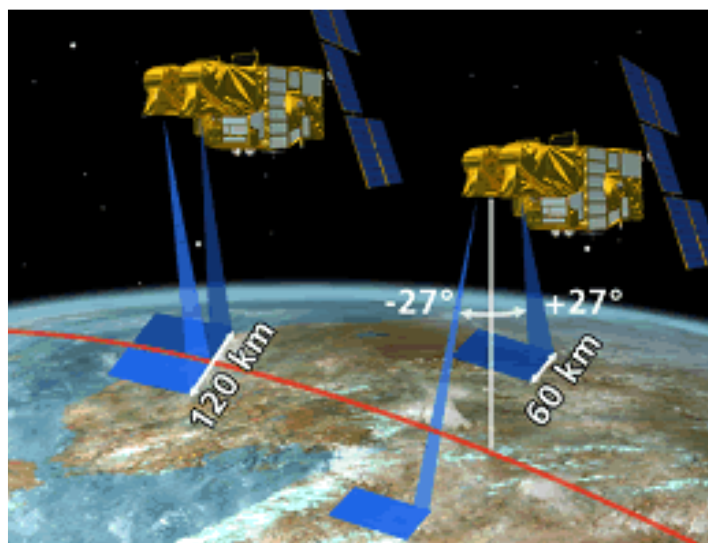


Figure 8 Champ de vue des instruments HRG 1&2 du satellite SPOT-5.

D'autres erreurs non systématiques, ne peuvent être modélisées de cette manière, et l'on doit effectuer la correspondance géométrique de l'image à un système connu de coordonnées au sol. Ceci se fait en identifiant des coordonnées de l'image (*e.g.* ligne et colonne) de plusieurs points clairement distincts, appelés points amers ou points d'appuis, sur l'image à corriger et en les assortissant à leurs véritables positions en coordonnées au sol (*e.g.* en latitude et longitude).

La correction géométrique réalisée au début de la thèse (au sein du CIRAD) sur les images SPOT dans nos séries, a été faite par ortho-rectification (correction qui prend en compte les distorsions dues au relief). Le modèle de déformation géométrique SPOT fourni par ERDAS IMAGINETM a été utilisé, en intégrant un MNT (BD ORTHO[®], IGN)¹⁴ et une trentaine de points amers répartis sur la surface de l'île de La Réunion. Les points amers ont été sélectionnés en se basant sur une image de référence, l'image SPOT-5 du 06/07/2002, première image ortho-rectifiée par le CNES disponible dans la BD Kalideos Isle-Réunion[©]. La résolution spatiale de cette image est de 2,5 mètres. Le modèle mathématique de ré-échantillonnage utilisé est celui du « plus proche voisin ».

La correction géométrique effectuée par le CNES durant la campagne de pré-traitement est basée sur une « Dalle3D » de référence, réalisée à partir d'une ortho-mosaïque et d'un MNT respectivement à 2,5 m et 25 m de résolution spatiale (Lafrance, 2006). Le système de coordonnées de référence choisi est l'UTM (zone 40 sud). L'erreur quadratique moyenne de la correction géométrique de chaque image a été inférieure à 0,5 pixel.

¹⁴ Les MNTs dans la BD ORTHO[®] ont une résolution de 25 m en x et y , et une précision de 5 m en altitude.

3.3. Correction radiométrique

La première étape dans le processus de correction radiométrique consiste à calculer pour chaque image les réflectances des pixels au sommet de l'atmosphère, *Top Of Atmosphere* (TOA). Cette étape permet d'obtenir, pour une même bande spectrale, des mesures physiques qui sont indépendantes des caractéristiques du capteur. L'étape suivante consiste à normaliser la radiométrie des images. En effet, les images dans les séries temporelles sont souvent acquises par différents capteurs et sous différentes conditions atmosphériques et géométriques (*e.g.* illumination solaire, angles de prise de vue). La normalisation radiométrique de ces images est donc nécessaire pour enlever les distorsions radiométriques et pour les rendre comparables. Les effets d'artefacts, de directionnalité de surface et de l'atmosphère peuvent être corrigés d'une manière absolue ou relative.

Il existe plusieurs algorithmes opérationnels permettant de corriger les perturbations atmosphériques tels que Modtran2 (Anderson *et al.*, 1993), 5S (Tanre, 1990), SMAC (basé sur 5S) (Rahman and Dedieu, 1994) et 6S (Vermote *et al.*, 1997). L'objectif principal de ces codes est d'estimer la réflectance au sol, *Top Of Canopy* (TOC), à partir de la réflectance au sommet de l'atmosphère (TOA), cette dernière étant dérivée du rayonnement mesuré par le capteur. Pour se faire, des informations sur la sensibilité spectrale du capteur et sur les propriétés atmosphériques (teneur en gaz et en aérosols), au moment de l'acquisition de l'image, sont nécessaires pour estimer les effets atmosphériques de dispersion et d'absorption.

D'autres méthodes basées sur la soustraction d'objets sombres, *Dark Object Substraction* (DOS), ont également été développées (Teillet and Fedosejevs, 1995; Chavez Jr, 1996; Song *et al.*, 2001); ces méthodes permettent de s'affranchir de mesures atmosphériques mais exigent des codes de transfert radiatif pour faire la correction radiométrique absolue.

Afin d'éviter le besoin à la fois de codes de transfert radiatif et des propriétés optiques de l'atmosphère (qui sont difficiles à acquérir en particulier pour des données historiques), beaucoup d'investigateurs ont eu recours à la normalisation radiométrique relative. Les méthodes proposées supposent toutes que la relation entre les rayonnements TOA d'une région de réflectance constante et homogène dans l'espace, enregistrés à deux instants différents, peuvent être approchées par une fonction linéaire. Le procédé de normalisation d'images peut alors être réduit à un calcul de régression linéaire pour chaque bande spectrale

(Schott *et al.*, 1988; Hall *et al.*, 1991; Salvaggio, 1993; Séguis and Puech, 1997; Furby and Campbell, 2001; Du *et al.*, 2002).

La difficulté principale des méthodes de normalisation relative réside dans la détermination des cibles dans la scène dont les réflectances sont presque constantes au cours du temps (*i.e.* cibles invariantes). Il est efficace de choisir manuellement ces cibles invariantes comme présenté par (Schott *et al.*, 1988) et (Salvaggio, 1993), mais cette approche est coûteuse en temps et pourrait aboutir à une normalisation radiométrique subjective. (Elvidge *et al.*, 1995) ont développé une méthode qui sélectionne automatiquement les pixels invariants, en utilisant des diagrammes bidimensionnels des réflectances dans le Proche Infrarouge des images à normaliser. Cette méthode est efficace (Yuan and Elvidge, 1998), mais elle est seulement applicable aux images acquises sous des géométries semblables d'éclairage solaire, et pour des conditions phénologiques similaires. Une autre méthode pour déterminer automatiquement les pixels invariants a été présentée par (Nielsen *et al.*, 1998). Cette méthode, *Multivariate Alteration Detection* (MAD), effectue des analyses traditionnelles de corrélations canoniques, pour trouver des combinaisons linéaires entre deux groupes de variables (*e.g.* les bandes spectrales des images à normaliser). Les principaux inconvénients de cette méthode sont l'aspect bruyant des variables aléatoires de sortie (*i.e.* les MAD), le temps de calcul, et le besoin de ressources informatiques importantes lors de son application sur des images acquises à haute résolution spatiale (*e.g.* images SPOT). Des extensions plus récentes de cette méthode ont été développées pour améliorer ses performances, mais le temps de calcul et le problème de consommation de ressources demeurent (Nielsen and Canty, 2005; Nielsen, 2007).

D'après ce qui précède, on remarque qu'il est nécessaire de développer et d'évaluer une nouvelle méthode de normalisation radiométrique, qui soit adaptée au pré-traitement des séries temporelles acquises à haute résolution spatiale, et en particulier aux images SPOT. Cette méthode devrait être autonome, rapide et objective ; elle devrait également pouvoir pré-traiter des images multi-temporelles acquises sous différentes conditions atmosphériques et géométriques, et à différentes saisons (*e.g.* pour différentes conditions phénologiques). Ce besoin représente l'une des principales préoccupations de cette thèse.

Dans les sous-sections qui suivent, nous décrivons tout d'abord comment ce fait le calcul des réflectances TOA dans les images SPOT. Ensuite nous proposons une nouvelle méthode pour la normalisation radiométrique relative des séries temporelles. Cette méthode est basée

sur des régressions linéaires, dérivées des réflectances de cibles invariantes sélectionnées automatiquement.

Nous présentons également la méthode de correction atmosphérique, utilisée par le CNES, qui emploie le code 6S (Second Simulation of the Satellite Signal in the Solar Spectrum) (Vermote *et al.*, 1997). Les résultats de cette méthode (*i.e.* les images de réflectances au sol) nous serviront de référence pour la validation de notre méthode de normalisation.

Le jeu de données satellite qui a servi au développement et à la validation de notre méthode de normalisation est constitué des séries temporelles acquises en 2003 et 2004 (*c.f.* Section 2.3.1).

3.3.1. Calcul de la réflectance au sommet de l'atmosphère

Cette étape de correction radiométrique consiste à calculer pour chaque image les réflectances au sommet de l'atmosphère, *i.e.* réflectance TOA. En effet, les images SPOT sont fournies en comptes numériques qui sont simplement égalisés (corrigés pour le comportement individuel de chaque détecteur de pixel) et sans étalement dynamique. Cette phase de correction prend en considération :

- a) les paramètres de calibration à la date d'acquisition ;
- b) l'angle zénithal solaire ;
- c) l'éclairement solaire exo-atmosphérique.

Afin d'obtenir des mesures physiques qui sont indépendantes des caractéristiques du capteur, les comptes numériques doivent être tout d'abord convertis en luminance. Dans une bande spectrale k , la luminance d'un pixel L_{TOA}^k au niveau TOA est liée au compte numérique X^k par la relation suivante :

$$L_{TOA}^k = \frac{X^k}{A_k \cdot G_m^k} \quad (1)$$

avec :

- A_k est le coefficient absolu de calibration pour la bande k , estimé pour la date de l'acquisition d'image. Ce coefficient a été fourni par le CNES (Meygret, 2007) pour chaque image ; il tient compte de la dégradation du capteur avec le temps.

- G_m^k est le gain analogique de l'amplificateur à bord pour la bande k (Meygret, 2007).

La luminance L_{TOA}^k de chaque pixel dans chaque bande doit ensuite être normalisée par le flux solaire incident au sommet de l'atmosphère, afin d'obtenir une réflectance de surface au niveau TOA, ρ_{TOA}^k . Ceci est calculé par :

$$\rho_{TOA}^k = \frac{\pi \cdot L_{TOA}^k}{E_S^k \cdot \cos \theta_S \cdot (d_0/d)^2} \quad (2)$$

avec :

- E_S^k est l'éclairement solaire moyen exo-atmosphérique pour chaque bande k. Il peut être récupéré à partir des méta-données fournies avec l'image ou à partir du code 6S de transfert radiatif (Vermote *et al.*, 1997).
- θ_S est l'angle zénithal solaire provenant des données auxiliaires.
- d_0/d est le rapport entre la distance Terre-Soleil à la date d'acquisition et la distance moyenne Terre-Soleil (par an) (Vermote *et al.*, 1997).

Enfin, pour prendre en considération les effets de la pente de surface, on remplace, en utilisant un MNT, le $\cos(\theta_S)$ dans l'équation 2 par β_S :

$$\beta_S = \cos(\theta_S) \cdot \cos(\theta_n) + \sin(\theta_S) \cdot \sin(\theta_n) \cdot \cos(\varphi_S - \varphi_n) \quad (3)$$

avec θ_n est l'angle zénithal de surface (la pente), φ_S est l'angle azimutal solaire et φ_n l'angle azimutal de surface (orientation). La pente et l'orientation sont extraites du MNT raster.

Ce passage du compte numérique à la réflectance TOA a été réalisé pour toutes les images SPOT dans notre série temporelle.

3.3.2. Normalisation radiométrique relative

Nous avons développé une méthode simplifiée pour la normalisation radiométrique relative des images satellites acquises à haute résolution spatiale. Cette méthode n'exige pas de données atmosphériques, et tend à minimiser uniformément les effets de changement des conditions atmosphériques et solaires, par rapport à une image de référence. Le processus est basé sur le calcul de régressions linéaires qui relient les bandes spectrales des images dans la série à celles d'une image de référence. Trois étapes principales sont identifiées : le choix de l'image de référence, la sélection des cibles invariantes, et le calcul et l'inversion des régressions linéaires.

3.3.2.1. Image de référence

Parmi l'ensemble des images de la série temporelle, une image de référence doit être choisie. C'est par rapport à cette image que toutes les autres seront normalisées. Il est donc important qu'elle réponde à plusieurs critères :

- Une couverture nuageuse minimale ;
- Un faible pourcentage de pixels saturés ;
- Une large dynamique de radiométrie ;
- Une date d'acquisition proche du milieu de la série.

Ces critères permettent de maximiser le nombre et la diversité de cibles invariantes détectables. L'image de référence que nous avons choisie pour normaliser les séries de 2003 et 2004 est celle acquise le 13 mai 2004.

3.3.2.2. Sélection automatique des cibles invariantes

Afin de rendre le procédé de sélection des cibles invariantes (CI) objectif et pour obtenir un nombre suffisant de CI couvrant une large gamme spectrale, nous avons développé une technique de sélection automatique. L'organigramme de cette technique est représenté dans la Figure 9. Pour chaque image TOA dans la série temporelle, nous avons d'abord calculé une image multi-bande de différence (IMD), par une soustraction pixel à pixel de l'image de référence. Ensuite, nous avons appliqué à chaque IMD un ensemble de masques afin d'éliminer le maximum de pixels qui changent :

- Le premier est un masque de nuage créé pour chaque date à partir de l'IMD (voir l'Annexe 4).
- Le deuxième est un masque de saturation fourni avec chaque image, il permet de repérer dans chaque bande spectrale les pixels affectés par une saturation radiométrique et de les ignorer.
- Le troisième masque sert à éliminer un nombre maximum de pixels liés aux zones de végétation qui varient au cours du temps. Ce masque doit être créé en principe à partir d'une carte d'occupation des terres agricoles ; en raison de l'absence de cette donnée dans notre application, nous avons utilisé le parcellaire de canne à sucre comme masque. Ceci nous a permis d'éliminer environ 60 % de pixels liées aux terres agricoles.
- Le quatrième permet de masquer l'océan afin de ne garder que les pixels terrestres.

Nous avons ensuite choisi, pour chaque date, les CI en se basant sur les histogrammes des pixels restant dans les quatre bandes spectrales de l'IMD.

La forme de l'histogramme de chaque bande après l'application des masques dépend des types de changements qui se produisent entre la date d'acquisition de l'image à normaliser et celle de l'image de référence. En effet, chaque bande spectrale est sensible à différentes sortes de changement ; ainsi un changement d'occupation du sol pourrait engendrer une modification significative des valeurs des pixels dans une certaine bande, mais pas dans les autres. Dans toutes les bandes spectrales de l'IMD, les pixels avec des changements relativement faibles sont groupés autour des modes des histogrammes. Ceci signifie que « la majorité » des pixels restant après l'application des masques, est considérée comme des cibles sans aucun changement (ou presque). Les valeurs de différence correspondant à ces pixels sont différentes de zéro en raison du changement des conditions d'acquisition d'image d'une date à une autre. Le reste de l'histogramme correspond aux pixels avec de vrais changements. Ces pixels peuvent être affectés par les différences des conditions d'acquisition d'image, mais leur effet comparé au vrai changement est négligeable. La Figure 10 montre, à titre d'exemple, l'évolution de l'histogramme de l'IMD calculée entre les deux images SPOT acquises le 26 février 2003 et le 13 mai 2004, dans le Vert.

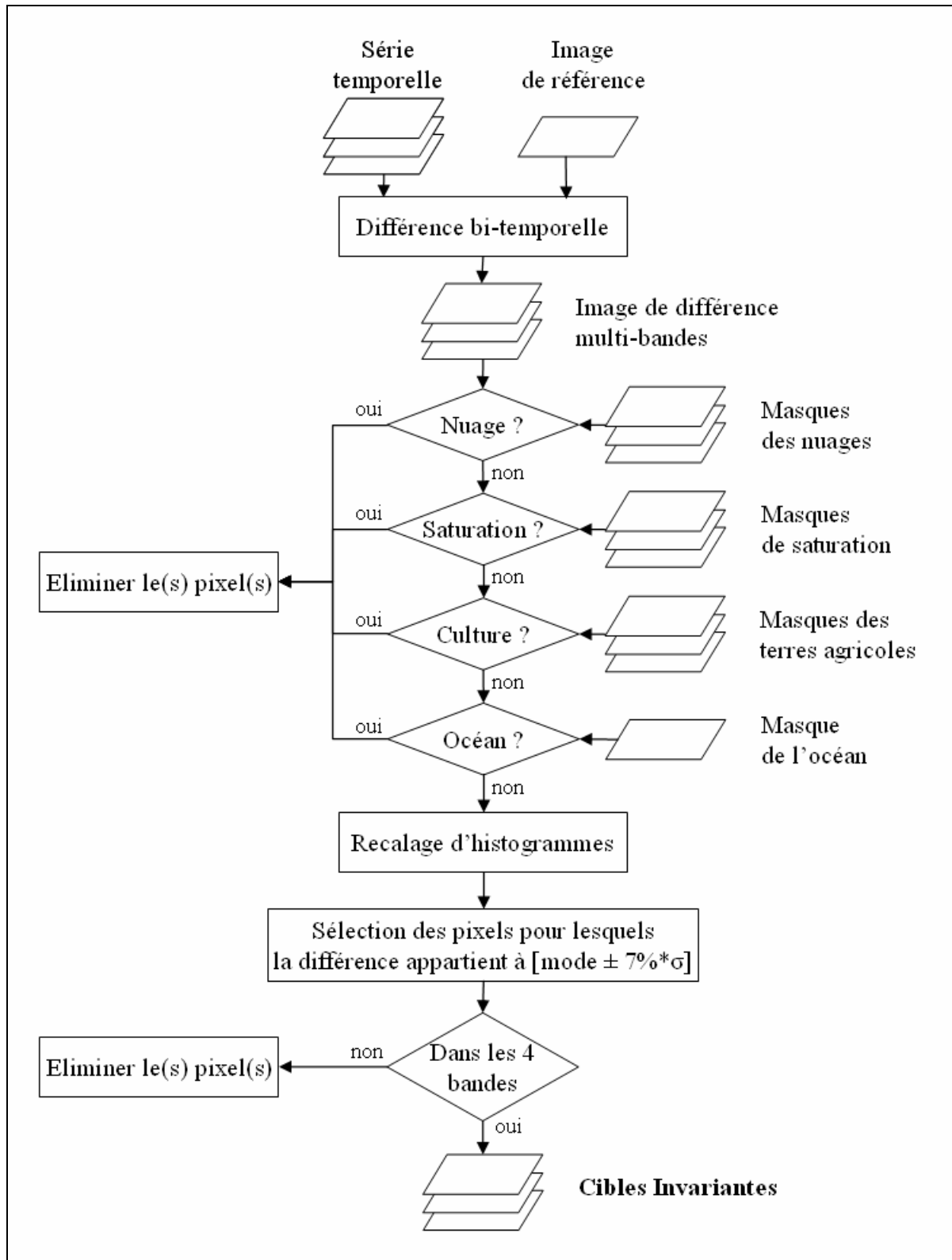


Figure 9. Organigramme de la technique de sélection automatique des cibles invariantes.

Pour chaque date, les histogrammes dans les quatre bandes de l'IMD sont décalés, de telle sorte que les valeurs de différence assignées aux majorités (modes) soient ramenées à zéro. Finalement, les pixels dans chaque IMD avec des valeurs proches de zéro simultanément dans les quatre bandes sont considérés comme invariants. Le terme « proche » a été traduit par 7% de l'écart type autour du mode de chaque histogramme. Ce pourcentage a été choisi après plusieurs essais sur une gamme allant de 2% à 20% ; il permet la sélection d'un nombre optimal de CI.

La validation de cette technique de sélection automatique des CI est représentée dans la section 3.3.4.1. Nous montrons, à titre d'exemple, dans la Figure 11 les CI sélectionnées automatiquement sur l'ensemble des images à normaliser. Nous signalons que les CI sélectionnées sur toutes les dates de la série représentent une moyenne de 0,044 % des pixels de la surface de l'île (10928 pixels), et incluent une large diversité de type d'occupation du sol.

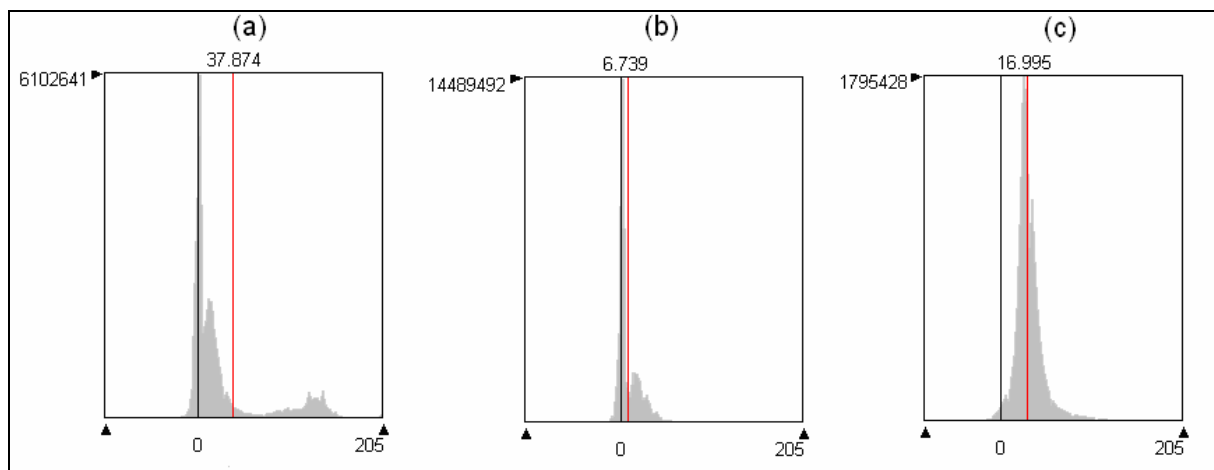


Figure 10. Histogramme et valeur moyenne obtenus pour le canal Vert de l'image de différence calculée à partir des images acquises le 26 février 2003 et le 13 mai 2004 (référence) : a) avant l'application des masques, b) après l'application du masque des nuages, c) après l'application de tous les masques (nuages, saturation, parcelles de canne, océan).

3.3.2.3. Calculs et inversion des régressions

Pour les cibles invariantes sélectionnées, nous avons extrait à partir des images TOA les valeurs de réflectance dans les quatre bandes spectrales. En utilisant ces valeurs, nous avons établi, pour chaque date, des régressions linéaires dans chaque bande de la forme :

$$y = ax + b \quad (4)$$

où y est l'image de référence et x les autres images alternativement. Dans la Figure 12, nous montrons un exemple de régression linéaire calculée dans le Vert pour normaliser l'image du 06 novembre 2004 par rapport à l'image du 13 mai 2004. Il est à signaler que ces images sont acquises dans des conditions phénologiques différentes, comme le montre les couleurs différentes des parcelles de canne à sucre autour du littoral.

Les régressions linéaires calculées pour toutes les dates ont ensuite été appliquées aux images en réflectance TOA, pour construire une nouvelle série d'images normalisées entre elles.

Présentons maintenant la méthode de correction atmosphérique développée par le CNES pour produire les images TOC (*i.e.* images de réflectances au sol).

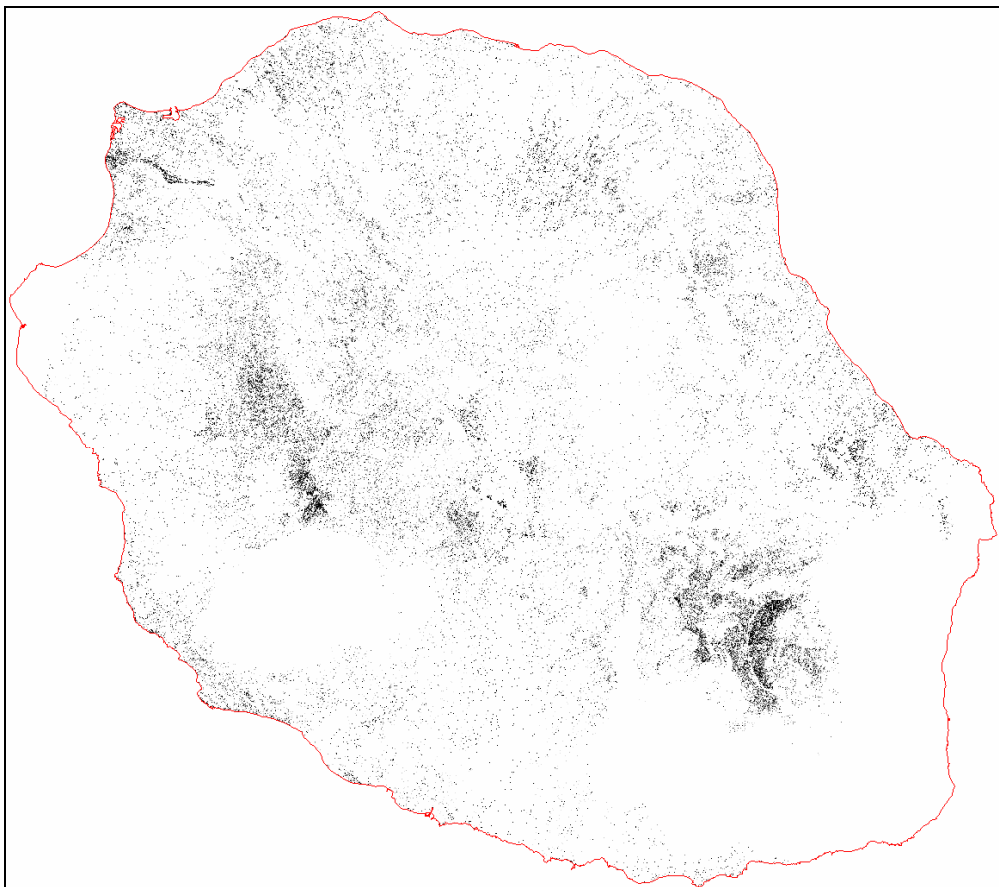


Figure 11. Cibles invariantes sélectionnées automatiquement sur l'ensemble des images : chaque point noir correspond à une cible.

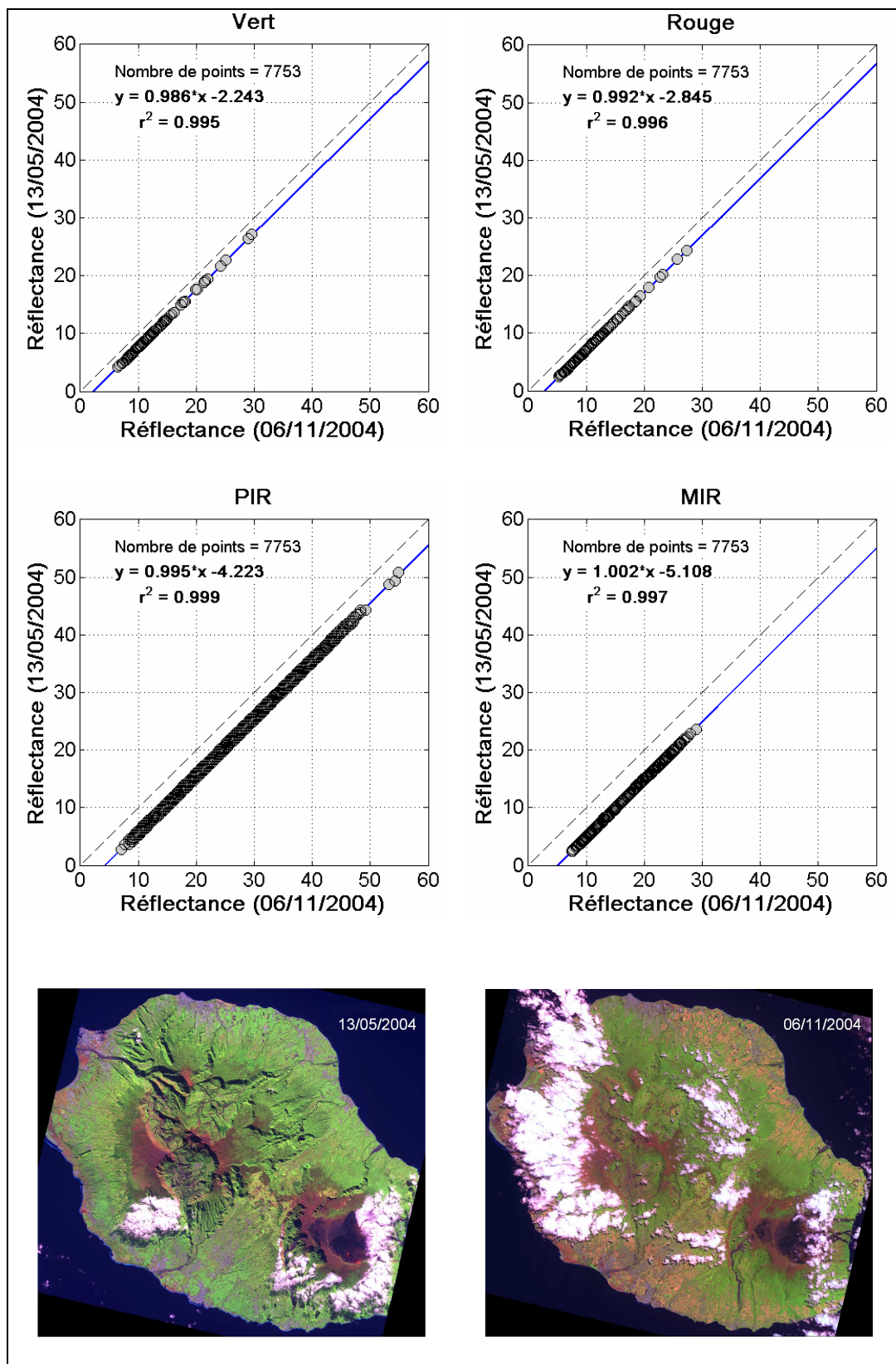


Figure 12. Régressions linéaires calculées automatiquement dans les quatre bandes spectrales (Vert, Rouge, PIR et MIR) pour normaliser l'image du 06 novembre 2004 par rapport à celle du 13 mai 2004.

3.3.3. Correction atmosphérique

Afin de calculer pour chaque image les réflectances TOC (réflectances au sol) des pixels dans chaque bande spectrale, des simulations basées sur le code 6S de correction atmosphérique ont été utilisées (Vermote *et al.*, 1997). Ce code simule le signal satellite de 0,25 à 4,0 μm , en supposant une atmosphère sans nuages. La correction atmosphérique par 6S nécessite des informations sur les paramètres atmosphériques à la date d'acquisition d'image : a) la pression atmosphérique, b) les quantités de gaz (vapeur d'eau, dioxyde de carbone, oxygène et ozone) pour la transmission atmosphérique due à l'absorption des molécules et c) les caractéristiques d'aérosols (type et concentration).

3.3.3.1. Données atmosphériques

Quantité de gaz et pression atmosphérique

Les bandes spectrales des images SPOT sont différemment « affectées » par l'absorption et la dispersion atmosphériques des gaz (Figure 13). L'absorption stratosphérique de l'ozone, a un maximum autour de 0,6 μm et s'étend approximativement entre 0,4 et 0,8 μm ; elle affecte ainsi les bandes B1 et B2 (*i.e.* Vert et Rouge). L'oxygène présente une absorption forte à 0,69 μm et à 0,763 μm , et affecte la bande B2 et légèrement la bande B3 (*i.e.* Proche Infrarouge, PIR). La vapeur d'eau présente un spectre d'absorption au delà de 0,6 μm , avec des lignes d'absorption plus ou moins importantes qui affectent principalement la bande B3 et un peu moins la bande B2. La sensibilité spectrale de la bande B4 (*i.e.* Moyen Infrarouge, MIR) aux absorptions gazeuses (gaz carbonique, vapeur d'eau ou méthane) est très faible.

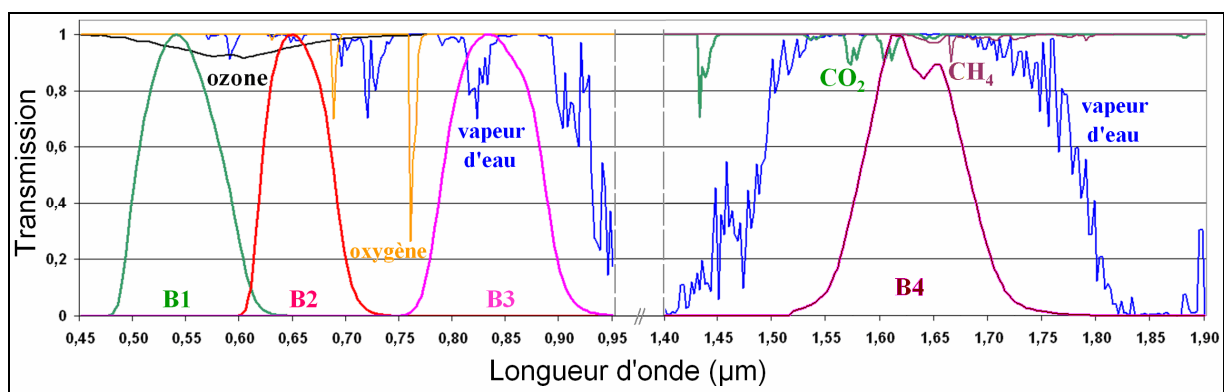


Figure 13. Les bandes spectrales des satellites SPOT-4&5 (B1 = Vert ; B2 = Rouge ; B3 = PIR ; B4 = MIR) et les transmissions des gaz entre 0,45 μm et 1,9 μm (source : le modèle atmosphérique US62 avec une teneur en vapeur d'eau de 1,424 $\text{g}\cdot\text{cm}^{-2}$ et une quantité d'ozone de 0,344 $\text{cm}\cdot\text{atm}$).

Pour les gaz stables, tel que l'oxygène, le gaz carbonique et le méthane, les simulations utilisent un profil vertical standard. Ce profil est obtenu par le model US62 inclus dans le code 6S. Pour les gaz avec plus de fluctuations, les valeurs quotidiennes sont nécessaires. La quantité de vapeur d'eau, intégrée sur la colonne atmosphérique, est obtenue à partir de l'information météorologique. Les modèles de NCEP¹⁵ (*National Centers for Environmental Prediction*) ont été utilisés pour déterminer ce paramètre. La quantité d'ozone est accessible à partir des données des satellites tels que TOMS et TOAST.

Le rayonnement moléculaire, résultant de la dispersion de la lumière par les molécules, règne pour les bandes à courtes longueurs d'onde (B1 et B2). Ce rayonnement dépend de l'épaisseur optique moléculaire. Cette dernière est bien connue pour la pression atmosphérique standard (1013 mb) et doit être équilibrée par la vraie pression correspondant à la date d'acquisition de l'image. La pression atmosphérique est également accessible à partir des données NCEP.

Les aérosols

Les aérosols troposphériques sont les composants atmosphériques les plus difficiles à caractériser. Le code de correction atmosphérique 6S nécessite en entrée l'épaisseur optique d'aérosols à 550 nm, et une description du type de particules afin de calculer leurs propriétés spectrales.

Étant donné que l'île de La Réunion est située dans l'Océan Indien, elle est fortement influencée par des conditions océaniques. Les informations requises pour la caractérisation des aérosols sont fournies à la fois par les produits SeaWifs (L3 Standard Mapped Image), acquis sur l'océan entourant l'île, et par les données météorologiques issues des modèles NCEP. Les produits SeaWIFs donnent l'épaisseur optique des aérosols pour une longueur d'onde de 865 nm, alors que les données de NCEP fournissent l'humidité relative de l'air. En supposant un modèle maritime pour les aérosols, le modèle de Shettle and Fenn (Shettle and Fenn, 1979) a été utilisé pour calculer l'épaisseur optique à 550 nm, à partir de l'humidité relative de l'air et de l'épaisseur optique à 865 nm.

¹⁵ www.ncep.noaa.gov.

Correction de l'effet « altitude »

En raison de son origine volcanique, la topographie de l'île de La Réunion est fortement hétérogène avec des fortes pentes et une altitude excédant 3 000 m au dessus du niveau de la mer, sur une surface de 2 500 km² seulement. Etant donné que l'épaisseur optique de l'atmosphère dépend du relief, l'altitude du sol doit être prise en compte dans le processus de correction atmosphérique.

Comme les paramètres atmosphériques sont indiqués pour une altitude donnée z_{ref} , leurs valeurs V doivent être ajustées aux altitudes simulées z_i . Les concentrations des composants atmosphériques sont ajustées en utilisant une longueur d'échelle H appliquée à leurs profils verticaux :

$$V(z_i) = V(z_{ref}) \times \exp\left(-\frac{z_i - z_{ref}}{H}\right) \quad (5)$$

La correction d'altitude $V(z_i)$ (équation 5) est appliquée à la pression atmosphérique, à l'épaisseur optique d'aérosol et à la teneur en vapeur d'eau intégrée sur la colonne verticale. La pression atmosphérique est corrélée avec la densité moléculaire (longueur d'échelle de 8 km). La vapeur d'eau et les aérosols sont la plupart du temps concentrés dans la troposphère basse. Bien que leurs profils soient fortement variables, on peut les approcher en assumant une diminution exponentielle de concentration avec une longueur d'échelle de 2 km.

3.3.3.2. Calcul de la réflectance au sol

Le principe de calcul des réflectances TOC est basé sur des *Look-Up Tables* (LUTs) qui regroupent des couples de réflectances TOC et de réflectances TOA simulées. Le code 6S a été adapté pour permettre des boucles internes sur les différentes valeurs de réflectance TOC, et ceci pour calculer facilement une LUT par bande spectrale. Les valeurs de réflectance TOC par bande varie entre zéro, pour des objets noirs, et 0,8, pour des objets clairs ou lumineux, avec un pas d'incrémentations prédéfini de 0,01. À partir de la réflectance TOA du pixel, une interpolation linéaire dans les LUTs a été effectuée pour obtenir la réflectance TOC correspondante.

Afin de tenir compte de l'altitude, un MNT qui donne la vraie altitude z de chaque pixel a été utilisé. A partir des LUTs simulés par 6S pour les altitudes z_i et z_{i+1} autour de z , de nouvelles LUTs (réflectance TOC / réflectance TOA simulée) ont été calculées pour l'altitude

z par interpolation linéaire. Une fois que ces nouvelles LUTs sont disponibles pour le pixel à l'altitude z , elles sont utilisées pour inverser la réflectance TOC correspondant à la réflectance TOA de ce pixel.

La Figure 14 récapitule les différentes phases de correction radiométrique effectuées sur une image SPOT pour passer des comptes numériques aux réflectances TOC.

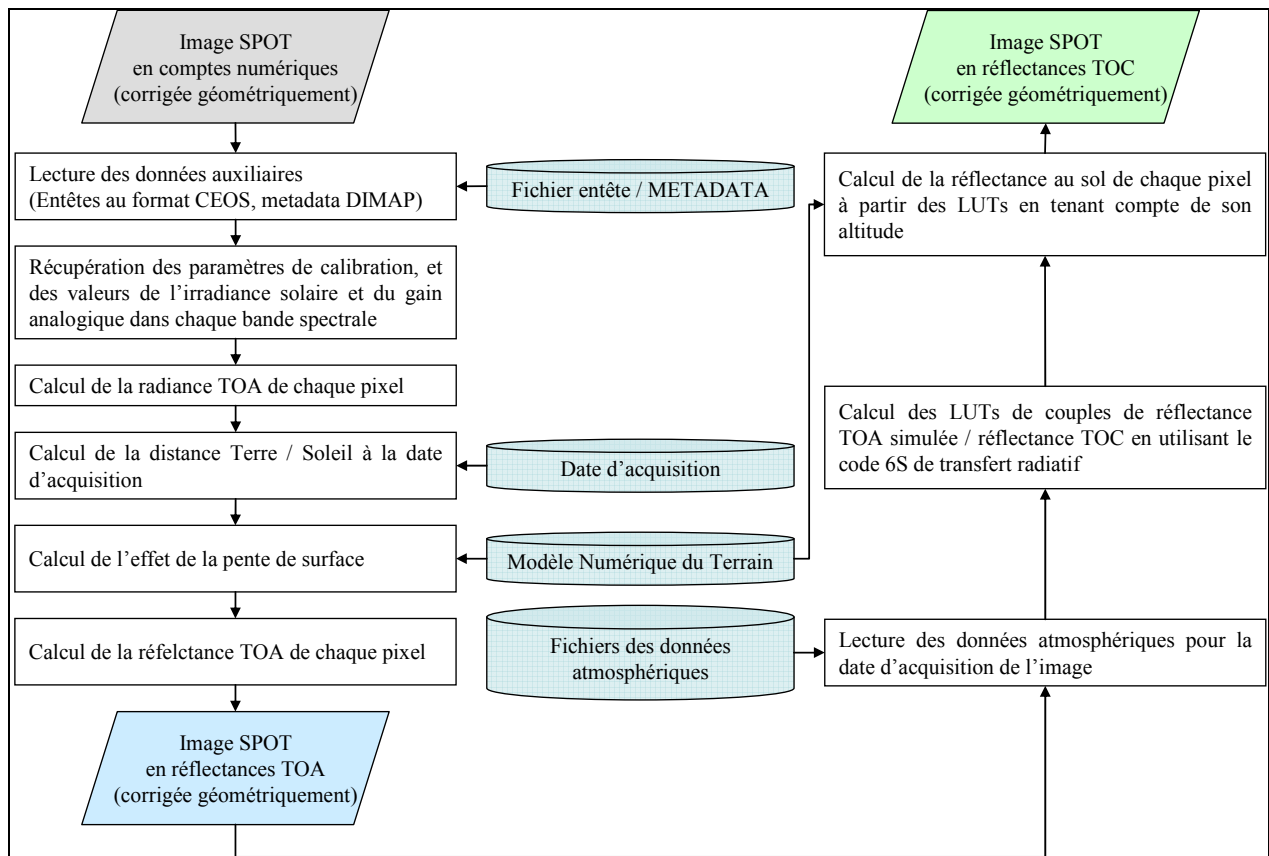


Figure 14. Illustration des phases de correction radiométrique d'une image SPOT pour passer du compte numérique à la réflectance au sol.

3.3.4. Résultats et discussion

Dans cette section, nous nous intéressons dans un premier temps à la validation de la technique de sélection automatique des cibles invariantes introduite dans la section 3.3.2.2. Dans un deuxième temps, nous comparons la méthode de normalisation relative, que nous avons développée avec la méthode de correction atmosphérique utilisée par le CNES pour produire les images TOC. Ces deux tâches sont réalisées en s'appuyant sur les 18 images SPOT des séries temporelles acquises en 2003 et 2004 ; les caractéristiques d'acquisition de ces images sont présentées dans le Tableau 2 du chapitre précédant ; les paramètres atmosphériques des dates d'acquisition sont dans le Tableau 4. On peut remarquer d'après ces deux tableaux, que les conditions géométriques de prise de vue et les caractéristiques atmosphériques varient considérablement d'une image à l'autre, ce qui affirme le besoin de normalisation radiométrique.

Dates	Pression (mbar)	H ₂ O (g.cm ⁻²)	Ozone (cm.atm)	τ_{550}
10/01/2003	1014	2,783	0,264	0,538
26/02/2003	1013	5,469	0,259	0,322
24/04/2003	1017	4,24	0,253	0,246
04/05/2003	1015	2,649	0,252	0,262
21/07/2003	1022	2,332	0,263	0,112
21/08/2003	1024	2,151	0,272	0,273
01/09/2003	1026	1,954	0,276	0,277
08/10/2003	1018	2,671	0,297	0,432
19/12/2003	1017	3,093	0,272	0,357
17/03/2004	1014	2,761	0,255	0,176
11/04/2004	1014	4,793	0,252	0,26
13/05/2004	1018	1,986	0,246	0,22
18/06/2004	1024	2,357	0,237	0,11
09/07/2004	1020	2,674	0,242	0,221
19/08/2004	1027	2,126	0,254	0,197
26/10/2004	1018	2,946	0,275	0,329
06/11/2004	1021	2,541	0,286	0,351
07/12/2004	1021	1,962	0,276	0,355

Tableau 4. Paramètres atmosphériques aux dates d'acquisition d'images SPOT : Pression atmosphérique (Pression), taux de vapeur d'eau (H₂O), taux d'Ozone (Ozone) et épaisseur optique à 550 nm (τ_{550}).

3.3.4.1. Validation de la sélection automatique des cibles invariantes

La technique de sélection automatique des cibles invariantes (CI) dans notre méthode de normalisation représente son atout principal. Afin de valider cette technique, nous avons choisi manuellement, par une photo-interprétation basée sur une connaissance du site d'étude, un ensemble de CI. Cet ensemble de CI sélectionnées manuellement (CI_{Man}), contient 70 cibles de dimensions 20 m x 20 m, réparties sur la surface de l'île (Figure 15). Il comporte de grands bâtis, des forêts denses, des laves volcaniques, des sols nus, des tarmacs d'aéroport, des parkings, etc. Les CI_{Man} ont été choisies de manière à couvrir une gamme spectrale étendue dans chacune des quatre bandes spectrales de SPOT.

Nous avons tout d'abord examiné la stabilité de chaque CI_{Man} , en évaluant l'écart type des signatures temporelles de ses réflectances dans la série temporelle d'images TOA. Les valeurs moyennes des écarts types pour toutes les CI_{Man} , dans les quatre bandes, ont été inférieures à 4%, ainsi l'ensemble a été considéré comme acceptable.

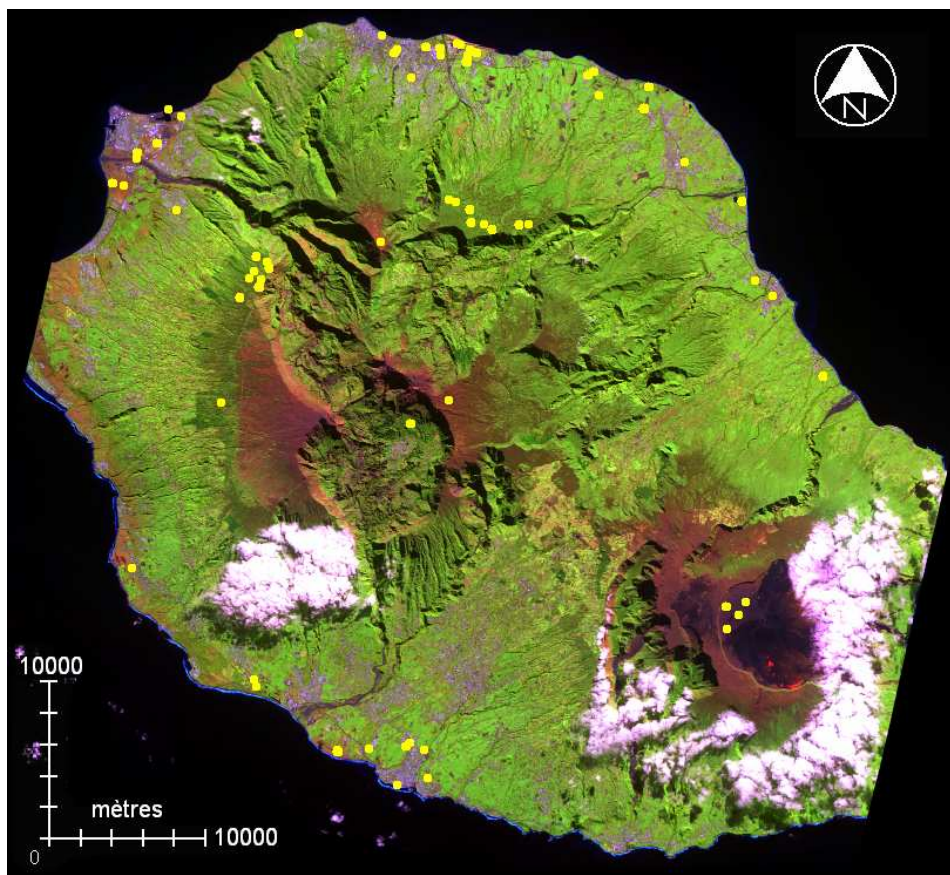


Figure 15. Localisation (en jaune) des cibles invariantes sélectionnées manuellement (CI_{Man}) sur l'image SPOT-5 (Rouge : B4 ; Vert : B3 ; Bleu : B1) acquise le 13/05/2004.

Cinquante CI_{Man} parmi les 70 ont été sélectionnées pour établir, pour chaque date, un nouveau jeu de régressions linéaires selon l'équation 4. L'image de référence que nous avons choisie est celle du 13 mai 2004 (Figure 15).

Les régressions obtenues par la sélection manuelle des invariants et celles obtenues par la technique automatique (en utilisant la même image de référence) ont ensuite été utilisées pour normaliser séparément les réflectances des 20 CI_{Man} restantes ; ces cibles couvrent une large gamme spectrale (Figure 16).

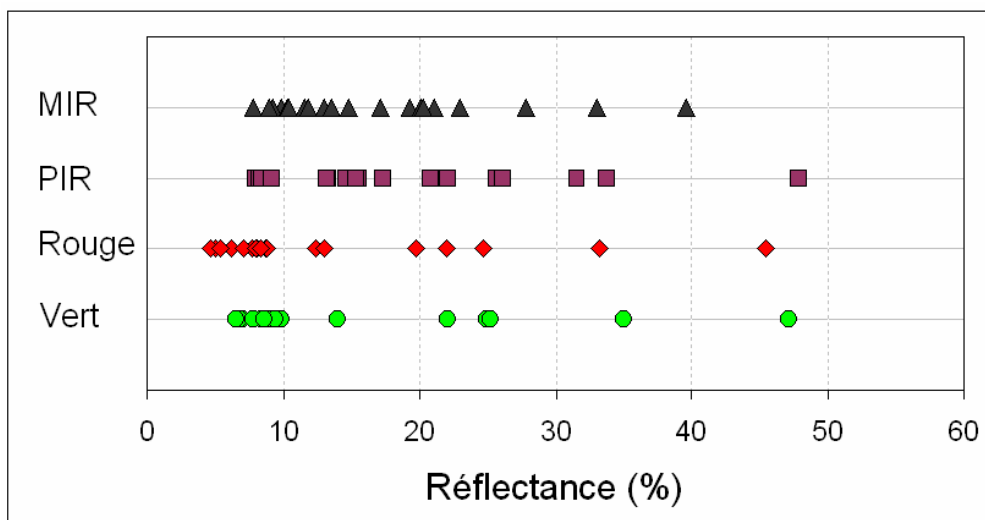


Figure 16. Valeurs moyennes des réflectances de 20 cibles invariants sélectionnées manuellement utilisées pour valider la technique de sélection automatique.

Les résultats des deux normalisations ont ensuite été comparés en se basant sur une analyse de régression linéaire et sur une évaluation de RMSE (erreur quadratique moyenne) et du biais donnés par :

$$RMSE = \sqrt{\frac{1}{n} \sum (\rho_{Man} - \rho_{Aut})^2} \quad (6)$$

$$Biais = \frac{1}{n} \sum (\rho_{Man} - \rho_{Aut}) \quad (7)$$

avec n est le nombre de cibles, et ρ_{Man} et ρ_{Aut} sont les réflectances normalisées des cibles, obtenues respectivement par la sélection manuelle et automatique des CI.

La Figure 17 montre les régressions obtenues entre les réflectances des CI_{Man} normalisées avec les deux méthodes, dans chacune des bandes spectrales. La corrélation entre la normalisation résultante des deux approches de sélection de CI est très forte : r^2 excède 0,98 dans toutes les bandes. D'autre part, les valeurs de RMSE varient entre 0,843 et 1,205, avec un biais très faible qui varie entre -0,081 et 0,285 (les valeurs de RMSE et du biais que l'on montre ici sont exprimés en pourcentage de réflectance).

Ces résultats montrent que la technique de sélection automatique des CI proposée est très satisfaisante, et représente une excellente alternative à une sélection manuelle fastidieuse et délicate.

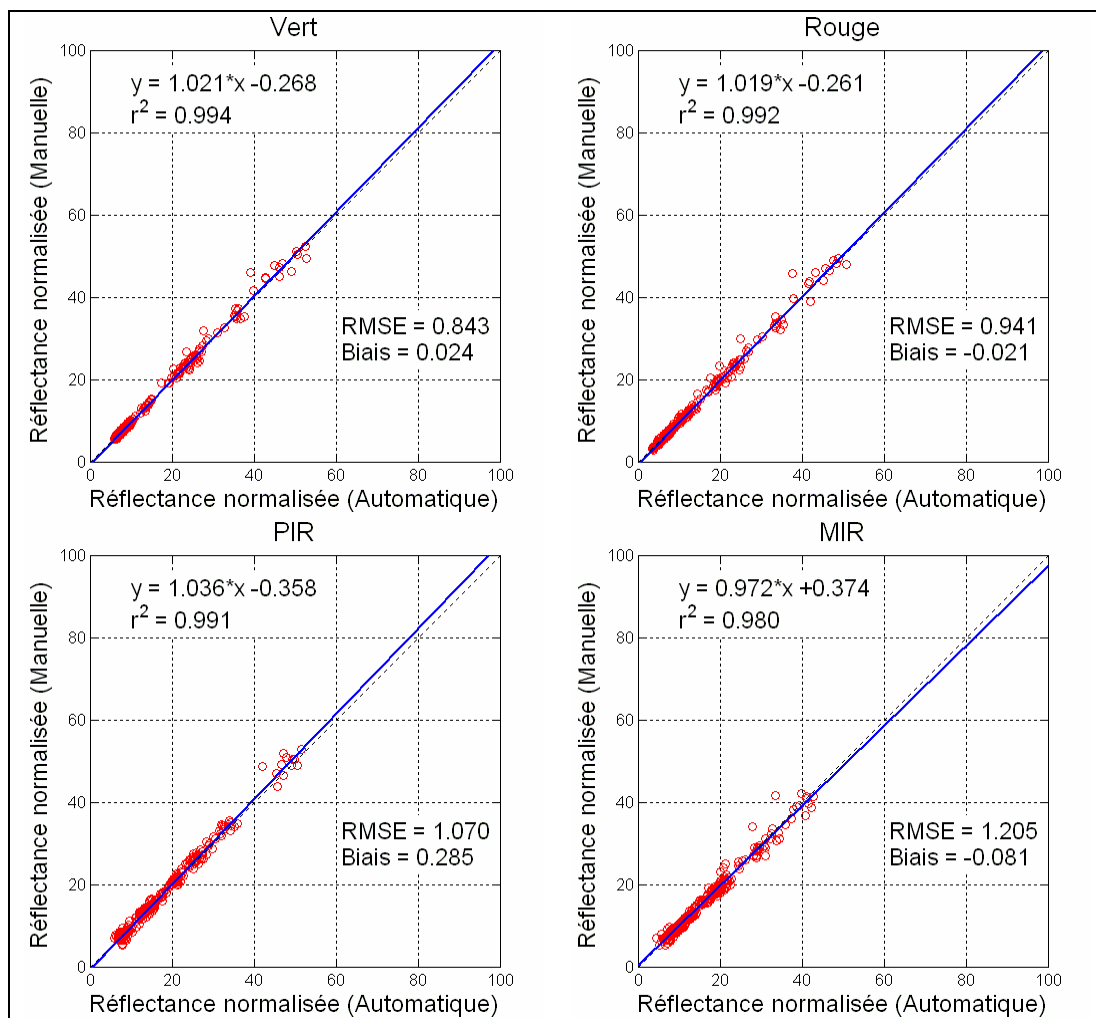


Figure 17. Comparaison des réflectances normalisées obtenues par les deux méthodes de sélection des cibles invariantes (automatique et manuel) pour un jeu de données de 20 cibles indépendantes sélectionnées par interprétation visuelle.

3.3.4.2. Comparaison entre normalisation relative et correction atmosphérique

La comparaison des performances de la normalisation relative et de la correction atmosphérique a été effectuée en évaluant l'impact de chaque méthode sur les réflectances des cibles invariantes sélectionnées manuellement (CI_{Man}).

Pour chaque CI_{Man} , nous avons extrait les valeurs moyennes des réflectances corrigées atmosphériquement (réflectances TOC) à chaque date, ainsi que les réflectances normalisées en utilisant la sélection automatique des CI. La Figure 18 montre les relations entre ces réflectances dans toutes les bandes spectrales ; les corrélations obtenues sont excellentes : r^2 est égal à 0,992, 0,994, 0,983 et 0,960 respectivement dans les bandes Vert, Rouge, PIR et MIR. Par ailleurs, les valeurs du RMSE et du biais sont faibles : le premier varie entre 1,544 à 3,221, et le dernier entre -2,649 et 1,302.

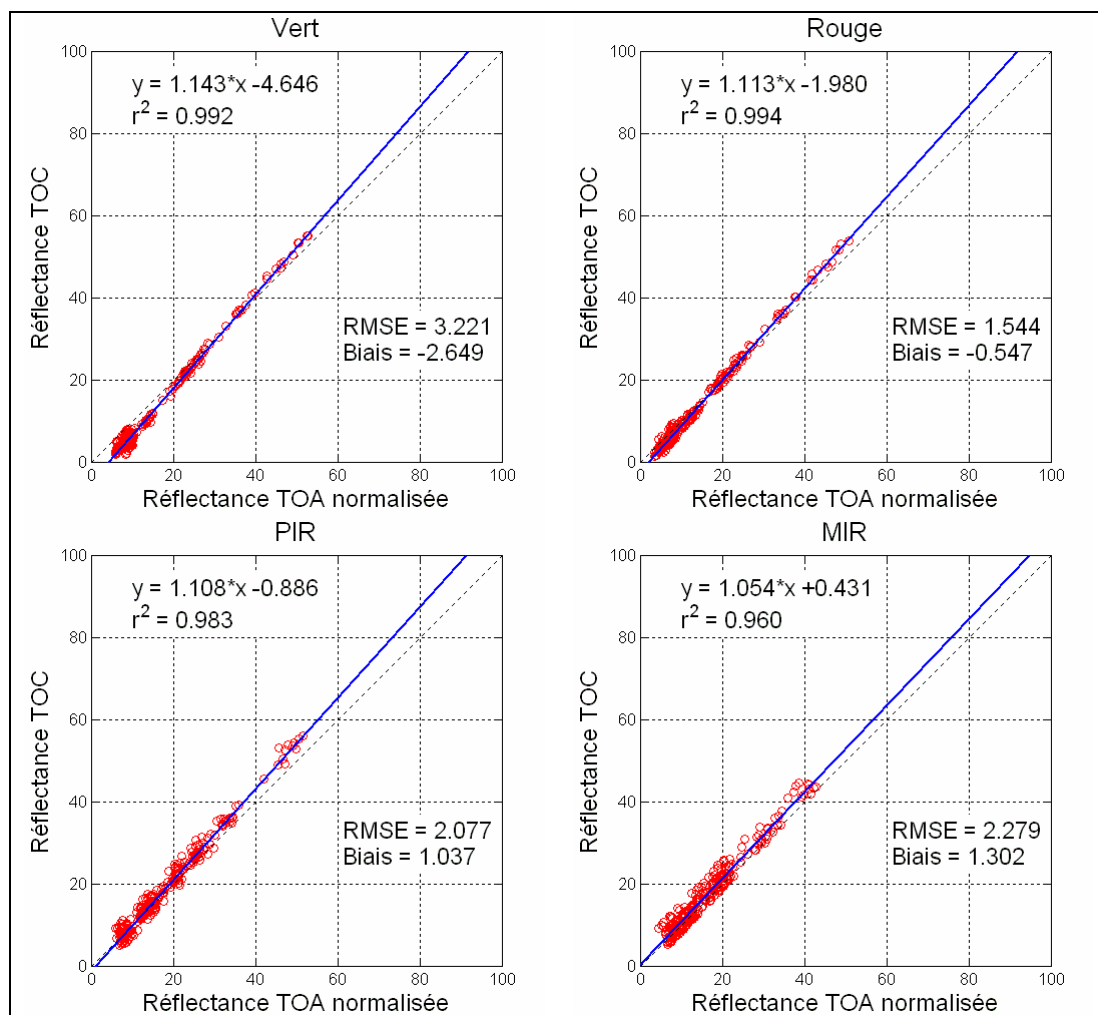


Figure 18. Comparaison des réflectances des cibles invariantes sélectionnées manuellement corrigées par 6S (réflectances TOC) et celles corrigées par la normalisation radiométrique relative.

Nous avons également évalué l'impact de chaque méthode sur le comportement temporel des réflectances des CI_{Man} (Figure 19). Pour cela, nous avons calculé l'écart type du profil temporel de chaque CI_{Man} pour les 18 dates d'acquisition, et ce a) avant la normalisation (réflectances TOA), b) après la normalisation relative, et c) après la correction atmosphérique (réflectances TOC). Le Tableau 5 montre les valeurs moyennes et maximales de l'écart type obtenues à chaque niveau et dans chaque bande. En regardant les moyennes des écarts types, nous remarquons que les deux méthodes (normalisation relative et correction atmosphérique) lissent les profils temporels des réflectances des CI_{Man} dans les quatre bandes spectrales. Cependant, la normalisation relative lisse mieux, et diminue également la valeur maximale des écarts types. Nous pouvons donc dire que les deux méthodes réduisent les perturbations radiométriques de la série temporelle, et que la normalisation relative est plus performante. Ceci pourrait être dû au fait que la normalisation relative corrige non seulement une partie des effets atmosphériques, mais également une partie des effets directionnels (ce qui n'est pas le cas avec la correction atmosphérique).

	TOA				TOA + Normalisation relative				TOA + Correction atmosphérique (TOC)			
	B1	B2	B3	B4	B1	B2	B3	B4	B1	B2	B3	B4
Valeur moyenne de l'écart type	1,34	1,59	1,80	2,08	0,88	1,10	1,25	1,25	1,23	1,46	1,69	1,98
Valeur maximale de l'écart type	3,43	3,54	2,95	4,21	3,04	3,21	2,63	2,59	3,51	3,57	2,86	4,08

Tableau 5. Valeurs moyennes et maximales des écarts types (sur 18 dates) calculées pour les réflectances des cibles invariantes de validation dans les quatre bandes spectrales (B1 : Vert ; B2 : Rouge ; B3 : PIR ; B4 : MIR).

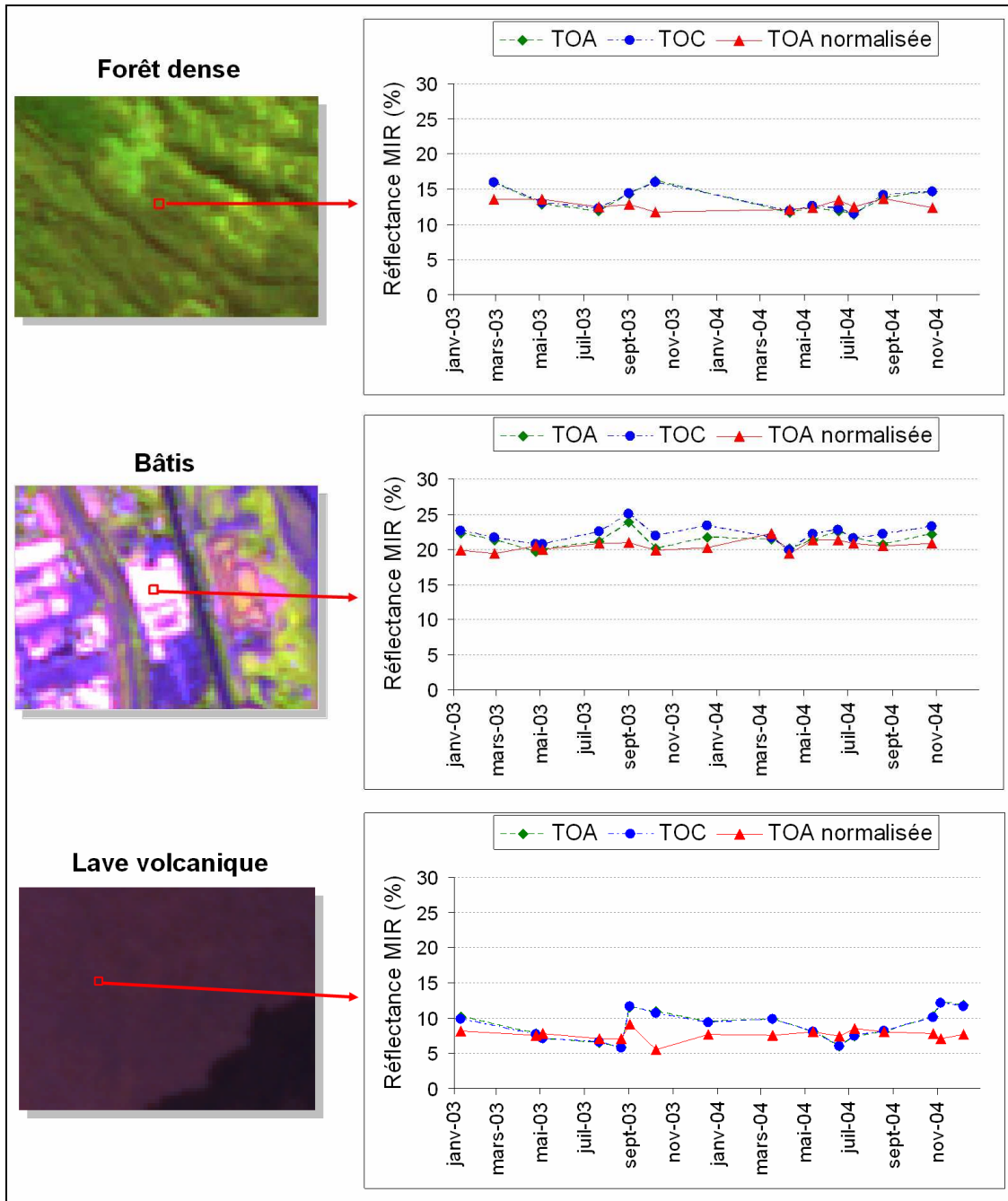


Figure 19. Exemples de profils temporels de cibles invariantes (forêt dense, bâtis, lave volcanique) dans le MIR obtenus à partir des images en réflectance TOA, des images en réflectance TOC et des images en réflectance TOA normalisée.

3.3.4.3. Impact sur les propriétés spectrales de la canne à sucre

Pour évaluer l'impact de la correction radiométrique dans le cadre de notre application agricole, nous avons comparé les deux méthodes de corrections (normalisation relative et correction atmosphérique) sur une série temporelle d'images de NDVI (*Normalized Difference Vegetation Index*). Le NDVI est calculé par :

$$NDVI = \frac{\rho_{PIR} - \rho_{Rouge}}{\rho_{PIR} + \rho_{Rouge}} \quad (8)$$

avec ρ_{PIR} et ρ_{Rouge} respectivement les réflectances dans la bande PIR et la bande Rouge.

Vingt parcelles de canne à sucre à La Réunion d'une surface moyenne de 4,3 ha ont été choisies pour réaliser la comparaison. Pour chaque parcelle, nous avons calculé les valeurs moyennes de NDVI à chaque date en utilisant les images normalisées et les images corrigées atmosphériquement (TOC).

La Figure 20 montre la régression obtenue entre les NDVI normalisés et les NDVI TOC, calculés pour toutes les dates et pour toutes les parcelles. Une forte corrélation ($r^2 = 0,959$) existe entre les deux méthodes. Cependant, les valeurs de RMSE et du biais sont non-négligeables (RMSE = 0,098 ; biais = 0,092). L'erreur relative estimée est de 12,66%, ce qui est également significatif.

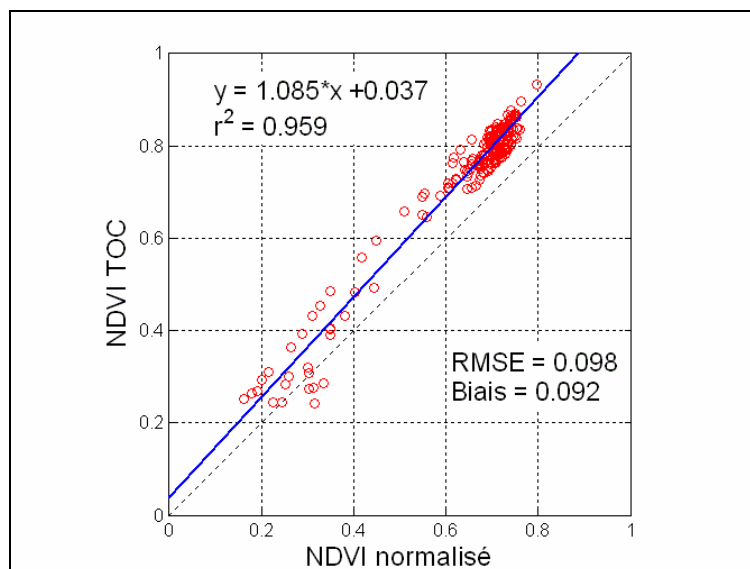


Figure 20. Comparaison des valeurs de NDVI calculées à l'échelle de la parcelle à partir des images corrigées par 6S (NDVI TOC) et des images normalisées (à toutes les dates et pour toutes les parcelles).

La normalisation relative n'étant pas prévue pour obtenir des valeurs radiométriques absolues, il nous a semblé plus approprié de comparer l'évolution relative (augmentation ou diminution) du NDVI d'une date à une autre. Ainsi, nous avons comparé les valeurs de pente de NDVI calculées pour chaque couple de dates consécutives pour toutes les parcelles (Figure 21.a). Nous avons remarqué une bonne corrélation ($r^2 = 0,898$) et des basses valeurs de RMSE et du biais (respectivement 0,00134 et -0,00028). Seuls 4% des points ont été significativement dispersés autour de la régression (avec une erreur absolue $> 0,002$). Ceci signifie que globalement les deux méthodes mènent à un comportement très similaire du NDVI.

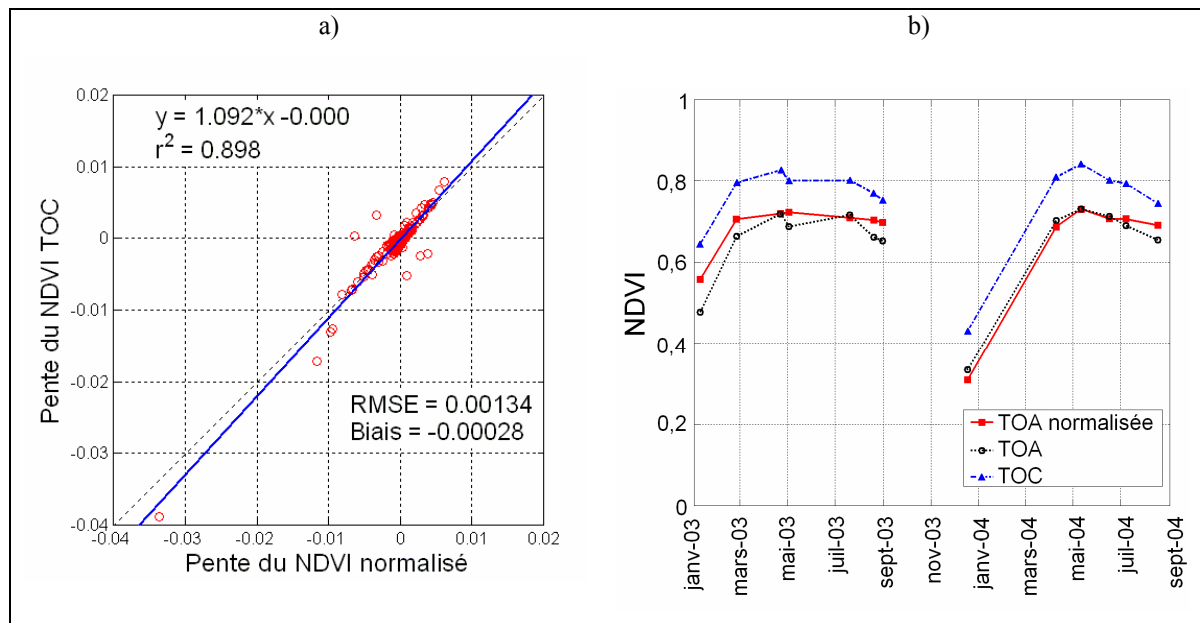


Figure 21. a) Comparaison des valeurs de pente dans les profils temporels de NDVI corrigés par 6S (NDVI TOC) et ceux normalisés relativement. b) Profils temporels de NDVI calculé pour une parcelle de canne à partir des TOA réfléchances, des réfléchances TOA relativement normalisée et des réfléchances TOC.

En regardant plus en détail les résultats obtenus pour chaque parcelle, nous avons remarqué que les profils temporels de NDVI, bien que très comparables, peuvent parfois montrer de légères différences à des phases critiques du cycle de la canne. La Figure 21.b montre le profil de NDVI obtenu pour une parcelle de canne en utilisant les réfléchances TOA, les réfléchances TOA normalisées et les réfléchances TOC. Nous pouvons constater que la normalisation relative lisse davantage le profil que la correction atmosphérique. Par exemple, la fluctuation de NDVI à 10 jours d'intervalle qui apparaît en mai 2003 sur les profils TOA et TOC

n'apparaît pas sur le profil normalisé. De légères différences sont également observées pendant la phase de sénescence (*e.g.* de juillet à septembre 2003). Ces différences subtiles peuvent mener à des interprétations diverses quant à l'état de la culture. Par exemple, le profil du NDVI TOC entre juin et octobre 2004 est le signe d'un fort dessèchement de la canne, avec une diminution du NDVI de 0,1, alors que le NDVI normalisé ne chute que de 0,04 traduisant une sénescence naturelle.

Comme aucune vérité terrain n'est disponible pour cette comparaison, nous ne pouvons pas déterminer quelle méthode est la plus fiable. Nous pouvons dire à ce stade que les profils temporels de NDVI obtenus par les deux méthodes sont très comparables, mais que les faibles différences existantes pourraient avoir un impact sur les indicateurs de production (*e.g.* intégrale de NDVI, max de NDVI, taux de décroissance). En ce qui concerne notre application agricole (la détection des coupes des parcelles), on peut considérer que les informations issues du profil de NDVI, que ce soit normalisé ou corrigé atmosphériquement, sont imprécises.

3.4. Conclusion

La problématique du pré-traitement des séries temporelles d'images satellites a été abordée dans ce chapitre. La contribution de cette thèse porte essentiellement sur les méthodes de normalisation radiométrique.

Nous avons proposé une méthode de normalisation radiométrique relative des séries temporelles basée sur les réflectances des cibles invariantes. La sélection des cibles invariantes étant une étape primordiale et délicate, nous avons développé et validé une technique de sélection automatique de ces cibles. Les avantages principaux de cette méthode sont : sa simplicité, son automaticité, son applicabilité aux images acquises dans différentes saisons, et le fait qu'elle n'ait pas besoin de données atmosphériques. Cependant, les corrections ainsi effectuées restent « relatives », et les valeurs corrigées dépendent du choix de l'image de référence.

Nous avons également présenté la méthode de correction atmosphérique utilisée par le CNES pour calculer les images de réflectance au sol (TOC). Cette méthode, basée sur le modèle 6S, corrige la radiométrie des effets atmosphériques quelque soit la date et l'endroit. Néanmoins, sa principale restriction réside dans la difficulté d'estimer les paramètres atmosphériques. En effet, ces paramètres sont souvent fournis à une échelle globale, et leur estimation à une échelle locale n'est pas aisée. Une autre limitation de cette méthode est qu'elle ne corrige pas les effets bidirectionnels de réflectance.

La comparaison des performances des deux méthodes, normalisation et correction atmosphérique, à partir d'un ensemble de cibles invariantes sélectionnées manuellement, a montré une très forte corrélation et un faible taux d'erreur. L'analyse du NDVI des parcelles de canne à sucre a également montré une forte corrélation, bien que le taux d'erreur observé entre NDVI normalisé et NDVI TOC ait été relativement élevé. Des profils temporels très comparables de NDVI à l'échelle de la parcelle ont été obtenus, mais de légères différences ont été remarquées, différences qui pourraient influencer le calcul des indicateurs de production (*e.g.* intégrale de NDVI, max de NDVI, taux de décroissance).

D'après les différents résultats, nous pouvons conclure que la méthode de normalisation relative développée dans cette thèse peut être une bonne alternative à la correction atmosphérique, pour le pré-traitement de l'imagerie multi-temporelle à haute résolution spatiale.

Chapitre 4 :

Systeme d'aide à la détection des coupes de canne à sucre

Sommaire

4.1. Introduction	80
4.2. Sources d'information	81
4.2.1. Séries temporelles	81
4.2.2. Modèle de culture.....	86
4.2.2.1. Simulations du LAI	87
4.2.2.2. Simulation du NDVI.....	87
4.2.2.3. Indicateur de coupe.....	89
4.2.3. Connaissances expertes.....	92
4.3. Formalisme.....	94
4.3.1. Ensembles flous et termes linguistiques.....	95
4.3.2. Inférence à base de règles	96
4.4. Structure et composants.....	98
4.4.1. Etage d'entrée	98
4.4.2. Sortie.....	105
4.4.3. Base de règles.....	106
4.4.3.1. Génération manuelle par expertise	107
4.4.3.2. Génération automatique par apprentissage	112
4.4.4. Moteur d'inférence.....	114
4.5. Prise de décision.....	116
4.5.1. Systeme à base de règles expertes.....	116
4.5.2. Systeme à base de règles induites	118
4.6. Conclusion.....	119

4.1. Introduction

Après la présentation des phases de pré-traitement des séries temporelles dans le chapitre précédent, nous consacrons ce chapitre à la présentation d'une nouvelle approche de suivi des opérations culturales. Nous la décrivons sur la base de l'application agricole de cette thèse : le suivi des coupes de canne à sucre à La Réunion (*c.f.* Chapitre 2). Rappelons que cette application constitue un exercice particulièrement difficile de détection de changement dans les milieux agricoles, en raison de la très forte variabilité spatio-temporelle des parcelles de canne.

Nous présentons un système d'aide à la détection des coupes de parcelles de canne à sucre. Nous définissons tout d'abord, les sources d'informations qui alimentent ce système. Les séries temporelles d'images sont la source principale. Deux autres sources sont intégrées afin de faire face au nombre insuffisant d'images dégagées et d'automatiser le processus de détection : un modèle de culture et des connaissances expertes.

Le formalisme de fusion et d'inférence sur lequel le système est basé sera présenté avant de détailler sa structure et ses composants. Nous terminons le chapitre par une description de quelques méthodes de prise de décisions, qui permettent d'exploiter les sorties du système pour détecter les coupes.

Les travaux présentés dans ce chapitre ont fait l'objet :

- d'un article "*Integrating SPOT-5 time series, crop growth modeling, and expert knowledge for monitoring agricultural practices - The case of sugarcane harvest on Reunion Island*" (El Hajj *et al.*, 2008b) accepté dans la revue *Remote Sensing of Environment* (Facteur d'impact = 3,013 ; 2007) ;
- d'un article "*Multi-source information fusion for crop monitoring*" qui sera soumis à la revue *Agriculture Ecosystems & Environment* (Facteur d'impact = 2,308 ; 2007) ;
- de 3 communications dans les colloques internationaux : MultiTemp-2007 (El Hajj *et al.*, 2007a) avec un poster, Fusion-2008 (El Hajj *et al.*, 2008a) avec une présentation orale et LFA (El Hajj *et al.*, 2008d) avec une présentation orale.

Ces publications sont fournies dans l'Annexe P.2.

4.2. Sources d'information

Nous avons vu dans les Chapitres 1 et 2 que les méthodes développées jusqu'à présent pour suivre les opérations culturales, et notamment les coupes de canne à sucre, souffrent de l'irrégularité des données satellite due à la nébulosité ou au nombre d'images acquises. Nous avons également vu que ces méthodes sont limitées par la subjectivité et par le manque d'automatisme. Afin de répondre à ces problématiques et de pouvoir prendre des décisions fiables et crédibles sur les changements à partir des séries temporelles, il est nécessaire de mobiliser d'autres sources d'information.

Dans la méthode que nous proposons, nous utilisons un modèle de culture et des connaissances expertes comme sources d'information exogènes. Les informations extraites à partir de ces deux sources ainsi que celles des séries temporelles d'images sont présentées dans les sous-sections qui suivent.

4.2.1. Séries temporelles

Les séries temporelles d'images SPOT représentent la source principale d'information. En utilisant le parcellaire de canne à sucre, les profils temporels de réflectances dans les quatre bandes spectrales, Vert, Rouge, Proche Infra-Rouge (PIR) et Moyen Infra-Rouge (MIR), ont été extraits à partir des images TOC¹⁶ pour chaque parcelle. Cette extraction a été réalisée à l'échelle de la parcelle en effectuant à chaque date la moyenne de tous ses pixels. Les pixels mixtes correspondant aux bords des parcelles ont été éliminés ainsi que les pixels affectés par les nuages ou les ombres (utilisation des masques de nuages). En utilisant les valeurs de réflectance dans le Rouge (ρ_{Rouge}) et dans le PIR (ρ_{PIR}), nous avons calculé les profils temporels de NDVI (*Normalized Difference Vegetation Index*) :

$$NDVI = \frac{\rho_{PIR} - \rho_{Rouge}}{\rho_{PIR} + \rho_{Rouge}} \quad (9)$$

¹⁶ Images de réflectances au sol (*Top Of Canopy*).

Propriétés spectrales et temporelles à l'échelle de la parcelle

Dans la Figure 22 nous illustrons un exemple de profils temporels de réflectances et de NDVI calculés pour une parcelle de canne en repousse et pour une parcelle nouvellement plantée (*i.e.* canne vierge).

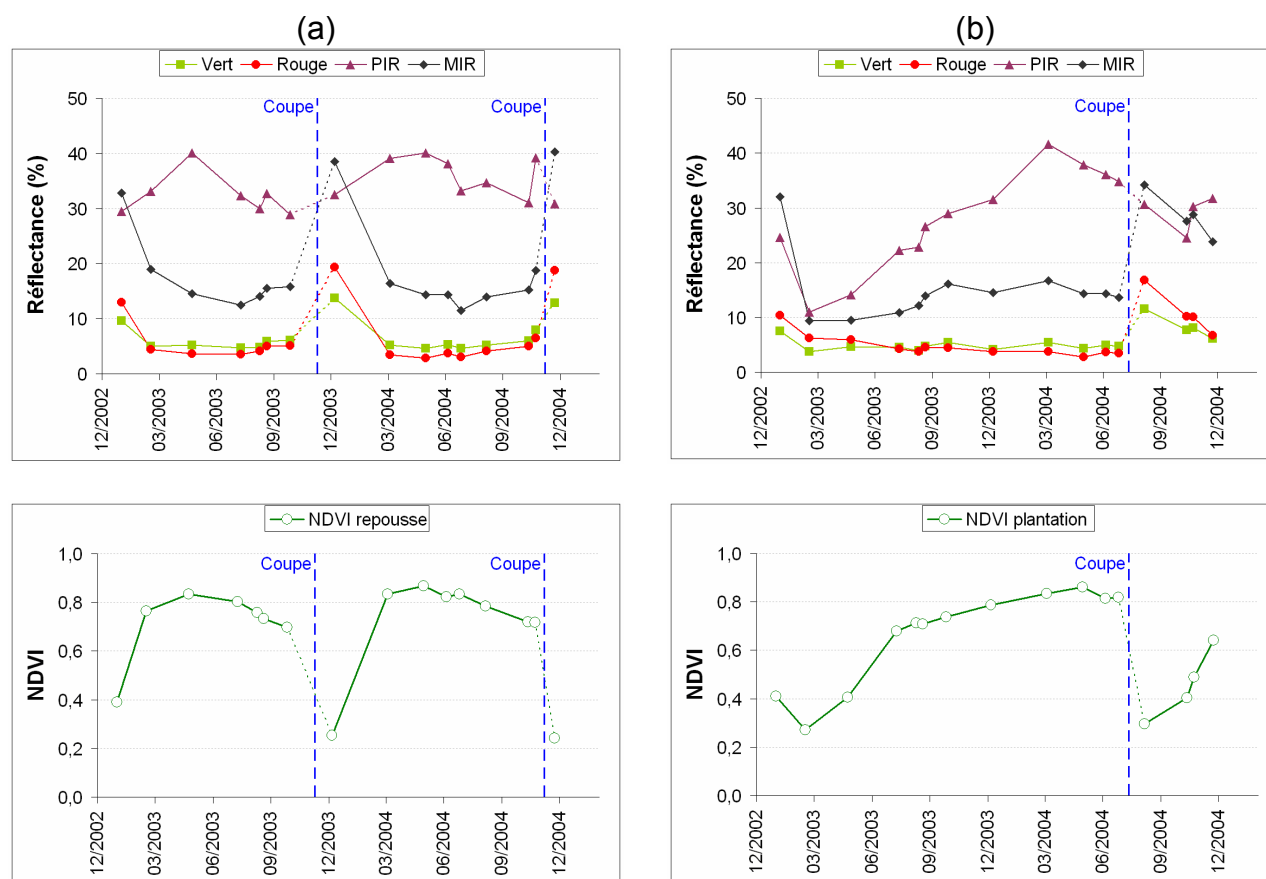


Figure 22. Exemple de profils temporels de NDVI et de réflectances calculés à l'échelle de la parcelle pour a) une repousse et b) une plantation.

- Dans le visible (Vert et Rouge) on note, lors de la coupe d'une parcelle, une augmentation de réflectance de 15% entre deux acquisitions proches. L'amplitude du signal pendant la phase de croissance est relativement faible.
- Les profils temporels dans le PIR, reflétant généralement les évolutions de la structure du couvert, ne permettent pas la détection d'une coupe suivie d'une repousse (ce qui est le cas d'environ 85% des coupes à La Réunion) : les valeurs de réflectance des résidus de récolte, qui couvrent la parcelle après la coupe, sont à peu près égales à celles d'une canne mature. Par contre, on observe une forte discrimination entre la canne verte et le sol nu

(Figure 22.b), ce qui permet la détection des plantations dans le cas d'acquisitions fréquentes.

- Dans le MIR, la réponse spectrale des résidus de récolte est très forte et très différente de celle d'un couvert de canne (Bégué *et al.*, 2008). Lorsque la fréquence d'acquisition est élevée, une augmentation significative des valeurs de MIR (*e.g.* 25% dans la Figure 22) est observée quand la parcelle est coupée.
- Le profil temporel de NDVI fournit des informations sur l'état actuel de la parcelle ainsi que sur les différents stades phénologiques. Généralement, ce profil peut être divisé en deux périodes : une période pendant laquelle les valeurs de NDVI augmentent, correspondant à la phase du développement végétatif de la canne, et une autre période avec des valeurs quasi-constantes ou décroissantes, correspondant à la phase de maturation ou de sénescence. Dans le cas d'une repousse après une coupe, le NDVI d'une parcelle de canne augmente très rapidement pour atteindre son maximum en moins de 4 mois. Dans le cas des plantations, l'augmentation du NDVI est plus lente, avec un maximum qui est atteint en 8 ou 9 mois, du fait de la mise en place du système racinaire de la plante. Dans tous les cas (repousse ou plantation), la canne atteint une valeur élevée de NDVI qui va se stabiliser jusqu'à la coupe, ou décroître si la plante subit un stress hydrique (*i.e.* sénescence des feuilles).

Variabilité spatio-temporelle

Les profils temporels de NDVI des parcelles de canne à sucre, à l'échelle d'une exploitation ou d'un bassin de production, sont affectés par une forte variabilité spatio-temporelle due à l'étalement de la campagne de récolte (Figure 23). En effet, la longueur de la campagne de récolte conduit à des variations phénologiques intra-parcellaire importante, et des conditions environnementales de croissance différentes entre les cannes coupées en début ou en fin de la campagne (Bégué *et al.*, 2008).

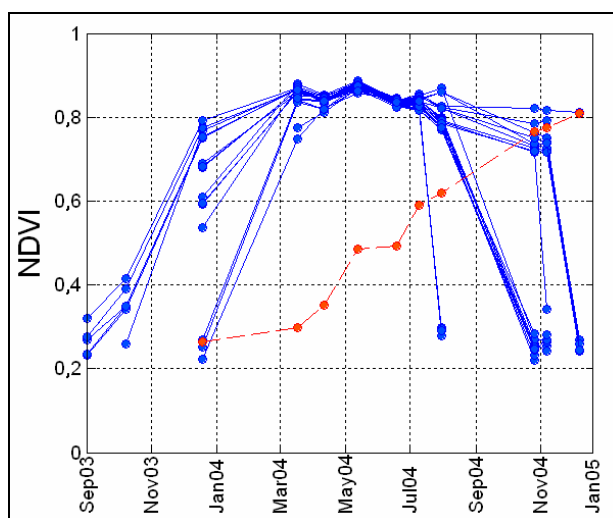


Figure 23. Variabilité spatio-temporelle du NDVI à l'échelle de l'exploitation de M. Caruel. En bleu les profils temporels de NDVI pour quelques parcelles en repousse et en rouge le profil pour une parcelle nouvellement plantée.

- La campagne de récolte durant plusieurs mois (6 environ), les dates de repousse ou de plantation des parcelles sont différentes d'une parcelle à une autre, ainsi, les profils temporels de NDVI correspondant peuvent être décalés dans le temps (Figure 23).
- Les conditions climatiques, comme la pluie et la température, au cours de la campagne de récolte influent sur le développement de la canne ; à titre d'exemple, la phase de croissance d'une repousse suite à une coupe tardive, en décembre, est généralement plus courte que celle d'une repousse en début de campagne, en juillet. En effet, en début de campagne, les températures sont relativement basses, aux alentours de 20°C, alors qu'en fin de campagne, elles sont plus élevées, aux environs de 30°C. Les parcelles coupées en début de campagne vont avoir des conditions de reprise plus difficiles que celles coupées en fin de campagne. L'état des réserves hydriques du sol influe dans le même sens que les températures ; le climat plus sec en début de campagne accentue les difficultés pour la reprise du développement de la canne (Bappel, 2005). Toutefois, les facteurs climatiques jouent le rôle de régulateurs qui synchronisent le maximum de croissance végétative des repousses avec un pic de NDVI observé en fin de saison des pluies (autour de Mai-Juin). À ce pic de croissance, seules les nouvelles plantations sont encore à un stade précoce de développement, et leurs valeurs de NDVI sont beaucoup plus faibles (Figure 23).

Conclusion de l'analyse des profils

Conformément à ce qui est rapporté dans la littérature (Bappel, 2005; Lebourgeois *et al.*, 2007; Bégué *et al.*, 2008), nous avons remarqué que parmi les profils temporels extraits à partir des images SPOT, ceux du NDVI et du MIR sont les plus intéressants pour le suivi des coupes.

Une augmentation significative des valeurs de MIR entre deux acquisitions proches semble être un indice de coupe de la parcelle. Quant au NDVI, il paraît comme un bon indicateur à deux titres :

- La chute du NDVI entre deux acquisitions proches peut renseigner sur la coupe de la parcelle ;
- L'analyse du profil temporel de NDVI permet de connaître le stade phénologique et l'état de la parcelle à instant donné, et d'informer ainsi sur la possibilité de sa coupe.

Les informations issues des profils temporels de NDVI et de réflectances, notamment les indicateurs bi-dates, subissent deux limitations majeures : le manque d'acquisition d'images et l'imprécision. En effet, les profils temporels sont souvent incomplets à cause des lacunes d'acquisition ou des conditions atmosphériques, ce qui fait que les indicateurs bi-dates (*e.g.* chute de NDVI, augmentation de MIR) ne sont pas toujours fiables (Figure 24) ; lorsque l'intervalle qui sépare la vraie date de coupe d'une acquisition satellite est supérieur à 2 mois, il est difficile de détecter cette coupe en utilisant les images comme unique source d'information (la différence de NDVI ou de MIR dans ce cas n'est plus significative). De plus, l'addition de la nébulosité dans les images réduit considérablement la surface exploitable avec les indicateurs bi-dates.

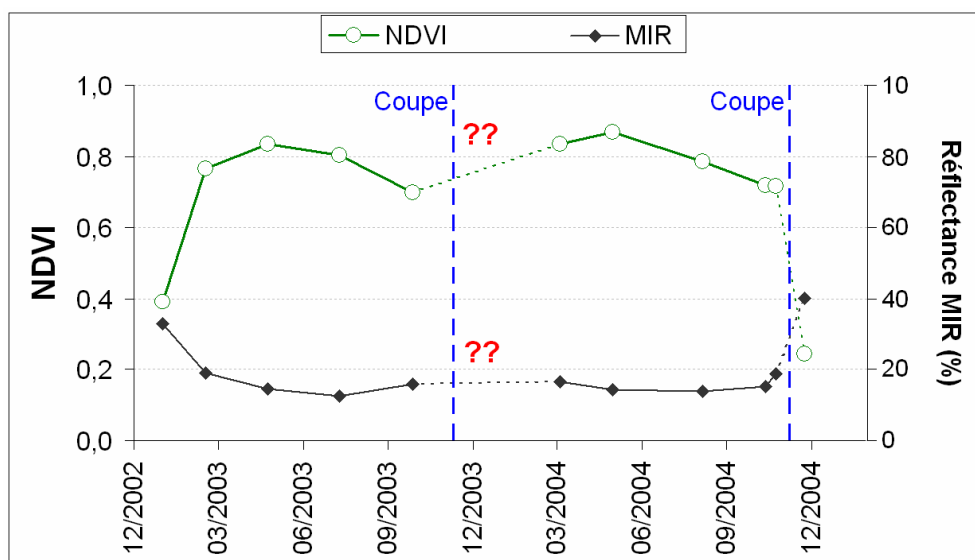


Figure 24. Exemples de profils temporels incomplets de NDVI et de MIR pour une repousse. La coupe de 2003 de cette parcelle ne peut pas être détectée par les différences bi-dates de NDVI et de MIR.

La deuxième limitation, l'imprécision, est due à la fois à la radiométrie des images (*e.g.* imperfections des corrections) et à la différence d'état de couvert d'une parcelle à une autre. A titre d'exemple, la différence de NDVI avant et après une coupe sera beaucoup plus forte pour un couvert dense et photosynthétiquement actif que pour un couvert stressé, jaunissant. En conséquence de cette imprécision, la définition de seuils génériques sur les indicateurs bi-dates issus des images n'est pas aisée.

Afin de répondre à ces limitations, il s'avère donc nécessaire (1) d'intégrer d'autres sources d'information, (2) d'utiliser l'information temporelle (*e.g.* valeurs historiques et comportement du NDVI) et (3) de travailler avec des indicateurs qualitatifs plutôt que quantitatifs.

4.2.2. Modèle de culture

La première source d'information mobilisée pour limiter les difficultés liées à la présence de nuages et au manque d'acquisitions dans les séries temporelles est un modèle de culture, basé sur les données environnementales. Cette source est utilisée pour simuler des profils temporels de NDVI afin de construire un indicateur synthétique indépendant des séries temporelles, qui fournit une information sur la possibilité de coupe d'une parcelle entre deux dates données (*i.e.* entre deux dates d'acquisitions satellites t et t').

4.2.2.1. Simulation du LAI

Nous avons utilisé le modèle ecophysiologique MOSICAS (Martiné and Todoroff, 2002) pour simuler la croissance de la canne à sucre dans chaque parcelle sur les sites d'études à un pas de temps journalier. Les simulations sont alimentées avec les données climatiques, enregistrées dans des stations météorologiques près de chaque site, et les paramètres biophysiques du sol prédéfinis dans le modèle.

MOSICAS est un modèle de type thermo-radiatif qui tient compte du stress hydrique. Il est composé de deux modules dynamiques : un module de croissance et un module de bilan hydrique. Le module de croissance est basé sur l'interception du rayonnement, la conversion de ce dernier en biomasse, et ensuite la partition de la biomasse en ses différentes composantes (racines, feuilles, tiges et sucre). Le module de bilan hydrique est un modèle simplifié (WB1, pFactor). Il a été adapté à partir du module inclus dans le modèle Cers (*Crop Environment Resources Synthesis*), développé pour le maïs puis utilisé pour des plantes dans des conditions tropicales (Martiné, 2003). Le sol dans ce module est représenté par un réservoir dont la profondeur est limitée par la zone d'enracinement.

Parmi d'autres variables d'état, MOSICAS simule l'indice foliaire, *i.e.* le LAI, de la canne à sucre. La modélisation du LAI est faite selon le formalisme « *big leaf* » ; elle est basée sur un bilan entre les surfaces foliaires quotidiennement élaborée et sénescence. Ce bilan est principalement mené par la température. Il est limité par la source de rayonnement correspondant à la masse des feuilles et au stress hydrique.

4.2.2.2. Simulation du NDVI

Les profils temporels de LAI simulés par MOSICAS ont été transformés en profils temporels de NDVI en utilisant la relation suivante (Figure 25) :

$$NDVI = 0,146 \times \ln(LAI) + 0,571 \quad (10)$$

Cette relation a été calculée à partir de mesures *in situ* de LAI réalisées en 2002, 2003 et 2007 sur des parcelles de canne à sucre à La Réunion (variétés R570 et R579), et des valeurs de NDVI des images SPOT interpolées aux dates de mesures terrain. Les mesures en 2002 et 2003 de LAI ont été effectuées par (Bappel, 2005) sur une trentaine de parcelles, entre juillet et octobre en 2002 et entre avril et juillet en 2003. Une nouvelle campagne de mesures a été

mise en place sur six parcelles, en octobre 2007, afin d'enrichir la base de données et notamment celles concernant les jeunes cannes (*i.e.* âge inférieur à 3 mois). Le protocole et la méthode utilisée pour faire les mesures en 2002 et 2003 sont décrits dans (Bappel, 2005). Les valeurs de LAI obtenues en 2007 sont basées sur une relation biométrique (équation 11) dont les paramètres sont d'une part les hauteurs $htvd$ des tiges¹⁷ dans une placette définie de surface S , et d'autre part le nombre de limbes verts nl jusqu'à la dernière ligule visible, dite TDV (*Terminal Visible Diulap*) (Martiné, 2004). Les coefficients a et b dans l'équation 11 varient en fonction de la variété. Leurs valeurs sont recensées dans le Tableau 6.

$$LAI = \sum_{tiges} \frac{a \times [1 - \exp(-b \times htdv_{cm} \times nl/100)]}{S_{cm^2}} \quad (11)$$

Variété de canne à sucre	a	b
R579	8024,1136	0,11378854
R570	10747,546	0,097933269

Tableau 6. Valeurs des coefficients a et b nécessaires pour l'estimation du LAI selon la variété de canne à la Réunion (Source : Martiné, 2004).

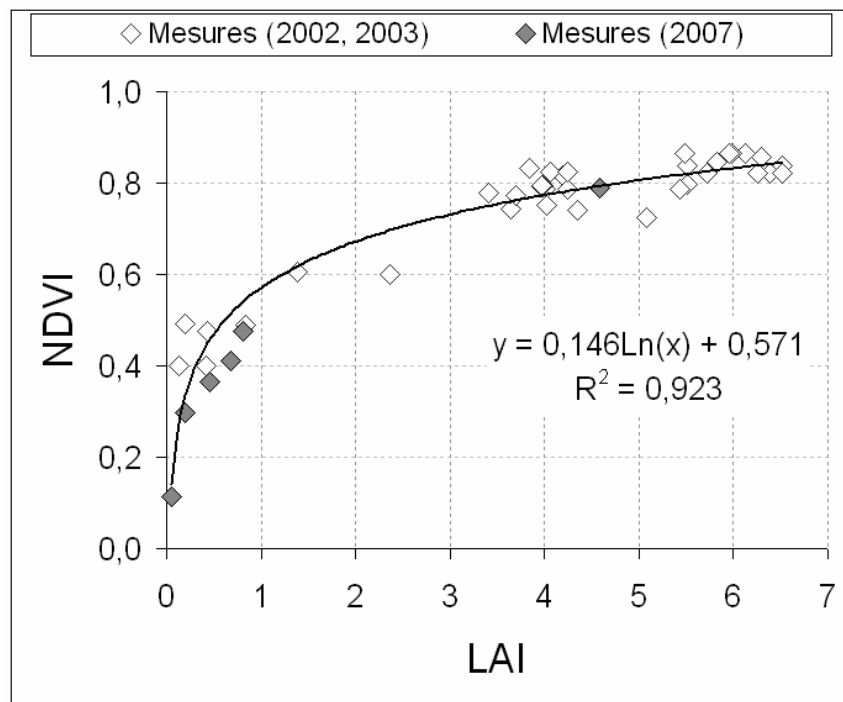


Figure 25. Relation entre le LAI (mesuré *in situ*) et le NDVI (SPOT).

¹⁷ La hauteur $htvd$ d'une tige est la distance entre le sol et la dernière ligule visible.

La relation (équation 10) est établie uniquement à partir des données acquises au cours de la phase de croissance de la canne (avant et au cours de la période maximale de NDVI), ainsi elle est valable pour le LAI vert.

La Figure 26 montre un exemple de profil temporel de NDVI obtenu avec la relation (équation 10) à partir des simulations journalières de LAI par MOSICAS. En comparant ce profil à celui issu des images SPOT, on observe une cohérence importante dans les phases de croissance (intervalle de 4 mois environ) et pour le plateau de croissance (intervalle de 4 mois environ). Quant à la phase de sénescence, une différence est observée entre le profil de NDVI « images » et celui de NDVI « simulé ». En effet, cette différence est due au fait que MOSICAS simule le LAI « vert » et ne tient pas compte de la sénescence (*i.e.* jaunissement) de la canne à sucre, ce qui fait que le profil de NDVI calculé à partir des images chute remarquablement dans cette phase contrairement au NDVI simulé.

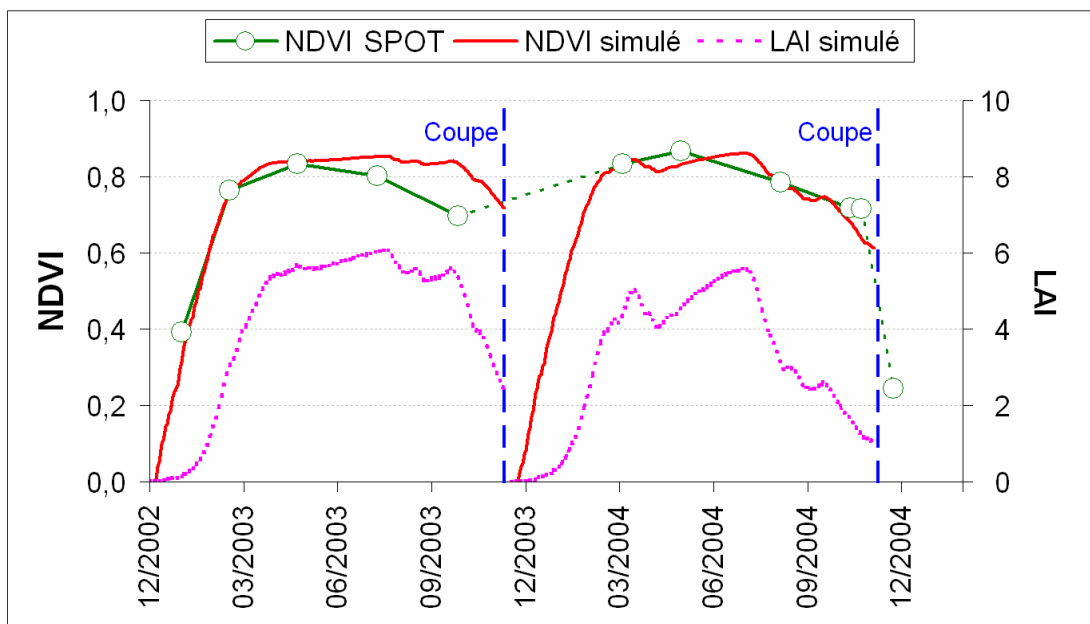


Figure 26. Des profils temporels de LAI, de NDVI simulé en utilisant le modèle MOSICAS et de NDVI calculé à partir des images SPOT pour la même parcelle.

4.2.2.3. Indicateur de coupe

A partir des profils temporels de NDVI simulé, nous avons construit un indicateur, T_n , utile pour la détection de coupe de parcelle ; cet indicateur représente le Temps nécessaire (en nombre de jours) pour que le NDVI d'une parcelle de canne, coupée à une date donnée, atteigne un certain seuil.

La Figure 27 illustre un exemple des relations calculées entre le Tn et la date de coupe d'une parcelle (*i.e.* date de début de simulation de LAI par MOSICAS), et ce pour différents seuils de NDVI. Selon cet exemple, pour atteindre un NDVI de 0,7 une parcelle de canne à sucre aura besoin de 133 jours ($Tn = 133$) de croissance si elle a été coupée en début juillet (pendant la saison fraîche et sèche), et de 89 jours seulement ($Tn = 89$) si elle a été coupée en début décembre (pendant la saison chaude et humide). Nous avons observé que pour des seuils élevés de NDVI (0,7 dans notre exemple), le modèle est très sensible aux variations climatiques, telle que la quantité de précipitations.

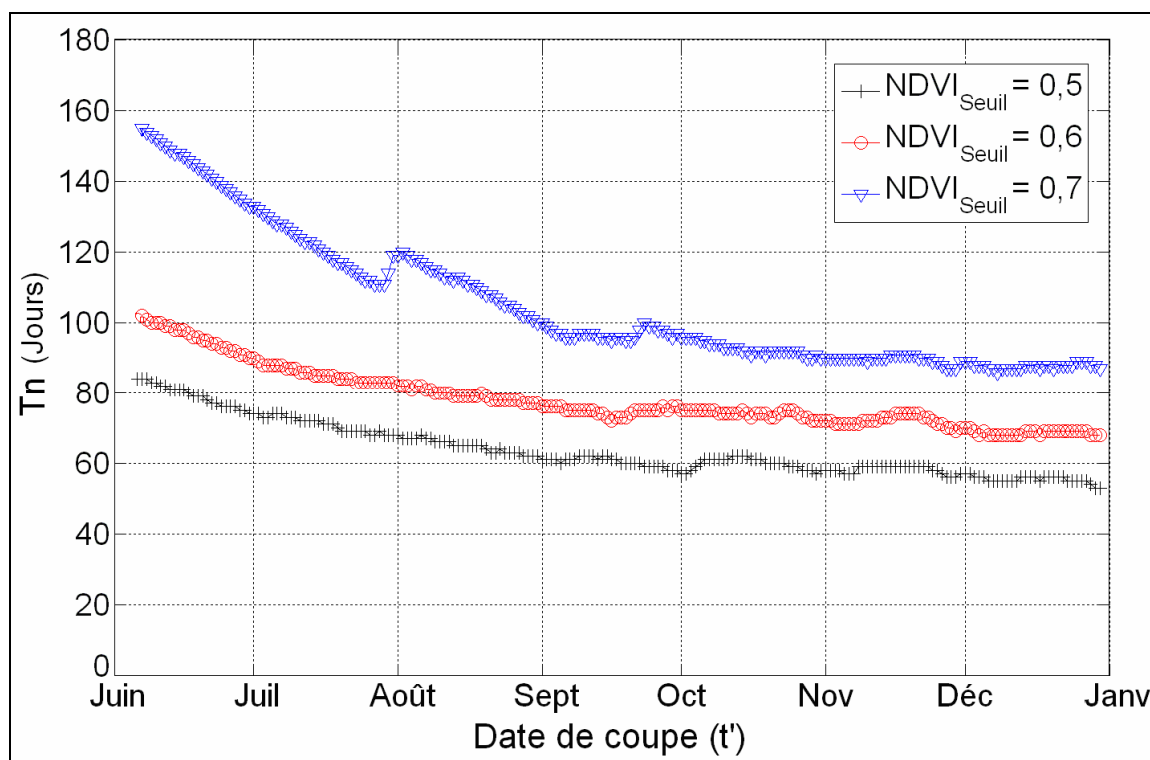


Figure 27. Exemple de relations entre la date de coupe (date de début de simulation de NDVI par MOSICAS) et le temps nécessaire (Tn) pour qu'une parcelle de canne atteigne un niveau NDVI_{Seuil} de NDVI. La date de coupe correspond dans le système expert à la date d'acquisition satellite t' .

En utilisant le modèle MOSICAS, on peut réaliser, si besoin, des simulations de NDVI entre chaque couple d'acquisitions satellites (*e.g.* Figure 28) afin d'évaluer la possibilité de coupe entre ces acquisitions. Dans l'exemple de la Figure 28, nous pouvons constater l'importance de l'information que le modèle de culture peut apporter en cas de manque d'acquisitions : en se basant sur les données environnementales, on peut conclure d'après les simulations qu'il est possible que la canne soit coupée entre les acquisitions du 08/10/2003 et du 17/03/2004.

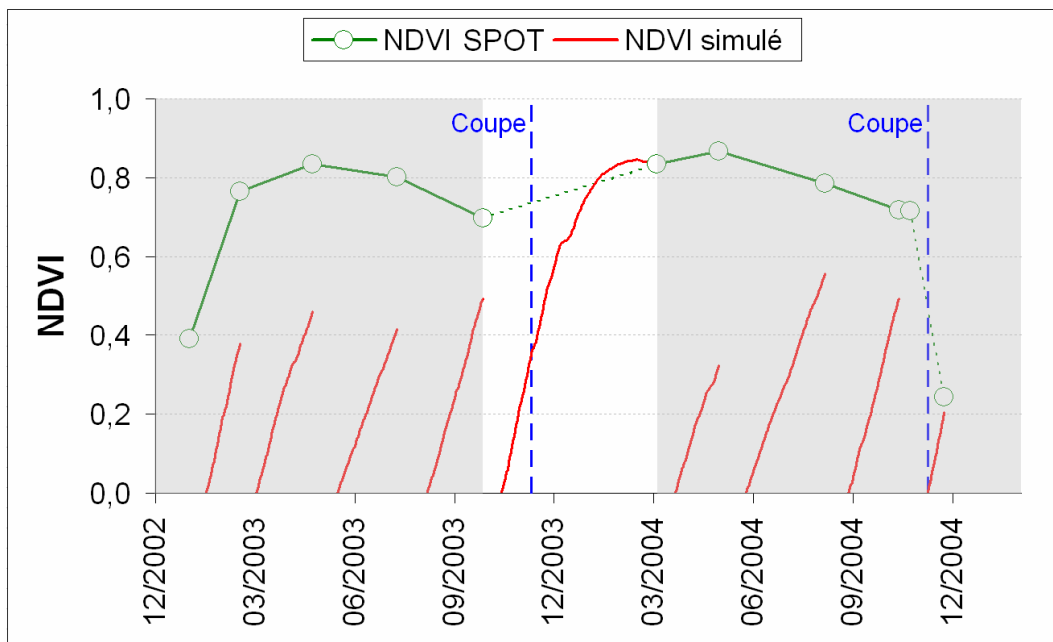


Figure 28. Un exemple illustrant comment les simulations de MOSICAS peuvent fournir une information utile en cas de manque d'acquisitions.

Les simulations du modèle de culture restent tout de même une source d'information imprécise et incertaine. L'imprécision de cette source est liée à sa nature (*i.e.* simulation) et à l'imprécision des paramètres environnementaux ; son incertitude est due principalement à l'incertitude du modèle. En ce qui concerne ce dernier point, nous tenons à signaler que le modèle MOSICAS a été calé pour la variété R570 cultivée à la Réunion en repousse sur des cycles maximum de 13 mois. Ainsi, les simulations faites par ce modèle pour les autres variétés sont incertaines, et l'information issue de ses simulations n'est fiable que pour les repousses.

Dans la Figure 29, nous illustrons un exemple de simulations de NDVI réalisées pour une parcelle appartenant à l'exploitation de M. Caruel. Cette parcelle est en repousse dans le cycle de 2003 et en plantation en 2004. En comparant les profils simulés pour cette parcelle avec ceux calculés à partir des images SPOT, nous pouvons constater que MOSICAS ne peut pas simuler correctement la croissance d'une plantation. Néanmoins, on peut se servir des simulations de ce modèle pour tirer une information sur la possibilité de coupe à la fin du cycle, que ce soit pour une parcelle en repousse ou en plantation.

D'autres exemples de profils temporels de NDVI simulés par MOSICAS et de NDVI SPOT sont recensés dans l'Annexe 5.

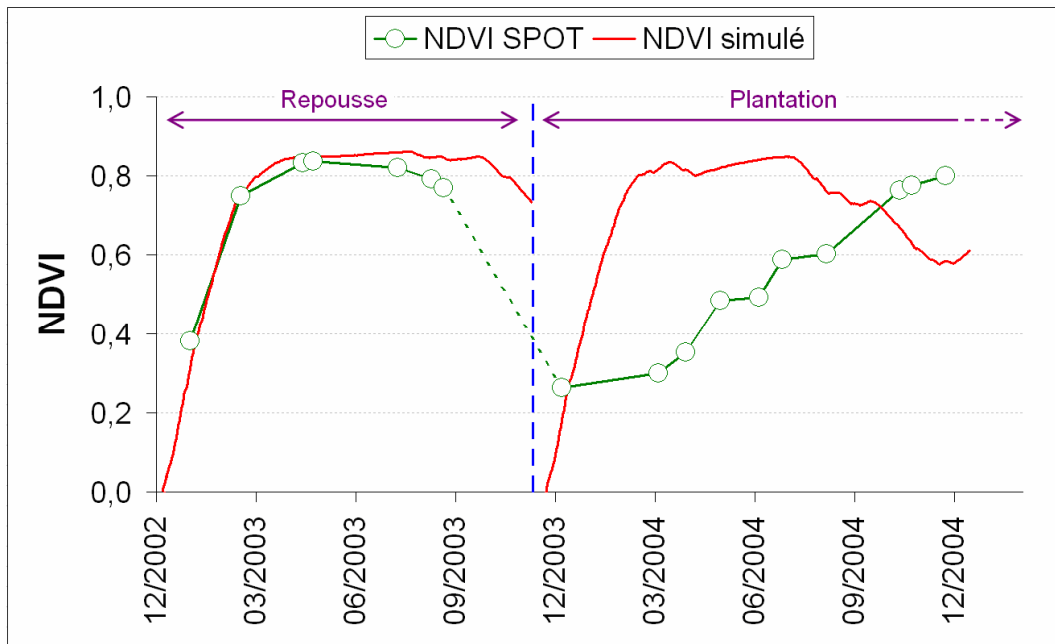


Figure 29. Comparaison entre les profils temporels de NDVI SPOT et ceux du NDVI simulés par MOSICAS pour une parcelle en repousse puis en plantation.

4.2.3. Connaissances expertes

La troisième source d'information est l'ensemble des connaissances agronomiques sur la canne à sucre. Ces connaissances permettent de mieux comprendre la relation entre les dynamiques des signaux extraits à partir des séries temporelles d'images et les changements des états de parcelle ; elles apportent également des contraintes temporelles très utiles, notamment en cas de manque d'information radiométrique. Des exemples de paramètres définissant ces contraintes temporelles sont la longueur nominale du cycle de canne (Figure 30.a) et les dates d'ouverture et de fermeture des usines sucrières qui déterminent les campagnes de récolte (Figure 30.b).

Une description du système de culture de la canne à sucre a été faite dans le chapitre 2. Nous rappelons ci-après les éléments déterminants pour la compréhension des règles de décision du système. La canne à sucre est une graminée vivace qui est cultivée sur une même parcelle durant plusieurs années consécutives. Après une première plantation, la canne est coupée au bout de 18 à 24 mois ; les tiges sont coupées au ras du sol et les souches laissées en place produisent de nouvelles pousses (*i.e.* des repousses). La récolte des repousses est annuelle dans la plupart des cas, et ce cycle annuel s'appelle cycle de récolte. Dans certains cas particuliers, et suite à des conditions climatiques exceptionnelles ou à des problèmes techniques affectant les usines, la coupe de certaines parcelles de canne (repousses et/ou

plantations) est reportée d'une année à une autre. A titre d'exemple, en 2007 le début de la campagne sucrière a été retardé de 3 semaines environ à cause de problèmes techniques dans les deux usines de l'île, par conséquent une dizaine de milliers de tonnes de cannes sont restées sur pied après la campagne, et leur coupe a été reportée à 2008. Au bout de 5 à 6 cycles de récoltes successives qui se font à partir de la même plantation, le rendement de la parcelle tend à diminuer et le champ est généralement replanté avec de nouvelles boutures, laissant place à un nouveau cycle cultural.

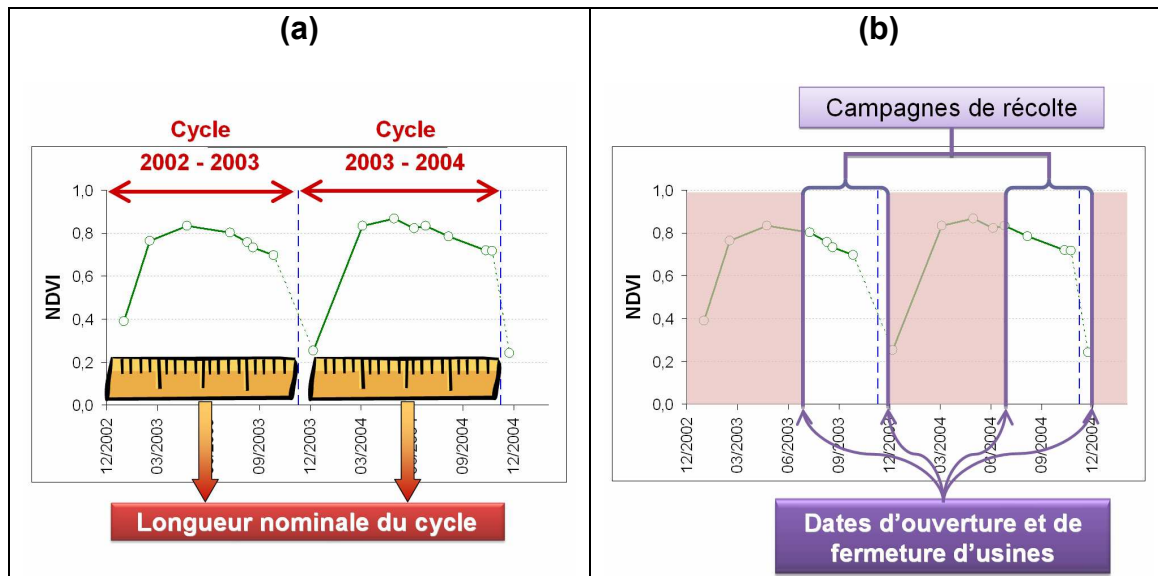


Figure 30. Des exemples de connaissances expertes qui permettent de définir des contraintes temporelles : a) la longueur nominale du cycle de canne ; b) les dates d'ouverture et de fermeture des usines sucrières.

Les connaissances expertes agronomiques et radiométriques sur la canne sont impliquées dans différentes parties du système d'aide à la décision que nous proposons. Elles sont utilisées pour :

- (1) définir les variables d'entrée qui sont nécessaires pour détecter la coupe ;
- (2) configurer les partitions de ces variables (Section 4.4.1);
- (3) intégrer des informations sur le cycle de la canne à sucre ;
- (4) définir des règles de décision (Section 4.4.3.1).

4.3. Formalisme

La gestion des informations multi-sources pour la prise des décisions nécessite la conception d'un cadre approprié qui définit l'utilisation de chaque source, identifie la technique de fusion d'informations et caractérise la sortie.

Dans cette thèse, les informations provenant des trois sources (série temporelle, modèle de culture et connaissances expertes) sont hétérogènes et de natures différentes. Les données issues du modèle de culture sont des « estimations » basées sur des paramètres qui ne sont pas directement reliés au phénomène étudié (*i.e.* la coupe). Les informations extraites à partir des images satellite sont des « mesures » de réflectance. Les connaissances expertes ne sont ni des « mesures » ni des « estimations », elles sont d'une autre nature.

En raison de sa capacité bien connue pour modéliser des concepts linguistiques et pour permettre le traitement d'informations de nature différentes, la logique floue a été choisie pour concevoir notre système d'aide à la décision. Ce formalisme permet de gérer l'imprécision des données, l'incertitude des sources et l'approximation du raisonnement mis en œuvre. En outre, grâce à l'inférence floue, des facteurs de confiance peuvent être associés aux sorties du système.

Les théories des ensembles flous et de possibilité (Zadeh, 1965; Zadeh, 1978) ont commencé à apparaître dans la littérature en télédétection il y a environ 20 ans. L'utilisation principale des systèmes flous a été dans la segmentation et la classification d'image *e.g.* (Gopal *et al.*, 1999; Tapia *et al.*, 2005; Li *et al.*, 2007), et dans l'implémentation des indices flous d'évaluation *e.g.* (Gill *et al.*, 2000; Laba *et al.*, 2008), basés sur le travail de (Woodcock and Gopal, 2000).

Le système d'aide à la décision que nous présentons dans ce chapitre est un système d'inférence floue. Les systèmes d'inférences sont l'une des applications les plus utilisées de la logique floue et des ensembles flous. Leur force se situe dans leur identité double (Guillaume, 2001) :

- d'une part, ils sont capables de modéliser et de manipuler des concepts linguistiques basés sur un langage de tous les jours, ce qui fait que ces systèmes sont faciles à comprendre et à interpréter même par des opérateurs non experts ;

- et d'autre part, ils sont des approximateurs universels qui peuvent réaliser une cartographie non linéaire entre les variables d'entrée et de sortie du système. Le processus de cartographie non linéaire peut être effectué par des procédures manuelles ou automatiques, tout en gardant à l'esprit l'originalité de la logique floue, *i.e.* son interprétabilité¹⁸.

Avant de présenter la structure du système d'inférence floue et de détailler ses composants, nous rappelons les concepts de base de la modélisation linguistique floue. Par ailleurs, un aperçu sur l'historique, les bases et les avantages de la logique floue est fourni dans l'Annexe 6.

4.3.1. Ensembles flous et termes linguistiques

Les ensembles flous permettent d'évaluer à quel point la réalisation d'un événement donné est possible. La définition des ensembles flous se fait par des distributions de possibilité (*i.e.* des fonctions d'appartenance); chaque ensemble flou a sa fonction d'appartenance. Ainsi, chaque point, x , dans l'univers appartient à un ensemble flou, A , avec un degré d'appartenance $\mu_A(x)$ qui est entre 0 et 1.

Les ensembles flous permettent de modéliser des informations imprécises et/ou incertaines avec des concepts linguistiques. Si A est un ensemble qui représente les valeurs *Fortes* de NDVI, le degré d'appartenance d'une valeur x de NDVI, $\mu_A(x)$, à cet ensemble peut être interprété comme étant le niveau jusqu'auquel cette valeur peut être considérée comme *Forte*. Plusieurs ensembles flous, *e.g.* *Faible*, *Moyenne* et *Forte*, peuvent être définis dans le même univers de la variable, comme illustré dans la Figure 31.

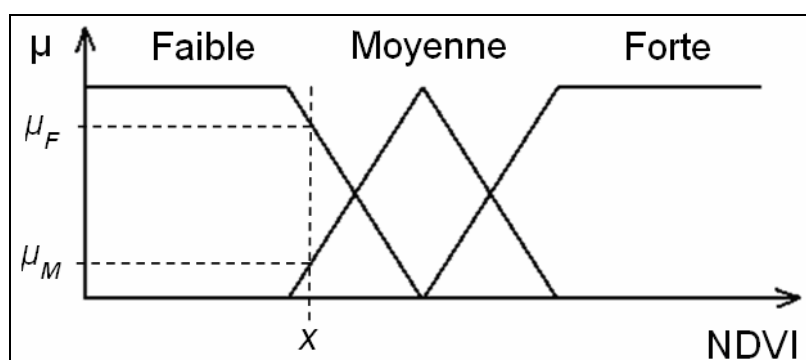


Figure 31. Exemple d'une partition de NDVI.

¹⁸ L'interprétabilité est la facilité de compréhension des relations d'entrées-sorties d'un système par un expert humain.

Les ensembles flous d'une variable se recouvrent habituellement, ainsi, une valeur donnée de cette variable est susceptible d'appartenir à plusieurs ensembles flous à la fois. Dans la partition représentée dans la Figure 31, la valeur x appartient à la fois aux ensembles flous *Faible* et *Moyenne* avec des degrés d'appartenance μ_F et μ_M respectivement. Dans cet exemple, une valeur donnée peut appartenir, avec un degré non nul, à deux ensembles flous maximum. De même, pour chaque point dans l'univers, la somme des degrés d'appartenance à tous les ensembles flous est égale à un. Ces deux caractéristiques ont été respectées lors de la définition des partitions dans notre système (Section 4.4.1).

4.3.2. Inférence à base de règles

L'utilisation des ensembles flous permet de gérer des informations de natures hétérogènes. En effet, en se basant sur les labels linguistiques de ces ensembles, on va pouvoir établir des règles de la forme « *Si je rencontre telle situation, Alors j'en tire telle conclusion* ». La situation, appelée prémisses ou antécédent de la règle, est définie par une combinaison de relations de la forme $x \text{ est } A$ pour chacune des informations hétérogènes en entrée. La partie conclusion de la règle est appelée sortie de la règle, et elle est de la forme $y \text{ est } C$:

$$\text{Règle } r : \text{ Si } x_1 \text{ est } A_1^{r'} \text{ et } x_2 \text{ est } A_2^{r'} \dots \text{ et } x_p \text{ est } A_p^{r'} \text{ Alors } y \text{ est } C^{r'} \quad (12)$$

avec $A_k^{r'}$ est l'ensemble flou de la variable d'entrée k utilisée dans la règle r et $C^{r'}$ est la conclusion de la règle.

Pour chaque règle du système d'inférence, un poids p^r (entre 0 et 1) est attribué pour désigner la certitude associée à cette règle. Par défaut, la valeur de p^r est égale à 1 pour toutes les règles dans la base. Une valeur de p^r de 0 signifie que la règle est absolument incertaine et donc non utilisable dans le processus d'inférence. Une valeur entre 0 et 1 signifie que la règle est utilisable mais incertaine ; plus la valeur est proche de 1, plus la règle est certaine.

Chaque règle participe au processus d'inférence selon son degré d'activation α^r . Ce dernier est calculé à partir des degrés d'appartenance des variables d'entrée, dans la prémisses de la règle, aux ensembles flous de ces variables.

Grâce au recouvrement des ensembles flous dans l'univers, une entrée donnée est susceptible d'activer plusieurs règles ($\alpha^r > 0$). Une méthode d'inférence doit être définie pour déterminer la manière dont le système attribue des poids aux conclusions des règles actives, et

la manière dont les poids de toutes les conclusions sont agrégés et affectés à la sortie du système.

Du fait que la logique floue autorise l'utilisation des concepts linguistiques via des ensembles flous, son utilisation dans les systèmes experts, sous la forme de règles de raisonnement flou, a permis de dépasser les limites de la logique du « tout ou rien », et cela grâce à la gradualité : on ne passe pas brutalement de « jeune » à « vieux » le jour de notre anniversaire, et c'est pareil pour beaucoup de phénomènes, de mesures ou d'estimations. Un autre atout du formalisme utilisé dans ce type de systèmes est la modularité ; en effet, il est possible d'ajouter des règles ou d'en enlever simplement selon les besoins de l'utilisateur. En définitive, ce formalisme fournit un outil intéressant pour fusionner des informations hétérogènes et tirer des conclusions associées à des facteurs de confiance.

4.4. Structure et composants

La structure du système d'aide à la décision est illustrée dans la Figure 32. Cette structure comporte quatre composants principaux : un étage d'entrée, une base de règles, un moteur d'inférence et une sortie. Le rôle et le fonctionnement de chacun de ces composants sont décrits dans les sous-sections suivantes.

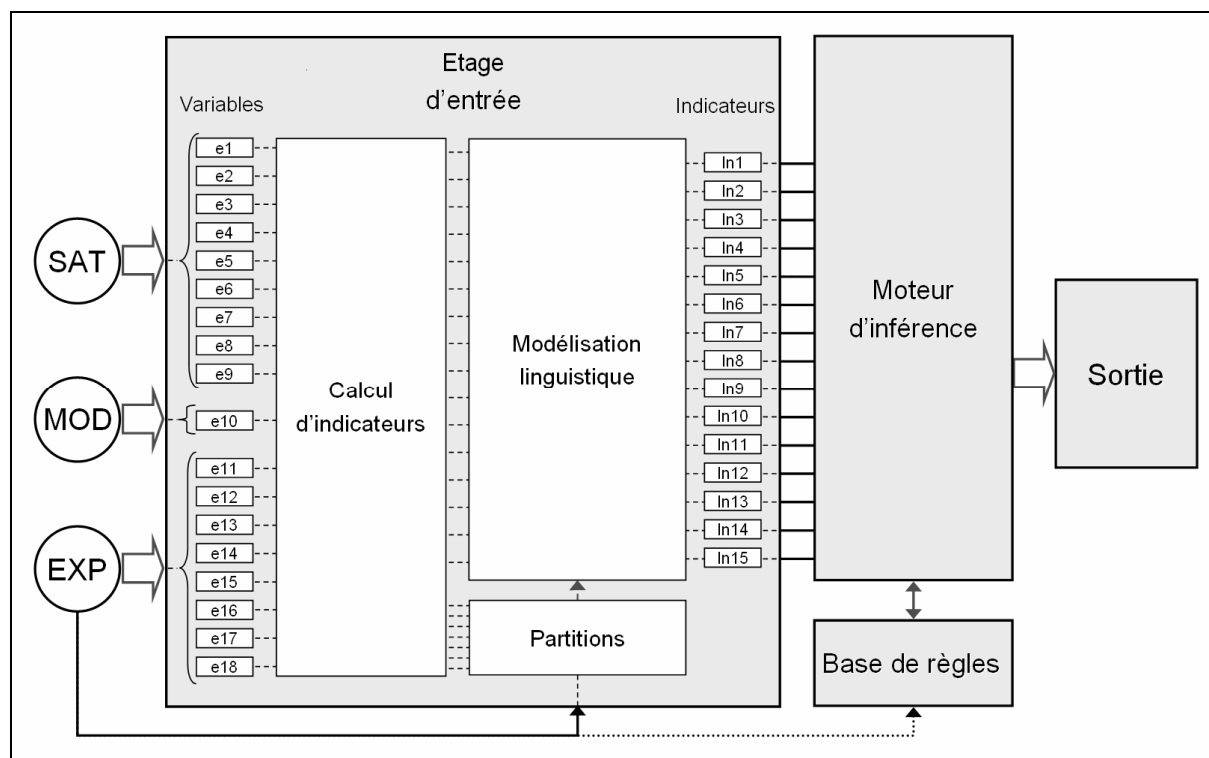


Figure 32. Structure et composants du système d'aide à la décision.

De manière schématique, on peut dire qu'à partir d'un jeu de données issu de la série temporelle d'images satellite (SAT), du modèle de culture (MOD) et des connaissances expertes (EXP), le système construira des indicateurs linguistiques et inférera des conclusions sur la possibilité de coupe d'une parcelle entre deux dates d'acquisitions satellites t et t' . Cette inférence se fait en utilisant une base de règles qui peut être créée manuellement, par expertise, ou automatiquement, avec un jeu d'apprentissage.

4.4.1. Etage d'entrée

L'étage d'entrée du système est le composant responsable de la construction d'indicateurs utiles pour la détection automatique des coupes. Dix-huit variables extraites à partir des trois

sources hétérogènes (SAT, MOD et EXP) sont utilisées dans ce module afin de créer quinze indicateurs associés à des labels linguistiques.

Les variables

Les variables extraites des trois sources sont recensées dans le Tableau 7. Le choix de ces variables a été fait par expertise¹⁹. Celles issues de SAT sont principalement liées aux profils temporels de NDVI et de MIR (valeurs et dates d'acquisitions) ; en effet ces profils s'avèrent être les plus intéressants, parmi les profils extraits à partir de SAT, pour la détection des coupes (Bappel, 2005; Lebourgeois *et al.*, 2007; Bégué *et al.*, 2008) (*c.f.* Section 4.2.1). Le MOD génère l'indicateur *Tn* afin de fournir une information sur la possibilité de coupe qui est principalement basée sur les données climatiques. Quant aux EXP, elles fournissent des seuils sur les différences bi-dates de NDVI et de MIR (utiles pour la détection des coupes) ainsi que des variables temporelles sur le cycle de la parcelle (*e.g.* la dernière date de coupe qui permet de calculer l'âge de la canne) et sur les campagnes de récolte (*i.e.* dates d'ouverture et de fermeture d'usines sucrières) ; en cas de non disponibilité de ces variables expertes, des valeurs par défaut sont prédéfinies (*e.g.* les dates d'ouverture et de fermeture des usines sucrières prédéfinies par défaut sont respectivement le 1^{er} juillet de l'année en cours et le 1^{er} janvier de l'année d'après).

Les indicateurs

Les indicateurs créés à partir des dix-huit variables, ou des combinaisons de ces variables sont détaillés dans le Tableau 8. Ces indicateurs peuvent être divisés en trois groupes principaux :

- Le premier groupe (In1 à In4) est basé sur les valeurs de NDVI et de MIR calculées à l'échelle de la parcelle à chaque date d'acquisition.
- Le second groupe (In5 à In9) concerne les dynamiques des profils temporels de NDVI et de MIR.
- Quant au troisième groupe (In10 à In14), il représente les contraintes temporelles issues du modèle agronomique et des connaissances expertes.

¹⁹ Expertise sur les propriétés optiques des couverts végétaux, sur le cycle de la canne à sucre et sur la détection des changements.

- Le dernier indicateur In15 indique tout simplement si la parcelle est couverte ou non par des nuages à la dernière acquisition satellite (à la date t).

Variable	Définition	Valeur	Source
e_1	Valeur de NDVI à l'échelle de la parcelle calculée à la date d'acquisition t .	$NDVI_t$	SAT
e_2	Valeur de NDVI à l'échelle de la parcelle calculée à la date d'acquisition t' .	$NDVI_{t'}$	
e_3	Valeurs de NDVI à l'échelle de la parcelle calculées aux dates d'acquisition t'' ($t'' < t'$).	$NDVI_{t''}$ (vecteur)	
e_4	Valeur du MIR à l'échelle de la parcelle calculée à la date d'acquisition t .	MIR_t	
e_5	Valeur du MIR à l'échelle de la parcelle calculée à la date d'acquisition t' .	$MIR_{t'}$	
e_6	Variable qui indique si la parcelle est dégagée ou non	$Nuage_t$	
e_7	Date d'acquisition t .	t	
e_8	Date d'acquisition t' . Il s'agit de la date d'acquisition dégagée qui précède t .	t'	
e_9	Dates d'acquisition t'' ($t'' < t'$).	t'' (vecteur)	
e_{10}	Valeur de l'indicateur de coupe issue du modèle de culture.	T_n	MOD
e_{11}	Seuil de la différence bi-temporelle de NDVI. (0,3 par défaut)	$\Delta NDVI_{Seuil}$	EXP
e_{12}	Seuil de la différence bi-temporelle de MIR. (15 par défaut)	ΔMIR_{Seuil}	
e_{13}	Date d'ouverture de l'usine sucrière déclenchant le début de la campagne de coupe de canne à sucre. (1 ^{er} juillet de l'année en cours par défaut)	T_O	
e_{14}	Date de fermeture de l'usine sucrière mettant fin à la campagne de coupe de canne à sucre. (1 ^{er} janvier de l'année d'après par défaut)	T_F	
e_{15}	Date d'ouverture précédente de l'usine sucrière déclenchant le début de la campagne de coupe précédente. (1 ^{er} juillet de l'année précédente par défaut)	T_{OP}	
e_{16}	Date de fermeture précédente de l'usine sucrière mettant fin à la campagne de coupe précédente. (1 ^{er} janvier de l'année en cours par défaut)	T_{FP}	
e_{17}	Dernière date de coupe de la parcelle. (1 ^{er} juillet de l'année précédente par défaut)	DDC	
e_{18}	Longueur nominale d'un cycle de canne. Il s'agit de l'âge à partir duquel une parcelle en repousse peut être coupée. (270 jours par défaut)	LNC	

Tableau 7. Définition des variables extraites à partir de la série temporelle d'images satellite (SAT), du modèle de culture (MOD) et des connaissances expertes (EXP).

La modélisation linguistique et les partitions

La modélisation linguistique des indicateurs est réalisée en se basant sur les connaissances expertes. En effet, ces connaissances permettent de faire une partition fiable et interprétable de l'univers de chaque indicateur par des ensembles (flous ou classiques) associés à des labels linguistiques.

A titre d'exemple, les ensembles flous des indicateurs liés aux valeurs de NDVI ont été conçus selon les connaissances expertes²⁰ sur le comportement du NDVI par rapport au changement des stades phénologiques et des états de parcelle de canne à sucre. Voici les éléments de raisonnement utilisés :

- Une valeur *Faible* de NDVI (inférieure à 0,30) correspond généralement à des résidus (*i.e.* pailis), à un sol nu après la coupe de la parcelle, ou à un couvert qui démarre (Bappel, 2005).
- Une valeur *Moyenne* de NDVI (entre 0,30 et 0,75) est observée pendant les phases de croissance et de sénescence. La phase de croissance est caractérisée par une accumulation de biomasse et par une croissance de la longueur des tiges. La sénescence est provoquée par le vieillissement de la canne, la diminution de la température et le manque d'eau (Martiné, 2003).
- Une valeur *Forte* de NDVI (supérieure à 0,75) est observée à la fin de la phase de croissance et avant la phase de sénescence. Dans cette phase on retrouve des parcelles de canne avec un couvert bien développé et vert.

La Figure 33 illustre les ensembles flous de cette partition ainsi qu'un exemple de profils de NDVI tracés en fonction du temps thermique²¹ (somme des degrés Celsius accumulés par jour depuis la coupe précédente) pour 28 parcelles dans l'exploitation de M. Caruel, en utilisant les séries d'images de 2003 et 2004. Le temps thermique renseigne de manière pertinente sur le stade de croissance de la canne à sucre (Fauconnier and Bassereau, 1970). Il est calculé selon l'expression suivante :

²⁰ Cette expertise a été développée en analysant des profils temporels de NDVI de plusieurs parcelles dans différents cycles pour faire le lien avec les évolutions phénologiques et les changements d'état de parcelle. L'analyse des profils a été accompagnée d'une phase de photo-interprétation d'images.

²¹ Nous avons choisi la dimension du temps thermique pour synchroniser les profils de NDVI et faciliter l'illustration de la partition. Nous rappelons que la forte variabilité spatio-temporelle des parcelles de canne entraîne une asynchronisation des profils temporels de NDVI.

$$Temps\ thermique = \sum_{j=t_0}^t \frac{T_{max}(j) - T_{min}(j)}{2} - T_{ref} \quad (13)$$

avec t_0 est la date de coupe précédente de la parcelle, t est la date d'acquisition de l'image à partir de laquelle le NDVI est calculé, $T_{max}(j)$ et $T_{min}(j)$ sont respectivement les températures maximales et minimales enregistrées par les stations météorologiques les plus proches du site d'étude à la date j (j allant de t_0 à t avec un pas journalier), et T_{ref} est la température de végétation de la canne à sucre ($T_{ref} = 12^\circ$).

Certaines partitions d'indicateur font appel à des variables extraites des sources d'information (e.g. T_n). Le Tableau 9 montre les intitulés et les descriptions des labels linguistiques de chaque indicateur (certains ayant des partitions similaires) ainsi que les marges d'ambiguïtés associées aux ensembles des indicateurs flous. En ce qui concerne ce dernier point, nous soulignons, à titre d'exemple, l'importance de la marge d'ambiguïté de ± 1 mois associée aux indicateurs utilisant T_n ; cette marge permet de faire face à l'imprécision du modèle et aux effets de variation climatique observés pour les valeurs élevées du seuil de NDVI (c.f. Section 4.2.2). Une illustration de toutes les partitions des indicateurs flous est faite dans l'Annexe 7. La définition des intervalles temporels utilisés pour la localisation des dates d'acquisitions satellite est présentée dans la Figure 34.

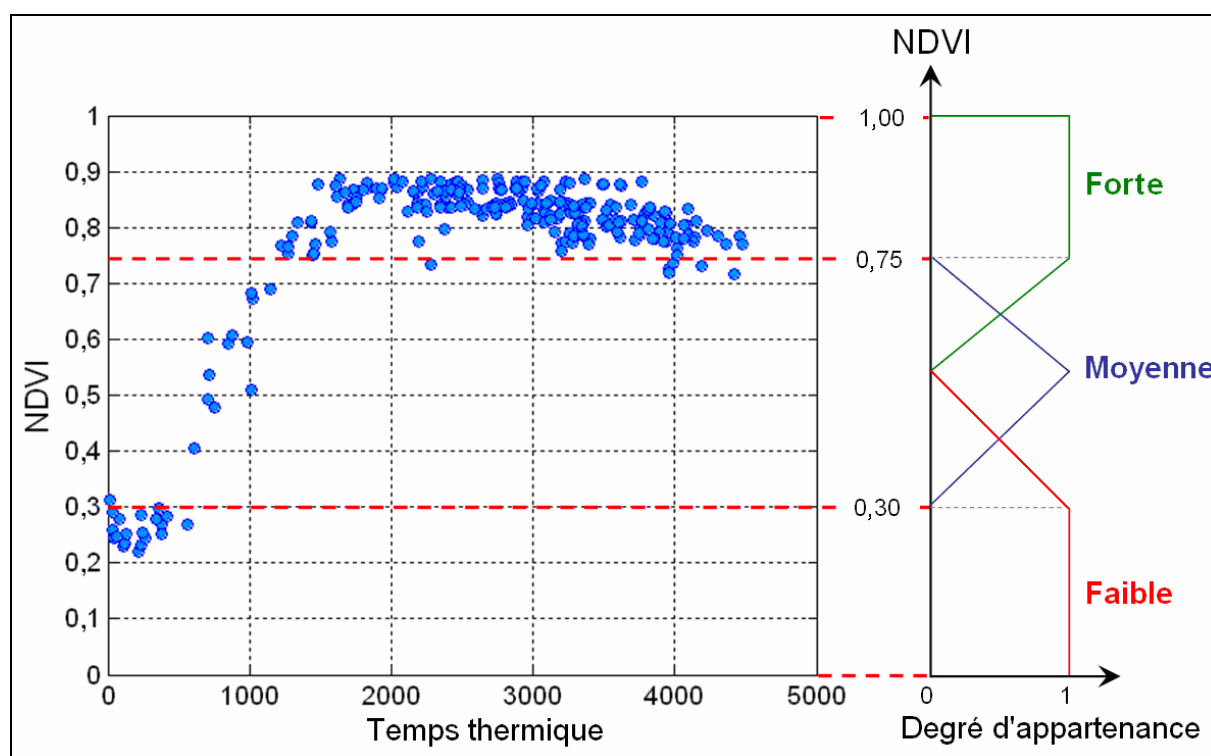


Figure 33. Ensembles flous du NDVI (i.e. indicateurs In1 et In2) avec des profils de NDVI de plusieurs parcelles en fonction de la somme de degrés-jour depuis la date de coupe.

Indicateur	Définition	Formule	Variable(s) impliquée(s)	Source(s)
In1	Valeurs de NDVI à l'échelle de la parcelle, calculées respectivement aux dates t et t' entre lesquelles le système doit décider si la parcelle a été coupée ou non.	$NDVI_t$	e_1	SAT
In2		$NDVI_{t'}$	e_2	SAT
In3	Valeurs de MIR à l'échelle de la parcelle, calculées respectivement aux dates t et t' .	MIR_t	e_4	SAT
In4		$MIR_{t'}$	e_5	SAT
In5	Résultat de la comparaison de la différence bi-temporelle de valeurs de NDVI calculées aux deux dates t et t' avec un seuil $\Delta NDVI_{Seuil}$.	$(NDVI_t - NDVI_{t'})$ vs. $\Delta NDVI_{Seuil}$	e_1, e_2, e_{11}^*	SAT + EXP
In6	Résultat de la comparaison de la différence bi-temporelle de valeurs de MIR calculées aux deux dates t et t' avec un seuil ΔMIR_{Seuil} .	$(MIR_t - MIR_{t'})$ vs. ΔMIR_{Seuil}	e_4, e_5, e_{12}^*	SAT + EXP
In7	Quantité d'acquisitions t'' précédant t' qui, à partir de la date de début de campagne, ont un gradient de NDVI négatif.	$Card(t'') \mid \frac{NDVI_{t'} - NDVI_{t''}}{t' - t''} < 0$	e_2, e_3, e_8, e_9	SAT
In8	Idem qu'In7, sauf que cet indicateur concerne le gradient positif.	$Card(t'') \mid \frac{NDVI_{t'} - NDVI_{t''}}{t' - t''} > 0$	e_2, e_3, e_8, e_9	SAT
In9	Quantité d'acquisitions t'' précédant t' qui, à partir de la date de début de campagne, ont une valeur <i>Fort</i> de NDVI.	$Card(t'') \mid NDVI_{t''}$ est " <i>Fort</i> "	e_3, e_9	SAT
In10	localisations des dates d'acquisition t et t' par rapport aux dates d'ouverture et de fermeture des campagnes.	t vs. $[T_{FP}; T_O] \cup [T_O; T_F]$	$e_7, e_{13}^*, e_{14}^*, e_{16}^*$	SAT + EXP
In11		t' vs. $[T_{OP}; T_{FP}] \cup [T_{FP}; T_O] \cup [T_O; T_F]$	$e_8, e_{13}^*, e_{14}^*, e_{15}^*, e_{16}$	SAT + EXP
In12	Résultat de la comparaison de la distance temporelle entre t et la dernière date de coupe DDC avec la longueur nominale d'un cycle de canne LNC .	$(t - DDC)$ vs. LNC	e_7, e_{17}, e_{17}^*	SAT + EXP
In13	Résultat de la comparaison de la distance temporelle séparant t et t' avec l'indicateur Tn issu du modèle de croissance. Si cette distance est supérieure à Tn alors la coupe de la parcelle est possible. Cet indicateur est utile lorsque t et t' sont dans la campagne de coupe.	$(t - t')$ vs. Tn	e_7, e_8, e_{11}^*	SAT + MOD
In14	Résultat de la comparaison de la différence entre t et la date du début de campagne courante T_O avec Tn . Si cette distance est supérieure à Tn alors la coupe de la parcelle est possible. Cet indicateur est utile lorsque t est dans la campagne de coupe et t' est avant la campagne.	$(t - T_O)$ vs. Tn	e_7, e_{13}, e_{11}^*	SAT + MOD + EXP
In15	Indicateur de présence de nuages sur l'image de la parcelle à la date d'acquisition t .	$Nuage$ vs. " <i>Oui</i> " ou " <i>Non</i> "	e_6	SAT

Tableau 8. Définition et formulation des indicateurs utilisés dans le système d'aide à la décision. Figurent aussi les variables impliquées dans chaque définition d'une façon directe ou indirecte (*), *i.e.* dans la partition, ainsi que les sources d'informations utilisées (SAT : Série temporelle d'images satellite ; MOD : Modèle de culture ; EXP : Connaissances expertes).

Indicateur	Labels linguistiques	Description	Marge d'ambiguïté
In1	« Faible » « Moyenne » « Forte »	Label de la valeur de NDVI de parcelle calculée à la date t . L'intitulé de ce label est «Faible » si la valeur de NDVI est inférieure à 0,3, il est « Moyenne » si elle est entre 0,3 et 0,75, et « Forte » si elle excède 0,75.	+/- 0,125 ; +/- 0,1
In2		Idem que pour In1 mais cette fois pour la valeur de NDVI de parcelle calculée à la date t' .	
In3	« Faible » « Forte »	Label de la valeur de MIR de parcelle calculée à la date t . L'intitulé de ce label est «Faible » si la valeur de MIR est inférieure à 20, et il est « Forte » si elle excède 22,5.	+/- 1,25
In4		Idem que pour In4 mais cette fois pour la valeur de MIR de parcelle calculée à la date t' .	
In5	« Inférieur au seuil de Δ NDVI » « Supérieur au seuil de Δ NDVI »	Label de la différence bi-temporelle de NDVI.	+ / - 0,1
In6	« Inférieur au seuil de Δ MIR » « Supérieur au seuil de Δ MIR »	Label de la différence bi-temporelle de MIR.	+ / - 2
In7	« Pas de t'' »	Label de la quantité des dates t'' ayant un gradient de NDVI négatif.	-
In8	« Pour au moins un t'' »	Label de la quantité des dates t'' ayant un gradient de NDVI positif.	
In9	« Pour la majorité de t'' » « Pour tout t'' »	Label de la quantité des dates t'' dont la valeur de NDVI est <i>Forte</i> .	
In10	« Entre campagnes » « Campagne actuelle »	Label de la localisation de la date d'acquisition t par rapport aux campagnes de coupe.	-
In11	« Entre campagnes » « Campagne actuelle » « Campagne précédente »	Label de la localisation de la date d'acquisition t' par rapport aux campagnes de coupe.	-
In12	« Inférieur à la longueur du cycle » « Supérieur à la longueur du cycle »	Label de la distance temporelle entre la date t et la dernière date de coupe.	+ / - 1 mois
In13	« Inférieur à T_n »	Label de la distance temporelle entre les dates t et t' .	+ / - 1 mois
In14	« Inférieur à T_n »	Label de la distance temporelle entre le début de campagne et la dates t .	+ / - 1 mois
In15	« Oui » « Non »	Label de la présence de nuages sur la parcelle.	-

Tableau 9. Intitulés et descriptions des labels linguistiques associés aux partitions des indicateurs, et marges d'ambiguïté (entre les ensembles) associées aux indicateurs flous.

En ce qui concerne la contribution des trois sources dans la préparation des indicateurs, remarquons que :

- SAT intervient dans la définition de tous les indicateurs (valeurs de NDVI et de MIR, dates d'acquisitions et présence de nuages) ;
- MOD contribue dans la partition de deux indicateurs (In12 et In13) importants pour la détection de coupe lors du manque d'acquisitions ;
- EXP fournit des variables qui sont utilisées directement dans la définition de la majorité des indicateurs (les dates de début et de fin de la campagne de coupe actuelle, T_O et T_F respectivement ; les dates de début et de fin de la campagne de coupe précédente, T_{OP} et T_{FP} respectivement ; la longueur nominale d'un cycle, LNC ; la dernière date de coupe, DDC), ainsi que des informations essentielles pour la définition des partitions.

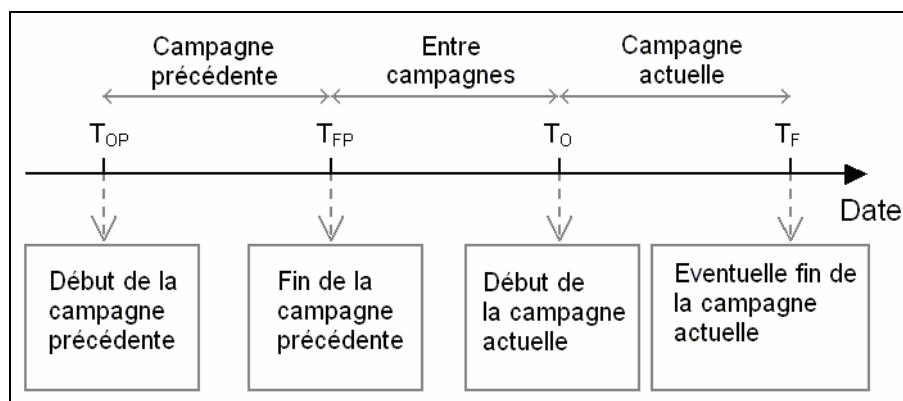


Figure 34. Intervalles temporels utilisés pour la localisation des dates d'acquisitions satellites.

4.4.2. Sortie

La variable de sortie du système devrait comporter principalement deux possibilités : « Coupe » et « Non coupe ». La première possibilité indique que la parcelle de la canne à sucre a été coupée entre les deux dates d'acquisitions satellites t et t' , et la deuxième indique qu'elle n'a pas été coupée entre ces deux dates.

Lors de la génération manuelle par expertise de la base de règles (*c.f.* Section 4.4.3.1), l'ajout de possibilités supplémentaires s'est avéré nécessaire pour donner une conclusion fiable pour tous les scénarios : il s'agit des possibilités « Je ne sais pas », « Plutôt coupe » et

« Plutôt non coupe ». Les deux dernières possibilités ont pu être remplacées par les possibilités basiques « Coupe » et « Non coupe » respectivement tout en attribuant aux règles qui les utilisent des poids p' inférieurs à 1 (c.f. Section 4.3.2). La valeur de ces poids a été fixée à 0,75. En d'autres termes, l'incertitude de la conclusion de l'expert contenue dans le mot « Plutôt » a été attribuée à la règle toute entière, tout en utilisant les conclusions de bases « Coupe » et « Non coupe ». La possibilité « Je ne sais pas » a été adoptée. Ainsi, les conclusions sur le statut d'une parcelle de canne dans la version du système d'aide à la décision, employant la base de règles expertes, sont : « Coupe » (C), « Non coupe » (NC) et « Je ne sais pas » (JSP).

Dans tous les cas, la sortie du système d'aide à la décision est constituée de niveaux de possibilité des différentes conclusions : e.g. dans la version avec des règles expertes, la sortie est μ_C , μ_{NC} et μ_{JSP} (Figure 35).

Les niveaux de possibilités de la sortie varient entre 0 et 1. Ils peuvent être considérés comme des valeurs de confiance assignées aux différentes conclusions.

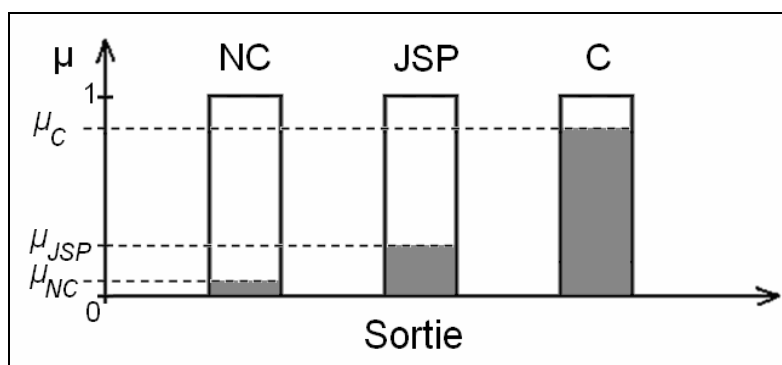


Figure 35. Sortie du système d'aide à la décision avec la base de règles expertes (NC = Non Coupe ; JSP = Je ne Sais Pas ; C = Coupe).

4.4.3. Base de règles

La base de règles du système oriente le processus d'inférence en utilisant les termes linguistiques associés aux indicateurs d'entrée et ceux associés à la variable de sortie. Cette base regroupe une collection de règles de type «Si-Alors» (équation 12).

Deux voies ont été abordées pour la construction de la base de règles. La première consiste à générer les règles manuellement en s'appuyant sur les connaissances expertes. La deuxième consiste à générer les règles automatiquement en utilisant un jeu d'apprentissage.

4.4.3.1. Génération manuelle par expertise

Nous avons défini une base de règles linguistiques en nous basant sur les connaissances expertes sur le profil temporel de NDVI d'une parcelle de canne. Les règles dans cette base, incluent des scénarios différents qui visent à couvrir « tous » les cas possibles que le système peut confronter pour décider si une coupe de parcelle a eu lieu entre les acquisitions t et t' . La génération des règles a été faite en utilisant 12 indicateurs parmi les 15 développés : les indicateurs liés au MIR (In3, In4 et In6) ont été ignorés à ce stade parce que les informations qu'ils apportent ont été considérées comme redondantes avec celles issues du NDVI.

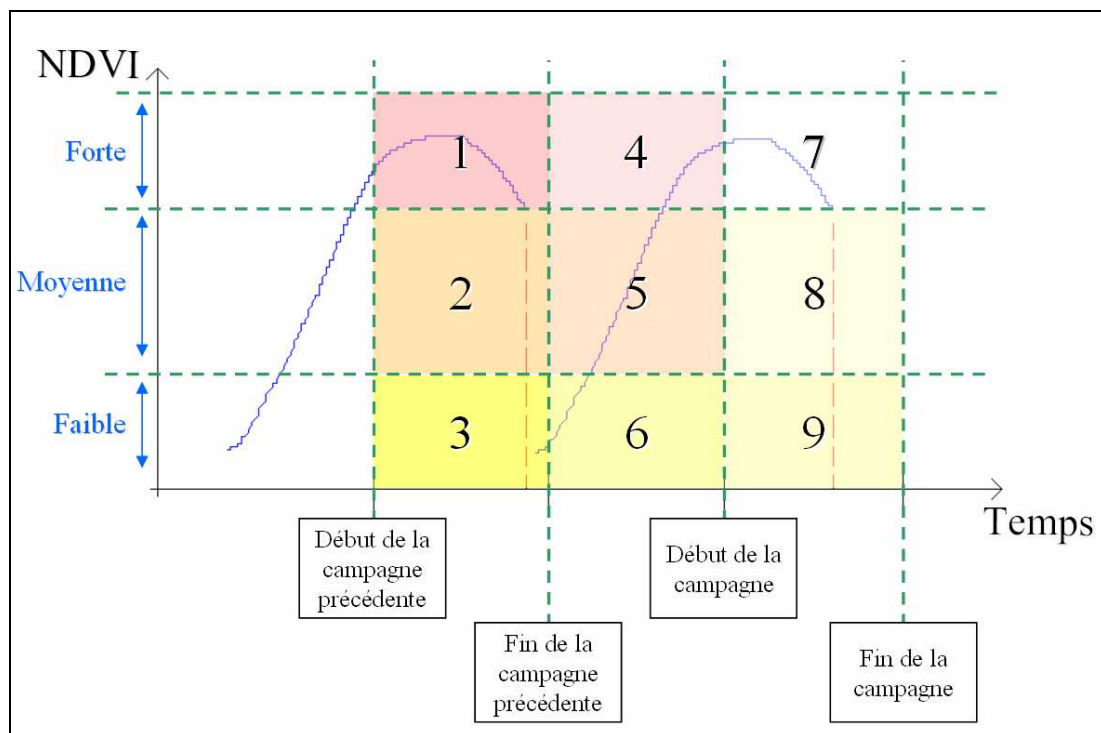


Figure 36. Sous-zones (1 à 9) de profil temporel de NDVI.

Conception des règles

Afin de créer les règles, nous avons découpé le profil temporel de NDVI en neuf zones (Figure 36) en nous basant sur les trois classes de valeur de NDVI (*Faible*, *Moyenne* et *Forte*) et sur les trois intervalles temporels correspondant à la campagne de récolte précédente, la campagne actuelle et à la période entre les campagnes. Cette décomposition a permis d'imaginer des scénarios qui prennent en compte toutes les positions possibles que le couple de points NDVI (t) et NDVI (t') peut avoir. A chaque combinaison, un ensemble de

règles a été défini en faisant appel à tout ou partie des indicateurs pour conclure sur la possibilité de coupe.

Un couple de points (NDVI (t), NDVI (t')) peut avoir 45 combinaisons possibles dans la grille définie dans la Figure 36 : 9 possibilités²² dans l'espace (NDVI x NDVI), multiplié par 5 possibilités²³ dans l'espace (Temps x Temps). Deux règles en moyenne ont été générées pour chaque combinaison, ce qui a produit un total de 90 règles.

D'autres règles ont été définies pour gérer le cas où la parcelle à la date t est couverte par un nuage (*i.e.* masquée). Ces règles sont basées sur l'historique du NDVI, et sur les informations provenant du modèle de culture et/ou des connaissances expertes. Leurs conclusions sont incertaines, *i.e.* elles sont toujours associées au terme « Plutôt ». Le nombre de ces règles est 26 : il s'agit d'environ 2 règles en moyenne pour 15 possibilités de base (3 possibilités dans la dimension NDVI multiplié par 5 possibilités dans la dimension temporelle).

Le nombre total de règles définies par expertise est 116 : 56 règles concernent des scénarios de coupe, 37 correspondent à des scénarios de non coupe, et 23 sont liées à des situations dont les conclusions sont inconnues (*i.e.* « Je ne sais pas »).

La base de règles expertes définie dans cette thèse ne peut pas être considérée comme unique. Un autre expert aurait en effet pu concevoir les scénarios différemment et imaginé d'autres règles.

Exemples de règles

La base de règles expertes est fournie dans l'Annexe 8. Nous montrons dans le Tableau 10 quelques exemples de règles avec des illustrations indiquant les positions des points NDVI (t), NDVI (t') et dans certain cas NDVI (t''). Ci-après les explications :

- Le premier exemple, r1, considère le scénario banal dans lequel les deux points NDVI (t) et NDVI (t') sont dans la période séparant deux campagnes de récolte ; dans ce cas, on peut directement conclure que la parcelle ne peut pas être coupée entre t et t'.

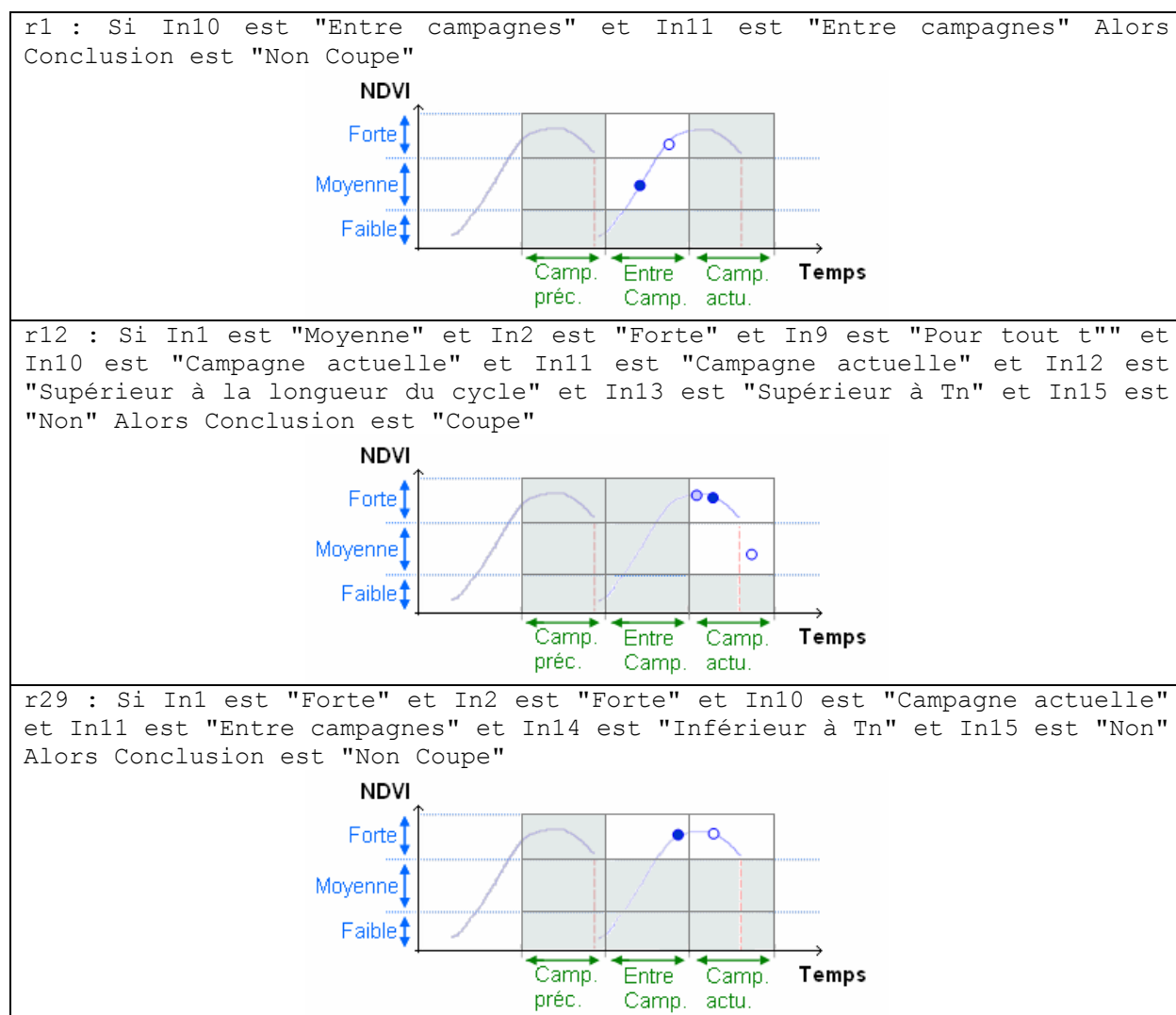
²² Chaque point NDVI(t) et NDVI(t') peut appartenir aux sous-espaces « Faible », « Moyenne » ou « Forte ».

²³ NDVI(t') peut appartenir aux trois sous-espaces temporels, et NDVI(t) à deux seulement (à partir de « Fin de campagne précédente »).

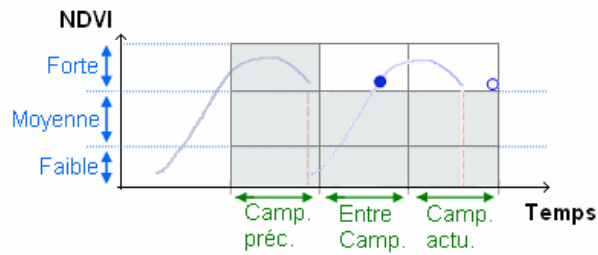
- Le deuxième exemple, r12, représente un scénario dans lequel les deux dates t et t' sont dans une campagne de récolte, et les valeurs de NDVI de la parcelle à ces deux dates sont respectivement *Moyenne* et *Forte* ; pour conclure dans une telle situation, il est important de faire appel aux indicateurs concernant l'historique du NDVI, à l'indicateur issu du modèle de culture et à celui provenant des connaissances expertes en ce qui concerne l'âge de la canne. Ainsi, si l'historique montre que la canne était dans une phase de maximum de croissance avant la date t' , et si le modèle dit qu'il est possible que la canne soit coupée entre t et t' , et que la connaissance experte sur l'âge de la canne le confirme, la conclusion logique sera qu'il s'agit d'une coupe.
- Les troisième et quatrième exemples, r29 et r30, concernent des situations dans lesquelles t est dans une campagne de récolte, t' dans la période entre campagnes et les valeurs de NDVI à ces deux dates sont *Fortes*. On remarque dans les illustrations correspondant à ces règles, que malgré le fait que les positions des points par rapport aux zones prédéfinies de l'espace NDVI-temps (Figure 36) sont les mêmes, l'espacement (*i.e.* la distance temporelle) des points peut engendrer des conclusions différentes. Dans le premier cas (*i.e.* r29), les deux points sont proches, ainsi, il suffit que le modèle de culture affirme qu'une coupe n'est pas possible entre t et t' pour adopter cette conclusion. Dans le deuxième cas (*i.e.* r30), les points sont plus éloignés, donc ainsi, pour conclure sur la coupe, il faut que le modèle l'affirme et que la connaissance experte sur l'âge de la canne le confirme.
- Le cinquième exemple, r68, considère la situation où t et t' sont respectivement dans une période entre campagnes et dans la campagne de récolte précédente, et les valeurs de NDVI à ces dates sont respectivement *Forte* et *Moyenne* ; pour pouvoir conclure, il faut regarder l'historique de NDVI, consulter l'avis du modèle de culture ainsi que celui de la connaissance experte sur l'âge de la canne. Si l'historique indique que la canne était dans une phase de sénescence, et le modèle dit qu'une coupe entre le début de campagne et la date t est possible et que l'âge de la canne le confirme, la conclusion est coupe.
- Le sixième exemple, r88, simule une situation dans laquelle on ne peut pas conclure par coupe ou non coupe ; le point NDVI (t) est dans la période entre campagnes avec une valeur *Faible*, le point NDVI (t') est dans la campagne de récolte précédente avec une valeur *Moyenne*, si la connaissance experte sur l'âge considère que la canne n'est pas

encore prête pour être coupée, on ne peut pas conclure à cause de la contradiction entre l'information images et l'information expertise.

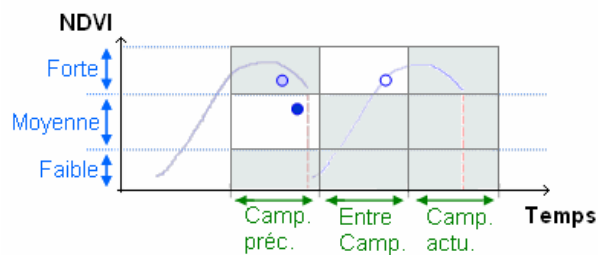
- Le septième exemple, r113, montre une règle qui propose une conclusion malgré la présence d'un nuage à la date t . Dans le scénario de cette règle, nous supposons que lors d'une acquisition à la date t entre les campagnes la parcelle est couverte par un nuage, et la valeur de NDVI dans l'image d'avant, qui est acquise dans la campagne précédente, est *Moyenne*. Dans ce cas nous considérons que si l'âge de la canne jusqu'à la date t est supérieur à la longueur du cycle et s'il existe au moins une image à t' dans laquelle la valeur de NDVI est *Forte*, alors la parcelle est plutôt coupée entre t et t' .



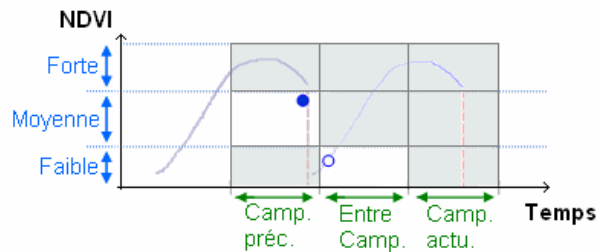
r30 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"



r68 : Si In1 est "Forte" et In2 est "Forte" et In9 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"



r88 : Si In1 est "Faible" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Je ne sais pas"



r113 : Si In2 est "Moyenne" et In7 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

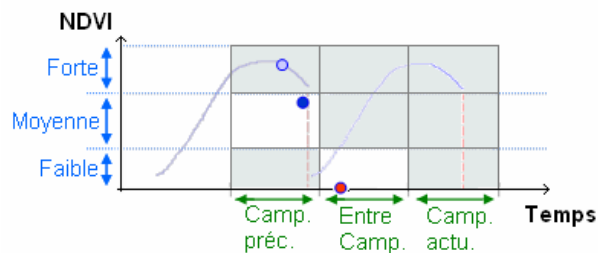


Tableau 10. Exemples de règles définies par expertise, avec des illustrations indiquant respectivement en blanc, bleu foncé et bleu clair les positions des points NDVI (t), NDVI (t') et NDVI (t''), et en rouge les acquisitions avec nuages.

4.4.3.2. Génération automatique par apprentissage

La génération automatique des règles (*i.e.* induction) a été faite avec le logiciel Fispro[®] (Guillaume *et al.*, 2002) en tenant compte de tous les indicateurs. Cette induction est fondée sur la construction d'un arbre de décision flou en utilisant un jeu d'apprentissage.

Les arbres de décision flous sont une extension des arbres de décision classiques (Breiman *et al.*, 1984; Quinlan, 1986). Ils sont composés d'une racine, qui est le sommet ou point de départ de l'arbre, et de nœuds. Les nœuds terminaux sont appelés feuilles de l'arbre. Chaque nœud correspond à un test sur un indicateur d'entrée (variable explicative). Ces éléments sont déterminés de façon à avoir une homogénéité maximale des exemples appartenant au nœud, par rapport à la variable à expliquer (variable de sortie). La mesure d'hétérogénéité est l'entropie dans les nœuds que l'algorithme cherche à minimiser. Les chemins allant de la racine aux feuilles peuvent être interprétés de façon naturelle comme des règles de décision floues (Figure 37). Les arbres de décision flous proposés dans Fispro sont basés sur une implémentation floue de l'algorithme ID3 (Ichihashi *et al.*, 1996). Le partitionnement de chaque indicateur d'entrée doit être réalisé avant de lancer l'algorithme. Nous avons utilisé les mêmes partitions que précédemment (Section 4.4.1).

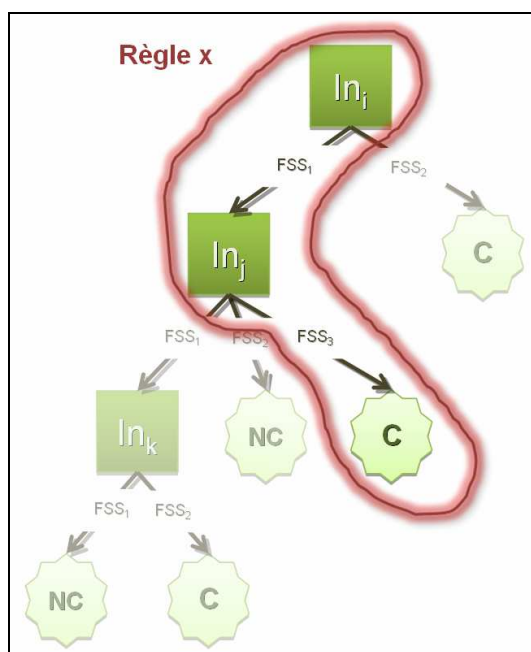


Figure 37. Schéma général d'un arbre de décision flou. Les nœuds In_i, \dots, In_k sont des sous-ensembles de valeurs d'indicateurs, les FSS sont les sous-ensembles flous utilisés dans la partition de chaque indicateur, et C (coupe) et NC (Non Coupe) sont les possibilités de la sortie. Chaque chemin entre la racine et une feuille constitue une règle de décision (*e.g.* Règle x : Si In_i est FSS_1 et In_j est FSS_3 Alors C).

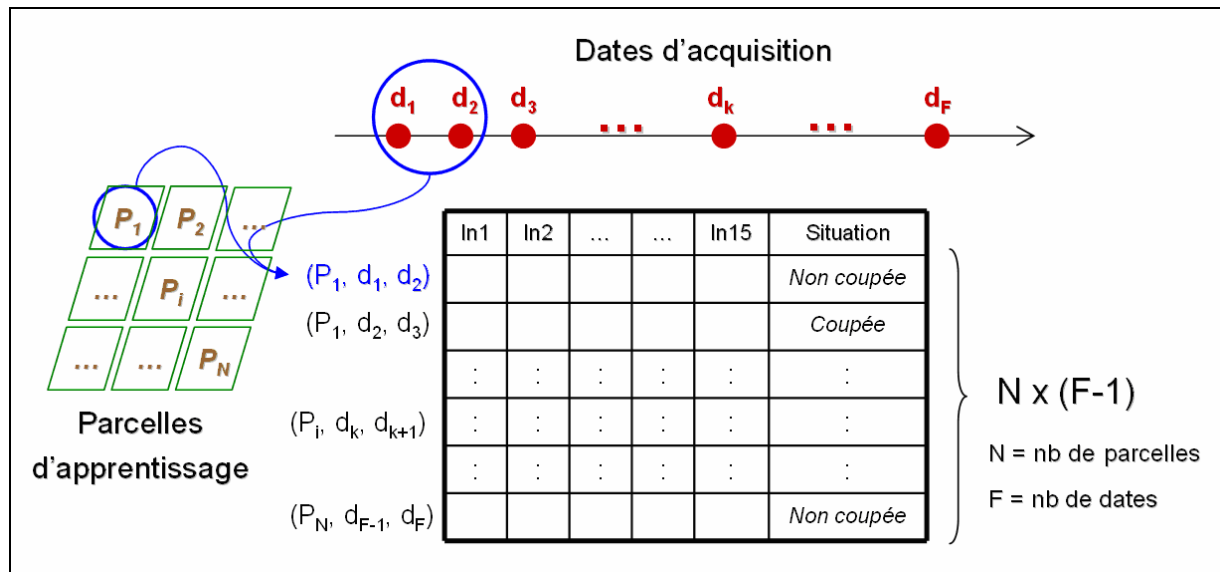


Figure 38. Illustration du processus de construction de jeu d'apprentissage pour la génération automatique des règles.

En ce qui concerne le jeu d'apprentissage, il s'agit d'un tableau de données qui contient pour une série d'images donnée, les valeurs des indicateurs calculés pour un certain nombre de parcelles, dites d'apprentissage, pour chaque couple de dates d'acquisition (Figure 38) ; ce tableau contient également la situation de la parcelle (« coupée » ou « non coupée ») entre les deux dates d'acquisition en question. La situation de la parcelle peut être issue d'une base de vérité terrain, ou bien définie par photo-interprétation.

Malgré le fait qu'elles soient générées automatiquement, les règles induites avec les arbres de décision flous sont interprétables. Elles sont de la même forme et définies avec les mêmes partitions que les règles expertes. La Figure 39 montre un exemple d'une règle induite par Fispro[®] ; cette règle peut être interprétée facilement grâce aux labels linguistiques associés aux indicateurs d'entrée et à la variable de sortie. En interprétant chaque règle induite on peut comprendre les scénarios auxquels elles sont liées, et évaluer ainsi leur aspect générique.

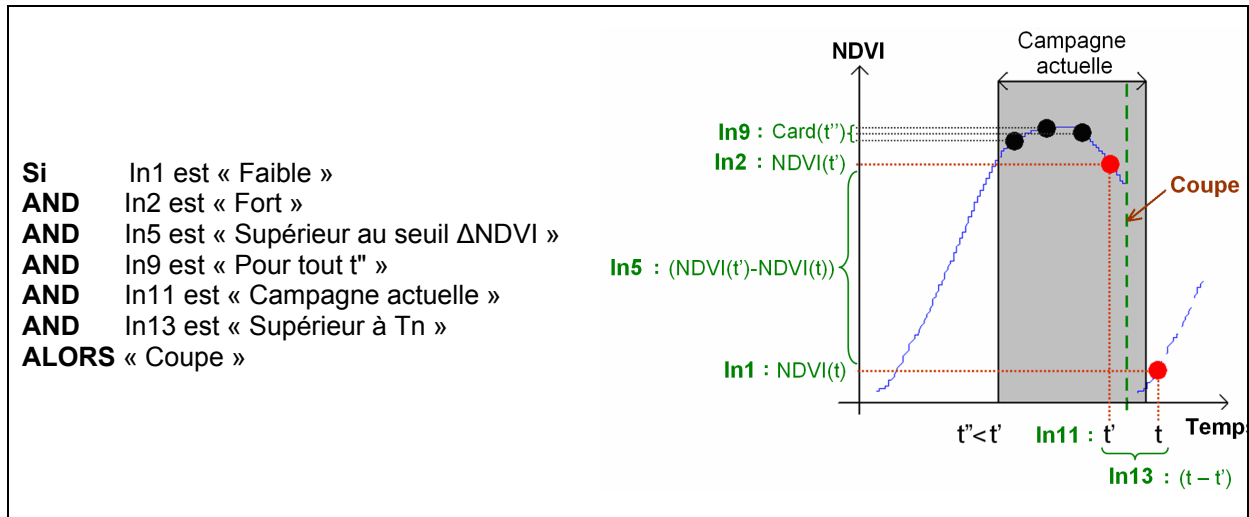


Figure 39. Exemple d'une règle induite par apprentissage et illustration du scénario correspondant.

4.4.4. Moteur d'inférence

Comme indiqué dans la section 4.3.2, un jeu de données en entrée est susceptible d'activer plusieurs règles, et ce grâce au recouvrement des ensembles flous dans les partitions des indicateurs. Le degré d'activation α^r d'une règle r est calculé en combinant les degrés d'appartenance de la prémisse de la règle d'une manière conjonctive :

$$\alpha^r = \mu_{A_1^r}(x_1) \wedge \mu_{A_2^r}(x_2) \wedge \dots \wedge \mu_{A_p^r}(x_p) \quad (14)$$

avec $\mu_{A_j^r}(x_j)$ est le degré d'appartenance de la valeur x_j de l'indicateur In_j à l'ensemble flou A_j^r , et \wedge un opérateur conjonctif. Les opérateurs conjonctifs les plus utilisés dans la logique floue sont le *minimum* (min) et le *produit*.

Le poids w^r attribué à la conclusion C^r d'une règle active r est calculé en prenant le minimum entre le poids p^r associée à cette règle (c.f. Section 4.3.2) et son degré d'activation α^r :

$$w^r(C^r) = \min(\alpha^r, p^r) \quad (15)$$

L'agrégation des poids associés aux conclusions de toutes les règles actives pour les m conclusions distinctes du système est faite d'une manière disjonctive :

$$\begin{cases} \forall j = 1, \dots, m \\ \mu_j = \left\{ \vee [w_r(C^r)] \mid C^r = j \right\} \end{cases} \quad (16)$$

avec \vee est un opérateur disjonctif comme le *maximum* (max) ou la *somme bornée*²⁴.

La technique d'inférence utilisée dans notre système est basée sur la méthode de Mamdani (Mamdani and Assilian, 1975). Ainsi, l'opérateur de conjonction employé pour calculer le degré d'activation de chaque règle est le *minimum*, et celui utilisé pour faire la disjonction des poids des conclusions est le *maximum*.

La sortie de l'inférence est le niveau de possibilité μ_j garantie par la base des règles et affecté à chaque conclusion possible j . Le processus d'inférence selon la méthode de Mamdani est illustré sur un exemple simple dans la Figure 40.

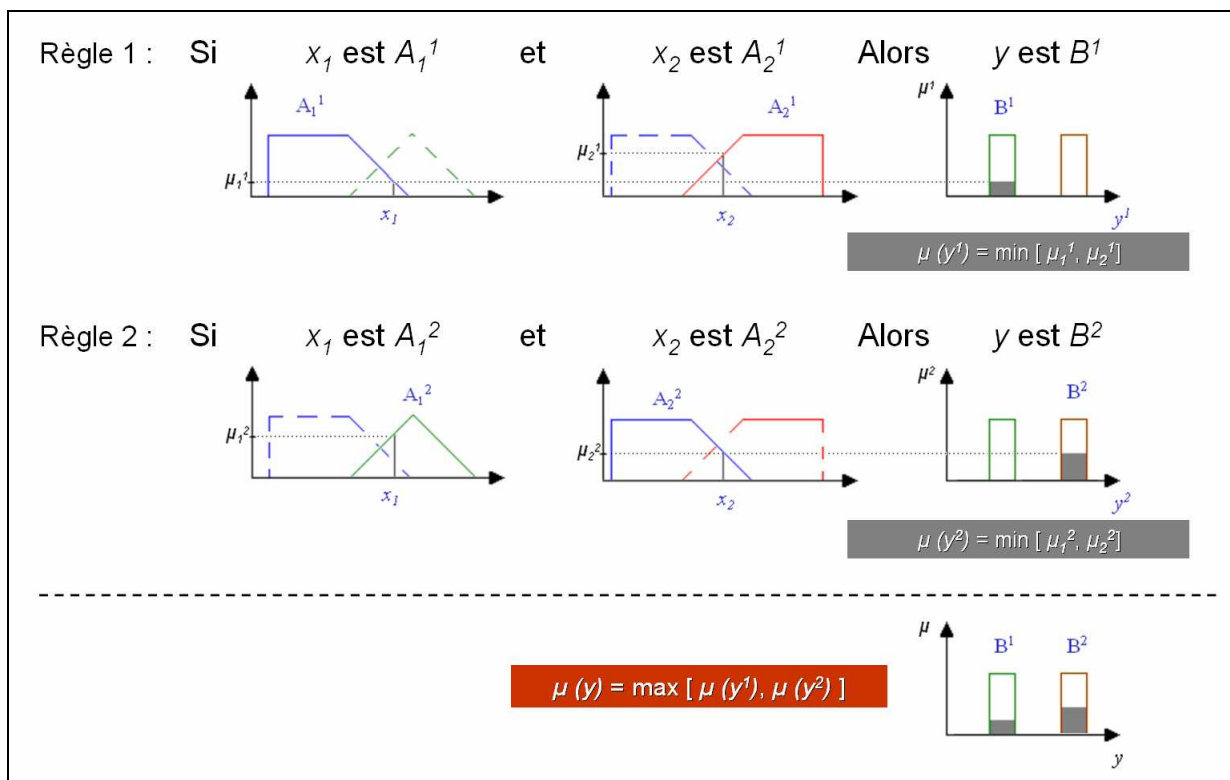


Figure 40. Illustration de la méthode d'inférence de Mamdani sur un exemple avec deux règles seulement et deux indicateurs d'entrée dont les valeurs sont respectivement x_1 et x_2 .

²⁴ $\max(1, \sum w_r(C^r))$.

4.5. Prise de décision

En utilisant la sortie du système, *i.e.* les différents niveaux des possibilités, l'utilisateur a le choix entre plusieurs méthodes de décision. Ce choix doit être guidé par l'application qui est en jeu.

4.5.1. Système à base de règles expertes

Avec la première version du système (à base de règles expertes), une première méthode de décision peut consister à choisir la possibilité qui a le niveau le plus élevé (équation 17) ; dans ce cas l'utilisateur est considéré comme étant « prudent » car il préfère prendre la décision la plus possible même s'il s'agit d'une « Je ne sais pas » (JSP).

$$Decision = \arg \max(\mu_C, \mu_{NC}, \mu_{JSP}) \quad (17)$$

L'utilisateur peut aussi être « exigeant » et définir une méthode de décision qui garantit un certain seuil de confiance μ_{conf} (équation 18). Selon cette méthode, pour décider si une parcelle a été coupée ou non, il faut tout d'abord que μ_C ou μ_{NC} soit supérieur à μ_{JSP} , *i.e.* qu'il y ait plus d'évidence pour soutenir une décision de coupe ou de non coupe que pour marquer le statut de la parcelle comme inconnu. Il faut en plus que μ_C ou μ_{NC} soit supérieur au seuil μ_{conf} choisi.

$$Decision = \begin{cases} \arg \max(\mu_C, \mu_{NC}) & \text{si } (\max(\mu_C, \mu_{NC}) \geq \max(\mu_{JSP}, \mu_{conf})) \\ JSP & \text{ailleurs} \end{cases} \quad (18)$$

Dans le cas où l'utilisateur est « pragmatique » et qu'il ne souhaite pas obtenir une carte de coupes avec des parcelles dont le statut est « Je ne sais pas », il peut tout simplement ignorer la possibilité « Je ne sais pas » et choisir entre « Coupe » et « Non coupe » celle qui a le maximum de niveau de possibilité (équation 19).

$$Decision = \arg \max(\mu_C, \mu_{NC}) \quad (19)$$

La méthode de décision que nous utiliserons dans les applications de cette version dans le chapitre suivant sera basée sur le cas de l'utilisateur « exigeant » (équation 18). Nous considérons le cas où il faut produire des cartes de coupe avec un seuil de confiance garanti. Cette méthode autorise des décisions de type « Je ne sais pas ».

Continuité temporelle

Les décisions prises par le système peuvent servir, selon le souhait de l'utilisateur²⁵, à actualiser la dernière date de coupe (DDC) de la parcelle pour les prochaines décisions ; cette mise à jour va concerner l'indicateur In12 lié à l'âge de la canne dans la parcelle et à la longueur nominale du cycle (Tableau 8).

Lorsqu'une décision « Je ne sais » est prise, deux cas, selon la vérité terrain, sont possibles :

- La parcelle n'est pas coupée entre t et t' : dans ce cas la seule erreur commise par le système est une omission de la détection d'une « Non coupe ». Cette omission n'aura aucun impact sur la suite.
- La parcelle est coupée entre t et t' : dans ce cas le système (i) omet la détection de la « Coupe », (ii) commettra une fausse détection de « Coupe » dans les images qui suivent parce que l'information sur la DDC n'est pas mise à jour, ce qui favorise la détection de coupe, et (iii) une fois que la fausse détection de coupe est faite, la mise à jour de la DDC sera biaisée, mais ceci ne doit normalement pas trop impacter les prochaines décisions. En effet, grâce à la partition floue concernant la longueur du cycle, le système dispose d'une certaine flexibilité (on n'est pas à 1 mois près). D'autre part, vu que les règles tiennent compte de la variabilité spatio-temporelles des parcelles (on sait que la date de coupe peut avoir lieu à n'importe quel moment pendant les 6 mois que dure la récolte) le système va favoriser les bonnes possibilités (peut être avec un niveau moins important que dans le cas normal) en se basant sur les autres indicateurs qui eux ne sont pas biaisés.

Nous signalons qu'à tout moment, l'utilisateur peut intervenir pour mettre à jour la DDC ou n'importe quelles variables liées aux connaissances expertes (Tableau 7).

²⁵ L'utilisateur a également la possibilité de fournir au système la dernière date de coupe réelle. Dans ce cas, cette information fait partie des connaissances expertes qui alimentent le système.

4.5.2. Système à base de règles induites

La sortie de la deuxième version du système ne contient que deux possibilités « Coupe » (C) ou « Non coupe » (NC). La seule approche décisionnelle logique dans ce cas consiste à choisir la possibilité qui a le degré d'appartenance le plus élevé (équation 19).

4.6. Conclusion

Dans ce chapitre, nous avons présenté un système d'aide à la décision pour la détection automatique de coupe de canne à sucre à La Réunion. Ce système traite les informations issues des séries temporelles d'images satellite en intégrant les sorties d'un modèle de croissance de canne à sucre et des connaissances expertes.

La façon dont nous utilisons le modèle de culture et les connaissances expertes représente une innovation dans le domaine du traitement de l'imagerie multi-temporelle. En effet, le rôle de ces sources exogènes consiste à aider le suivi des coupes à partir des images satellite (1) en compensant le manque de données, (2) en imposant des contraintes temporelles et « environnementales », et (3) en automatisant le raisonnement expert humain.

Le formalisme que nous avons choisi pour combiner les informations multi-sources est basé sur la logique floue. Ce formalisme a permis au système de manipuler des données imprécises, incertaines et de natures hétérogènes. En outre, grâce aux ensembles flous, la modélisation des concepts linguistiques a été rendue possible ; cela a permis de construire des règles de décision interprétables, faciles à comprendre. Concernant ce dernier point, deux méthodes de construction de base de règles ont été proposées. La première repose sur les connaissances expertes, et la deuxième sur un apprentissage.

Contrairement au large éventail de méthodes de détection de changement qui ont été proposées jusqu'à présent pour l'analyse des séries temporelles (*c.f.* Chapitre 1), notre système répond à la question de dépendances temporelles entre les images satellites. Ces dépendances sont prises en compte par l'intégration des indicateurs sur la dynamique des profils temporels issus des images, et des contraintes temporelles imposées par le modèle de culture et les connaissances expertes. Elles sont plus concrètement considérées dans les règles expertes de décision.

Concernant son utilisation, notre système peut assurer le suivi des coupes dans un cadre opérationnel. A chaque nouvelle acquisition satellite, il permet d'identifier les parcelles qui ont été coupées entre cette image et la dernière image dégagée (sans nuage) dans la série. La flexibilité du système, autorisée par la gradualité des indicateurs et sa méthode d'inférence, permet d'assurer une continuité temporelle et de limiter la propagation des erreurs.

De point de vue technique, les logiciels utilisés pour développer le système sont :

- ERDAS Imagine[®] pour l'extraction des profils temporels de réflectances et de NDVI à partir des séries temporelles d'images ;
- Matlab[®] pour l'extraction de l'information à partir du modèle de culture, le calcul des indicateurs, et la conception de l'ensemble du système à base de règles expertes, notamment avec la boîte à outil « Fuzzy Logic » ;
- Fispro[®] pour la conception du système à base de règles induites par apprentissage.

Le chapitre suivant est consacré à l'évaluation des performances du système sur les sites d'études à La Réunion. Les deux versions, *i.e.* celle à base de règles expertes et celle à base de règles induites par apprentissage, sont appliquées sur différentes séries temporelles et sont comparées entre elles.

Chapitre 5 :

Système : résultats et discussion

Sommaire

5.1. Introduction	122
5.2. Système à base de règles expertes.....	123
5.2.1. Performances.....	123
5.2.1.1. Analyse des matrices de confusion.....	125
5.2.1.2. Stabilité des décisions.....	128
5.2.1.3. Validité pluriannuelle	129
5.2.2. Analyse des erreurs	130
5.2.3. Contribution des ensembles flous	131
5.2.4. Analyse de la robustesse par rapport au nombre d'images disponibles	132
5.3. Système à base de règles induites par apprentissage.....	135
5.3.1. Performances.....	135
5.3.2. Analyse de la robustesse par rapport au nombre d'images disponibles	139
5.3.3. Impact des partitions	140
5.3.4. Contribution des différentes sources d'informations	142
5.3.5. Validation pluriannuelle.....	147
5.4. Comparaison.....	148
5.5. Conclusion.....	151

5.1. Introduction

Dans ce chapitre nous évaluons et comparons les deux versions du système d'aide à la décision développées pour la détection automatique des coupes de canne à sucre. Nous rappelons que la première version est caractérisée par une base de règles expertes, quant à la deuxième, elle comporte une base de règles induites automatiquement par apprentissage. Pour chacune des deux versions, nous réalisons une analyse de performances sur plusieurs sites d'études à La Réunion en utilisant différentes séries temporelles d'images SPOT.

Afin d'évaluer la robustesse de chaque version, nous effectuons des analyses de sensibilité du système par rapport au manque d'images satellite dans les séries temporelles. Cette évaluation nous permettra de conclure sur le caractère opérationnel de chacune des deux versions en termes de nombre d'images par an nécessaire pour assurer des bonnes performances globales. La validité pluriannuelle des deux versions sera aussi analysée.

Nous nous intéressons également à l'étude de la contribution des ensembles flous dans les performances globales, et à l'impact du changement de configuration des partitions d'indicateurs d'entrée (avec la deuxième version).

La version avec génération automatique des règles sera utilisée pour étudier la contribution des trois sources d'informations (série temporelle, modèle de culture et connaissances expertes) dans les performances du système. Cette étude serait laborieuse à mettre en place avec la première version étant donné la difficulté de définir une nouvelle base de règles expertes à chaque fois que les paramètres d'entrée changent.

Pour finir, une comparaison entre les caractéristiques des deux versions est effectuée.

Les résultats présentés dans ce chapitre ont fait l'objet :

- d'un article "*Integrating SPOT-5 time series, crop growth modeling, and expert knowledge for monitoring agricultural practices - The case of sugarcane harvest on Reunion Island*" (El Hajj *et al.*, 2008b) accepté dans la revue *Remote Sensing of Environment* (Facteur d'impact = 3,013 ; 2007) ;
- d'un article "*Multi-source information fusion for crop monitoring*" qui sera soumis à la revue *Agriculture Ecosystems & Environment* (Facteur d'impact = 2,308 ; 2007) ;
- de 3 communications dans les colloques internationaux : MultiTemp-2007 (El Hajj *et al.*, 2007a) avec un poster, Fusion-2008 (El Hajj *et al.*, 2008a) avec une présentation orale et LFA (El Hajj *et al.*, 2008d) avec une présentation orale.

Ces publications sont fournies dans l'Annexe P.2.

5.2. Système à base de règles expertes

La première version du système, comportant une base de règles expertes, a été évaluée sur les exploitations de M. Barau (Nord), M. Caruel (Nord) et M. Isautier (Sud) (c.f. Section 2.3.2.1), en utilisant quatre séries temporelles d'images acquises en 2003, 2004, 2006 et 2007. La Figure 41 illustre les dates d'acquisitions dans chaque série temporelle, ainsi que les sites d'étude sur lesquels le système a été appliqué. Le choix des combinaisons site d'étude - série temporelle a été guidé par la disponibilité des données de terrain.

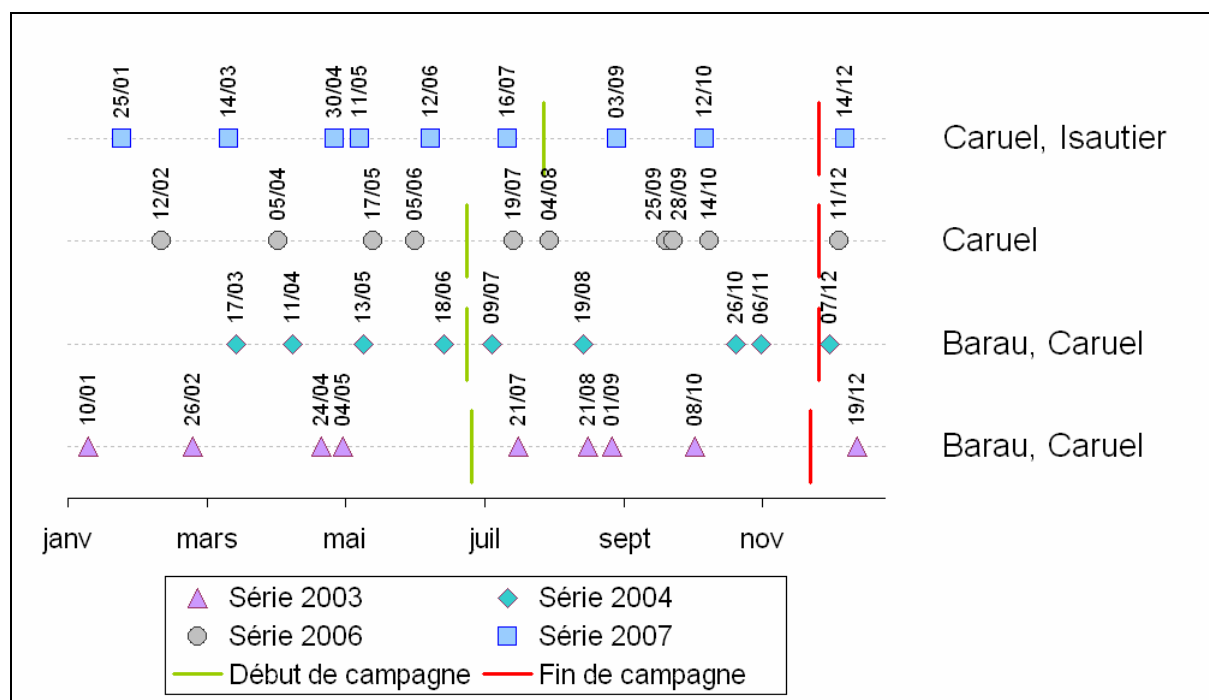


Figure 41. Séries temporelles utilisées pour évaluer le système à base de règles expertes sur les sites d'étude, et dates de début et de fin des campagnes de récolte.

5.2.1. Performances

L'évaluation du système en utilisant les différentes séries temporelles d'images a été réalisée avec le seul historique de l'année en cours ; e.g. pour détecter les coupes de l'exploitation de M. Caruel en 2004, seules les images dans la série de 2004 ont été utilisées. Pour chaque combinaison illustrée dans la Figure 41, le système à base de règles expertes a été appliqué, et ses sorties ont été utilisées pour prendre une décision sur la coupe de chaque parcelle entre chaque couple d'images satellite consécutives. La méthode de décision utilisée est celle d'un utilisateur « exigeant » (c.f. Chapitre 4, Section 4.5.1). Elle autorise trois conclusions « Coupe », « Non coupe » et « Je ne sais pas », en tenant compte d'un seuil de

confiance μ_{conf} défini par l'utilisateur (ce seuil permet de garantir une certaine valeur de confiance aux décisions). Les décisions obtenues avec différentes valeurs de μ_{conf} ont ensuite été comparées aux données de vérité terrain, et des matrices de confusion ont été calculées (Congalton, 1991). Une matrice de confusion permet d'évaluer :

- La précision globale du système, *i.e.* le pourcentage du nombre de décisions correctes par rapport au nombre total de décisions prises par le système.
- La précision du producteur pour les classes « Coupe » et « Non coupe ». La précision du producteur pour une classe donnée est le pourcentage de décisions correctes correspondant à cette classe par rapport à l'effectif de cette classe dans la vérité terrain.
- La précision de l'utilisateur pour les classes « Coupe » et « Non coupe ». La précision de l'utilisateur pour une classe donnée est le pourcentage de décisions correctes correspondant à cette classe par rapport à l'effectif de cette classe dans les décisions du système.
- Les taux d'erreurs d'omission pour les classes « Coupe » et « Non coupe ». L'erreur d'omission pour une classe donnée est le pourcentage de parcelles appartenant dans la vérité terrain à cette classe et qui ont été affectés par le système à d'autres classes. Par exemple, une erreur de « Coupe » est omise lorsque la vérité terrain d'une parcelle est « Coupe » et le système décide « Non Coupe » ou « Je ne sais pas ».
- Les taux d'erreurs de commission pour les classes « Coupe » et « Non coupe ». L'erreur de commission pour une classe donnée est le pourcentage de parcelles affectées par le système à cette classe qui appartiennent dans la vérité terrain à d'autres classes. Par exemple, une erreur de « Coupe » est commise lorsque le système décide « Coupe » et la vérité terrain est « Non Coupe ».
- Le pourcentage de décisions dont le label est « Je ne sais pas » par rapport au nombre total de décisions.

Pour ne pas alourdir la lecture de ce chapitre, nous analysons en détail uniquement les résultats obtenus pour les exploitations de M. Barau et Caruel (74 parcelles) en 2003 et 2004. Les performances du système obtenues pour les autres combinaisons site d'étude – série temporelle sont recensées dans l'Annexe 9 et synthétisées dans la section 5.2.1.3.

5.2.1.1. Analyse des matrices de confusion

Deux séries temporelles d'images SPOT acquises en 2003 et 2004 (Figure 41), comprenant chacune 9 images, ont été utilisées pour détecter automatiquement les coupes dans les exploitations de M. Barau et M. Caruel. Nous nous sommes mis dans un contexte opérationnel : en partant du début de chaque série (*i.e.* premier couple d'images après la fin de la dernière campagne), à chaque nouvelle acquisition d'image des décisions sont prises pour les parcelles dégagées sur les sites d'étude. Au total 1001 décisions ont été prises²⁶. Nous illustrons dans l'Annexe 10 un exemple de cartographie dynamique réalisée pour les exploitations de M. Barau et M. Caruel en 2004.

En calculant des matrices de confusions, nous avons évalué les précisions globales (PG) du système pour des valeurs de μ_{conf} allant de 0 à 1 avec un pas de 0,1. Nous avons également calculé les précisions du producteur pour les classes « Coupe » et « Non coupe » (respectivement PP Coupe et PP Non coupe) et les pourcentages des décisions « Je ne sais pas » (JSP).

Le Tableau 11 montre la matrice de confusion calculée pour μ_{conf} égal à zéro. Les colonnes du tableau correspondent aux décisions du système (« Coupe », « Non coupe » ou « Je ne sais pas ») prises pour toutes les parcelles entre chaque couple d'images dégagées et consécutives, et les lignes correspondent à la vérité terrain (« Coupe » ou « Non coupe ») issues de la base de données des agriculteurs. Les matrices de confusion calculées pour les différentes valeurs de μ_{conf} (entre 0 et 1 avec un pas de 0,1) sont recensées dans l'Annexe 11.

Nous récapitulons dans la Figure 42 les évolutions de PG, JSP, PP Coupe et PP Non coupe en fonction de μ_{conf} . On remarque que ces évolutions sont monotones ; elles sont décroissantes pour PG, PP Coupe et PP Non Coupe, et croissante pour JSP. On observe deux plages de quasi-stabilité ; la première pour $0 \leq \mu_{conf} \leq 0,6$ et la deuxième pour $0,7 \leq \mu_{conf} \leq 0,9$. Cette quasi-stabilité est due à l'utilisation des ensembles flous dans les partitions des indicateurs d'entrée.

Les performances du système sont très satisfaisantes pour $0 \leq \mu_{conf} \leq 0,6$ avec PG entre 98,30% et 98,80%, PP Coupe entre 93,75% et 96,06%, PP Non coupe entre 98,97% et 99,08%, et JSP entre 0,80% et 1,20%. Ces performances diminuent pour $0,7 \leq \mu_{conf} \leq 0,9$,

²⁶ Le nombre de décisions dans le cas où toutes les images sont sans nuages est égal à 1184 (74 parcelles x 8 couples de dates x 2 années).

avec une chute moyenne de 7,31% pour la PG, 3,57% pour la PP Coupe, 7,84% pour la PP Non coupe, et une augmentation moyenne de 7,36% pour le JSP. Cette chute dans les performances est due au fait que 5,26% de décisions « Coupe » et « Non Coupe » ont des valeurs de confiance entre 0,6 et 0,7, ainsi pour un μ_{conf} de 0,7 ces décisions passent à « Je ne sais pas ». Pour une valeur de μ_{conf} de 1, la PG, la PP Coupe, la PP Non coupe et le JSP sont 75,75%, 89,06%, 73,80% et 23,95% respectivement.

		Décisions			Total ligne	Précision du producteur (%)	Erreur d'omission (%)
		Coupe ($\mu_C > 0$)	Non Coupe ($\mu_{NC} > 0$)	Je ne sais pas			
Vérité terrain	Coupe	123	2	3	128	96,09	3,91
	Non Coupe	2	866	5	873	99,20	0,80
Total colonne		125	868	8		Précision globale	Pourcentage de "Je ne sais pas"
Précision de l'utilisateur (%)		98,40	99,77			98,80 %	0,80 %
Erreur de commission (%)		1,60	0,23				

Tableau 11. Matrice de confusion obtenue en comparant les décisions prises à l'aide du système à base de règles expertes (pour μ_{conf} égal à 0) avec la vérité terrain sur les exploitations de M. Barau et M. Caruel en 2002 et 2003.

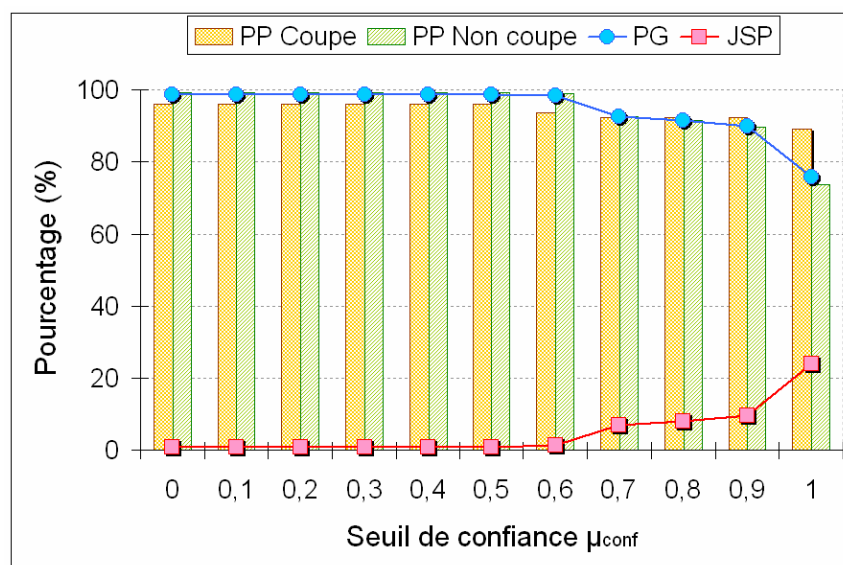


Figure 42. Performances du système à base de règles expertes obtenues pour les parcelles de M. Barau et M. Caruel en 2003 et 2004 avec différents seuils de confiance μ_{conf} . (PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement ; PG : précision globale ; JSP : décisions « Je ne sais pas »).

Dans la Figure 43, nous illustrons les évolutions des taux d'erreurs de commission et d'omission pour les classes « Coupe » et « Non Coupe » en fonction de μ_{conf} . Les erreurs de commissions sont quasi-constantes et faibles pour les deux classes avec une moyenne de 2,36% pour « Coupe » et 0,19% pour « Non Coupe ». Quant aux erreurs d'omission, elles sont stables jusqu'à μ_{conf} égal à 0,5 pour la classe « Coupe » avec une moyenne de 3,91%, et elles augmentent progressivement pour atteindre 10,94% pour μ_{conf} égal à 1; pour la classe « Non Coupe », elles sont constantes avec une moyenne de 0,92% pour μ_{conf} entre 0 et 0,6, et augmentent d'environ 7,86% entre 0,6 et 0,9 et d'environ 16% entre 0,9 et 1. L'augmentation des erreurs d'omission est due principalement au passage des décisions « Coupe » ou « Non Coupe » en décisions « Je ne sais pas ». Une discussion sur les origines de ces erreurs est présentée dans la section 5.2.2.

D'après ce qui précède, nous pouvons remarquer que plus le décideur (*i.e.* l'utilisateur) se fonde sur le système en acceptant des décisions « Coupe » et « Non Coupe » avec moins d'exigence de confiance, plus les performances globales sont meilleures. Ceci est dû au fait que la base de règles expertes utilisée pour faire l'inférence est performante et qu'elle a pratiquement couvert toutes les situations que le système a rencontré ; *i.e.* même les conclusions qui sont associées à un niveau de confiance non élevé ont été correctes. Une autre raison principale est que les partitions floues, construites en se basant sur les connaissances expertes, ont du sens. Ce dernier point sera évoqué plus en détail dans les sections 5.2.3 et 5.3.3.

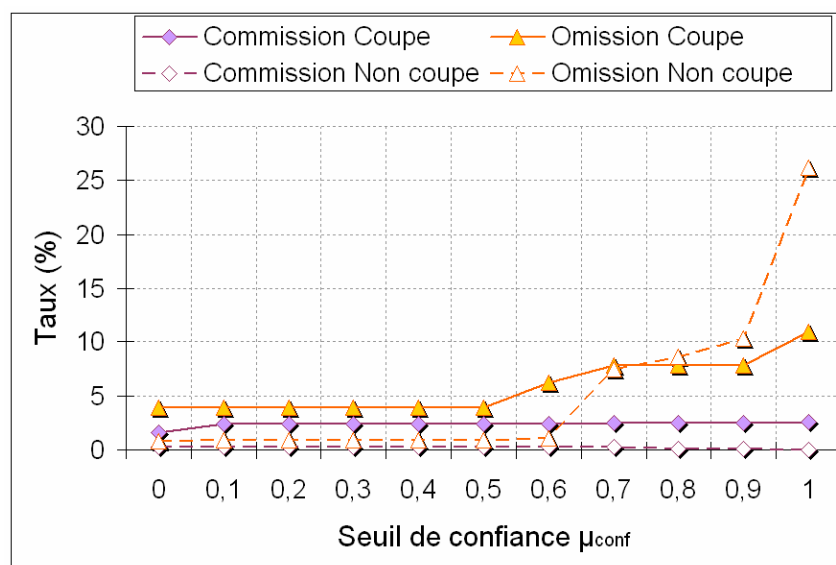


Figure 43. Taux d'erreurs de commission et d'omission des classes « Coupe » et « Non Coupe » obtenues pour les parcelles de M. Barau et M. Caruel en 2003 et 2004 avec différents seuil de confiance μ_{conf} .

5.2.1.2. Stabilité des décisions

Etant donnée la présence de la possibilité « Je ne sais pas » parmi les conclusions des règles du système, nous sommes amenés à étudier la stabilité des décisions « Coupe » et « Non Coupe » prises en utilisant les sorties μ_C , μ_{NC} et μ_{JSP} du système. En d'autres termes, il faut s'assurer que les degrés d'appartenance μ_C et μ_{NC} des décisions « Coupe » et « Non Coupe » soient significativement supérieurs au degré d'appartenance μ_{JSP} de la possibilité « Je ne sais pas ». La stabilité de ces décisions a donc été calculée par la formule suivante :

$$\forall x = \arg \max(\mu_C, \mu_{NC}) \quad Stabilité(x) = \mu_x - \mu_{JSP} \quad (20)$$

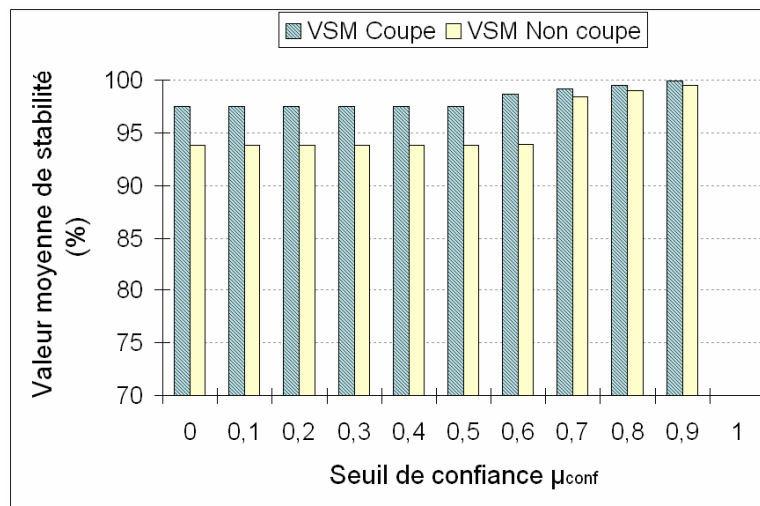


Figure 44. Valeurs moyennes de stabilité des décisions « Coupe » et « Non Coupe » obtenues lors de l'application du système à base de règles expertes sur les parcelles de M. Barau et M. Caruel en 2003 et 2004.

La Figure 44 montre les valeurs de stabilité moyenne (à chaque niveau de μ_{conf}) calculées pour toutes les décisions « Coupe » et « Non Coupe » obtenues lors de l'application du système sur les exploitations de M. Barau et M. Caruel en utilisant les séries temporelles de 2003 et 2004. Il est à signaler qu'une valeur de stabilité moyenne (VSM) pour une classe de décision donnée représente tout simplement la moyenne (en %) des valeurs de stabilités obtenues pour cette classe :

$$VSM_{Coupe} = \frac{100}{TDC} \sum_{i=1}^{TDC} Stabilité(C_i) \quad (21)$$

$$VSM_{Non\ coupe} = \frac{100}{TDNC} \sum_{i=1}^{TDNC} Stabilité(NC_i) \quad (22)$$

où TDC est le nombre total des décisions « Coupe » et $TDNC$ est le nombre total des décisions « Non Coupe ».

En observant la Figure 44, on remarque que les VSM des décisions « Coupe » et « Non Coupe » diminuent légèrement avec μ_{conf} ; cependant, même les valeurs les plus faibles obtenues (97,52% et 93,84% pour VSM_{Coupe} et $VSM_{Non\ coupe}$ respectivement) reflètent une forte stabilité.

5.2.1.3. Validité pluriannuelle

Les performances globales obtenues en appliquant le système à base de règles expertes sur les différentes combinaisons site d'étude – série temporelle (Annexe 9), montrent que cette version du système est robuste et performante (Figure 45) : pour toutes les combinaisons, la précision globale du système atteint des valeurs supérieures à 96,73% (avec $\mu_{conf} = 0$), les précisions du producteur pour les classes « Coupe » et « Non coupe » excèdent respectivement 95% et 96,41%, et le pourcentage de décisions « Je ne sais pas » est inférieure à 1,96%, et ce avec des séries temporelles qui comportent 9 images par an (sauf pour 2006 : 10 images par an). Une analyse plus détaillée sur l'influence du nombre d'images par an et de leurs localisations par rapport aux campagnes de récolte sera présentée dans la section 5.3.2. Les valeurs moyennes de stabilité obtenues pour les différentes combinaisons sont également fortes : elles sont toutes supérieures à 83%.

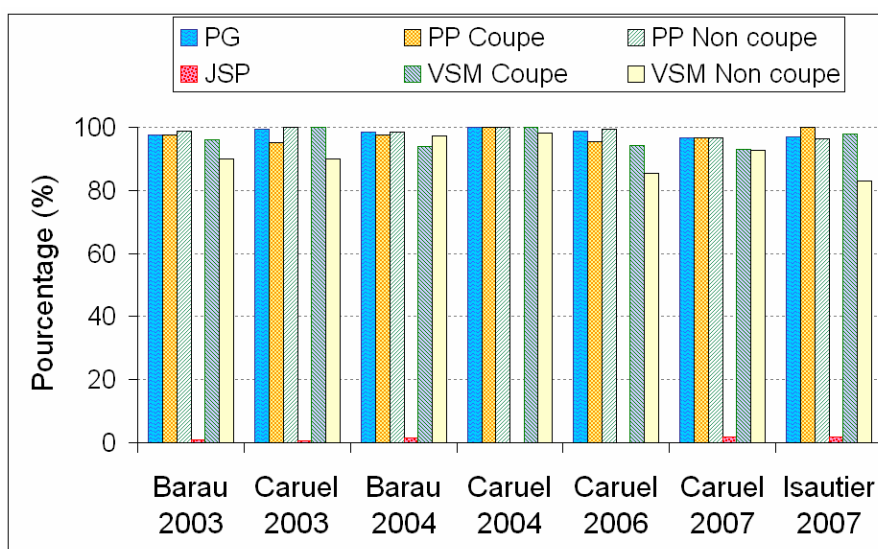


Figure 45. Performances globales du système à base de règles expertes lors de son application sur différents sites d'études et différentes années. (PG : précision globale ; JSP : décisions « Je ne sais pas » ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement ; VSM Coupe et VSM Non coupe : valeurs de stabilité moyenne des décisions « Coupe » et « Non coupe » respectivement).

5.2.2. Analyse des erreurs

Afin de comprendre les causes des différences obtenues entre les données de vérité terrain et les décisions prises, nous avons examiné toutes les règles actives impliquées dans l'inférence. Dans l'Annexe 12 nous montrons, à titre d'exemple, le numéro ainsi que le degré d'activation de chaque règle lors de l'application du système sur l'exploitation de M. Caruel avec la série de 2004. Etant donné que ces règles sont linguistiques, nous avons pu analyser l'adéquation de leur activation avec la situation rencontrée, et nous avons ainsi pu dégager l'origine des erreurs. Cette analyse d'erreurs a été réalisée pour les résultats obtenus sur les exploitations de M. Barau et M. Caruel en utilisant les séries de 2003 et 2004

Pour μ_{conf} égal à zéro, le pourcentage d'erreur est égal à 1,20% (12 décisions) (c.f. Tableau 11). Les erreurs (Figure 46) sont liées soit à un problème radiométrique affectant les images (i.e. saturation dans le PIR), soit à un comportement inhabituel d'une parcelle de canne (e.g. erreur sur la nature de la culture). L'inadéquation des partitions à ces exceptions génère de fausses décisions qui entraînent par la suite des futures erreurs éventuelles (d'omission et/ou de commission).

Nous considérons que ces erreurs sont « normales » car la configuration des partitions des indicateurs et les scénarios inclus dans la base de règles prétendent être généraux.

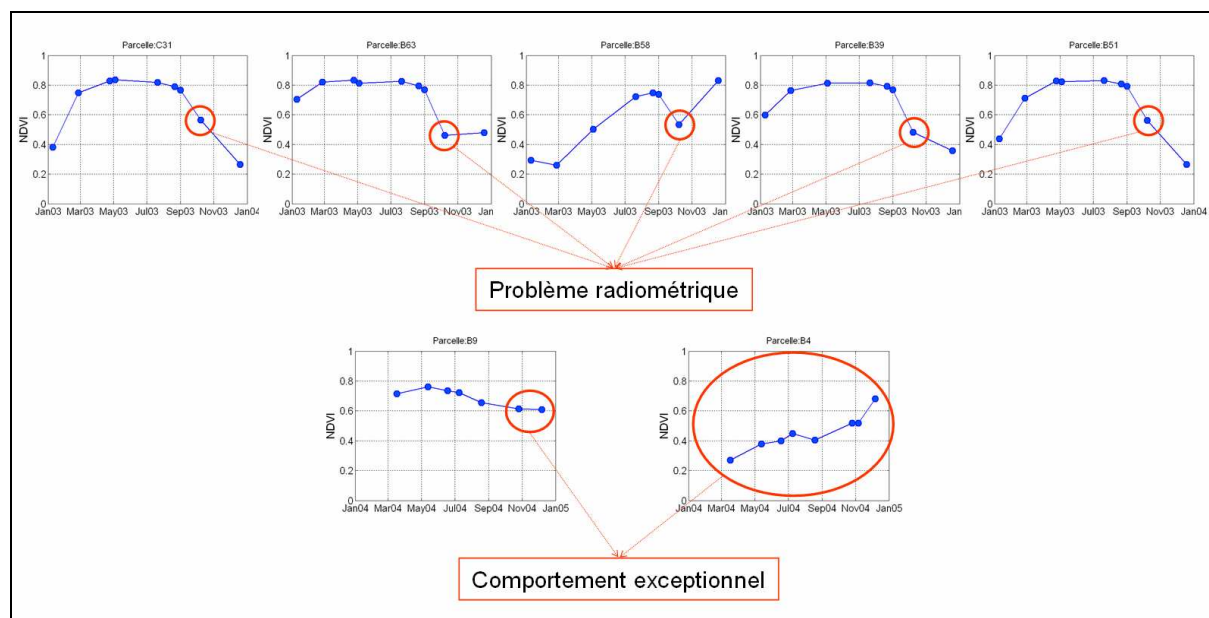


Figure 46. Profils temporels de NDVI des parcelles pour lesquelles le système à base de règles expertes a pris de mauvaises décisions.

Pour les valeurs plus élevées de μ_{conf} , l'augmentation du désaccord entre les décisions et la vérité terrain est principalement due à l'augmentation des décisions « Je ne sais pas », et par conséquent aux omissions de détection de coupe précédente.

Cette analyse permet d'envisager des pistes d'amélioration de cette version du système.

5.2.3. Contribution des ensembles flous

Étant donné que les partitions de la majorité des indicateurs d'entrée du système contiennent des ensembles flous, nous nous sommes intéressés à l'évaluation de la contribution de ces ensembles flous dans la performance globale du système. Pour ceci, nous avons créé une nouvelle configuration « non-floue » de partitions, et nous avons comparé les résultats obtenus par cette configuration avec ceux obtenus précédemment (en utilisant la configuration floue). La configuration non-floue a été définie en enlevant toute les marges d'ambiguïté dans les partitions des entrées floues (e.g Figure 47). Avec cette configuration, les degrés d'appartenance associées aux conclusions des règles actives et ainsi aux sorties du système (μ_C , μ_{NC} et μ_{JSP}) sont soit zéro, soit un. Par conséquent, l'intégration du paramètre μ_{conf} était sans signification dans ce cas-ci et toutes les décisions ont été prises selon l'équation 17, $\text{argmax}(\mu_C, \mu_{NC}, \mu_{JSP})$.

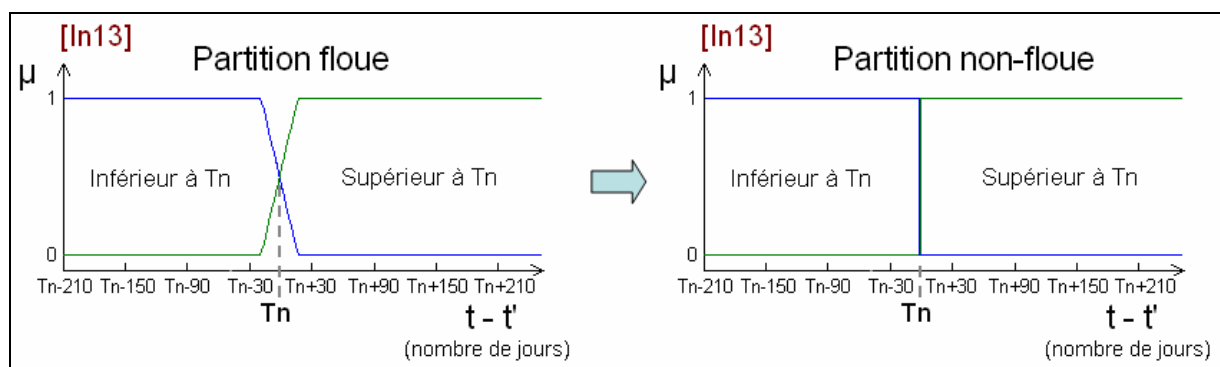


Figure 47. Un exemple illustrant le passage d'une partition floue à une partition non-floue.

La précision globale et le pourcentage de décisions « Je ne sais pas » obtenus pour les résultats de M. Barau et M. Caruel de 2003 et 2004 en utilisant la configuration non-floue sont respectivement 81,95% et 16,61%. En comparant ces valeurs à celles obtenues avec la configuration floue aux différents niveaux de μ_{conf} (Figure 42), nous constatons que plus le niveau de μ_{conf} est faible plus la contribution des ensembles flous est importante ; cette contribution atteint 16,85% de précision globale supplémentaire quand μ_{conf} est égale à zéro.

5.2.4. Analyse de la robustesse par rapport au nombre d'images disponibles

L'analyse de la robustesse du système à base de règles expertes a été réalisée sur deux séries temporelles acquises en 2003 et 2004 sur les exploitations de M. Barau et M. Caruel. Par « robustesse », nous entendons la sensibilité des performances globales du système par rapport au manque d'acquisitions satellite dans la série temporelle.

Pour réaliser cette analyse, nous avons fait des échantillonnages sur les deux séries « complètes » de 2003 et 2004 afin de créer des séries avec des fréquences d'acquisition de moins en moins importantes. A partir des séries de neuf images par an, des séries comportant cinq images par an ont été créées, puis des séries de quatre images, puis de trois, et finalement de deux images par an. L'échantillonnage a été réalisé, par rapport aux dates d'ouverture et de fermeture de la campagne de récolte, selon le protocole suivant :

- Pour cinq images par an : deux images avant la campagne de récolte, y compris une peu avant la campagne, deux images pendant la campagne, et une peu après.
- Pour quatre images par an : une image peu avant la campagne de récolte, deux pendant, et une peu après.
- Pour trois images par an : une image avant la campagne de récolte, une pendant, et une peu après.
- Pour deux images par an : une peu avant la campagne de récolte et une peu après.

La Figure 48 montre les dates d'acquisition sélectionnées dans les différentes séries. Il est à signaler que ces images ne sont pas complètement dégagées de nuages, donc toutes les parcelles sur les sites d'études ne sont pas « visibles » tout le long de chaque série.

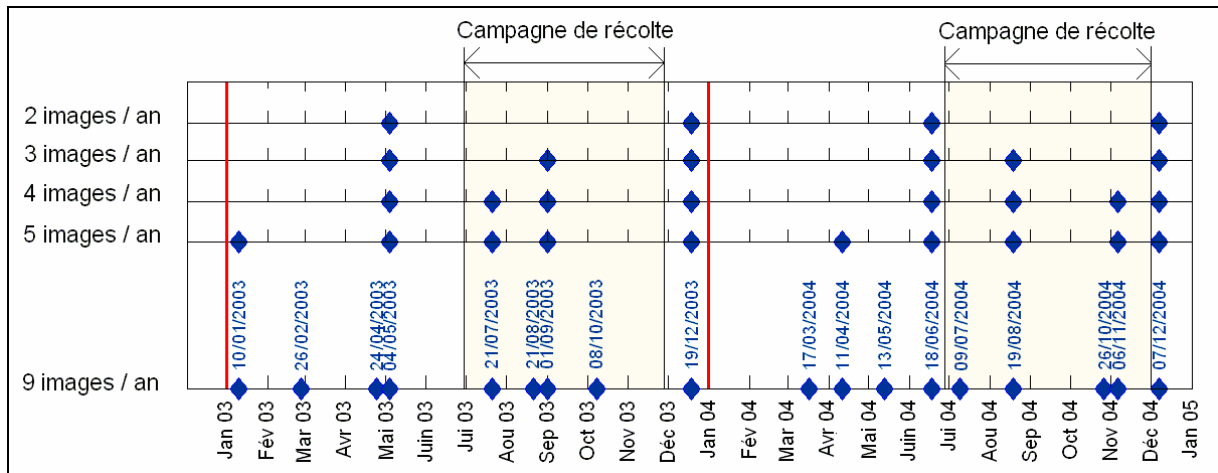


Figure 48. Sur l'axe le plus bas, les dates d'acquisition satellite de la série complète (9 images / an); Sur les autres axes, les échantillonnages réalisés pour construire des séries avec 5, 4, 3 et 2 images par an respectivement.

La décision sur la coupe de chaque parcelle entre chaque couple d'images dégagées (sans nuages) a été prise selon le critère utilisé dans la section 5.2.1 avec μ_{conf} égal à zéro. La Figure 49 montre l'évolution de la précision globale (PG) en fonction du nombre d'images par an dans la série, ainsi que les évolutions des précisions du producteur pour les classes « Coupe » et « Non coupe » (PP Coupe et PP Non coupe respectivement). Les valeurs de ces précisions ont été calculées à partir des matrices de confusions obtenues en 2003 et en 2004 pour les séries comprenant le même nombre d'images. D'après cette figure, on remarque que le système à base de règles expertes peut fournir à l'utilisateur des résultats satisfaisants avec des séries comprenant au moins trois images par an. Avec cette fréquence d'acquisition, la PG, la PP Coupe et la PP Non coupe sont respectivement 93,56%, 89,57% et 96,64%. Ces précisions augmentent avec le nombre d'images pour atteindre, avec neuf images par an, 98,80% pour la PG, 96,06% pour la PP Coupe et 99,20% pour la PP Non coupe. Quand deux images par an seulement sont utilisées, les performances du système diminuent : les valeurs de PG de PP Coupe et de PP Non coupe chutent à 82,40%, 83,33% et 72,73% respectivement. Nous attirons l'attention sur le fait que les performances représentées dans la Figure 49 ne dépendent pas seulement du nombre d'images par an dans la série, mais aussi de la position temporelle de ces images par rapport à la campagne de récolte.

Les résultats obtenus dans cette analyse soulignent l'influence de la variabilité spatio-temporelle élevée des parcelles de canne à sucre sur le procédé de détection de coupe. Avec deux images par an, une peu avant et une peu après la campagne de récolte, seules les parcelles coupées à la fin de la campagne peuvent être « facilement » détectées car la

végétation dans ces parcelles n'a pas eu le temps de repousser, et les propriétés spectrales des parcelles dans l'image acquise peu avant restent sensiblement différentes de celles dans l'image prise après la campagne. Ceci n'est pas le cas pour les parcelles qui sont coupées en début de campagne. Si des images acquises pendant la campagne de récolte sont utilisées, la distance temporelle entre les images sera réduite, et la probabilité de la confusion entre les parcelles non coupées et celles en repousse sera réduite au minimum ; dans ce cas, il y aura plus de parcelles coupées identifiées. L'intégration des images (au moins une) prises au début de l'année et pendant la phase de croissance (avant la campagne de coupe) enrichit l'information sur les étapes de développement de chaque parcelle et, par conséquent, améliore les résultats.

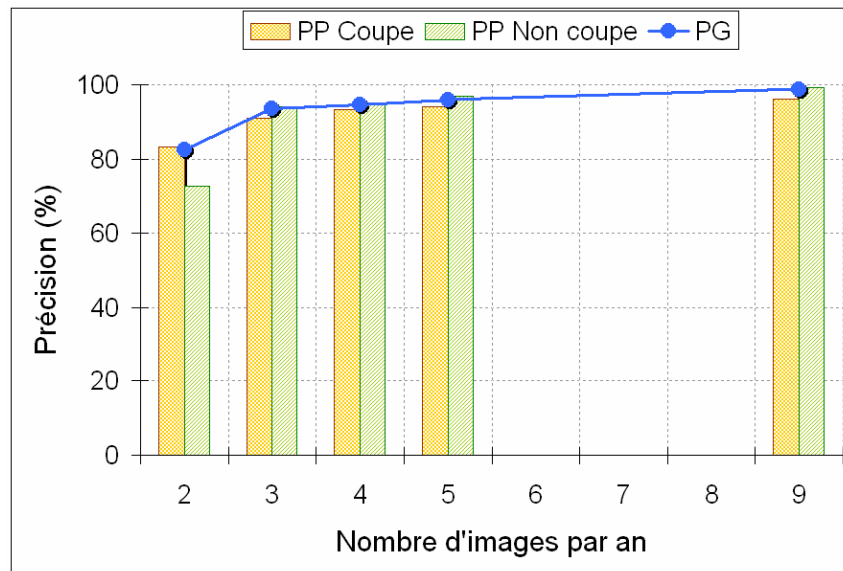


Figure 49. Performances du système à base de règles expertes en fonction du nombre d'images par an (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement).

5.3. Système à base de règles induites par apprentissage

La deuxième version du système d'aide à la décision a été évaluée sur les exploitations de M. Barau et de M. Caruel en utilisant les séries temporelles d'images SPOT acquises en 2003, 2004 et 2006. Les principales analyses ont été réalisées sur les données de 2003 et 2004 (Figure 41), comme pour la première version. Celles de 2006 ont été utilisées pour évaluer le caractère général (*i.e.* la validité pluriannuelle) des règles générées par apprentissage sur les données de 2003 et 2004. Les analyses ont porté sur :

- les performances du système et les règles générées ;
- la robustesse du système par rapport au manque d'images ;
- l'impact du changement de configuration des partitions des indicateurs d'entrée sur les performances ;
- la contribution des sources d'information.

Afin de réaliser ces analyses, nous avons testé cette deuxième version du système en faisant varier :

- les pourcentages de données d'apprentissage pour la génération des règles ;
- le nombre d'images acquises par an ;
- les configurations des partitions ;
- les sources d'informations.

Tous les résultats obtenus sont recensés dans l'Annexe 13. Pour ne pas alourdir la lecture de ce chapitre, nous nous limiterons ici à l'étude de quelques cas illustratifs.

5.3.1. Performances

Neuf pourcentages de jeu d'apprentissage ont été utilisés : 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% et 90%. Pour chaque pourcentage, une base de règle a été générée, et les performances du système ont été évaluées en l'appliquant sur les données qui ne sont pas utilisées pour l'apprentissage. Nous rappelons qu'un jeu d'apprentissage est une base de données qui contient, pour une série d'images, les valeurs des indicateurs du système calculées pour un certain pourcentage de parcelles ainsi que la situation (*i.e.* vérité terrain) de ces parcelles entre chaque couple d'acquisitions satellite (*c.f.* Chapitre 4, Section 4.4.3.2).

Afin que les conclusions des analyses soient générales et indépendantes d'une sélection particulière de jeu d'apprentissage, une validation croisée (Browne, 2000) de dix itérations a été réalisée pour chaque cas. Pour chaque itération, une matrice de confusion a été calculée en comparant les décisions du système à la vérité terrain.

Dans cette section, sont discutés les résultats obtenus en appliquant le système sur les deux séries temporelles « complètes » de 9 images par an acquises en 2003 et 2004, en utilisant toutes les sources d'informations, et en adoptant la configuration experte de partition (*c.f.* Chapitre 4, Section 4.4.1).

La Figure 50 montre les valeurs moyennes et les écarts-types (pour dix itérations) de la précision globale (PG) du système et des précisions du producteur pour les classes « Coupe » et « Non coupe » (PP Coupe et PP Non Coupe respectivement), obtenus avec des pourcentages de jeu d'apprentissage (PJA) allant de 10 à 90% avec un pas de 10%. D'après cette figure, nous pouvons constater que les performances du système sont satisfaisantes même avec de faibles PJA ; *e.g.* en utilisant 10% de données seulement pour l'apprentissage, *i.e.* environ 6 parcelles par image, la PG est de 96,84% et les PP Coupe et PP Non coupe sont respectivement 86,46% et 98,20%. Les performances du système s'améliorent avec l'augmentation du PJA : les valeurs de PG, PP Coupe et PP Non coupe peuvent atteindre 98,42%, 96,01% et 99,07% respectivement.

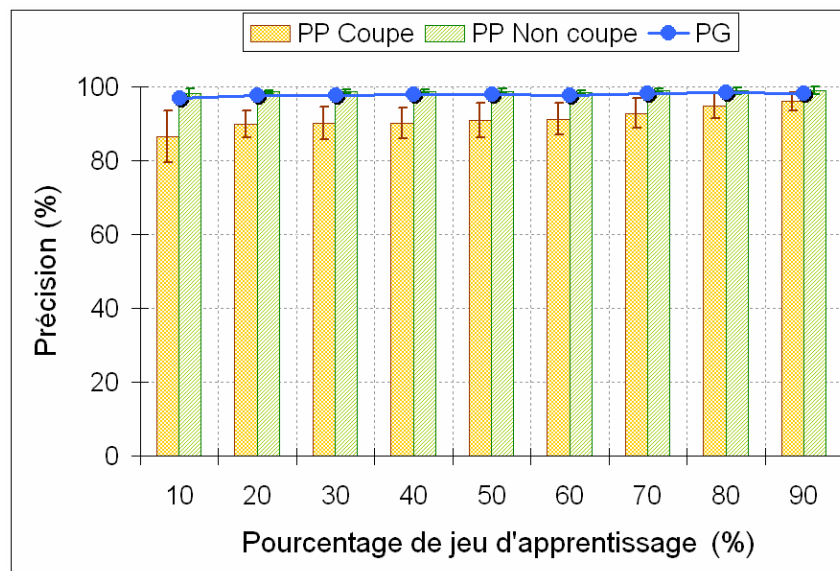


Figure 50. Performances du système à base de règles induites par apprentissage obtenues pour les parcelles de M. Barau et M. Caruel en 2003 et 2004, en utilisant différents pourcentages de jeu d'apprentissage (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement). Les valeurs représentées correspondent à la moyenne obtenue pour 10 itérations et la barre d'erreur correspond à l'écart type.

La Figure 51 illustre le nombre de règles générées en utilisant les différentes valeurs de PJA. Le nombre moyen de règles générées (pour les 10 itérations) avec 10% de données est égal à 9 seulement. Ce nombre augmente avec l'augmentation du PJA pour atteindre une moyenne de 51 règles pour un PJA de 90%. L'écart type du nombre de règles (pour les 10 itérations) varie entre 4 et 9 règles.

Les règles générées sont toutes interprétables grâce aux termes linguistiques associés aux indicateurs d'entrée et à la sortie du système. Le Tableau 12 montre à titre d'exemple une base de règles générée en utilisant 10% de données pour l'apprentissage. La première chose que nous remarquons dans cette base est le faible nombre de règles, 6, et le faible nombre d'indicateurs utilisés, 5. Nous remarquons également que la différence dans le MIR a été préférée aux indicateurs de NDVI (ce qui justifie en quelque sorte la redondance signalée par l'expert entre les informations apportées par le NDVI et le MIR). En analysant le contenu de ces règles, comme nous le faisons dans le Tableau 12, l'utilisateur peut comprendre à quels scénarios ces règles correspondent, et éventuellement corriger, rajouter et/ou supprimer quelques règles.

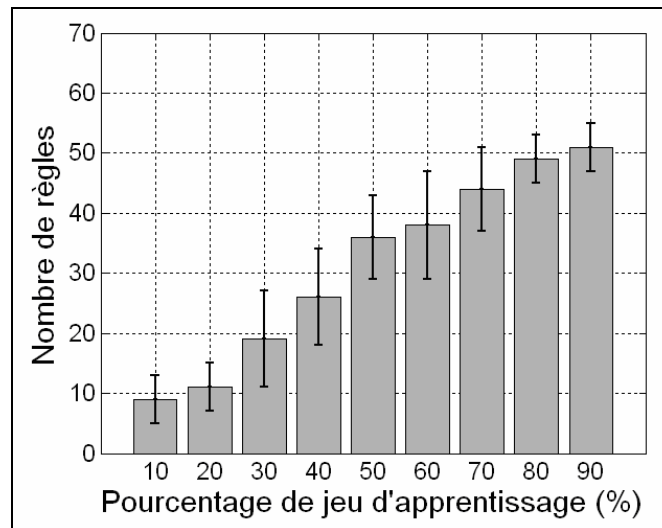


Figure 51. Nombres de règles générées automatiquement en utilisant les différents pourcentages de jeu d'apprentissage. Chaque barre représente le nombre moyen des règles obtenues pour 10 itérations et la barre d'erreur correspond à l'écart type.

Règles																	
Règles Affichage																	
Règle	Active	SI In1	ET In2	ET In3	ET In4	ET In5	ET In6	ET In7	ET In8	ET In9	ET In10	ET In11	ET In12	ET In13	ET In14	ET In15	ALORS ...
1	<input checked="" type="checkbox"/>						Faible						Inf. à LNC	Inf. à Tn			Non Coupe
2	<input checked="" type="checkbox"/>						Faible					Camp. pre.	Sup. à LNC				Coupe
3	<input checked="" type="checkbox"/>						Faible					Entre Camp.	Inf. à LNC		Inf. à Tn		Non Coupe
4	<input checked="" type="checkbox"/>						Faible					Camp. act.	Sup. à LNC	Sup. à Tn			Coupe
5	<input checked="" type="checkbox"/>						Faible							Inf. à Tn			Non Coupe
6	<input checked="" type="checkbox"/>						Forte										Coupe

Règle 1	Si In6 est « Faible » et In12 est « Inférieur à la longueur du cycle » et In13 « Inférieur à Tn » Alors Non Coupe <i>Si la différence entre MIR(t) et MIR(t') est faible, et l'âge de la canne est inférieur à la longueur nominale du cycle de la canne, et la distance temporelle entre les dates t et t' est inférieure à Tn (i.e. la distance entre les deux dates est insuffisante d'après MOSICAS pour qu'il y ait une coupe), alors Non coupe.</i>
Règle 2	Si In6 est « Faible » et In11 est « Campagne précédente » et In12 est « Supérieur à la longueur du cycle » Alors Coupe <i>Si la différence entre MIR(t) et MIR(t') est faible, et la date t' est dans la campagne de coupe précédente (dans ce cas la date t est soit dans la campagne actuelle soit dans la période entre deux campagnes), et l'âge de la canne est supérieur à la longueur nominale du cycle de la canne alors coupe.</i>
Règle 3	Si In6 est « Faible » et In11 est « Entre Campagnes » et In12 est « Inférieur à la longueur du cycle » et In14 « Inférieur à Tn » Alors Non Coupe <i>Si la différence entre MIR(t) et MIR(t') est faible, et la date t' est entre deux campagnes (dans ce cas t est soit dans la campagne actuelle soit dans la période entre deux campagnes), et l'âge de la canne est inférieur à la longueur nominale du cycle, et la distance temporelle entre les dates t et le début de campagne est inférieure à Tn (i.e. Coupe impossible d'après MOSICAS), alors Non coupe.</i>
Règle 4	Si In6 est « Faible » et In11 est « Campagne actuelle » et In12 est « Supérieur à la longueur du cycle » et In13 « Supérieur à Tn » Coupe <i>Si la différence entre MIR(t) et MIR(t') est faible, et la date t' est dans la campagne actuelle (la date t aussi), et l'âge de la canne est supérieur à la longueur du cycle, et In13 est supérieur à Tn (i.e. Coupe possible d'après MOSICAS), Alors Coupe</i>
Règle 5	Si In6 est « Faible » et In13 « Inférieur à Tn » Alors Non Coupe <i>Si la différence entre MIR(t) et MIR(t') est faible, et la distance temporelle entre les dates t et t' est inférieure à Tn (i.e. Coupe impossible d'après MOSICAS), alors Non coupe</i>
Règle 6	Si In6 est « Forte » Alors Coupe <i>Si la différence entre MIR(t) et MIR(t') est forte alors Coupe</i>

Tableau 12. Exemple d'une base de règles générée par le logiciel Fispro[®] en utilisant 10% de données de Barau et Caruel en 2003 et 2004 pour faire l'apprentissage (6 parcelles)

5.3.2. Analyse de la robustesse par rapport au nombre d'images disponibles

L'analyse de robustesse de la deuxième version du système par rapport au manque d'acquisition a été effectuée en comparant les résultats obtenus en utilisant des séries temporelles comprenant des fréquences d'acquisitions différentes (Figure 48). La sélection de ces images a été faite selon la même méthode décrite dans la section 5.2.4.

La Figure 52 montre l'évolution des performances en fonction du nombre d'images par an. Avec deux images par an seulement (une avant la campagne de récolte et une après), la PG moyenne pour les différentes configurations d'apprentissage est de 88,20%, et les PP Coupe et PP Non Coupe moyennes sont 85,41% et 92,91% respectivement. A partir de trois images par an, les performances s'améliorent nettement, avec une PG supérieure à 94,22%, une PP Coupe qui stagne autour de 91% et une PP Non coupe supérieure à 95,49% ; les valeurs de ces précisions atteignent respectivement 97,81%, 91,06% et 98,72% avec une série de neuf images par an. Nous signalons que le nombre d'images par an influe sur la qualité des indicateurs multi-temporels issus des images (*e.g.* différence bi-temporelle de NDVI) et de celui provenant du modèle de culture.

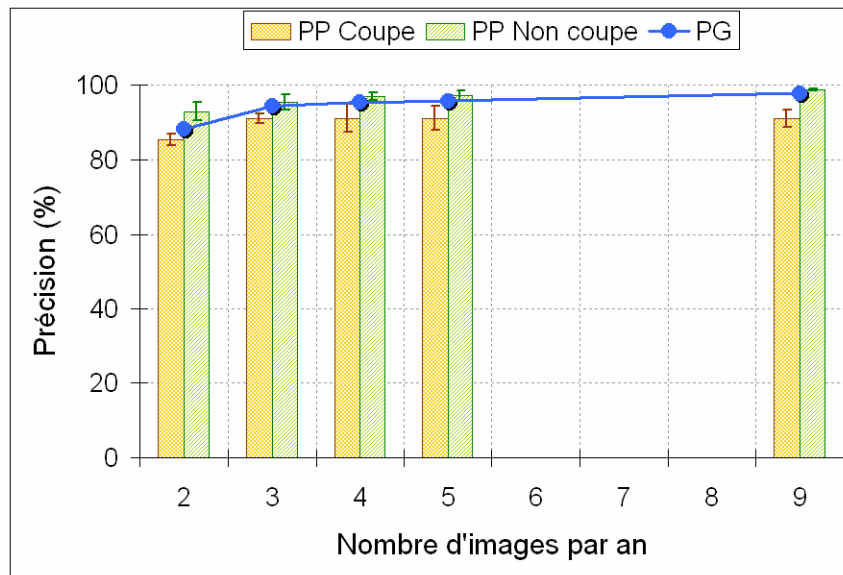


Figure 52. Performances du système à base de règles induites par apprentissage en fonction du nombre d'images par an (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement). Les valeurs représentées correspondent à la moyenne obtenue pour les différentes configurations d'apprentissage ; la barre d'erreur correspond à l'écart-type.

En ce qui concerne le nombre de règles générées pour obtenir ces résultats, nous illustrons dans la Figure 53 la moyenne (pour les différentes configurations d'apprentissage) et l'écart-type de ce nombre pour chaque fréquence d'acquisition. En moyenne, le nombre de règles générées avec cette version du système est inférieur à 40 quelque soit la fréquence d'images dans la série.

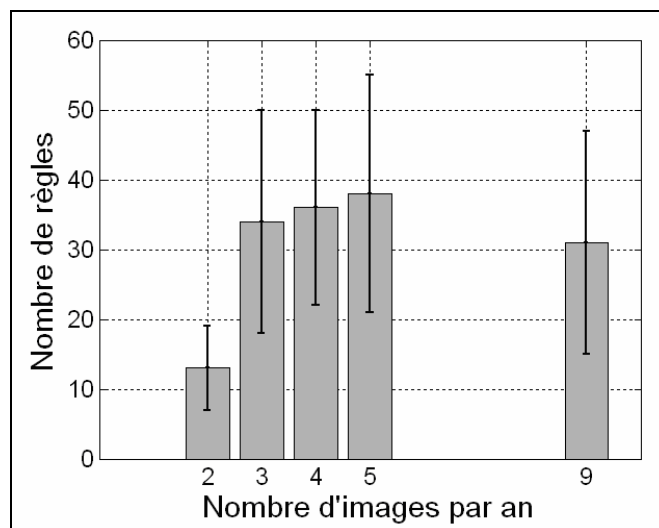


Figure 53. Variation du nombre de règles générées par apprentissage en fonction du nombre d'images dans la série ; chaque barre correspond à la moyenne obtenue pour toutes les configurations d'apprentissage, et les barres d'erreur représentent les écarts-types.

5.3.3. Impact des partitions

Comme pour la première version, nous avons réalisé une analyse de la contribution des partitions floues dans les performances globales du système, mais cette fois avec plus de configurations. Quatre configurations de partitions d'indicateurs d'entrée ont été définies : deux floues, « Config_Fuz1 » et « Config_Fuz2 », et deux non floues, « Config_Crisp1 » et « Config_Crisp2 ».

- La configuration « Config_Fuz1 » est la configuration experte du système décrite dans le chapitre précédent.
- La configuration « Config_Fuz2 » est identique à « Config_Fuz1 » sauf pour les partitions des indicateurs liés au NDVI (*i.e.* In1 et In2) (Figure 54).
- La configuration « Config_Crisp1 » est une défuzzification de la « Config_Fuz1 » (version nette) ; en d'autres termes, elle est issue de cette configuration en enlevant toutes les marges d'ambiguïté dans les partitions des indicateurs.

- La configuration « Config_Crisp2 » est une défuzzification de la « Config_Fuz2 ».

La Figure 54 montre les partitions de NDVI définies dans chacune des quatre configurations.

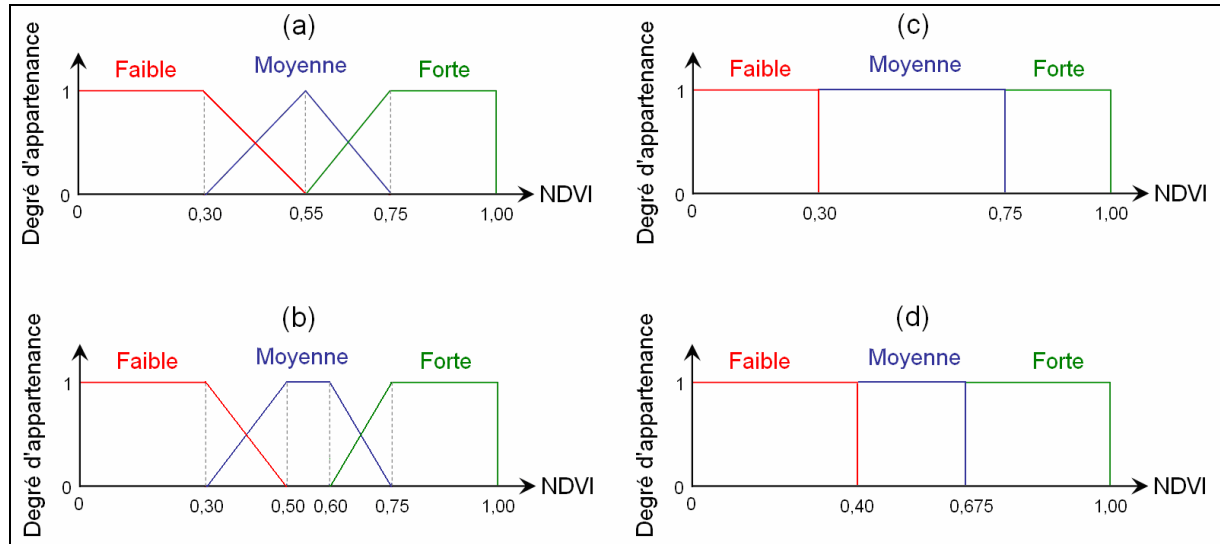


Figure 54. Partitions du NDVI selon les configurations floues (a) « Config_Fuz1 » et (b) « Config_Fuz2 » et les configurations nettes (c) « Config_Crisp1 » et (d) « Config_Crisp2 ».

Pour chaque série d'images, nous avons comparé les valeurs de précision globale (PG) obtenues avec les différentes configurations de partition. La Figure 55 montre les différences de PG pour chaque fréquence d'acquisition d'images entre la configuration floue experte « Config_Fuz1 » et les autres configurations « Config_Fuz2 » (floue), « Config_Crisp1 » (nette) et « Config_Crisp2 » (nette). D'après cette figure, nous constatons que le passage de la configuration floue basique « Config_Fuz1 » aux deux configurations non-floues « Config_Crisp1 » et « Config_Crisp2 » engendre une perte de performance quelque soit le nombre d'images dans la série. La perte de PG varie en moyenne (pour les différentes configurations d'apprentissage) entre 0,44% et 2,59% lorsqu'on passe de « Config_Fuz1 » à « Config_Crisp1 », et entre 0,37% et 1,66% lorsqu'on passe de « Config_Fuz1 » à « Config_Crisp2 ».

Le changement dans les ensemble flous, tout en gardant l'aspect flou (passage de « Config_Fuz1 » à « Config_Fuz2 »), engendre également une perte de performances, mais cette perte est inférieure à celles dues à la défuzzification, et ce quelque soit le nombre d'images dans la série : en moyenne la PG baisse entre 0,10% et 0,74%. Parfois ce changement est légèrement bénéfique pour certaines configurations d'apprentissage (le cas de quatre images par an), mais pas pour le reste des cas.

La conclusion de cette étude est qu'il vaut toujours mieux définir des partitions floues que des partitions non-floues, et que la définition de ces partitions floues doit respecter le mieux possible la relation entre les valeurs de l'indicateur à partitionner et le phénomène étudié.

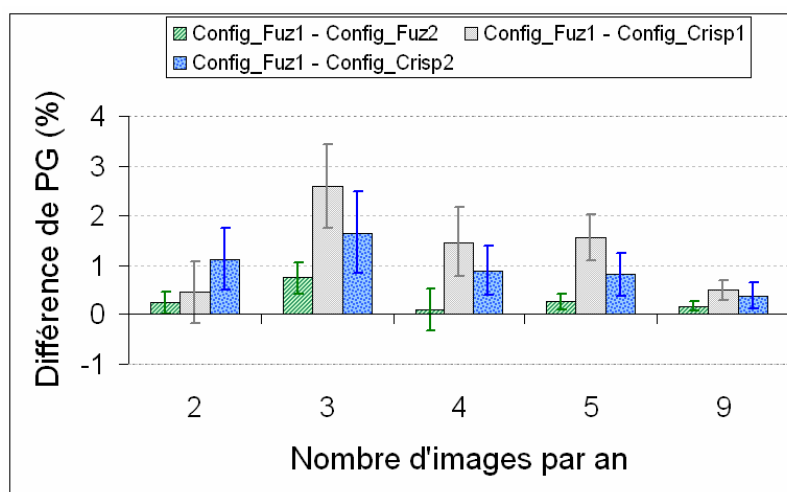


Figure 55. Différences dans les valeurs de précision globale (PG) obtenues en comparant la configuration experte de partition « Config_Fuz1 » aux configurations « Config_Fuz2 », « Config_Crisp1 » et « Config_Crisp2 ». Chaque barre représente la valeur moyenne calculée pour les différentes configurations d'apprentissage, et la barre d'erreur représente l'écart-type.

5.3.4. Contribution des différentes sources d'informations

Etant donné que la génération automatique des règles par apprentissage permet de supprimer ou d'intégrer facilement un ou plusieurs indicateurs d'entrée, nous pouvons évaluer la contribution des différentes sources aux performances du système.

L'isolement de la contribution du modèle, pour son évaluation, est facile à réaliser ; il suffit d'ignorer les indicateurs qui sont liés directement au modèle et de générer ensuite une base de règles en utilisant les autres indicateurs.

Concernant la contribution des connaissances expertes, l'évaluation de son intégralité est impossible. En effet, ces connaissances interviennent dans différentes parties du système :

- dans le choix des indicateurs utiles pour répondre à la problématique ;
- dans la définition des partitions de ces indicateurs ;
- dans l'intégration de certaines informations sur le cycle de culture ;
- dans la définition des règles de décision (le cas de la première version).

Par conséquent, nous nous sommes limités à l'évaluation de la partie des connaissances expertes que nous avons pu isoler facilement : il s'agit de l'information liée à la longueur du cycle nominal d'une parcelle de canne à sucre.

Etant donné que les indicateurs liés au MIR ont été ignorés dans la première version du système (leur information nous semblait redondante avec celles issue du NDVI), nous sommes également intéressés à l'évaluation de leur contribution spécifique.

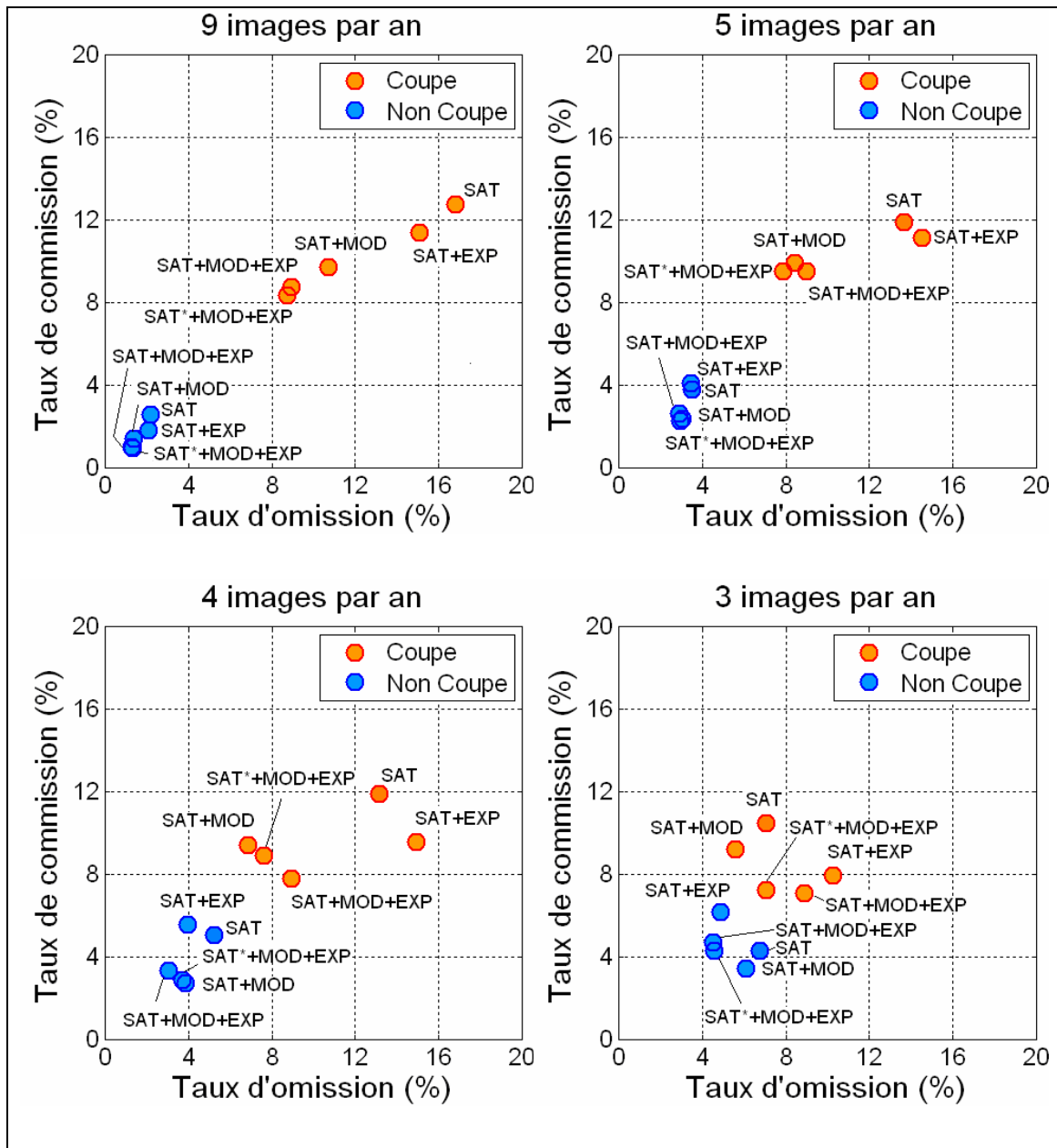


Figure 56. Performances de la deuxième version du système pour les différentes fréquences d'acquisition en utilisant les différentes combinaisons de sources d'informations (SAT : série temporelle d'images satellite ; SAT* : idem que SAT mais sans les indicateurs liés au Moyen Infra-rouge; MOD : modèle de canne ; EXP : une partie des connaissances expertes).

Ainsi, nous avons appliqué le système sur les différentes séries d'images, en 2003 et 2004, en considérant cinq combinaisons de sources d'informations (SAT pour série d'images, MOD pour modèle de canne et EXP pour la partie des connaissances expertes) :

- la première combinaison (SAT + MOD + EXP) tient compte de tous les indicateurs issus des trois sources ;
- la deuxième combinaison (SAT + EXP) ignore les indicateurs liés au modèle de culture (*i.e.* In13 et In14) ;
- la troisième combinaison (SAT + MOD) ignore l'indicateur lié directement aux connaissances expertes sur la longueur nominal du cycle de la canne (*i.e.* In 12) ;
- la quatrième combinaison (SAT* + MOD + EXP) ignore, comme lors de la construction de la base de règles experte de la première version, les indicateurs liés au MIR issus de la série temporelle (*i.e.* In3, In4 et In6) ;
- La cinquième combinaison considère seulement les indicateurs provenant de SAT.

La Figure 56 montre les performances, en termes de taux d'omission et de commission des décisions « Coupe » et « Non Coupe », obtenues avec les cinq combinaisons de sources d'information. Plus le point est proche de zéro, meilleurs sont les résultats. En ce qui concerne la classe « Coupe », nous pouvons constater, en comparant les résultats de SAT avec ceux des autres combinaisons, que :

- MOD réduit le taux de commission (entre 1,5% et 3,1%) et surtout le taux d'omission (entre 1,6% et 6,4%) quel que soit le nombre d'images par an ;
- EXP diminue le taux de commission (entre 0,8% et 2,9%) et tend à augmenter le taux d'omission (entre 0,9% et 3,4%) pour un nombre d'images par an inférieure ou égale à 5 ;
- L'intégration simultanée de MOD et d'EXP avec SAT réduit le taux de commission (entre 2,2% et 4,1%) et d'omission (entre 4,3% et 7,8%, sauf pour 3 images par an).
- La contribution du MIR est quasi-nulle voire négative dans le taux de commission (entre -1,1% et 0,5) et positive dans le taux d'omission (entre 0,3% et 2%).

Pour la classe « Non Coupe » :

- MOD réduit à la fois le taux de commission (entre 0,4% et 1,3%) et d'omission (entre 0,4% et 1,3%) quel que soit le nombre d'images par an ;
- EXP contribue faiblement (inférieure à 1%) mais positivement dans la réduction des taux d'omission et de commission avec 9 images par. Pour des fréquences d'acquisition inférieurs ou égaux à 5, elle réduit le taux d'omission (entre 0,1% et 0,7%) mais augmente le taux de commission (entre 0,2% et 2,6%) ;
- L'intégration simultanée de MOD et d'EXP avec SAT réduit le taux d'omission (entre 0,6% et 2,3%) et de commission (entre 1,1% et 2%, sauf pour 3 images par an).
- La contribution du MIR est très faible et parfois négative dans le taux de commission (entre -0,2% et 0,9%), et quasi-nulle dans le taux d'omission.

La Figure 57 illustre la contribution du modèle de canne, de la partie des connaissances expertes liée au cycle de la canne et du MIR dans la précision globale du système. En ce qui concerne la contribution du modèle de canne, nous remarquons que sa moyenne (pour toutes les configurations d'apprentissage) est positive quel que soit le nombre d'images par an : elle varie entre 1,27% (3 images par an) et 2,5% (4 images par an). Au delà de 4 images par an, cette contribution baisse jusqu'à 1,44% (avec 9 images par an).

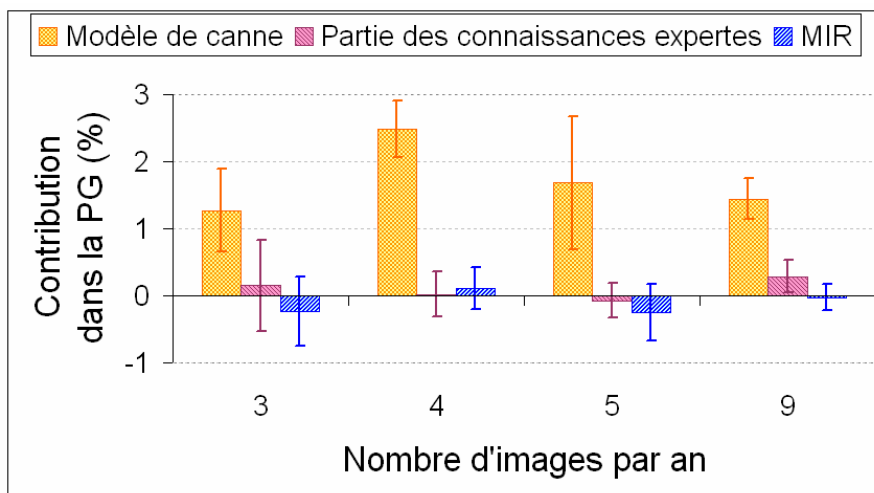


Figure 57. Contributions du modèle de canne à sucre, d'une partie des connaissances expertes et du MIR dans la précision globale (PG) du système à base de règles induites par apprentissage.

Quant à la contribution de la partie des connaissances expertes, sa moyenne est faible, voire parfois négative (avec 5 images par an). Cette moyenne varie entre -0,07% et 0,29%. L'écart-type de cette contribution par rapport aux différentes configurations d'apprentissage est relativement important, et ce quelle que soit la fréquence d'acquisition ; sa valeur varie entre 0,24% et 0,68%.

La contribution du MIR dans la précision globale est très faible et aléatoire : sa moyenne varie entre -0,25% et 0,23%, et son écart-type entre 0,20% et 0,51%. Cette contribution quasi nulle du MIR est due au fait que l'information apportée par le NDVI dans les zones qui ne subissent pas de stress hydrique est suffisante (le cas de nos exploitations). Dans les zones où ce facteur est présent, l'apport du MIR peut devenir important (Bégué *et al.*, 2008).

Analysons maintenant la contribution du modèle de canne à sucre dans la précision de chacune des deux classes « Coupe » et « Non coupe » séparément (Figure 58). Nous constatons que pour la classe « Coupe » la contribution est positive quel que soit le rythme d'acquisition : elle varie entre 1,57% (pour 3 images par an) et 6,39% (5 images par an). Cette contribution est en effet primordiale pour une utilisation opérationnelle du système car elle permet de passer en dessous du seuil des 5% d'erreur sur les surfaces attendu par la filière. Pour la classe « Non coupe » la contribution est toujours positive mais moins importante : sa valeur varie entre 0,4% (5 images par an) et 1,28% (4 images par an).

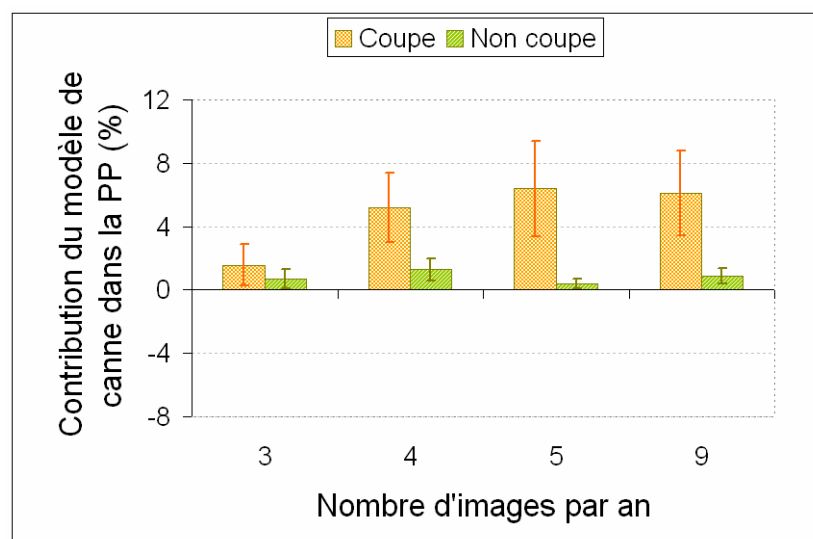


Figure 58. Contribution du modèle de canne à sucre dans les précisions du producteur pour la classe « Coupe » (PP Coupe) et « Non Coupe ». Les barres d'histogramme correspondent à la moyenne obtenue pour les différentes configurations d'apprentissage ; la barre d'erreur représente l'écart-type.

5.3.5. Validation pluriannuelle

Une étude préliminaire de généralité des règles induites automatiquement a été réalisée sur les bases générées à partir des données de 2003 et 2004 (en utilisant différents pourcentages de jeu d'apprentissage). Nous avons utilisé ces bases de règles pour détecter les coupes sur l'exploitation de M. Caruel à partir de la série d'images acquises en 2006 (Figure 41).

La valeur moyenne de précision globale du système obtenue pour les différentes configurations d'apprentissage est 94,12%. L'écart-type est égal à 0,56%. Nous constatons ainsi que les règles induites par apprentissage ont potentiellement un caractère général. Toutefois, ces règles dépendent de la disposition des images dans la saison et du climat influant sur l'indicateur issu de la modélisation. Afin de conclure définitivement sur la validité temporelle des règles induites, davantage analyses sont nécessaires.

5.4. Comparaison

Nous pouvons maintenant comparer les deux versions du système. Cette comparaison porte sur les performances, les bases de règles, le caractère opérationnel et la modularité de chacune des versions.

Performances

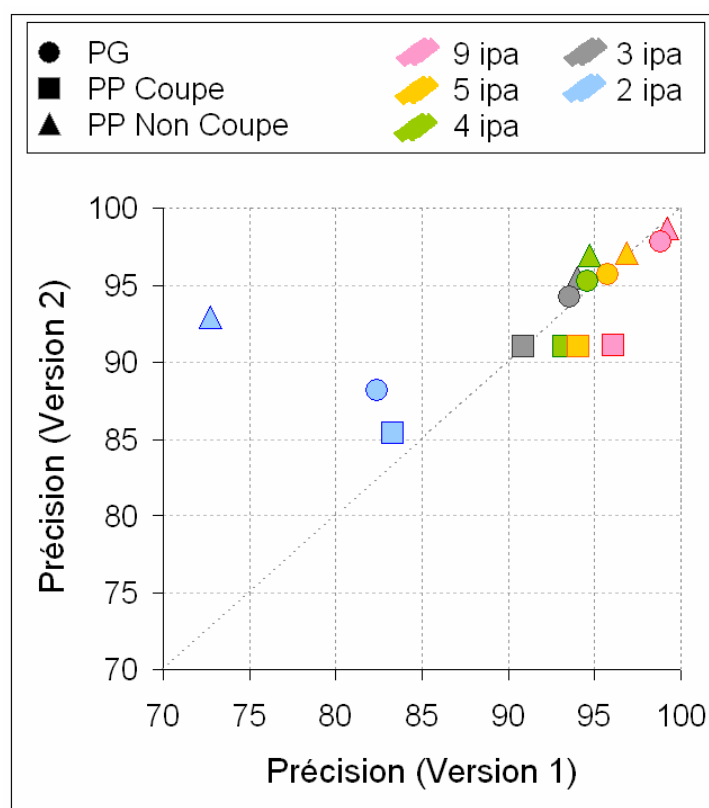


Figure 59 Comparaison entre les performances des deux versions du système (Version 1 : règles définies par expertise ; Version 2 : règles induites par apprentissage) en les appliquant sur les exploitations de M. Barau et M. Caruel en 2003 et 2004 avec des séries temporelles de 9, 5, 4, 3 et 2 images par an (ipa). (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement)

La Figure 59 compare les performances des deux versions en les appliquant sur des séries temporelles de 9, 5, 4, 3 et 2 images par an, sur les exploitations de M. Barau et M. Caruel en 2003 et 2004. D'après cette figure, on remarque que les deux versions fournissent des résultats avec des précisions globales (PG) très proches et très satisfaisantes en les appliquant sur des séries qui comportent au moins 3 images par an (une avant la campagne de récolte, une pendant et une après) ; les précisions du producteur pour la classe « Non coupe » (PP Non coupe) avec ce rythme d'acquisition sont également très voisines. Pour un rythme d'acquisition supérieur à 4 images par an, la précision du producteur pour la classe Coupe (PP

Coupe) est meilleure avec la première version. Avec des séries de 2 images par an seulement (une avant et une après la campagne de récolte) la deuxième version donne de meilleurs résultats, notamment en terme de PP Non coupe ; en effet cette différence est due au fait que la première version favorise, pour certain cas, la décision « Je ne sais pas » lorsque les acquisitions sont très éloignées, ce qui n'est pas le cas avec la deuxième version qui doit toujours décider « Coupe » ou « Non coupe ».

Règles

La différence majeure entre les bases de règles des deux versions est leur taille. La base de règle définie par expertise comprend 116 règles, alors que le nombre de règles générées avec la deuxième version est beaucoup moins important : à titre d'exemple, la taille des bases de règles générées pour faire la détection des coupes sur les exploitations de M. Barau et M. Caurel en 2003 et en 2004 varie entre 13 et 39 règles (Figure 53) pour les différentes séries d'images, ce qui est largement en dessous de 116.

Outre le nombre de règles, nous avons remarqué que la longueur des règles générées automatiquement par apprentissage est plus courte que celles des règles expertes. Leurs prémisses comprennent moins d'indicateurs, ainsi leur interprétation est plus facile.

Caractère opérationnel

Concernant le côté opérationnel, la première version du système a montré une légère supériorité sur la deuxième. En effet, sa base de règles est complète, ce qui fait qu'on peut l'appliquer quels que soient la série temporelle, le site d'étude et l'année, en changeant si nécessaire les partitions des indicateurs ; ceci a été prouvé par les bons résultats obtenus sur les différents sites d'étude pour les campagnes de 2003, 2004, 2006 et 2007. La deuxième version peut aussi être utilisée d'une façon opérationnelle ; afin de définir les règles de décision, deux choix sont possibles : soit l'utilisateur fait un apprentissage sur un faible pourcentage de données (10 à 20% sont suffisants d'après nos résultats), soit il utilise des règles qui sont déjà générées à partir des séries antérieures. La validation de ce dernier point nécessite des analyses complémentaires pour étudier la robustesse du système (1) par rapport à la différence du positionnement temporelle des images entre la série qui a servi pour générer les règles et celle sur laquelle il sera appliqué, et (2) par rapport aux différences des conditions climatiques entre les séries.

Modularité

La deuxième version du système est modulaire. Étant donné que la génération de règles de décisions dans cette version se fait par apprentissage, l'utilisateur peut inclure les indicateurs qu'il souhaite. Ceci n'est pas le cas avec la première version ; à chaque fois que le module d'indicateurs d'entrée est modifié, une nouvelle base de règles doit être redéfinie par expertise.

5.5. Conclusion

Dans ce dernier chapitre, nous avons évalué le système d'aide à la décision que nous avons présenté dans le Chapitre 4. Les deux versions du système, celle à base de règles expertes et celle à base de règles induites par apprentissage, ont été appliquées sur nos différents sites d'étude afin de détecter les coupes de canne à partir de séries temporelles d'images acquises à des années différentes.

Les comparaisons des décisions avec la vérité terrain ont montré que les deux versions du système fournissent des résultats satisfaisants en utilisant des séries qui comportent au moins 3 images par an (une avant la campagne de récolte, une pendant et une après) : la précision globale dans ce cas est supérieure à 93,5%. Elle s'accroît avec l'augmentation du nombre d'images dans la série pour atteindre une valeur supérieure à 97,8% avec 9 images par an (pour une année donnée). Les performances obtenues ne dépendent pas seulement du nombre d'images dans la série, mais aussi de la disposition de ces images dans le temps.

La contribution des ensembles flous dans la précision globale a également été étudiée : elle atteint 15% avec la première version et 2,5% avec la deuxième. La faible contribution dans la deuxième version est principalement due à la faible profondeur des arbres de décision construits par apprentissage. En effet, moins on a d'indicateurs dans les prémisses des règles, plus des variables floues avec des marges d'ambiguïté importantes (*e.g.* le NDVI) risquent d'être absentes, et moins l'apport des ensembles flous est remarquable. Toutefois, il vaut mieux toujours définir des partitions floues plutôt que non-floues, et cette définition de partitions doit respecter le mieux possible la relation entre les valeurs des indicateurs à partitionner et le phénomène étudié.

Concernant la contribution des différentes sources d'information, son évaluation n'a été réalisée que pour la deuxième version. En effet, pour la première version, la difficulté de cette évaluation est liée au fait que l'élimination d'une source nécessite la redéfinition manuelle (par expertise) d'une nouvelle base de règles ; cette tâche est facilitée avec la deuxième version grâce à la génération automatique des règles par apprentissage. Les résultats révèlent que la contribution du modèle de croissance atteint 2,5% dans la précision globale et 6,39% dans la précision de détection des coupes. Cette source contribue ainsi à ce que le système ne dépasse pas le seuil de 5% d'erreur sur les surfaces coupées, ce qui correspond à la demande de la filière. Cependant, connaissant le rôle important que joue cette source dans les règles

expertes de décision, nous pensons que sa contribution est beaucoup plus significative dans la première version ; nous laissons l'étude de ce point en perspective.

Quant à la contribution de la « source » connaissances expertes, elle a été difficile à évaluer du fait que ces connaissances interviennent dans toutes les parties du système.

La fin du chapitre a été consacrée à une comparaison entre les caractéristiques des deux versions du système. Nous avons conclu, entre autre, que la première version est plus générique que la deuxième. En effet, une fois la base de règles expertes définie, le système peut être appliqué sur différents sites et différentes années. Il suffit pour cela de modifier éventuellement les partitions des indicateurs pour s'adapter à des conditions locales. Le temps à investir pour la définition des règles expertes reste un point de discussion. La deuxième version quant à elle se distingue par davantage de modularité qui permet l'ajout et la suppression des sources d'information. Sa validation pluriannuelle reste un point à approfondir.

6. Conclusions et perspectives

Au moment où les technologies spatiales et les capacités d'acquisition ne cessent d'évoluer, la communauté des chercheurs en télédétection se retrouve devant un besoin grandissant en méthodes innovantes de traitement d'images et d'analyse de séries temporelles à haute résolution spatiale. Le riche contenu informatif de ces données est susceptible d'intéresser un large panel d'applications, notamment les applications agricoles.

Les travaux de cette thèse se sont appuyés sur des séries temporelles d'images SPOT de la base de données Kalideos ISLE-Réunion[©]. L'application agricole considérée est le suivi des coupes de canne à sucre à La Réunion ; le choix de cette application a été motivé d'une part par les besoins de la filière « canne » dans le monde, et d'autre part, par le défi méthodologique lié à la forte variabilité spatio-temporelle de cette culture.

Nous nous sommes intéressés à l'ensemble de la chaîne d'exploitation des séries temporelles, de la phase du pré-traitement des images jusqu'à la conception d'un système d'aide à la décision.

6.1. Contributions de la thèse

Pré-traitement des séries temporelles

Nous avons d'abord développé et validé une méthode de sélection automatique de cibles invariantes pour faire une normalisation relative des images par rapport à une date de référence. La normalisation relative tente de minimiser d'une manière uniforme, les différences (d'une date à une autre) de conditions atmosphériques, d'éclairement solaire, d'effets directionnels, et de calibration de capteur.

Comparée à une méthode de correction atmosphérique basée sur le code 6S, les résultats montrent que la méthode que nous proposons peut être une bonne alternative à la correction atmosphérique, qui reste difficile à mettre en œuvre.

Les avantages principaux de notre méthode de normalisation sont sa simplicité, son caractère automatique, son applicabilité aux images acquises dans différentes saisons, et le fait qu'elle n'ait pas besoin de données atmosphériques. Sa limitation principale est que les corrections restent « relatives » et dépendent donc du choix de l'image de référence.

Système d'aide à la décision

Pour aller au-delà des méthodes habituelles de détection du changement (*e.g.* classification), dont la mise en œuvre et l'interprétation souffrent de l'irrégularité des acquisitions et du manque d'automatisation, nous avons ensuite développé une méthode qui fusionne des informations issues des images satellite, des simulations de modèle de culture et des connaissances expertes. La mobilisation de ces diverses sources permet de :

- compenser l'absence de données affectant les informations provenant des images (*e.g.* profils temporels de NDVI). Cette absence est souvent liée soit à la présence des nuages soit au manque d'acquisition dû à la programmation satellite et au coût des images ;
- imposer des contraintes temporelles et environnementales qui facilitent le suivi des opérations culturales ;
- concevoir des règles de décision qui automatisent le raisonnement expert humain.

Nous avons présenté la méthode au travers d'un système d'inférence floue qui détecte les coupes de canne à sucre à partir des séries d'images SPOT. Ce système est basé sur la modélisation linguistique et sur la logique floue qui permettent, d'une part, de manipuler des données imprécises, incertaines et de nature hétérogène, et d'autre part, de construire des règles de décision interprétables. Deux méthodes de construction de base de règles ont été proposées : la première repose sur l'élaboration des règles par expertise et la deuxième par apprentissage.

Les performances des deux versions du système, celle à base de règles expertes et celle à base de règles apprises, ont été évaluées sur différents sites d'étude à La Réunion et sur

plusieurs années. Les résultats obtenus sont satisfaisants en utilisant des séries qui comportent au moins 3 images par an : la précision globale est en effet supérieure à 93,5% et le pourcentage de bonne détection des coupes est supérieur à 90%. L'augmentation du nombre d'images dans la série améliore les performances ; avec 9 images par an, la précision globale atteint 98,8% et le pourcentage de bonne détection des coupes 96,1%. Il est à signaler que les performances obtenues ne dépendent pas seulement du nombre d'images dans la série, mais aussi de leur position par rapport à la campagne de récolte ; il faut disposer d'une image peu avant la campagne, une peu après et quelques unes, si possible, réparties au cours de la campagne. Une image (ou plus) au début du cycle permet de renseigner davantage sur la croissance des parcelles et d'améliorer les performances du système.

Nous avons étudié la contribution des différentes sources d'information avec le système à base de règles induites automatiquement ; cette étude aurait été laborieuse à mettre en place avec la première version étant donné la difficulté de définir une nouvelle base de règles expertes dès que les paramètres d'entrée changent. Les résultats ont montré que le modèle de culture apporte jusqu'à 2,5% de plus dans la précision globale et 6,4% dans la précision de détection des coupes ; cette contribution est primordiale pour une utilisation opérationnelle du système car elle permet de passer en dessous du seuil des 5% d'erreur sur les surfaces attendu par la filière. Néanmoins, étant donné le rôle important que joue ce modèle dans les règles expertes de décision, nous pensons que sa contribution est beaucoup plus significative dans la première version. L'apport de la « source » connaissances expertes a été difficile à évaluer du fait que celles-ci sont impliquées dans toutes les parties du système et sont difficiles à isoler.

Enfin, une comparaison des caractéristiques des deux versions du système a été réalisée. Nous avons constaté que les précisions globales sont très similaires. La première version est plus générique que la deuxième, dans la mesure où, une fois la base de règles expertes définie, le système peut être appliqué sur différents sites et différentes années, en modifiant éventuellement les partitions des indicateurs pour s'adapter à des conditions locales. Cependant, le temps investi pour la définition des règles expertes reste le point faible de cette approche. La deuxième version se distingue par davantage de modularité qui permet l'ajout et la suppression des sources d'information. La principale faiblesse est ici la validité pluriannuelle de cette approche et le besoin de recourir à des données de terrain pour l'apprentissage.

6.2. Perspectives

Plusieurs axes de recherche et développement se dégagent de ces travaux. Ces axes portent d'une part sur l'amélioration des méthodes développées pour le suivi de la canne à sucre, et d'autre part sur l'applicabilité de ces méthodes sur d'autres problématiques agricoles et d'autres données satellite.

Suivi de la canne à sucre

En ce qui concerne le suivi de la canne à sucre, les résultats obtenus pour la détection des coupes sont encourageants et incitent à intégrer le système développé dans l'offre de service proposée par le CIRAD à l'ensemble de la filière canne à sucre à La Réunion (Bégué *et al.*, 2005). En effet, la précision atteinte par le système et son aspect opérationnel permettent de répondre au besoin exprimé par la filière.

Le champ d'application du système pourrait être également étendu pour traiter une autre application qui intéresse la filière : le suivi des plantations. Pour ce faire, il serait envisageable de construire de façon simple des indicateurs supplémentaires basés sur les profils temporels du Proche-Infrarouge afin de profiter du pouvoir discriminant de cette bande spectrale entre le sol nu et la canne verte.

L'évaluation du système développé pour la détection des coupes sur d'autres sites d'étude est souhaitée, notamment dans les régions où les ressources en eau sont limitées. Nous pensons que dans de telles zones, l'apport de l'information dans le Moyen-Infrarouge sera important pour réduire les confusions d'état de parcelle liées à la diminution significative de NDVI dû au stress hydrique des plantes. Il serait ainsi intéressant d'intégrer dans les règles de décision expertes (la première version du système) des indicateurs basés sur le Moyen-Infrarouge.

Un autre point important qui n'a pu être approfondi dans cette thèse mais qui semble pertinent à aborder, est la validation pluriannuelle des bases de règles induites par apprentissage. D'après les résultats obtenus, nous avons constaté que des performances globales très satisfaisantes peuvent être assurées avec des règles induites automatiquement en utilisant un faible jeu d'apprentissage (*e.g.* 10% de données dans une série). Cette génération automatique représente ainsi une excellente alternative à la définition manuelle, experte, des règles linguistiques interprétables. Cependant la validité pluriannuelle des règles induites

dépend de leur robustesse, d'une part par rapport à la diversité climatique d'une année sur l'autre, et d'autre part par rapport à la disposition des images dans la série. Les variations climatiques influent sur l'information provenant du modèle de culture ; les règles induites à partir d'une série d'images acquises à une certaine année climatique et qui intègrent l'indicateur du modèle risquent de ne pas être correctes pour d'autres années. Une solution envisageable consiste à rajouter une variable linguistique représentant le climat (exemple de labels « Année sèche », « Année humide », ou autre), puis à créer une base de règles « pluriannuelle » en faisant un apprentissage avec des données appartenant à plusieurs années climatiques. Pour assurer la robustesse par rapport à la disposition des images dans la série, un apprentissage sur différentes séries avec des images réparties différemment serait également opportun.

Il serait également intéressant de penser à une méthode hybride de génération des règles de décision. Etant donné l'aspect linguistique de ces règles, l'on peut faire une génération automatique par apprentissage en utilisant les deux sources série temporelle et connaissances expertes, et intégrer ensuite manuellement (par expertise) des prémisses liées au modèle de culture.

Enfin, la modularité du système à base de règles induites ouvre les portes à l'intégration facile d'autres sources d'information. Une source intéressante pour le suivi des coupes est l'imagerie radar qui est insensible à la couverture nuageuse. Les évolutions techniques en plein essor et l'arrivée sur le marché d'images radar d'une résolution spatiale métrique (e.g. Terra SAR-X et COSMO-SkyMed) font de cette source un outil prometteur pour le suivi des opérations culturales. L'alimentation du système avec des indicateurs radar permettrait de fournir des décisions pour les parcelles couvertes par les nuages.

Une méthodologie transférable

Les méthodes développées dans cette thèse sont applicables à d'autres jeux de données et transposables à d'autres applications agricoles.

La méthode de normalisation relative peut être utilisée pour traiter des images acquises à haute résolution spatiale avec d'autres capteurs que SPOT (e.g. IRS, KOMPSAT 2, IKONOS), sur différentes zones. La sélection automatique des cibles invariantes rend la mise en place de cette méthode simple et rapide. Toutefois la sélection fiable de ces cibles exige la disponibilité des masques de nuages et de zones végétatives changeantes ; ces masques

peuvent être définis d'une façon surdimensionnée spatialement étant donné le grand nombre de cibles invariantes qui peuvent être détectées dans une scène.

La conception du système de détection des changements proposée dans cette thèse est générique même si l'ensemble a été illustré par une application agricole. Le concept d'inférence à base de règles linguistiques sur lequel nous nous basons ouvre des perspectives pour l'adaptation de la méthode à la gestion d'autres applications : suivi en temps réel d'autres opérations culturales ou cartographie dynamique de l'occupation du sol. Cette dernière application permettrait de répondre à un besoin de base exprimé à la fois par les modélisateurs et par les gestionnaires.

Le système présenté s'appuie également sur un modèle de culture et sur des connaissances expertes. Nombre de modèles qui simulent la croissance des principales cultures, en faisant appel à des données climatiques et des paramètres biophysiques, existent, *e.g.* STICS (INRA, STICS), SARRA (Baron *et al.*, 2005). Ces modèles peuvent être utilisés à la place du modèle MOSICAS dans le système pour informer sur les processus les plus explicatifs de la croissance des cultures et pour fournir des indicateurs utiles pour le suivi des opérations culturales et/ou de l'occupation du sol. Les connaissances expertes concernant les cycles des principales cultures ainsi que les pratiques et les itinéraires culturaux peuvent être obtenues à partir des agriculteurs ou des bases de connaissance agricoles

Les méthodes développées dans cette thèse pourraient être appliquées aux produits des systèmes spatiaux conçus pour le suivi temporel, *e.g.* Formosat-2, Sentinel-2 (ESA-UE), RapidEye (DLR), etc. A terme, la généralisation de ces méthodes sur un grand nombre de paysages agricoles (européens, méditerranéens, etc.) pourrait offrir de nouvelles opportunités dans les prises de décision et le suivi des surfaces agricoles par télédétection.

Liste des tableaux

Tableau 1. Caractéristiques des principaux capteurs fournissant des séries temporelles d'images.....	18
Tableau 2. Caractéristiques des images SPOT issues de la base de données Kalideos ISLE-Réunion et utilisées dans cette thèse. (XI : mode multi-spectral de SPOT 4 ; J : mode multi-spectral de SPOT 5).	41
Tableau 3. Caractéristiques des trois exploitations de canne à sucre à La Réunion.....	46
Tableau 4. Paramètres atmosphériques aux dates d'acquisition d'images SPOT : Pression atmosphérique (Pression), taux de vapeur d'eau (H ₂ O), taux d'Ozone (Ozone) et épaisseur optique à 550 nm (τ_{550}).....	67
Tableau 5. Valeurs moyennes et maximales des écarts types (sur 18 dates) calculées pour les réflectances des cibles invariantes de validation dans les quatre bandes spectrales (B1 : Vert ; B2 : Rouge ; B3 : PIR ; B4 : MIR).	72
Tableau 6. Valeurs des coefficients a et b nécessaires pour l'estimation du LAI selon la variété de canne à la Réunion (Source : Martiné, 2004).....	88
Tableau 7. Définition des variables extraites à partir de la série temporelle d'images satellite (SAT), du modèle de culture (MOD) et des connaissances expertes (EXP).....	100
Tableau 8. Définition et formulation des indicateurs utilisés dans le système d'aide à la décision. Figurent aussi les variables impliquées dans chaque définition d'une façon directe ou indirecte (*), <i>i.e.</i> dans la partition, ainsi que les sources d'informations utilisées (SAT : Série temporelle d'images satellite ; MOD : Modèle de culture ; EXP : Connaissances expertes).....	103
Tableau 9. Intitulés et descriptions des labels linguistiques associés aux partitions des indicateurs, et marges d'ambiguïté (entre les ensembles) associées aux indicateurs flous.....	104
Tableau 10. Exemples de règles définies par expertise, avec des illustrations indiquant respectivement en blanc, bleu foncé et bleu clair les positions des points NDVI (t), NDVI (t') et NDVI (t''), et en rouge les acquisitions avec nuages.....	111
Tableau 11. Matrice de confusion obtenue en comparant les décisions prises à l'aide du système à base de règles expertes (pour μ_{conf} égal à 0) avec la vérité terrain sur les exploitations de M. Barau et M. Caruel en 2002 et 2003.	126
Tableau 12. Exemple d'une base de règles générée par le logiciel Fispro [®] en utilisant 10% de données de Barau et Caruel en 2003 et 2004 pour faire l'apprentissage (6 parcelles)	138

Liste des figures

Figure 1. Localisation de l'île de La Réunion.	31
Figure 2. Photo de canne à sucre à La Réunion (© V. Lebourgeois)	32
Figure 3. Classification des exploitations cannières selon leur surface en canne (Source : DAF Réunion, Février 2006).	33
Figure 4. Coupe mécanique d'une parcelle de canne à la Réunion (© M. El Hajj).	34
Figure 5. Dates d'acquisition des images SPOT dans les séries temporelles de 2003, 2004, 2006 et 2007.....	40
Figure 6. Localisation des exploitations de M.Barau, M. Caruel et M.Isautier sur l'île de la Réunion, et des stations météorologiques.....	43
Figure 7. Précipitations mensuelles enregistrées en 2003, 2004, 2006 et 2007 dans les stations météorologiques situées à coté des exploitations de M. Caruel (Nord), M. Barau (Nord) et de M. Isautier (Sud).	45
Figure 8 Champ de vue des instruments HRG 1&2 du satellite SPOT-5.....	51
Figure 9. Organigramme de la technique de sélection automatique des cibles invariantes.....	59
Figure 10. Histogramme et valeur moyenne obtenus pour le canal Vert de l'image de différence calculée à partir des images acquises le 26 février 2003 et le 13 mai 2004 (référence) : a) avant l'application des masques, b) après l'application du masque des nuages, c) après l'application de tous les masques (nuages, saturation, parcelles de canne, océan).....	60
Figure 11. Cibles invariantes sélectionnées automatiquement sur l'ensemble des images : chaque point noir correspond à une cible.....	61
Figure 12. Régressions linéaires calculées automatiquement dans les quatre bandes spectrales (Vert, Rouge, PIR et MIR) pour normaliser l'image du 06 novembre 2004 par rapport à celle du 13 mai 2004.	62
Figure 13. Les bandes spectrales des satellites SPOT-4&5 (B1 = Vert ; B2 = Rouge ; B3 = PIR ; B4 = MIR) et les transmissions des gaz entre 0,45µm et 1,9µm (source : le modèle atmosphérique US62 avec une teneur en vapeur d'eau de 1,424 g.cm ⁻² et une quantité d'ozone de 0,344 cm.atm).	63
Figure 14. Illustration des phases de correction radiométrique d'une image SPOT pour passer du compte numérique à la réflectance au sol.	66
Figure 15. Localisation (en jaune) des cibles invariantes sélectionnées manuellement (CI _{Man}) sur l'image SPOT-5 (Rouge : B4 ; Vert : B3 ; Bleu : B1) acquise le 13/05/2004.....	68
Figure 16. Valeurs moyennes des réflectances de 20 cibles invariantes sélectionnées manuellement utilisées pour valider la technique de sélection automatique.....	69
Figure 17. Comparaison des réflectances normalisées obtenues par les deux méthodes de sélection des cibles invariantes (automatique et manuel) pour un jeu de données de 20 cibles indépendantes sélectionnées par interprétation visuelle.....	70
Figure 18. Comparaison des réflectances des cibles invariantes sélectionnées manuellement corrigées par 6S (réflectances TOC) et celles corrigées par la normalisation radiométrique relative.....	71
Figure 19. Exemples de profils temporels de cibles invariantes (forêt dense, bâtis, lave volcanique) dans le MIR obtenus à partir des images en réflectance TOA, des images en réflectance TOC et des images en réflectance TOA normalisée.....	73
Figure 20. Comparaison des valeurs de NDVI calculées à l'échelle de la parcelle à partir des images corrigées par 6S (NDVI TOC) et des images normalisées (à toutes les dates et pour toutes les parcelles).....	74
Figure 21. a) Comparaison des valeurs de pente dans les profils temporels de NDVI corrigés par 6S (NDVI TOC) et ceux normalisés relativement. b) Profils temporels de NDVI calculé pour une parcelle de canne à partir des TOA réflectances, des réflectances TOA relativement normalisée et des réflectances TOC.	75
Figure 22. Exemple de profils temporels de NDVI et de réflectances calculés à l'échelle de la parcelle pour a) une repousse et b) une plantation.	82

Figure 23. Variabilité spatio-temporelle du NDVI à l'échelle de l'exploitation de M. Caruel. En bleu les profils temporels de NDVI pour quelques parcelles en repousse et en rouge le profil pour une parcelle nouvellement plantée.	84
Figure 24. Exemples de profils temporels incomplets de NDVI et de MIR pour une repousse. La coupe de 2003 de cette parcelle ne peut pas être détectée par les différences bi-dates de NDVI et de MIR.	86
Figure 25. Relation entre le LAI (mesuré <i>in situ</i>) et le NDVI (SPOT).	88
Figure 26. Des profils temporels de LAI, de NDVI simulé en utilisant le modèle MOSICAS et de NDVI calculé à partir des images SPOT pour la même parcelle.	89
Figure 27. Exemple de relations entre la date de coupe (date de début de simulation de NDVI par MOSICAS) et le temps nécessaire (T_n) pour qu'une parcelle de canne atteigne un niveau $NDVI_{Seuil}$ de NDVI. La date de coupe correspond dans le système expert à la date d'acquisition satellite t'	90
Figure 28. Un exemple illustrant comment les simulations de MOSICAS peuvent fournir une information utile en cas de manque d'acquisitions.	91
Figure 29. Comparaison entre les profils temporels de NDVI SPOT et ceux du NDVI simulés par MOSICAS pour une parcelle en repousse puis en plantation.	92
Figure 30. Des exemples de connaissances expertes qui permettent de définir des contraintes temporelles : a) la longueur nominale du cycle de canne ; b) les dates d'ouverture et de fermeture des usines sucrières.....	93
Figure 31. Exemple d'une partition de NDVI.	95
Figure 32. Structure et composants du système d'aide à la décision.	98
Figure 33. Ensembles flous du NDVI (<i>i.e.</i> indicateurs In_1 et In_2) avec des profils de NDVI de plusieurs parcelles en fonction de la somme de degrés-jour depuis la date de coupe.....	102
Figure 34. Intervalles temporels utilisés pour la localisation des dates d'acquisitions satellites.	105
Figure 35. Sortie du système d'aide à la décision avec la base de règles expertes (NC = Non Coupe ; JSP = Je ne Sais Pas ; C = Coupe).	106
Figure 36. Sous-zones (1 à 9) de profil temporel de NDVI.	107
Figure 37. Schéma général d'un arbre de décision flou. Les nœuds In_i, \dots, In_k sont des sous-ensembles de valeurs d'indicateurs, les FSS sont les sous-ensembles flous utilisés dans la partition de chaque indicateur, et C (coupe) et NC (Non Coupe) sont les possibilités de la sortie. Chaque chemin entre la racine et une feuille constitue une règle de décision (<i>e.g.</i> , Règle x : Si In_i est FSS_1 et In_j est FSS_3 Alors C).....	112
Figure 38. Illustration du processus de construction de jeu d'apprentissage pour la génération automatique des règles.	113
Figure 39. Exemple d'une règle induite par apprentissage et illustration du scénario correspondant.....	114
Figure 40. Illustration de la méthode d'inférence de Mamdani sur un exemple avec deux règles seulement et deux indicateurs d'entrée dont les valeurs sont respectivement x_1 et x_2	115
Figure 41. Séries temporelles utilisées pour évaluer le système à base de règles expertes sur les sites d'étude, et dates de début et de fin des campagnes de récolte.	123
Figure 42. Performances du système à base de règles expertes obtenues pour les parcelles de M. Barau et M. Caruel en 2003 et 2004 avec différents seuils de confiance μ_{conf} . (PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement ; PG : précision globale ; JSP : décisions « Je ne sais pas »).....	126
Figure 43. Taux d'erreurs de commission et d'omission des classes « Coupe » et « Non Coupe » obtenues pour les parcelles de M. Barau et M. Caruel en 2003 et 2004 avec différents seuil de confiance μ_{conf}	127
Figure 44. Valeurs moyennes de stabilité des décisions « Coupe » et « Non Coupe » obtenues lors de l'application du système à base de règles expertes sur les parcelles de M. Barau et M. Caruel en 2003 et 2004.	128
Figure 45. Performances globales du système à base de règles expertes lors de son application sur différents sites d'études et différentes années. (PG : précision globale ; JSP : décisions « Je ne sais pas » ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement ; VSM Coupe et VSM Non coupe : valeurs de stabilité moyenne des décisions « Coupe » et « Non coupe » respectivement). ..	129
Figure 46. Profils temporels de NDVI des parcelles pour lesquelles le système à base de règles expertes a pris de mauvaises décisions.	130
Figure 47. Un exemple illustrant le passage d'une partition floue à une partition non-floue.....	131
Figure 48. Sur l'axe le plus bas, les dates d'acquisition satellite de la série complète (9 images / an); Sur les autres axes, les échantillonnages réalisés pour construire des séries avec 5, 4, 3 et 2 images par an respectivement.	133
Figure 49. Performances du système à base de règles expertes en fonction du nombre d'images par an (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement).....	134
Figure 50. Performances du système à base de règles induites par apprentissage obtenues pour les parcelles de M. Barau et M. Caruel en 2003 et 2004, en utilisant différents pourcentages de jeu d'apprentissage (PG :	

précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement). Les valeurs représentées correspondent à la moyenne obtenue pour 10 itérations et la barre d'erreur correspond à l'écart type.	136
Figure 51. Nombres de règles générées automatiquement en utilisant les différents pourcentages de jeu d'apprentissage. Chaque barre représente le nombre moyen des règles obtenues pour 10 itérations et la barre d'erreur correspond à l'écart type.	137
Figure 52. Performances du système à base de règles induites par apprentissage en fonction du nombre d'images par an (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement). Les valeurs représentées correspondent à la moyenne obtenue pour les différentes configurations d'apprentissage ; la barre d'erreur correspond à l'écart-type.	139
Figure 53. Variation du nombre de règles générées par apprentissage en fonction du nombre d'images dans la série ; chaque barre correspond à la moyenne obtenue pour toutes les configurations d'apprentissage, et les barres d'erreur représentent les écarts-types.	140
Figure 54. Partitions du NDVI selon les configurations floues (a) « Config_Fuz1 » et (b) « Config_Fuz2 » et les configurations non-floues (c) « Config_Crisp1 » et (d) « Config_Crisp2 ».....	141
Figure 55. Différences dans les valeurs de précision globale (PG) obtenues en comparant la configuration experte de partition « Config_Fuz1 » aux configurations « Config_Fuz2 », « Config_Crisp1 » et « Config_Crisp2 ». Chaque barre représente la valeur moyenne calculée pour les différentes configurations d'apprentissage, et la barre d'erreur représente l'écart-type.	142
Figure 56. Performances de la deuxième version du système pour les différentes fréquences d'acquisition en utilisant les différentes combinaisons de sources d'informations (SAT : série temporelle d'images satellite ; SAT* : idem que SAT mais sans les indicateurs liés au Moyen Infra-rouge; MOD : modèle de canne ; EXP : une partie des connaissances expertes)	143
Figure 57. Contributions du modèle de canne à sucre, d'une partie des connaissances expertes et du MIR dans la précision globale (PG) du système à base de règles induites par apprentissage.....	145
Figure 58. Contribution du modèle de canne à sucre dans les précisions du producteur pour la classe « Coupe » (PP Coupe) et « Non Coupe ». Les barres d'histogramme correspondent à la moyenne obtenue pour les différentes configurations d'apprentissage ; la barre d'erreur représente l'écart-type.	146
Figure 59 Comparaison entre les performances des deux versions du système (Version 1 : règles définies par expertise ; Version 2 : règles induites par apprentissage) en les appliquant sur les exploitations de M. Barau et M. Caruel en 2003 et 2004 avec des séries temporelles de 9, 5, 4, 3 et 2 images par an (ipa). (PG : précision globale ; PP Coupe et PP Non coupe : précision du producteur pour les classes « Coupe » et « Non coupe » respectivement).....	148

Références

- Abdel-Rahman, E.M., Ahmed, F.B., 2008. The application of remote sensing techniques to sugarcane (*Saccharum spp. hybrid*) production: A review of the literature. *International Journal of Remote Sensing* 29, 3753-3767.
- Almeida, T.I.R., De Souza Filho, C.R., Rossetto, R., 2006. ASTER and Landsat ETM+ images applied to sugarcane yield forecast. *International Journal of Remote Sensing* 27, 4057-4069.
- Anderson, G.P., Chetwynd, J.H., Theriault, J.-M., Acharya, P., Berk, A., Robertson, D.C., Kneizys, F.X., Hoke, M.L., Abreu, L.W., Shettle, E.P., 1993. MODTRAN2: suitability for remote sensing. In: Kohnle, A., Miller Walter, B. (Eds.), *Proceedings of SPIE - The International Society for Optical Engineering*. Publ by Society of Photo-Optical Instrumentation Engineers, Orlando, FL, USA, pp. 514-525.
- Andrefouet, S., 2000. A fuzzy-possibilistic scheme of study for objects with indeterminate boundaries: Application to french polynesian reefs. *IEEE Transactions on Geoscience and Remote Sensing* 38, 257-270.
- Andréfouët, S., Muller-Karger, F.E., Hochberg, E.J., Hu, C., Carder, K.L., 2001. Change detection in shallow coral reef environments using Landsat 7 ETM+ data. *Remote Sensing of Environment* 78, 150-162.
- Azzali, S., Menenti, M., 2000. Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. *International Journal of Remote Sensing* 21, 973-996.
- Baldwin, J.F., 1991. A new approach to inference under uncertainty for knowledge based systems. *Symbolic and Quantitative Approaches to Uncertainty*, Marseille, pp. 107-114.
- Bappel, E., 2005. Apport de la télédétection aérospatiale pour l'aide à la gestion de la sole cannière réunionnaise. CIRAD. Thèse de doctorat de l'Université de La Réunion, p. 277.
- Bappel, E., Bégué, A., Martiné, J.F., Pellegrino, A., Siegmund, B., 2005. Assimilation of a biophysical parameter estimated by remote sensing using SPOT 4&5 data into a sugarcane yield forecasting model. ISSCT, Guatemala, p. 4.

- Baron, C., Sultan, B., Balme, M., Sarr, B., Traore, S., Lebel, T., Janicot, S., Dingkuhn, M., 2005. From GCM grid cell to agricultural plot: Scale issues affecting modelling of climate impact. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 2095-2108.
- Bastiaanssen, W.G.M., Ali, S., 2003. A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. *Agriculture, Ecosystems and Environment* 94, 321-340.
- Bastidas-Obando, E., Carbonell-Gonzalez, J., 2007. Evaluating the applicability of MODIS data for forecasting sugarcane yields in Colombia. *Proceedings of the International Society of Sugar Cane Technologists (ISSCT)*, Durban, South Africa, p. 4.
- Bayarjargal, Y., Karnieli, A., Bayasgalan, M., Khudulmur, S., Gandush, C., Tucker, C.J., 2006. A comparative study of NOAA-AVHRR derived drought indices using change vector analysis. *Remote Sensing of Environment* 105, 9-22.
- Bégué, A., Degenne, P., Baillarin, F., Lemonnier, H., Lebourgeois, V., Gargaros, D., 2005. *Projet SUCRETTE (SUivi de la Canne à sucRE par TélédéTEction)*, compte-rendu de fin de recherche. Laboratoire Geotrop du CIRAD AMIS, Montpellier, p. 114.
- Bégué, A., Degenne, P., Pellegrino, A., Todoroff, P., Baillarin, F., 2004. Application of remote sensing technology to monitor sugar cane cutting and planting in Guadeloupe (French West Indies). *Geomatica, Cuba*, p. 11.
- Bégué, A., Lebourgeois, V., Bappel, E., Todoroff, P., Pellegrino, A., Baillarin, F., Siegmund, B., 2008. Spatio-temporal variability of sugarcane fields and recommendations for yield forecast using NDVI. *International Journal of Remote Sensing (Accepted)*, 21p.
- Blackburn, G.A., Ferwerda, J.G., 2008. Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis. *Remote Sensing of Environment* 112, 1614-1632.
- Boles, S.H., Xiao, X., Liu, J., Zhang, Q., Munkhtuya, S., Chen, S., Ojima, D., 2004. Land cover characterization of Temperate East Asia using multi-temporal VEGETATION sensor data. *Remote Sensing of Environment* 90, 477-489.
- Bontemps, S., Bogaert, P., Titeux, N., Defourny, P., 2008. An object-based change detection method accounting for temporal dependences in time series with medium to coarse spatial resolution. *Remote Sensing of Environment* 112, 3181-3191.
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. *Classification and Regression Trees*. Belmont CA.
- Browne, M.W., 2000. Cross-validation methods. *Journal of Mathematical Psychology* 44, 108-132.
- Chavez Jr, P.S., 1996. Image-based atmospheric corrections - Revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62, 1025-1036.
- Chen, X., Vierling, L., Deering, D., 2005. A simple and effective radiometric correction method to improve landscape change detection across sensors and across time. *Remote Sensing of Environment* 98, 63-79.

- CIRAD, 2008. La canne aux multiples ressources. CIRAD.
- CNES, 2007. Images Spot: copyrigh CNES, Distribution Spot Image.
- Collins, J.B., Woodcock, C.E., 1996. An assessment of several linear change detection techniques for mapping forest mortality using multitemporal landsat TM data. *Remote Sensing of Environment* 56, 66-77.
- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37, 35-46.
- Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., 2004. Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing* 25, 1565-1596.
- Crist, E.P., Cicone, R.C., 1984. Application of the Tasseled Cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering & Remote Sensing* 50, 343-352.
- DeBoissezon, H., Sand, A., 2006. Reference Remote Sensing Data Bases: Temporal series of calibrated and ortho-rectified satellite images for scientific use. *Proceedings of Recent Advances in Quantitative Remote Sensing, Valencia, Spain*.
- Defries, R.S., Belward, A.S., 2000. Global and regional land cover characterization from satellite data: An introduction to the Special Issue. *International Journal of Remote Sensing* 21, 1083-1092.
- Delecalle, R., Maas, S.J., Guerif, M., Baret, F., 1992. Remote sensing and crop production models: present trends. *ISPRS Journal of Photogrammetry and Remote Sensing* 47, 145-161.
- Dempster, A.P., 1968. A generalisation of bayesian inference. *Journal of the Royal Statistical Society B* 30, 205-247.
- Desachy, J., Roux, L., Zahzah, E.H., 1996. Numeric and symbolic data fusion: A soft computing approach to remote sensing images analysis. *Pattern Recognition Letters* 17, 1361-1378.
- Du, Y., Teillet, P.M., Cihlar, J., 2002. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment* 82, 123-134.
- Dubois, D., Prade, H., 1979. Fuzzy real algebra: Some results. *Fuzzy Sets and Systems* 2, 327-348.
- Dubois, D., Prade, H., 1985. *Théorie des possibilités: applications á la representation des connaissances en informatique*. Masson, Paris.
- El Hajj, M., Bégué, A., Guillaume, S., 2007a. Multi-source information fusion: Monitoring sugarcane harvest using multi-temporal images, crop growth modelling, and expert knowledge. *MultiTemp-2007, Fourth International Workshop on the Analysis of Multi-Temporal Remote Sensing Images. IEEE 2007, Leuven, Belgium, p. 6.*

- El Hajj, M., Bégué, A., Guillaume, S., Martiné, J.F., 2008a. Combining multi-source information for crop monitoring. The 11th International Conference on Information Fusion, Cologne, Germany, p. 7.
- El Hajj, M., Bégué, A., Guillaume, S., Martiné, J.F., 2008b. Integrating SPOT-5 time series, crop growth modeling and expert knowledge to monitor agricultural practices - The case of sugarcane harvest in Reunion Island. *Remote Sensing of Environment* (Accepted).
- El Hajj, M., Bégué, A., Lafrance, B., Hagolle, O., Dedieu, G., Rumeau, M., 2008c. Relative radiometric normalization and atmospheric correction of a SPOT 5 time series. *Sensors* 8, 2774-2791.
- El Hajj, M., Guillaume, S., Bégué, A., 2008d. Fusion d'informations multi-sources pour le suivi des coupes de canne à sucre à La Réunion. LFA, Rencontres francophones sur la Logique Floue et ses Applications, Lens, p. 8.
- El Hajj, M., Rumeau, M., Bégué, A., Hagolle, O., G., D., 2007b. Radiometric normalization of high spatial resolution multi-temporal imagery: A comparison between a relative method and atmospheric correction. SPIE Europe Remote Sensing - The International Society for Optical Engineering, Florence, Italy, p. 10.
- Elvidge, C.D., Ding, Y., Weerackoon, R.D., Lunetta, R.S., 1995. Relative radiometric normalization of Landsat multispectral scanner (MSS) data using an automatic scattergram-controlled regression. *Photogrammetric Engineering & Remote Sensing* 61, 1255-1260.
- Fauconnier, B., Bassereau, D., 1970. La canne sucre. Paris, G-P Maisonneuve et Larose (Ed.).
- Fauvel, M., Chanussot, J., Benediktsson, J.A., 2005. Fusion of methods for the classification of remote sensing images from Urban areas. *International Geoscience and Remote Sensing Symposium (IGARSS)*, Seoul, pp. 2819-2822.
- Fortes, C., Dematte, J.A.M., 2006. Discrimination of sugarcane varieties using Landsat 7 ETM+ spectral data. *International Journal of Remote Sensing* 27, 1395-1412.
- Foucher, S., Boucher, J.M., Bénie, G.B., 2003. Multiscale Classification and Filtering of SAR Images Using Dempster-Shafer Theory. *International Geoscience and Remote Sensing Symposium (IGARSS)*, Toulouse, pp. 197-199.
- Fraser, R.H., Abuelgasim, A., Latifovic, R., 2005. A method for detecting large-scale forest cover change using coarse spatial resolution imagery. *Remote Sensing of Environment* 95, 414-427.
- Furby, S.L., Campbell, N.A., 2001. Calibrating images from different dates to 'like-value' digital counts. *Remote Sensing of Environment* 77, 186-196.
- Gers, C., Schmidt, E., 2001. Using SPOT4 satellite imagery to monitor area harvested by small scale sugarcane farmers at Umfolozi. 75th South African Sugar Technologists' Association (SASTA), pp. 28-33.
- Gers, C.J., 2003. Remotely sensed sugarcane phenological characteristics at Umfolozi South Africa. *International Geoscience and Remote Sensing Symposium (IGARSS)*, Toulouse, pp. 1010-1012.

Gill, S.J., Milliken, J., Beardsley, D., Warbington, R., 2000. Using a Mensuration Approach with FIA Vegetation Plot Data to Assess the Accuracy of Tree Size and Crown Closure Classes in a Vegetation Map of Northeastern California. *Remote Sensing of Environment* 73, 298-306.

Gopal, S., Woodcock, C.E., Strahler, A.H., 1999. Fuzzy Neural Network Classification of Global Land Cover from a 1° AVHRR Data Set. *Remote Sensing of Environment* 67, 230-243.

Guerif, M., Duke, C.L., 2000. Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation. *Agriculture, Ecosystems & Environment* 81, 57-69.

Guillaume, S., 2001. Designing Fuzzy Inference Systems from Data: An Interpretability-Oriented Review. *IEEE Transactions on Fuzzy Systems* 9, 426-443.

Guillaume, S., Charnomordic, B., Lablée, J.-L., 2002. FisPro: Logiciel open source pour les systèmes d'inférence floue. <http://www.inra.fr/bia/M/fispro>. INRA-Cemagref.

Hall, F.G., Strebel, D.E., Nickeson, J.E., Goetz, S.J., 1991. Radiometric rectification: toward a common radiometric response among multitemporal, multisensor images. *Remote Sensing of Environment* 35, 11-27.

Houlès, V., El Hajj, M., Bégue, A., 2006. Radiometric normalization of a spot 4 and spot 5 time series of images (ISLE-Reunion) for agriculture applications. *Revue Française de Photogrammétrie et de Télédétection*, 31-37.

Hubert-Moy, L., 2004. Occupation du sol et Télédétection : De l'inventaire à la modélisation prédictive. Mémoire d'habilitation à diriger des recherches. COSTEL UMR CNRS 6554 LETG. Université de Rennes 2 - Haute Bretagne, p. 278.

Hubert-Moy, L., Lecerf, R., Corpetti, T., Dubreuil, V., 2005. Monitoring winter vegetation cover using multitemporal modis data. *International Geoscience and Remote Sensing Symposium (IGARSS)*, Seoul, pp. 2113-2116.

Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., Hunsaker, C., 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment* 102, 63-73.

Ichihashi, H., Shirai, T., Nagasaka, K., Miyoshi, T., 1996. Neuro-fuzzy ID3: A method of inducing fuzzy decision trees with linear programming for maximizing entropy and an algebraic method for incremental learning. *Fuzzy Sets and Systems* 81, 157-167.

Inglada, J., 2001. Etat de l'art en détection de changement sur les images de télédétection. Note technique. CNES, Toulouse, p. 21.

Ingram, J.C., Dawson, T.P., 2005. Inter-annual analysis of deforestation hotspots in Madagascar from high temporal resolution satellite observations. *International Journal of Remote Sensing* 26, 1447-1461.

INRA, STICS. Simulateur multIdisciplinaire pour les Cultures Standard <http://www.avignon.inra.fr/stics/>.

Jensen, J.R., 1981. Urban change detection mapping using Landsat digital data. *American Cartographer* 8, 127-147.

Krishna Rao, P.V., Venkateswara Rao, V., Venkataratnam, L., 2002. Remote sensing : A technology for assessment of sugarcane crop acreage and yield. *Sugarcane Technology* 40, 97-101.

Laba, M., Downs, R., Smith, S., Welsh, S., Neider, C., White, S., Richmond, M., Philpot, W., Baveye, P., 2008. Mapping invasive wetland plants in the Hudson River National Estuarine Research Reserve using quickbird satellite imagery. *Remote Sensing of Environment* 112, 286-300.

Lafrance, B., 2006. Rapport technique. Traitement des produits SPOT 1A du projet ISLE REUNION. CNES, p. 103.

Largouët, C., 2000. Aide à l'interprétation d'une séquence d'images par la modélisation du système observé. Application à la reconnaissance de l'occupation du sol. Thèse de doctorat de l'Université de Rennes I, p. 188.

Launay, M., Guerif, M., 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. *Agriculture, Ecosystems & Environment* 111, 321-339.

Le Hégarat-Masclé, S., Seltz, R., 2004. Automatic change detection by evidential fusion of change indices. *Remote Sensing of Environment* 91, 390-404.

Le Hégarat-Masclé, S., Seltz, R., Hubert-Moy, L., Corgne, S., Stach, N., 2006. Performance of change detection using remotely sensed data and evidential fusion: Comparison of three cases of application. *International Journal of Remote Sensing* 27, 3515-3532.

Lebourgeois, V., Begue, A., Degenne, P., Bappel, E., 2007. Improving sugarcane harvest and planting monitoring for smallholders with geospatial technology: The Reunion Island experience. *International Sugar Journal* 109, 109-117.

Lee, T., Richards, J.A., Swain, P.H., 1986. Probabilistic and evidential approaches for multisource data analysis. *IEEE Transactions on Geoscience and Remote Sensing* GE-25, 283-293.

Li, Y., Vodacek, A., Zhu, Y., 2007. An automatic statistical segmentation algorithm for extraction of fire and smoke regions. *Remote Sensing of Environment* 108, 171-178.

Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., Benning, T.L., 2003. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. *Agriculture, Ecosystems & Environment* 94, 205-220.

Lucas, R., Rowlands, A., Brown, A., Keyworth, S., Bunting, P., 2007. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *Isprs Journal of Photogrammetry and Remote Sensing* 62, 165-185.

Lyon, J.G., Yuan, D., Lunetta, R.S., Elvidge, C.D., 1998. A change detection experiment using vegetation indices. *Photogrammetric Engineering and Remote Sensing* 64, 143-150.

- Maignan, F., Breon, F.M., Bacour, C., Demarty, J., Poirson, A., 2008. Interannual vegetation phenology estimates from global AVHRR measurements: Comparison with in situ data and applications. *Remote Sensing of Environment* 112, 496-505.
- Malila, W.A., 1980. Change vector analysis: An approach for detecting forest changes with Landsat. *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, 326-336.
- Mamdani, E.H., Assilian, S., 1975. Experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies* 7, 1-13.
- Martiné, J.F., 2003. Modélisation de la production potentielle de la canne à sucre en zone tropicale, sous conditions thermiques et hydriques contrastées. Applications du modèle. CIRAD. Thèse de doctorat de l'Institut national agronomique Paris-Grignon, Saint-Denis, La Réunion, p. 131.
- Martiné, J.F., 2004. Estimation de la surface totale de limbes verts totaux/ tige. Rapport technique. CIRAD, La Bretagne, p. 2.
- Martiné, J.F., Todoroff, P., 2002. Le modèle de croissance Mosaic et sa plateforme de simulation Simulex: état des lieux et perspectives. *Revue Agricole et Sucrière de l'île Maurice* 81, 133-147.
- Mas, J.F., 2000. A review of methods and techniques of remote detection of change. *Une revue des méthodes et des techniques de télédétection du changement* 26, 349-362.
- McCloy, K.R., Lucht, W., 2004. Comparative evaluation of seasonal patterns in long time series of satellite image data and simulations of a global vegetation model. *IEEE Transactions on Geoscience and Remote Sensing* 42, 140-153.
- Mercier, G., Hubert-Moy, L., Houet, T., Gouéry, P., 2005. Estimation and monitoring of bare soil/vegetation ratio with spot vegetation and HRVIR. *IEEE Transactions on Geoscience and Remote Sensing* 43, 348-353.
- Metternicht, G., 2001. Assessing temporal and spatial changes of salinity using fuzzy logic, remote sensing and GIS. *Foundations of an expert system. Ecological Modelling* 144, 163-179.
- Meygret, A., 2007. SPOT absolute calibration : synthesis. CNES technical report, S5-NT-0-2880-CN, Ed2, rev 0.
- Moisan, Y., Bernier, M., Dubois, J.M.M., 1999. Détection des changements dans une série d'images ERS-1 multitudes à l'aide de l'analyse en composantes principales. *International Journal of Remote Sensing* 20, 1149-1167.
- Moulin, S., Bondeau, A., Delécolle, R., 1998. Combining agricultural crop models and satellite observations: From field to regional scales. *International Journal of Remote Sensing* 19, 1021-1036.
- Muchoney, D.M., Haack, B.N., 1994. Change detection for monitoring forest defoliation. *Photogrammetric Engineering and Remote Sensing* 60, 1243-1251.

- Nielsen, A.A., 2007. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Transactions on Image Processing* 16, 463-478.
- Nielsen, A.A., Canty, M.J., 2005. Multi- and hyperspectral remote sensing change detection with generalized difference images by the IR-MAD method. *Proceedings of the Third International Workshop on the Analysis of Multi-Temporal Remote Sensing Images 2005*, Biloxi, MS, pp. 169-173.
- Nielsen, A.A., Conradsen, K., Simpson, J.J., 1998. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment* 64, 1-19.
- Price, K.P., Egbert, S.L., Nellis, M.D., Lee, R.Y., Boyce, R., 1997. Mapping land cover in a high plains agro-ecosystem using a multivariate landsat thematic mapper modeling approach. *Transactions of the Kansas Academy of Science* 100, 21-33.
- Quinlan, J.R., 1986. Induction of decision trees. *Machine Learning* 1, 81-106.
- Rahman, H., Dedieu, G., 1994. SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing* 15, 123-143.
- Ridd, M.K., Liu, J., 1998. A comparison of four algorithms for change detection in an urban environment. *Remote Sensing of Environment* 63, 95-100.
- Rudorff, B.F.T., Batista, G.T., 1990. Yield estimation of sugarcane based on agrometeorological-spectral models. *Remote Sensing of Environment* 33, 183-192.
- Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., Ohno, H., 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment* 96, 366-374.
- Salvaggio, C., 1993. Radiometric scene normalization utilizing statistically invariant features. *Proceedings of the workshop on atmospheric correction of Landsat imagery*, Torrance, California, pp. 155-159.
- Schmidt, E.J., Narciso, G., Frost, P., Gers, C., 2000. Application of remote sensing technology in the South African sugar industry : Review of recent research findings. *Proceedings of the 74th South African Sugar Technologists' Association (SASTA)*, Durban, South Africa, p. 19.
- Schott, J.R., Salvaggio, C., Volchok, W.J., 1988. Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment* 26, 1-16.
- Séguis, L., Puech, C., 1997. Méthode de détermination des invariants radiométriques adaptée au paysage semi-aride de l'Afrique de l'Ouest. *International Journal of Remote Sensing* 18, 255-271.
- Shafer, G., 1976. *A mathematical theory of evidence*. Princeton University Press. Princeton, NJ.

- Shettle, E.P., Fenn, R.W., 1979. Models for the Aerosols of the Lower Atmosphere and the Effects of Humidity Variation on their Optical Properties. Air Force Geophysics Laboratory.
- Shimoda, H., Hong, S.-p., Fukue, K., Sakata, T., 1991. A multi-temporal classification method using Dempster-Shafer model. In: Anon (Ed.), Digest - International Geoscience and Remote Sensing Symposium (IGARSS). Publ by IEEE, Espoo, Finl, pp. 1831-1834.
- Simoes, M.D.S., Rocha, J.V., Lamparelli, R.A.C., 2005. Spectral variables, growth analysis and yield of sugarcane. *Scientia Agricola* 62, 199-207.
- Simone, G., Farina, A., Morabito, F.C., Serpico, S.B., Bruzzone, L., 2002. Image fusion techniques for remote sensing applications. *Information Fusion* 3, 3-15.
- Singh, A., 1989. Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing* 10, 989-1003.
- Solaiman, B., 1999. Multisensor data fusion using fuzzy concepts: application to land-cover classification using ers-l/jers-1 sar composites. *IEEE Transactions on Geoscience and Remote Sensing* 37, 1316-1326.
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment* 75, 230-244.
- Swain, P.H., 1985. Approaches to computer reasoning in remote sensing and geographic information processing: a survey. In: Carver Keith, R. (Ed.), Digest - International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, Amherst, MA, USA, pp. 687-692.
- Tanre, D., 1990. Description of a computer code to simulate the satellite signal in the solar spectrum: the 5S code. *International Journal of Remote Sensing* 11, 659-668.
- Tapia, R., Stein, A., Bijker, W., 2005. Optimization of sampling schemes for vegetation mapping using fuzzy classification. *Remote Sensing of Environment* 99, 425-433.
- Teillet, P.M., Fedosejevs, G., 1995. On the dark target approach to atmospheric correction of remotely sensed data. *Canadian Journal of Remote Sensing* 21, 374-387.
- Ueno, M., Kawamitsu, Y., Sun, L., Taira, E., Maeda, K., 2005. Combined applications of NIR, RS and GIS for sustainable sugarcane production. Proceedings of the International Society of Sugar Cane Technologists (ISSCT), Guatemala, pp. 204-210.
- Van Cleynebreugel, J., Osinga, S.A., Fierens, F., Suetens, P., Oosterlinck, A., 1991. Road extraction from multi-temporal satellite images by an evidential reasoning approach. *Pattern Recognition Letters* 12, 371-380.
- Vermote, E.F., Tanre, D., Deuze, J.L., Herman, M., Morcrette, J.J., 1997. Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *IEEE Transactions on Geoscience and Remote Sensing* 35, 675-686.
- Wardlow, B.D., Egbert, S.L., 2003. A State-Level Comparative Analysis of the GAP and NLCD Land-Cover Data Sets. *Photogrammetric Engineering and Remote Sensing* 69, 1387-1397.

Wardlow, B.D., Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment* 112, 1096-1116.

Waser, L.T., Baltsavias, E., Ecker, K., Eisenbeiss, H., Feldmeyer-Christe, E., Ginzler, C., Küchler, M., Zhang, L., 2008. Assessing changes of forest area and shrub encroachment in a mire ecosystem using digital surface models and CIR aerial images. *Remote Sensing of Environment* 112, 1956-1968.

Westra, T., De Wulf, R.R., 2007. Monitoring Sahelian floodplains using Fourier analysis of MODIS time-series data and artificial neural networks. *International Journal of Remote Sensing* 28, 1595-1610.

Woodcock, C.E., Gopal, S., 2000. Fuzzy set theory and thematic maps: Accuracy assessment and area estimation. *International Journal of Geographical Information Science* 14, 153-172.

Woodcock, C.E., Macomber, S.A., Pax-Lenney, M., Cohen, W.B., 2001. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment* 78, 194-203.

Xavier, A.C., Rudorff, B.F.T., Shimabukuro, Y.E., Berka, L.M.S., Moreira, M.A., 2006. Multi-temporal analysis of MODIS data to classify sugarcane crop. *International Journal of Remote Sensing* 27, 755-768.

Xiao, X., Boles, S., Froking, S., Li, C., Babu, J.Y., Salas, W., Moore Iii, B., 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. *Remote Sensing of Environment* 100, 95-113.

Yan, H., Cao, M., Liu, J., Tao, B., 2007. Potential and sustainability for carbon sequestration with improved soil management in agricultural soils of China. *Agriculture, Ecosystems & Environment* 121, 325-335.

Young, S.S., Wang, C.Y., 2001. Land-cover change analysis of China using global-scale pathfinder AVHRR landcover (PAL) data, 1982-92. *International Journal of Remote Sensing* 22, 1457-1477.

Yuan, D., Elvidge, C., 1998. NALC land cover change detection pilot study: Washington D.C. Area experiments. *Remote Sensing of Environment* 66, 166-178.

Zadeh, L.A., 1965. Fuzzy sets. *Information and Control* 8, 338-353.

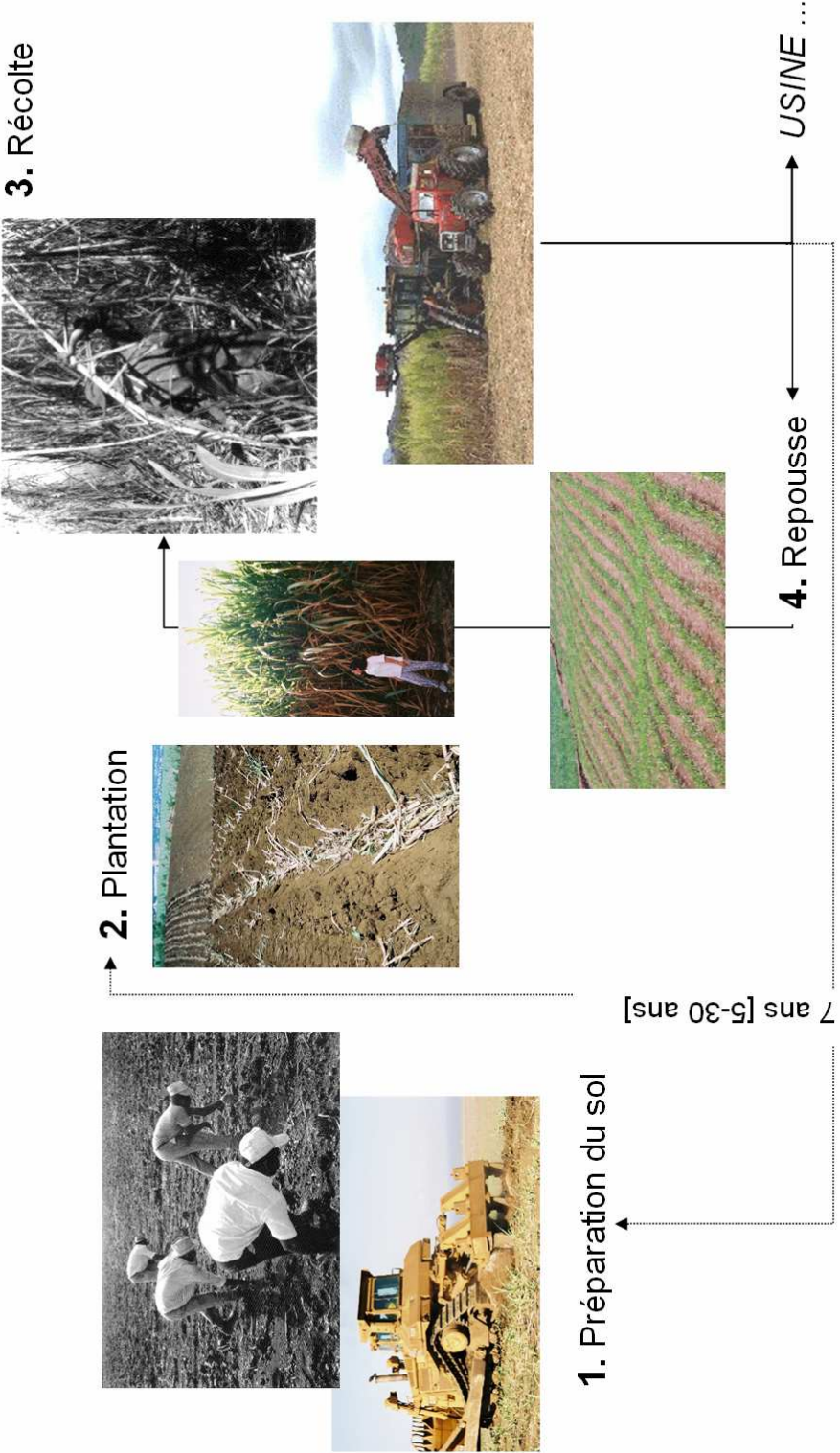
Zadeh, L.A., 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1, 3-28.

Annexes des chapitres

Annexe 1 : Cycles de la canne à sucre à La Réunion	179
Annexe 2 : Généralités sur la canne à sucre	181
Annexe 3 : Quick-looks des images SPOT	183
Annexe 4 : Création d'un masque des nuages	187
Annexe 5 : Exemples de profils temporels de NDVI à l'échelle de la parcelle.....	189
Annexe 6 : Logique floue	193
Annexe 7 : Partitions des indicateurs flous	195
Annexe 8 : Règles expertes de décision.....	197
Annexe 9 : Résultats du système à base de règles expertes.....	205
Annexe 10 : Cartographie dynamique des coupes sur les exploitations de M. Barau et M. Caruel en 2004	211
Annexe 11 : Matrices de confusion calculées à partir des résultats du système à base de règles expertes	215
Annexe 12 : Numéros et degrés d'activation des règles expertes impliquées dans la détection des coupes sur l'exploitation de M. Caruel en utilisant la série de 2003	217
Annexe 13 : Résultats du système à base de règles générées par apprentissage.....	219

Annexe 1 : Cycles de la canne à sucre à La Réunion

Dans cette annexe nous fournissons une schématisation du cycle total et du cycle de récolte de la canne à sucre à La Réunion (Source : Thèse d'Eric Bappel, 2005).



Annexe 2 : Généralités sur la canne à sucre

La canne à sucre, *Saccharum officinarum*, est une graminée tropicale originaire d'Asie. Chaque année, environ 20,1 millions d'hectares sont cultivés en canne à sucre dans plus de 100 pays tropicaux (FAOSTATS, 2007). La surface cultivée en canne augmente régulièrement, d'une part à cause de l'augmentation de la consommation de sucre dans les pays en développement, et d'autre part, en raison de l'essor des bio-énergies. Cultivée par des millions d'agriculteurs indépendants et par de grands complexes agro-industriels à travers le monde, elle produit les trois quarts du sucre mondial.

Les conditions de culture

La canne se développe dans des températures chaudes s'étalant en principe entre 10° et 30°C (les températures ne doivent en aucun cas descendre en dessous de 0°C). Lors de la phase de maturation (*i.e.* période au cours de laquelle la plante fabrique le plus de saccharose) les températures peuvent se situer entre 10°C et 20°C.

Les besoins en eau de la canne à sucre se situent en moyenne aux alentours de 1 500 mm par an, toutefois, la culture de la canne à sucre peut également avoir lieu dans des régions moins propices si un bon système d'irrigation est mis en place. Le sol doit être riche, lourd et assez poreux pour permettre une bonne pénétration et conservation de l'eau, mais il doit également être bien drainé, afin que le système racinaire de la plante, qui est assez important, se développe dans de bonnes conditions.

La croissance de la plante se fait de manière graduelle, assez lente au départ, elle s'accélère doucement jusqu'au début de la phase de maturation où le rythme de croissance ralentit de nouveau. La floraison de la canne dépend fortement des conditions environnementales, notamment de la quantité d'eau fournie, de la teneur en azote du sol et de la durée quotidienne d'ensoleillement. Cette étape ayant tendance à diminuer le rendement en sucre de la plante, les planteurs la préviennent souvent, soit en plantant des variétés hybrides qui ne fleurissent pas, soit, en pratiquant l'*écimage* de la plante.

Si ces diverses conditions sont réunies, les plants de canne à sucre arriveront à maturité au bout de 10 et 24 mois selon les régions (la moyenne se situant aux alentours de 16 mois). Ce moment est assez simple à déterminer, il intervient un mois après l'apparition de petites fleurs le long de la tige de la canne à sucre. Elle est alors coupée au plus près du sol en laissant

toutefois au moins un nœud apparent ; à savoir que c'est la partie inférieure de la tige qui contient le plus de saccharose. Cette opération peut se dérouler de manière manuelle quand c'est la tradition dans le pays ou quand le relief ne permet pas le recours aux machines. Une fois coupées, les tiges sont étêtées, débarrassées de leurs feuilles puis entassées sous forme d'andains. Lorsque la récolte se fait de manière mécanique, toutes les opérations de ramassage sont réalisées en un seul passage. Il arrive parfois que les champs soient brûlés avant d'être récoltés afin de les débarrasser des serpents et des rats, mais également pour éviter un important travail de main d'œuvre post-récolte (e.g. effeuillage et brûlage des débris végétaux).

La fragilité des plants de canne à sucre après ramassage explique la rapidité avec laquelle ils doivent être acheminés vers les sucreries et traités. En effet, plus le temps de stockage s'allonge, plus la teneur en sucre des plants diminue.

Les rendements

Plusieurs critères tels que le climat, la qualité des sols, la satisfaction des besoins en eau (soit naturellement, soit par le biais de l'irrigation) ainsi que la variété de plante cultivée ont une influence sur le rendement.

Le rendement mondial de la canne à sucre a augmenté à un rythme annuel de 0,6% entre 1961 et 2005 passant ainsi d'un rendement de 50 tonnes par hectare en 1961 à 65 tonnes en 2005 (FAOSTAT).

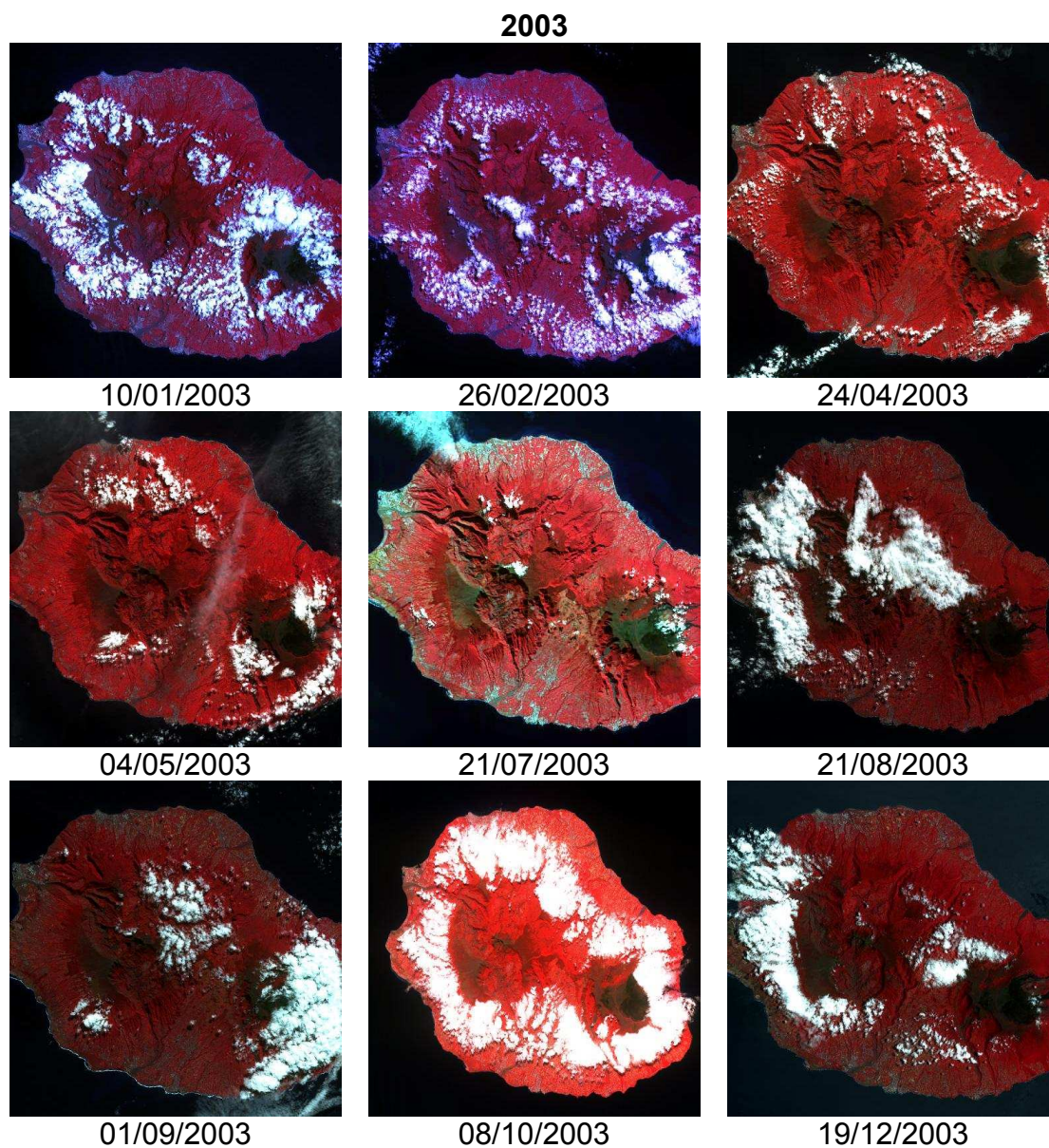
Coût de production du sucre de canne

Au niveau des coûts de production, la canne à sucre est plus compétitive que sa concurrente la betterave : son coût de revient est inférieur de 40%. Ce sont principalement les coûts de culture qui conduisent à cette différence car les coûts de production de la canne sont environ deux fois inférieurs à ceux de la betterave. Plusieurs facteurs jouent un rôle déterminant dans ce processus dont les plus importants sont le coût de l'énergie (la canne produit tout ou partie de l'énergie nécessaire à sa transformation) et le niveau du sucre tiré de la plante plus élevé pour la canne (Source : CNUCED²⁷).

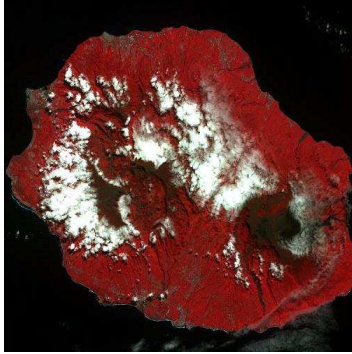
²⁷ Conférence des Nations Unies sur le Commerce Et le Développement

Annexe 3 : Quick-looks des images SPOT

Nous illustrons dans cette annexe les *quick-looks* des images SPOT utilisées dans cette thèse. Ces images sont issues de la base de données Kalideos Isle-Réunion[©].



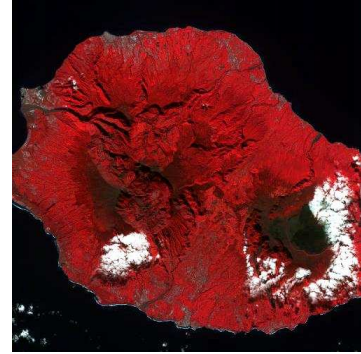
2004



17/03/2004



11/04/2004



13/05/2004



18/06/2004



09/07/2004



19/08/2004



26/10/2004

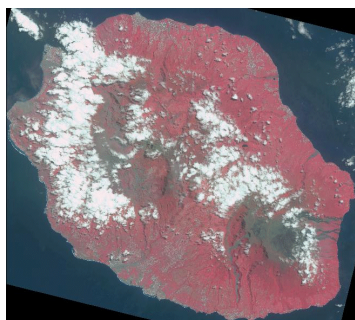


06/11/2004



07/12/2004

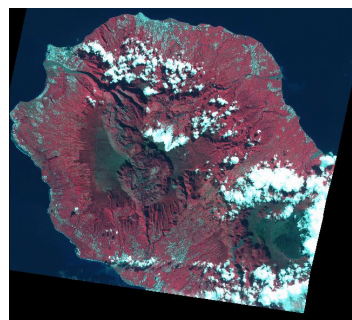
2006



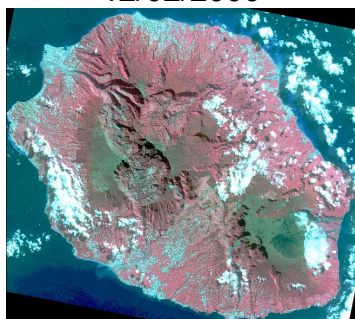
12/02/2006



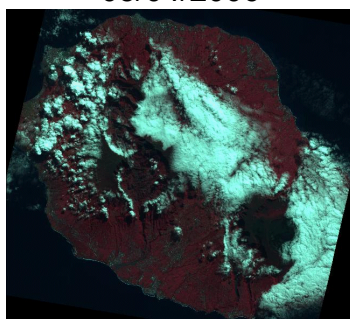
05/04/2006



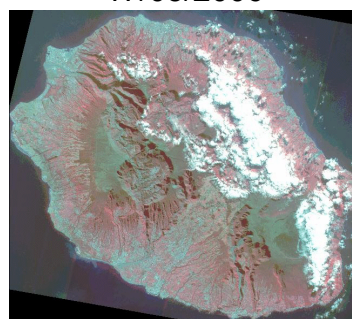
17/05/2006



05/06/2006



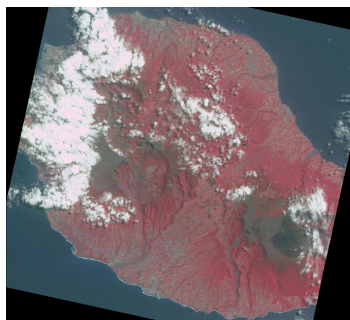
19/07/2006



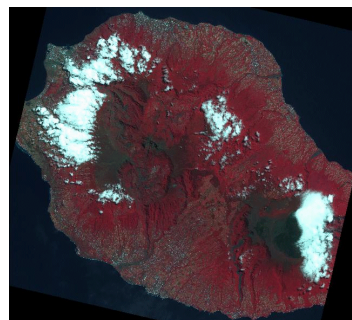
04/08/2006



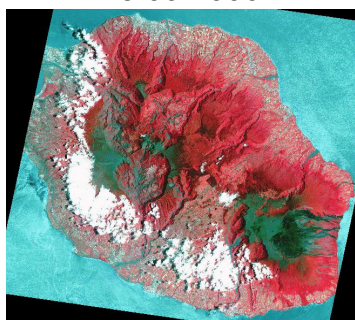
25/09/2006



28/09/2006

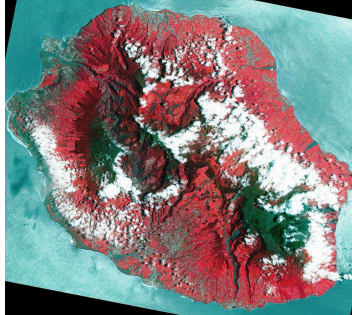


14/10/2006

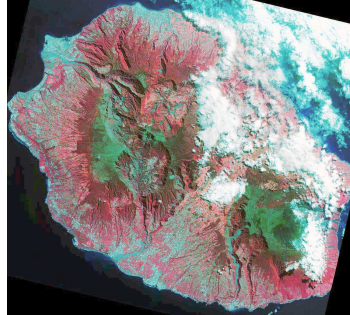


11/12/2006

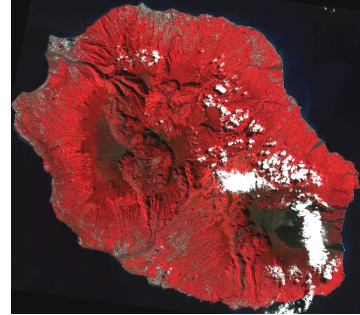
2007



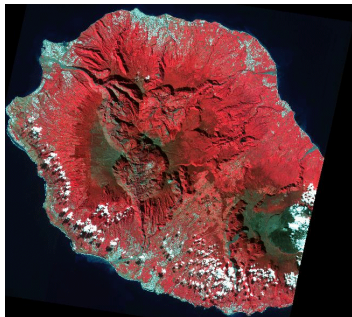
25/01/2007



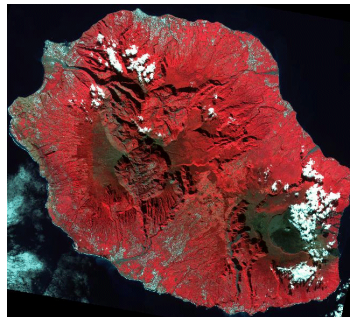
14/03/2007



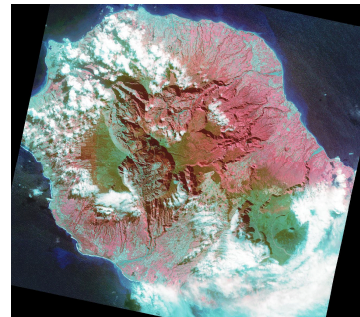
30/04/2007



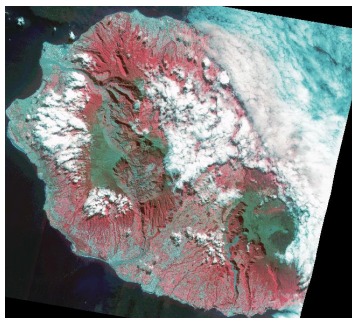
11/05/2007



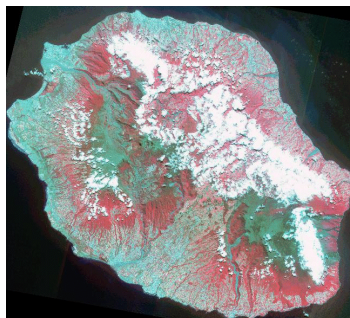
12/06/2007



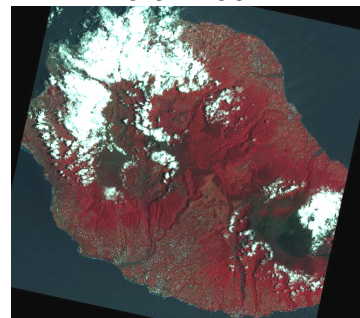
16/07/2007



03/09/2007



12/10/2007



14/12/2007

Annexe 4 : Création d'un masque des nuages

La détection de nuages sur une image satellite, I_t , constitue une problématique de segmentation complexe pour laquelle le choix de la méthode dépend du cadre et de l'objectif de la détection. Le but ici n'est pas d'obtenir un découpage précis du contour des nuages mais plutôt de supprimer rapidement les nuages et leur périphérie proche qui peut contenir des brumes et des ombres.

L'image I_t seule ne permet pas de différencier les nuages de certains éléments importants pour la sélection d'invariants comme les bâtis qui présentent les mêmes gammes de réflectances. Des algorithmes itératifs utilisant le voisinage de chaque pixel sont capables d'une telle différenciation mais s'avèrent très lourd à exécuter pour des images satellites dans ce contexte. La solution adoptée consiste à :

- faire une différence bi-temporelle entre l'image I_t et une (ou plusieurs) image(s) de référence. Les images de référence doivent être choisies de telle sorte que pour tout pixel p de I_t , il existe une image de référence telle que le pixel p ne soit pas un nuage ;
- effectuer une classification non-supervisée en deux classes (nuage, non nuage) pour chacune des images de différences ;
- faire l'intersection des différents masques obtenus à partir des différentes images de références (cas de plusieurs images de référence) ;
- faire une dilatation du masque résultant.

Exemple : Création du masque des nuages pour l'image SPOT du 26/02/2003

Les images de référence choisies sont celles du 13/05/2004 du 21/07/2003 et du 21/08/2003. Pour chacune de ces dates de référence, une image de différence avec I_t est créée puis une classification est effectuée comme décrit au dessus.

Pour chaque date t , il existe donc trois masques de nuage :

- Le masque comprenant les nuages de la date t et les nuages de la date de référence 1 (msq1),
- Le masque comprenant les nuages de la date t et les nuages de la date de référence 2 (msq2),

- Et le masque comprenant les nuages de la date t et les nuages de la date de référence 3 (msq_3).

Le masque de nuage de la date t est l'intersection de ces trois masques : si un pixel est classé nuage sur les trois masques, il est classé comme un nuage date t (valeur de pixel=1), sinon ce n'est pas un nuage (valeur de pixel=0).

Pour supprimer les brumes et les ombres présentes à proximité des contours des nuages, ce masque est dilaté pour obtenir le masque final. La dilatation revient à donner à un pixel la valeur maximale contenue dans un voisinage délimité par une matrice 7x7 pixels (ce choix est issu d'une série de tests). Le masque final pour la date du 26/02/2002 est représenté sur la figure suivante ainsi que les étapes intermédiaires.

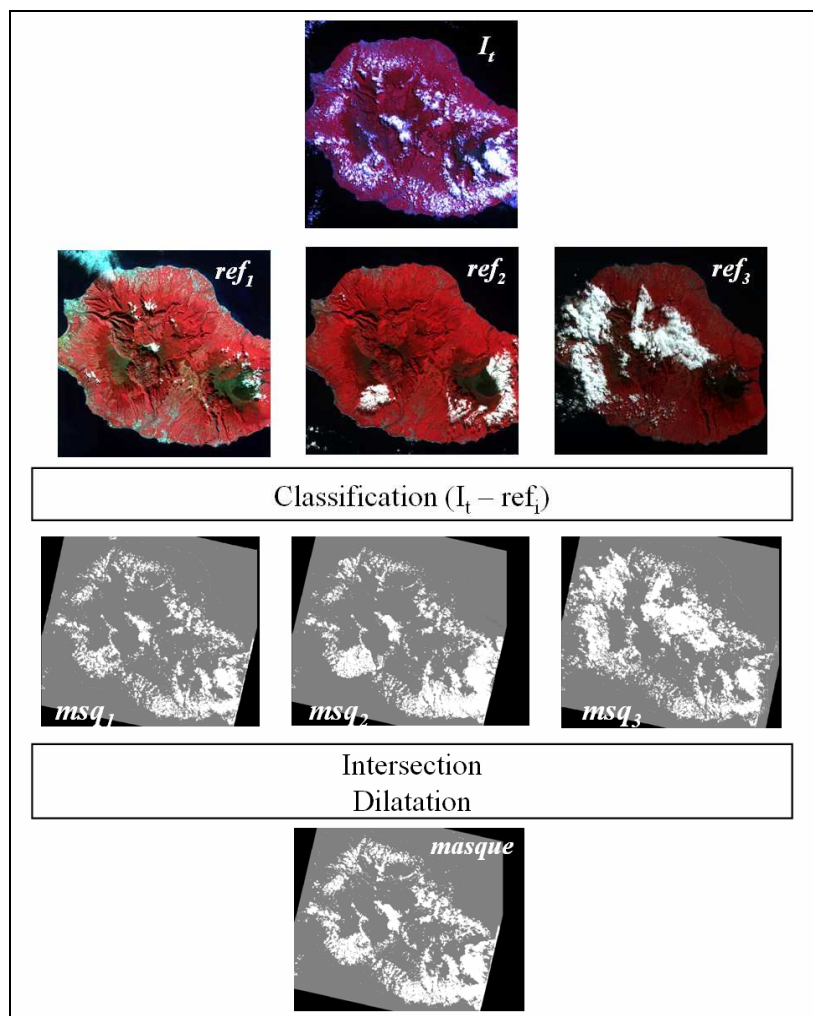


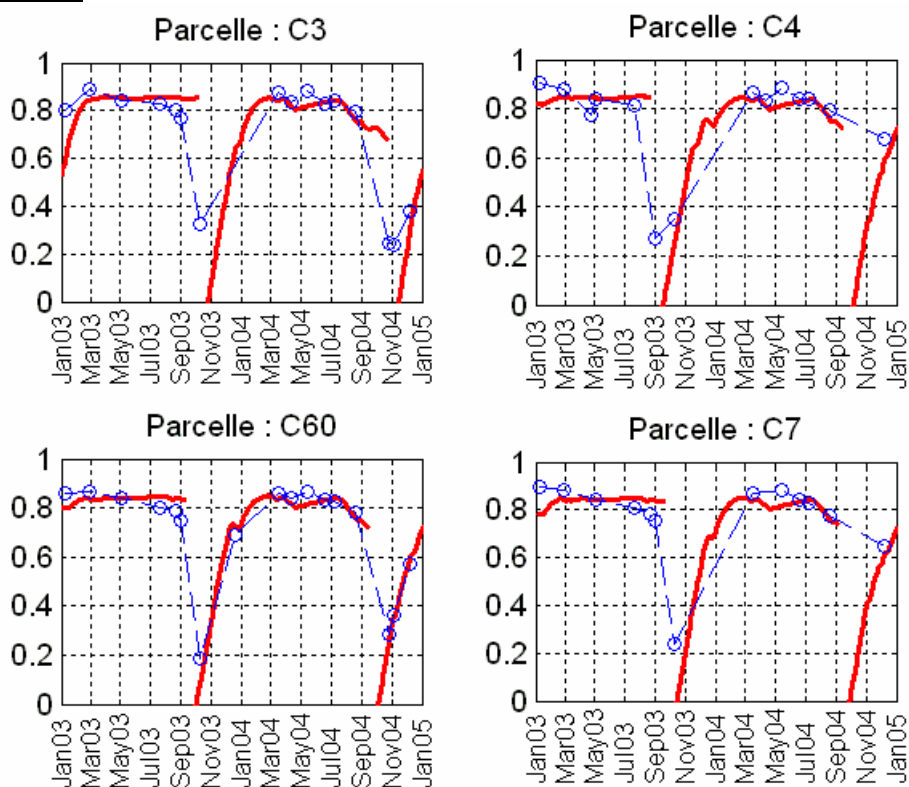
Figure A4. Extraction du masque de nuage pour l'image du 26/02/2003 (I_t) en utilisant les références 21/07/2003 (ref_1), 13/05/2004 (ref_2) et 21/08/2003 (ref_3).

Annexe 5 : Exemples de profils temporels de NDVI à l'échelle de la parcelle

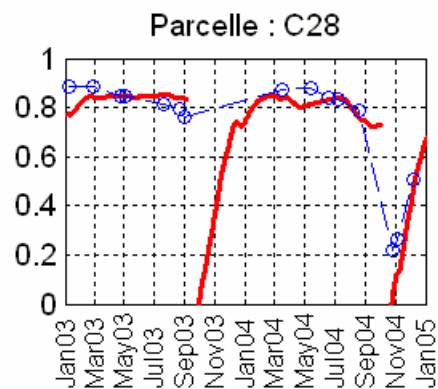
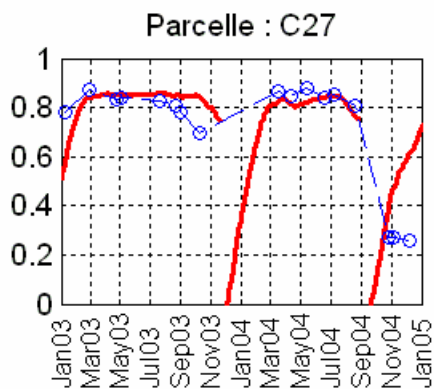
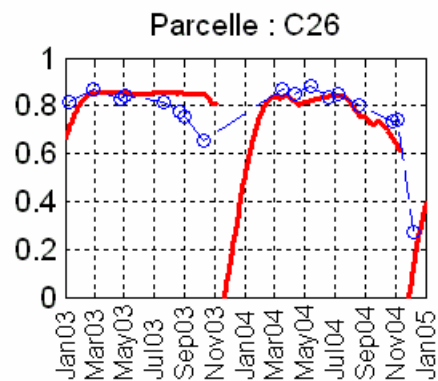
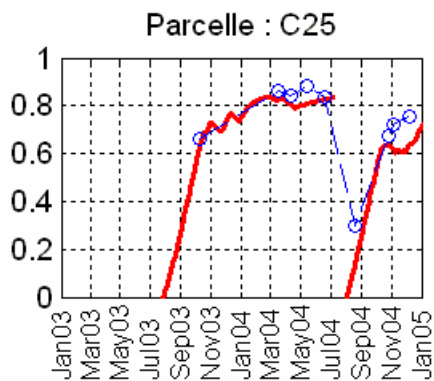
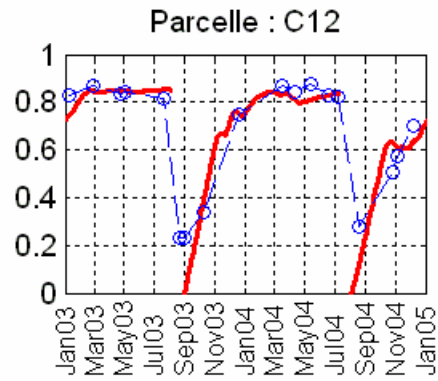
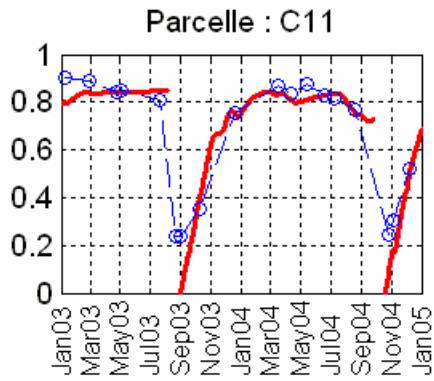
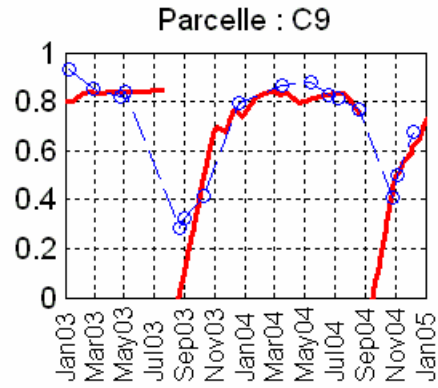
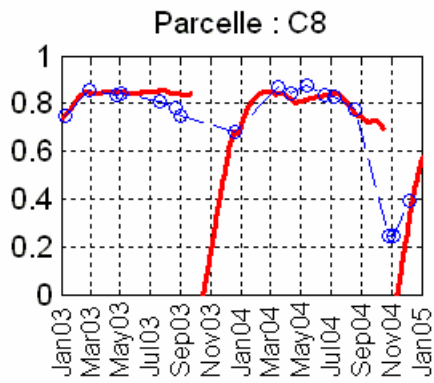
Nous illustrons dans cette annexe des profils temporels de NDVI calculés pour des parcelles de l'exploitation de M. Caruel en repousse et en plantation. Ces profils sont calculés :

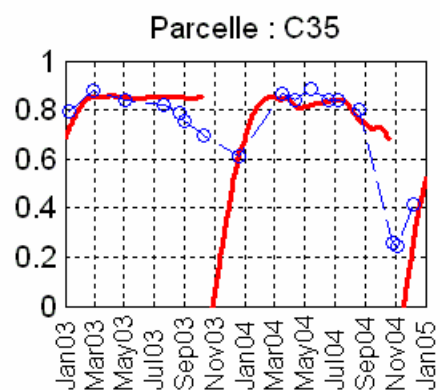
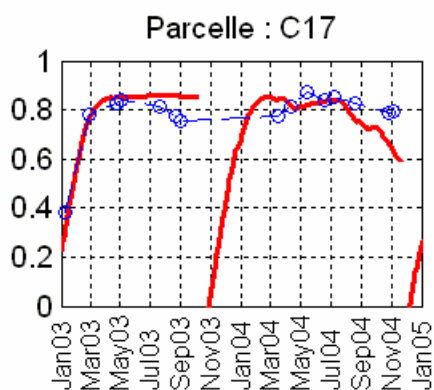
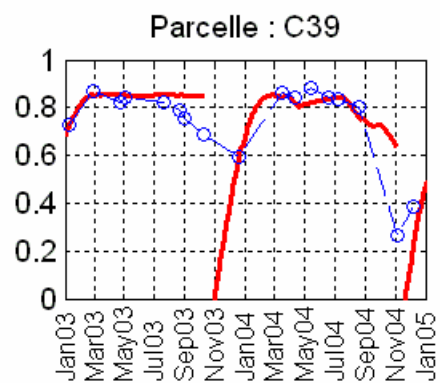
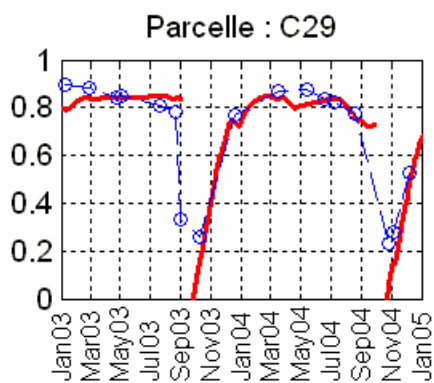
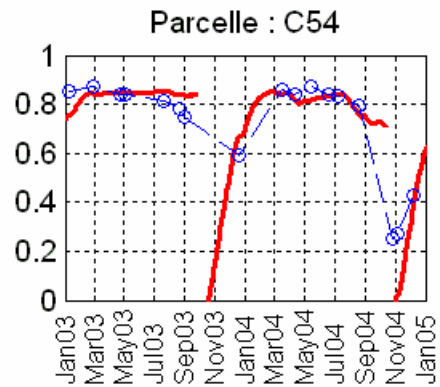
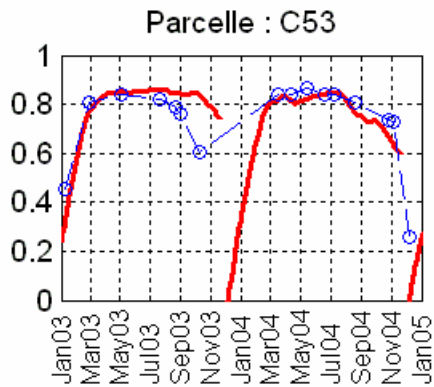
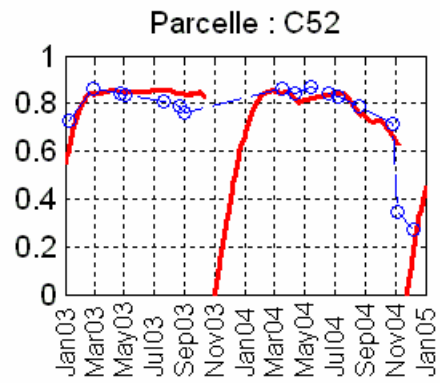
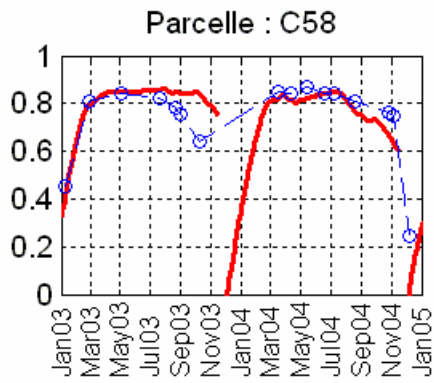
- à partir des images SPOT (en bleu) ;
- à partir des simulations faites par le modèle de croissance de canne à sucre MOSICAS (en rouge).

Repousses

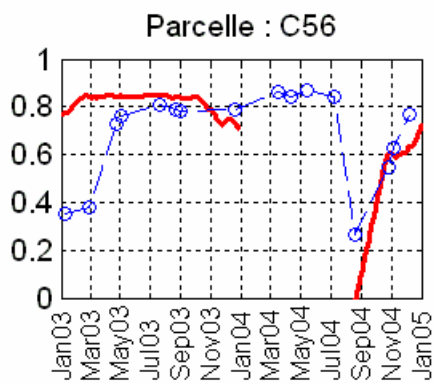
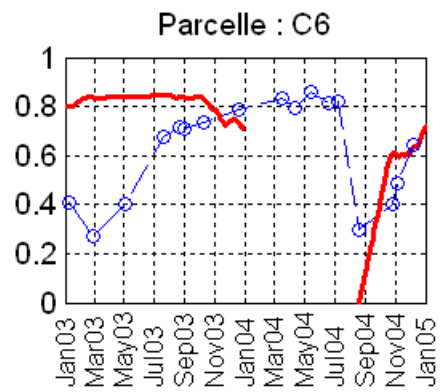
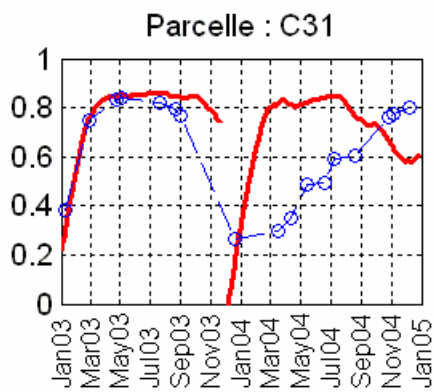


Annexe 5 : Exemples de profils temporels de NDVI à l'échelle de la parcelle





Plantations



Annexe 6 : Logique floue

Historique

La logique floue ou « Fuzzy logic » est née à l'université de Berkeley Californie aux États-Unis en 1963. Son père créateur est Lotfi Zadeh. Zadeh a fait une comparaison entre les ordinateurs de l'époque et le raisonnement humain : si l'ordinateur calcule beaucoup plus vite d'une façon rigoureuse, ses capacités de réflexion et d'apprentissage sont limitées. De plus sa rigidité en tant que machine et son fonctionnement binaire, le rend peu adapté à certaines tâches, qui pour un humain semble si simple. Zadeh a pris l'exemple du créneau réalisé par un automobiliste pour garer sa voiture ; si le conducteur réalise cette opération en une minute environ, la simulation de la même action sur un ordinateur demandait à l'époque plusieurs heures de calculs complexes. Le plus intrigant dans cette histoire, c'est que l'humain fait tout ceci de façon approximative que ce soit pour la conduite dans le cas présent, comme bon nombres d'actions de la vie quotidienne. Cela marche donc avec l'approximation et une expérience des fonctions à réaliser. Naissance de la logique floue, dont le concept sera complètement ignoré aux États-Unis, comme en Europe. Il faut attendre 25 ans pour voir apparaître les premières applications au Japon ; à noter que l'exemple de la voiture à garer de Zadeh a été repris, et mis en application avec une voiture truffée de capteurs, équipée d'un calculateur flou et qui parvient à se garer toute seule.

Bases

L'imprécision et l'approximation sont les bases de la la logique floue. Cette théorie remet en cause la notion d'appartenance d'un objet à un ensemble. Dans la théorie classique des ensembles, la notion d'appartenance est liée à une valeur binaire 0 ou 1. Pour les ensembles flous la notion d'appartenance, μ , est liée à une fonction pouvant prendre toutes les valeurs entre 0 et 1 indiquant le degré ou le taux d'appartenance. Entre le "tout ou rien", la logique floue prend en compte ainsi une infinité de degrés approximatifs comme : un peu, beaucoup, fort, faible, jeune, vieux, etc. On ne passe pas brutalement de « jeune » à « vieux » le jour de notre anniversaire. Ces concepts sont graduels et imprécis.

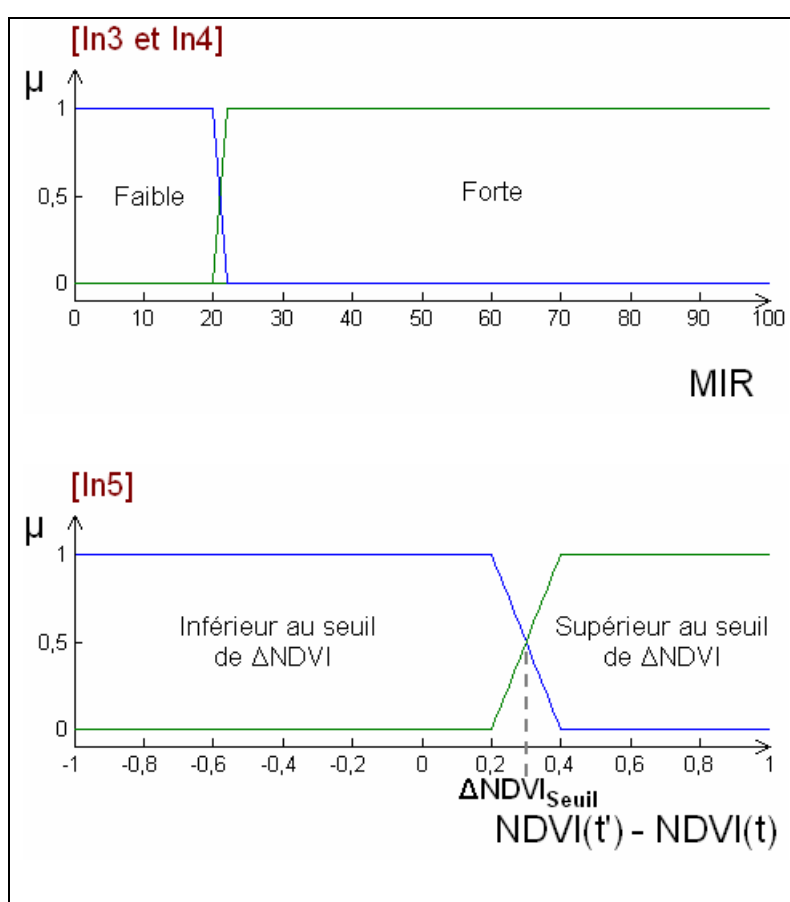
Avantages

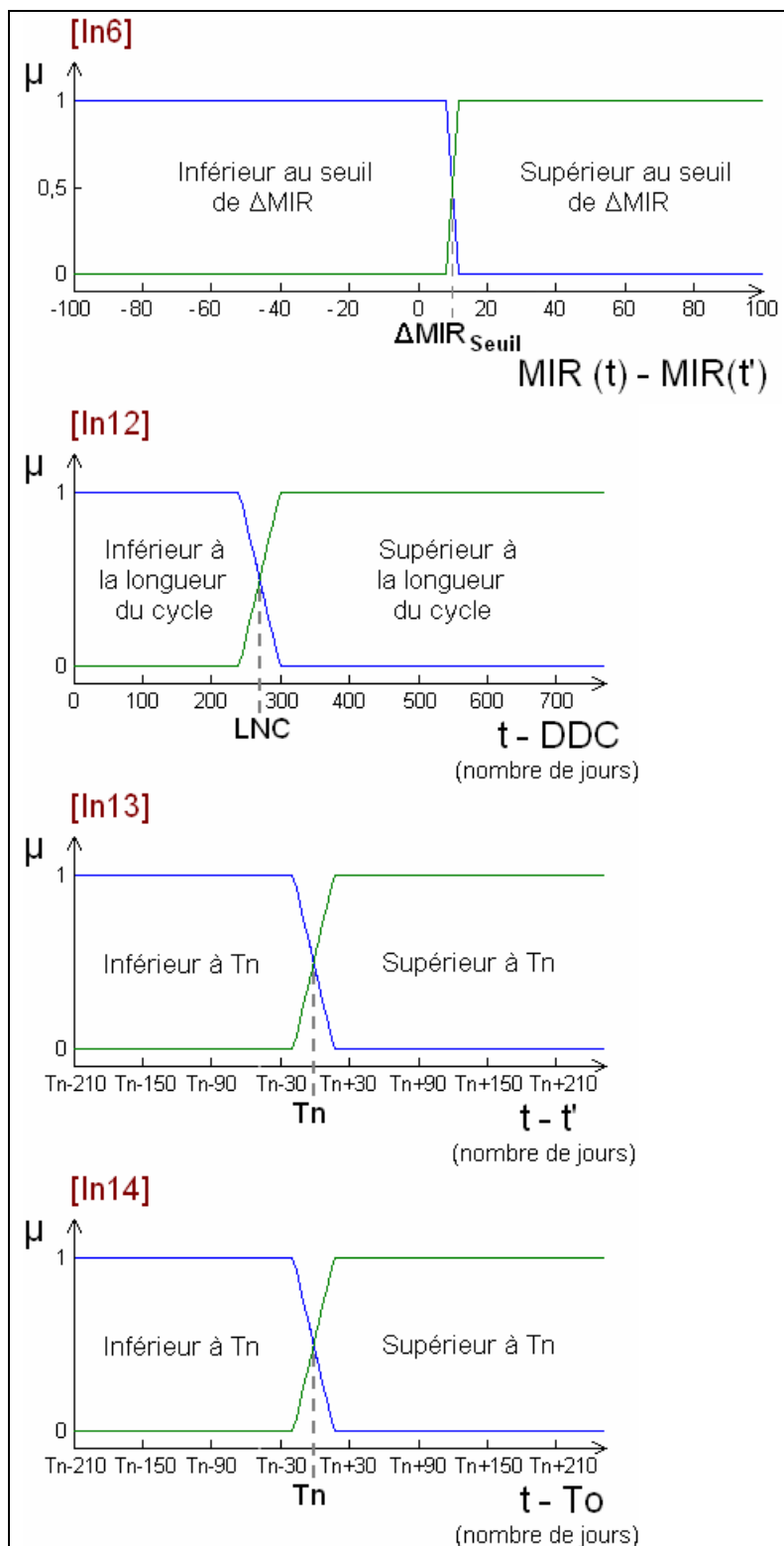
Fondamentalement, les applications utilisant la logique floue ne sont pas plus performantes mais tout simplement plus facile à réaliser et à utiliser. Des systèmes complexes peuvent devenir plus conviviaux ; on peut traiter des informations ayant des limites floues de l'entrée à la sortie en passant par le traitement, le calcul et la mémoire. La logique étant purement mathématique, les résultats ne sont pas flous, mais tout à fait clairs et précis. Par comparaison, dans un automatisme logique standard, un ensemble d'équations permet de déterminer une valeur de sortie à partir des données d'où la difficulté d'exprimer l'expérience humaine par les mathématiques. L'expression linguistique utilisée dans la formulation de la logique floue utilise un langage de tous les jours et donc l'ensemble du système flou reste compréhensible pour les opérateurs et les personnes non expertes. Ces derniers peuvent aisément interpréter les effets et les conséquences des règles de production.

Annexe 7 : Partitions des indicateurs flous

Nous représentons dans cette annexe les partitions floues des indicateurs In3, In4, In5, In12, In13 et In14 du système d'aide à la décision développé pour la détection automatique des coupes de canne à sucre.

(t et t' : dates d'acquisition ; $\Delta\text{NDVI}_{\text{Seuil}}$: seuil de la différence bi-temporelle de NDVI ; $\Delta\text{MIR}_{\text{Seuil}}$: seuil de la différence bi-temporelle de MIR ; DDC : dernière date de coupe ; LNC : longueur nominale du cycle ; T_n : indicateur issu du modèle de canne à sucre ; T_0 : date de début de campagne)





Annexe 8 : Règles expertes de décision

r1 : Si In10 est "Entre campagnes" et In11 est "Entre campagnes" Alors Conclusion est "Non Coupe"

r2 : Si In1 est "Forte" et In2 est "Forte" et In9 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r3 : Si In1 est "Forte" et In2 est "Forte" et In9 est "Pour la majorité de t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Non Coupe"

r4 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Non Coupe"

r5 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r6 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r7 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r8 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r9 : Si In1 est "Forte" et In2 est "Moyenne" et In7 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r10 : Si In1 est "Forte" et In2 est "Moyenne" et In7 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r11 : Si In1 est "Forte" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r12 : Si In1 est "Moyenne" et In2 est "Forte" et In9 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r13 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r14 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

Annexe 8 : Règles expertes de décision

r15 : Si In1 est "Moyenne" et In2 est "Moyenne" et In7 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r16 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r17 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r18 : Si In1 est "Moyenne" et In2 est "Faible" In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r19 : Si In1 est "Faible" et In2 est "Forte" et In9 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r20 : Si In1 est "Faible" et In2 est "Forte" et In9 est "Pour la majorité de t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r21 : Si In1 est "Faible" et In2 est "Forte" et In9 est "Pas de t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r22 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r23 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Supérieur au seuil de Δ NDVI" et In7 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r24 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Supérieur au seuil de Δ NDVI" et In8 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r25 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Supérieur au seuil de Δ NDVI" et In7 est "Pas de t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r26 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Inférieur au seuil de Δ NDVI" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In15 est "Non" Alors Conclusion est "Plutôt Non Coupe"

r27 : Si In1 est "Faible" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Non Coupe"

r28 : Si In1 est "Faible" et In2 est "Faible" In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r29 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In14 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r30 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r31 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In14 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r32 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r33 : Si In1 est "Forte" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In14 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r34 : Si In1 est "Forte" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r35 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In14 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r36 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r37 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In14 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r38 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r39 : Si In1 est "Moyenne" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In14 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r40 : Si In1 est "Moyenne" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In14 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r41 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r42 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r43 : Si In1 est "Faible" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

Annexe 8 : Règles expertes de décision

r44 : Si In1 est "Faible" et In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r45 : Si In1 est "Faible" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r46 : Si In1 est "Faible" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r47 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r48 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r49 : Si In1 est "Forte" et In2 est "Moyenne" et In7 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r50 : Si In1 est "Forte" et In2 est "Moyenne" In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r51 : Si In1 est "Forte" et In2 est "Moyenne" et In8 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r52 : Si In1 est "Forte" et In2 est "Faible" In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r53 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r54 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r55 : Si In1 est "Moyenne" et In2 est "Moyenne" et In7 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r56 : Si In1 est "Moyenne" et In2 est "Moyenne" In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r57 : Si In1 est "Moyenne" et In2 est "Moyenne" et In8 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r58 : Si In1 est "Moyenne" et In2 est "Faible" et In7 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est

"Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r59 : Si In1 est "Moyenne" et In2 est "Faible" In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r60 : Si In1 est "Moyenne" et In2 est "Faible" et In8 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r61 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r62 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r63 : Si In1 est "Faible" et In2 est "Moyenne" et In7 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r64 : Si In1 est "Faible" et In2 est "Moyenne" In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r65 : Si In1 est "Faible" et In2 est "Moyenne" et In8 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r66 : Si In1 est "Faible" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r67 : Si In1 est "Faible" et In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r68 : Si In1 est "Forte" et In2 est "Forte" et In9 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r69 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r70 : Si In1 est "Forte" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r71 : Si In1 est "Forte" et In2 est "Moyenne" et In9 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r72 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

Annexe 8 : Règles expertes de décision

r73 : Si In1 est "Forte" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r74 : Si In1 est "Forte" et In2 est "Faible" In10 est "Entre campagnes" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Non Coupe"

r75 : Si In1 est "Moyenne" et In2 est "Forte" et In9 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r76 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r77 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r78 : Si In1 est "Moyenne" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r79 : Si In1 est "Moyenne" et In2 est "Moyenne" et In9 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Coupe"

r80 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In13 est "Supérieur à Tn" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r81 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In13 est "Inférieur à Tn" et In15 est "Non" Alors Conclusion est "Non Coupe"

r82 : Si In1 est "Moyenne" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Non Coupe"

r83 : Si In1 est "Moyenne" et In2 est "Faible" In10 est "Entre campagnes" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Non Coupe"

r84 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r85 : Si In1 est "Faible" et In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r86 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Supérieur au seuil de ANDVI" et In7 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Coupe"

r87 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Supérieur au seuil de ANDVI" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12

est "Supérieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Plutôt Coupe"

r88 : Si In1 est "Faible" et In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r89 : Si In1 est "Faible" et In2 est "Moyenne" et In5 est "Inférieur au seuil de $\Delta NDVI$ " et In10 est "Entre campagnes" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Je ne sais pas"

r90 : Si In1 est "Faible" et In2 est "Faible" In10 est "Entre campagnes" et In11 est "campagne précédente" et In15 est "Non" Alors Conclusion est "Non Coupe"

r91 : Si In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r92 : Si In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r93 : Si In2 est "Moyenne" et In7 est "Pour tout t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r94 : Si In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r95 : Si In2 est "Moyenne" et In8 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r96 : Si In2 est "Moyenne" et In8 est "Pas de t" et In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r97 : Si In2 est "Faible" In10 est "Campagne actuelle" et In11 est "Campagne actuelle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r98 : Si In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r99 : Si In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r100 : Si In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r101 : Si In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r102 : Si In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

Annexe 8 : Règles expertes de décision

r103 : Si In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "Entre campagnes" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r104 : Si In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r105 : Si In2 est "Forte" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r106 : Si In2 est "Moyenne" et In7 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r107 : Si In2 est "Moyenne" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r108 : Si In2 est "Moyenne" et In8 est "Pour au moins un t" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r109 : Si In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r110 : Si In2 est "Faible" et In10 est "Campagne actuelle" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r111 : Si In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r112 : Si In2 est "Forte" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r113 : Si In2 est "Moyenne" et In7 est "Pour au moins un t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Coupe"

r114 : Si In2 est "Moyenne" et In7 est "Pas de t" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Supérieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Je ne sais pas"

r115 : Si In2 est "Moyenne" et In10 est "Entre campagnes" et In11 est "campagne précédente" et In12 est "Inférieur à la longueur du cycle" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

r116 : Si In2 est "Faible" In10 est "Entre campagnes" et In11 est "campagne précédente" et In15 est "Oui" Alors Conclusion est "Plutôt Non Coupe"

Annexe 9 : Résultats du système à base de règles expertes

Nous illustrons dans cette annexe les performances du système à base de règles expertes obtenues lors de son application sur nos sites d'étude en utilisant différentes séries temporelles d'images SPOT. Ces performances sont issues de la comparaison des décisions du système avec la vérité terrain.

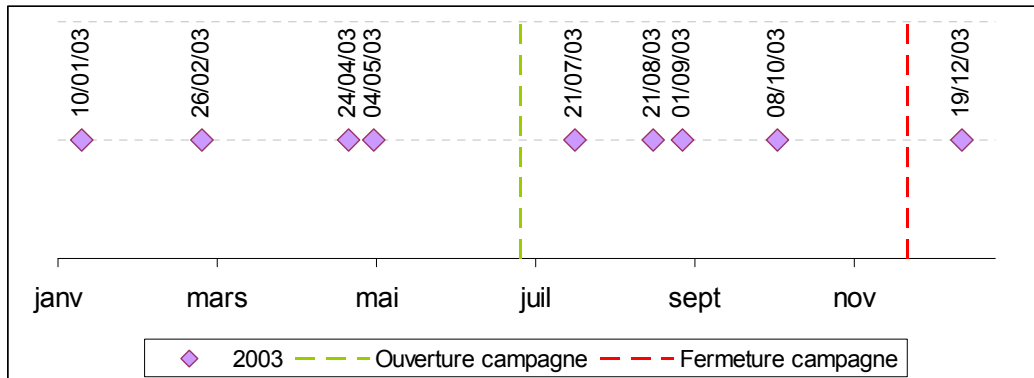
Nous montrons pour chaque résultat la disposition des images par rapport à la campagne de récolte²⁸, et les valeurs de précision globale (PG), du pourcentage de décisions « Je ne sais pas » (JSP) et des précisions du producteur pour les classes « Coupe » et « Non Coupe (PP Coupe et PP Non coupe respectivement). Ces performances sont obtenues en choisissant une approche de décision dans laquelle l'utilisateur exige un certain seuil de confiance (μ_{conf}) sur les degrés d'appartenance des possibilités « Coupe » et « Non Coupe ».

Nous rappelons que les exploitations de M. Barau de M. Caruel sont situées au Nord de La Réunion (côté « au vent ») et sont desservies par l'usine sucrière de Bois Rouge. Quant à l'exploitation de M. Isautier, elle est au Sud et elle est desservie par l'usine du Gol.

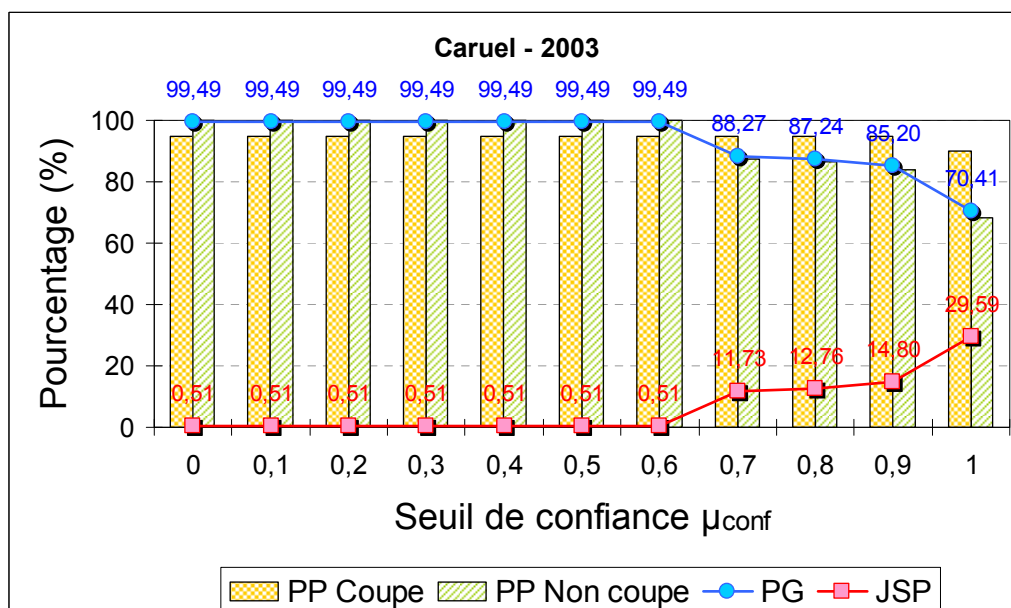
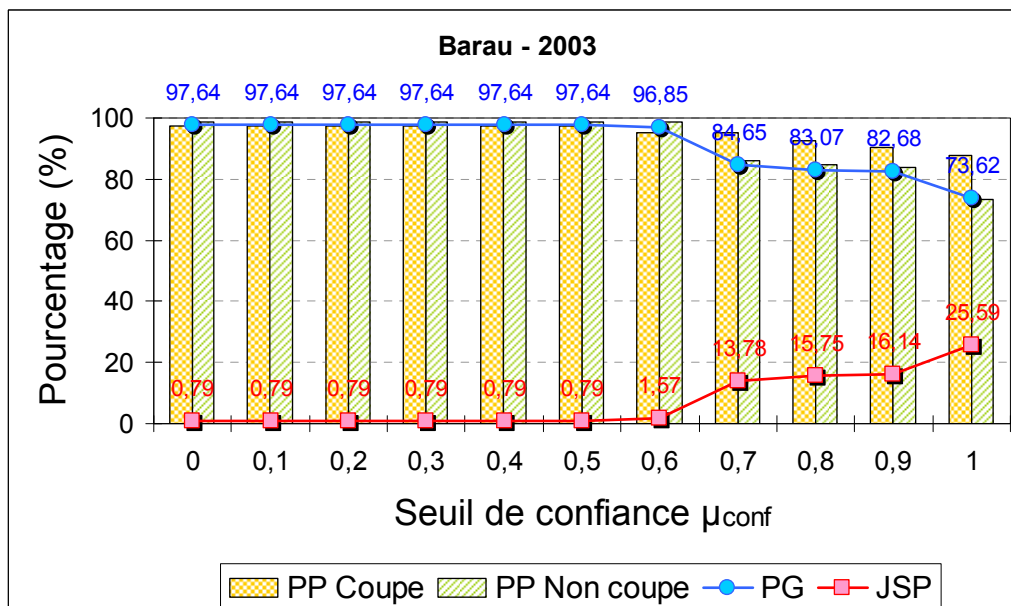
²⁸ La période de récolte est définie par les dates d'ouverture et de fermeture de l'usine sucrière.

En utilisant la série temporelle d'images en 2003

Dates d'acquisition, et dates d'ouverture et de fermeture de campagne (Bois Rouge) :

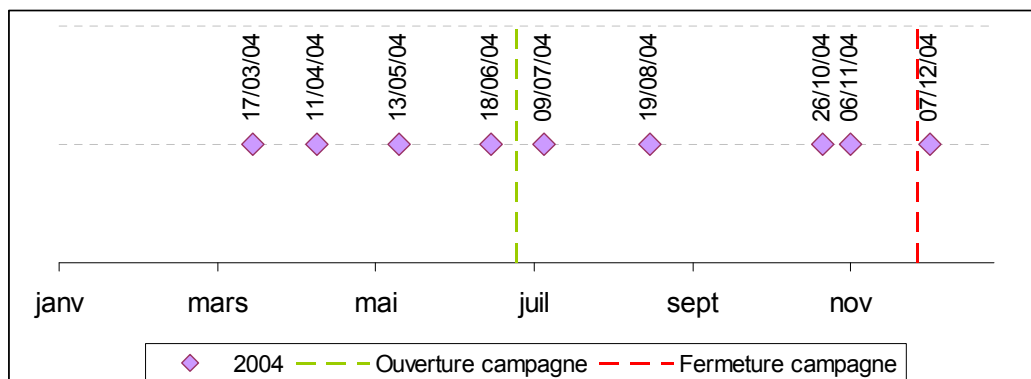


Performances du système sur les exploitations de M. Barau et M. Caruel :

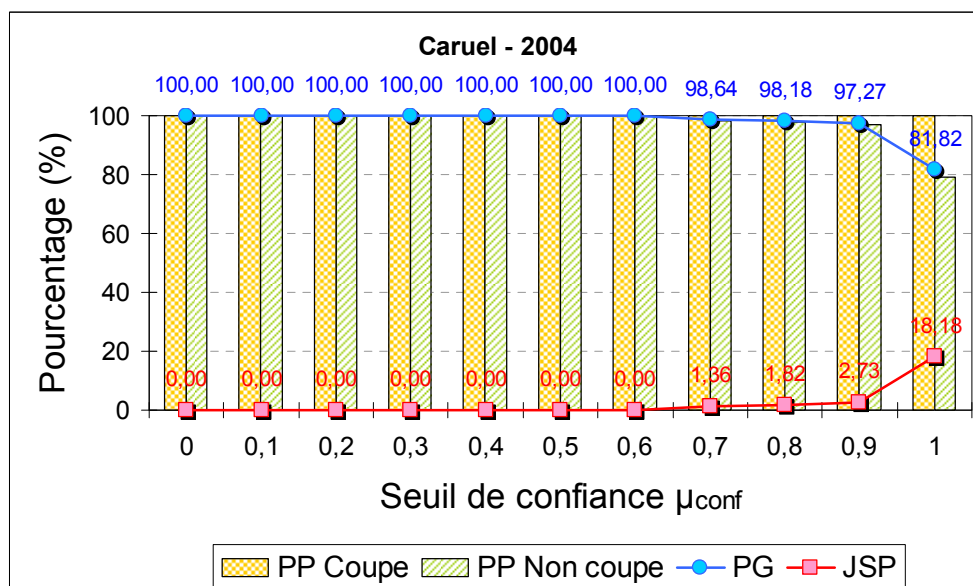
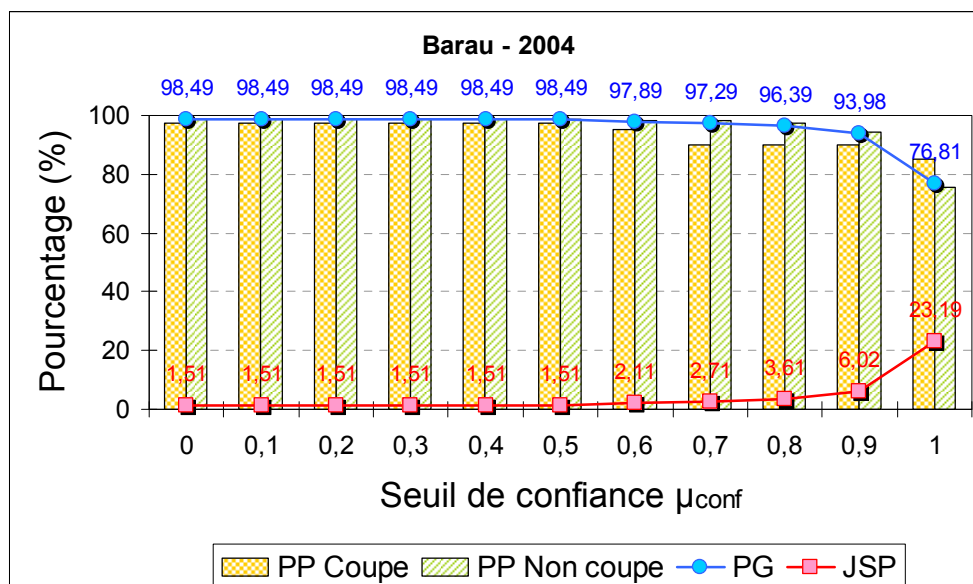


En utilisant la série temporelle d'images en 2004

Dates d'acquisition, et dates d'ouverture et de fermeture de campagne (Bois Rouge) :

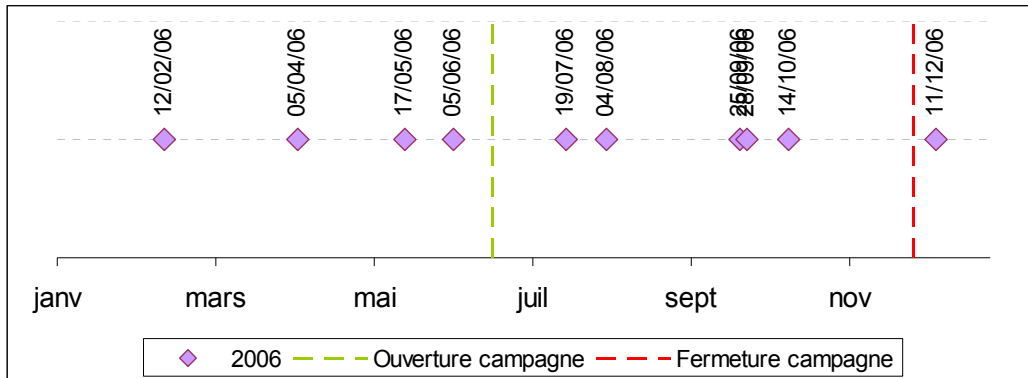


Performances du système sur les exploitations de M. Barau et M. Caruel :

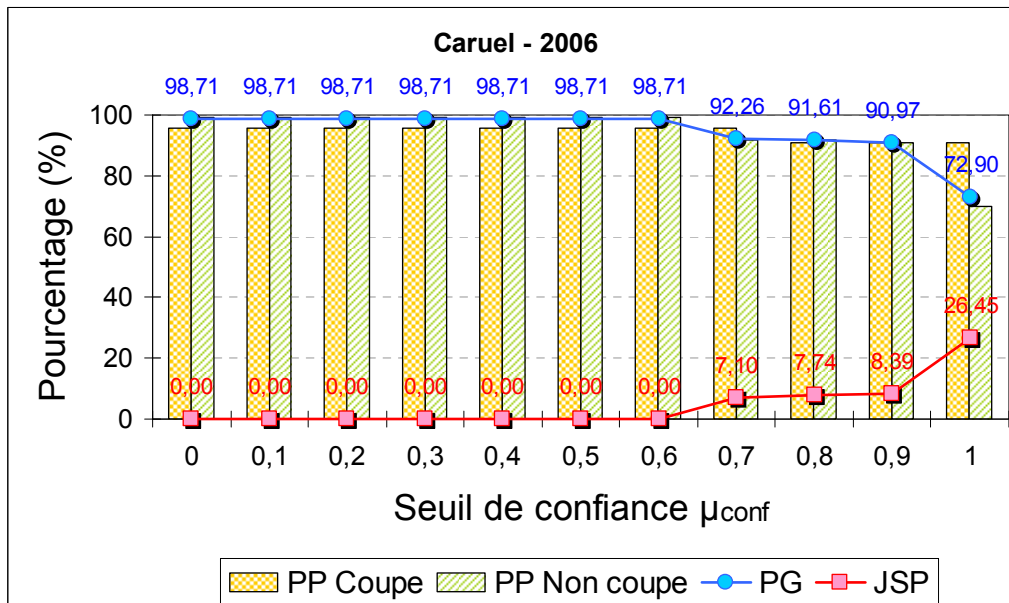


En utilisant la série temporelle d'images en 2006

Dates d'acquisition, et dates d'ouverture et de fermeture de la campagne (Bois Rouge) :

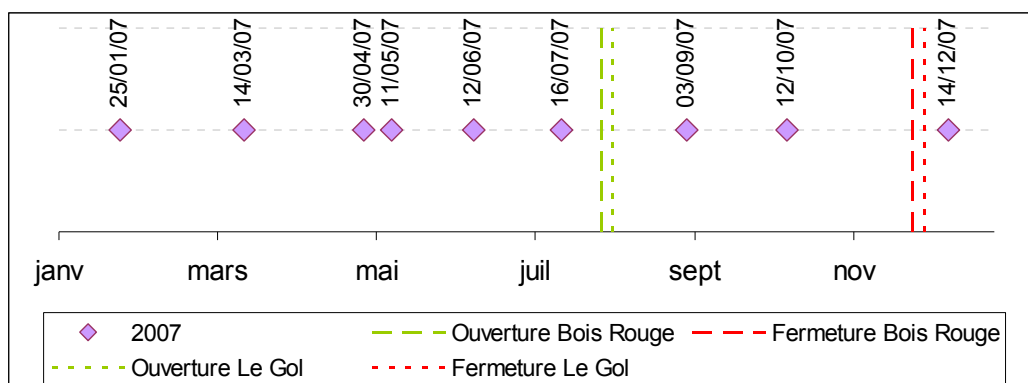


Performances du système sur l'exploitation de M. Caruel :

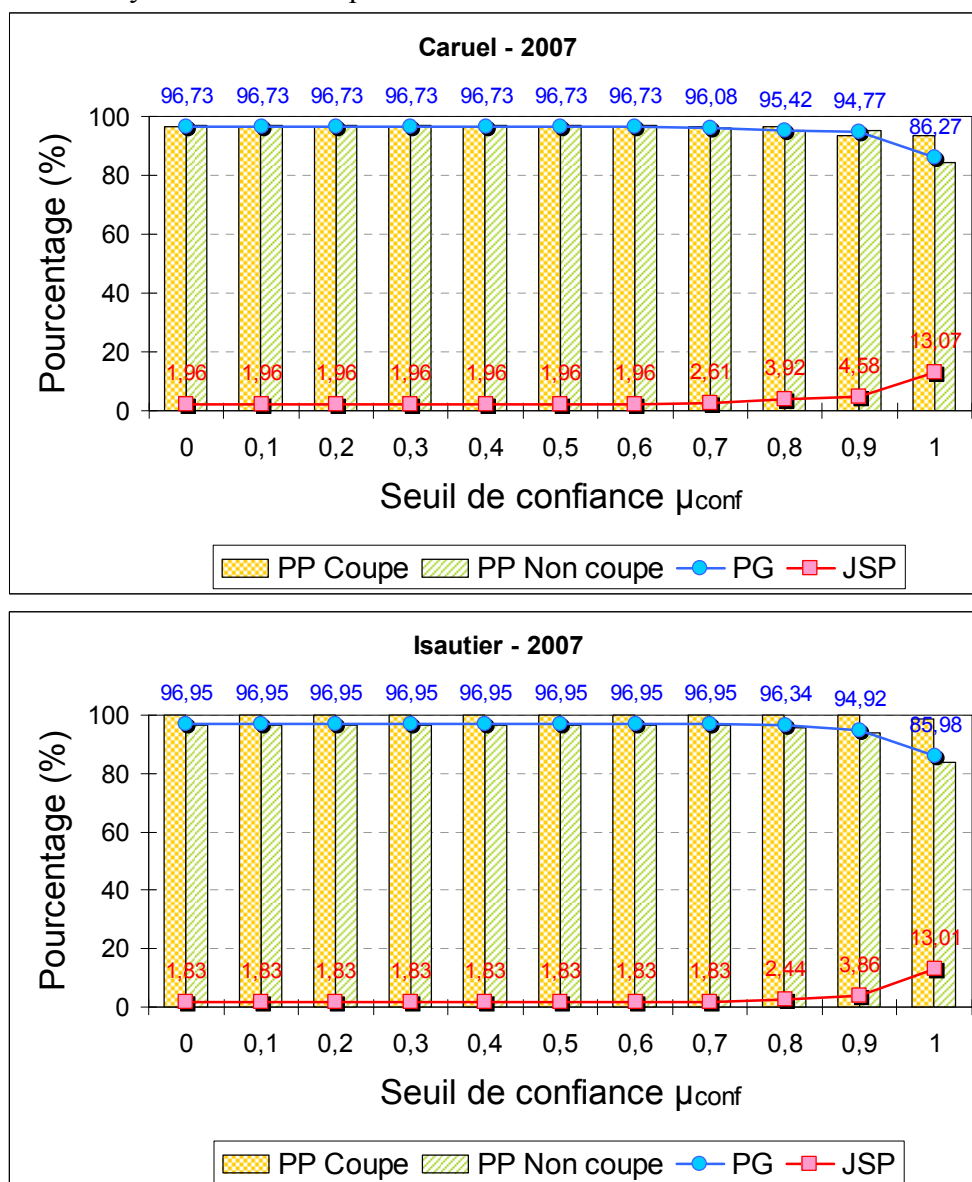


En utilisant la série temporelle d'images en 2007

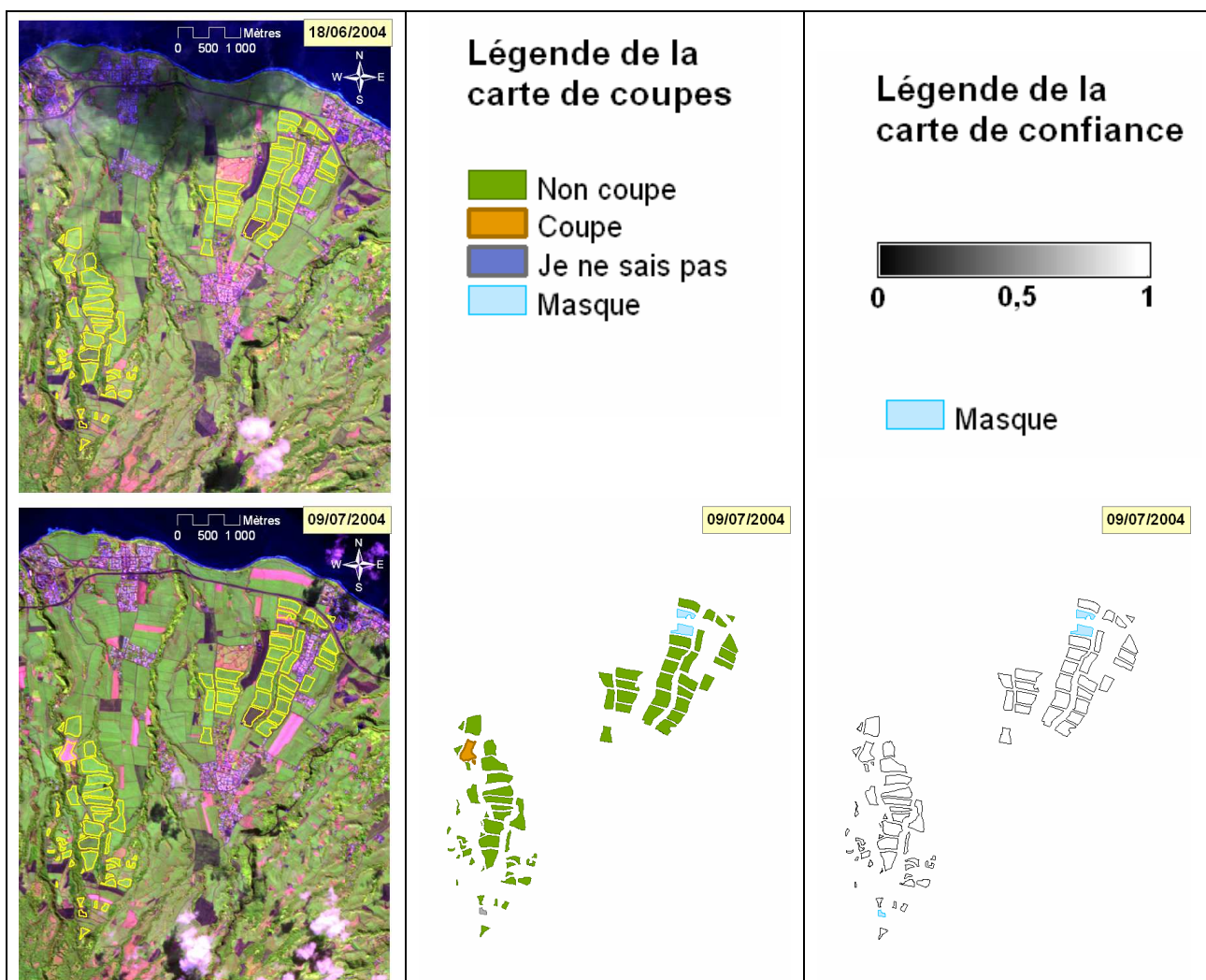
Dates d'acquisition, et dates d'ouverture et de fermeture des campagnes (Bois Rouge/Le Gol):



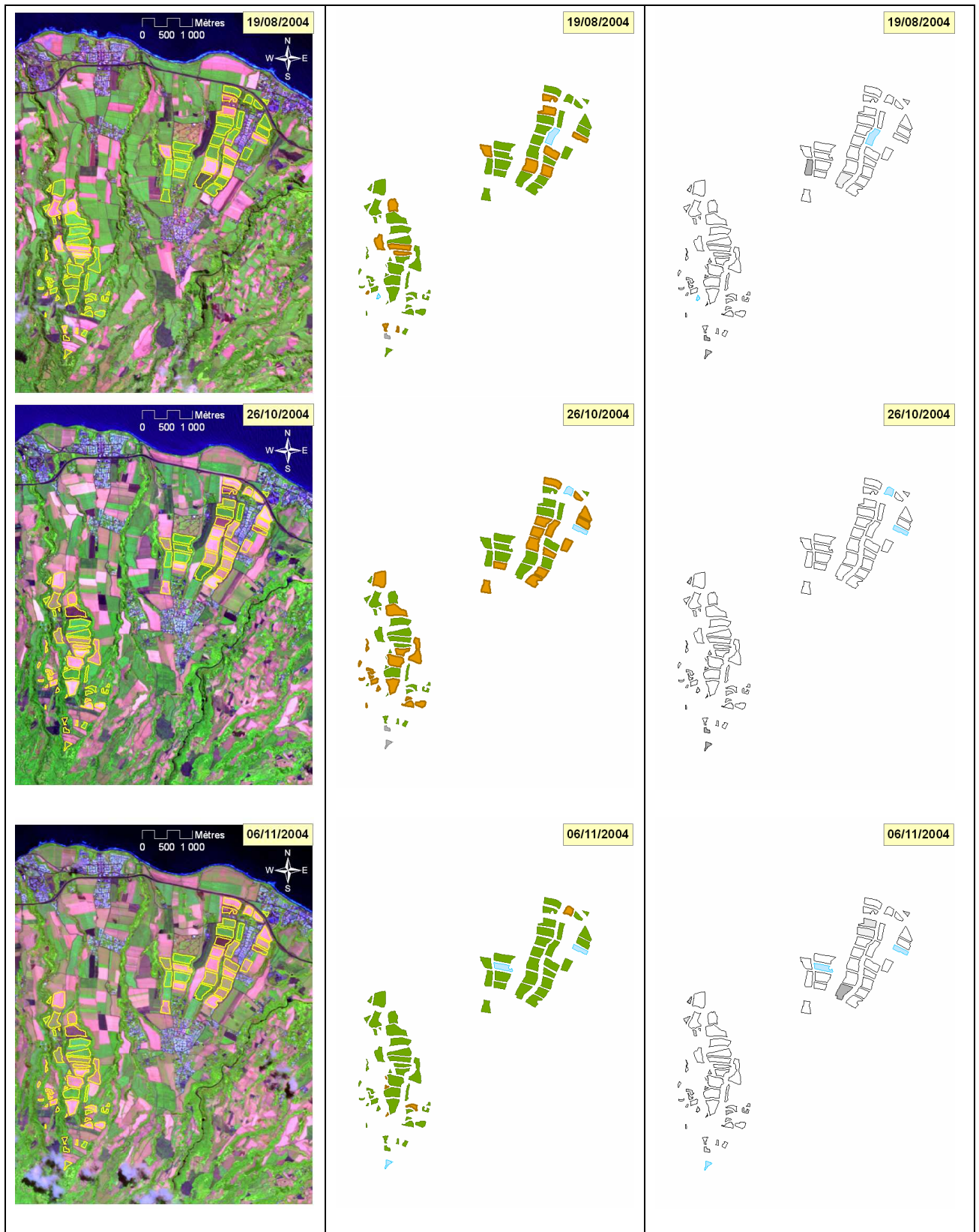
Performances du système sur les exploitations de M. Caruel et M. Isautier :

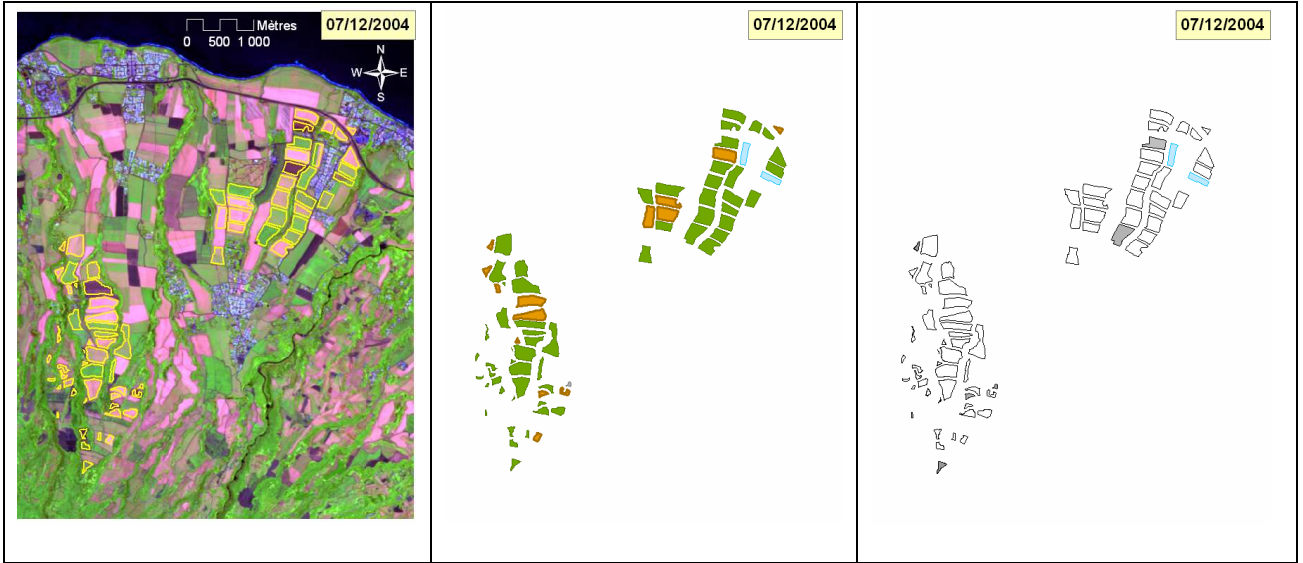


Annexe 10 : Cartographie dynamique des coupes sur les exploitations de M. Barau et M. Caruel en 2004



Annexe 10 : Cartographie dynamique des coupes sur les exploitations de M. Barau et M. Caruel en 2004





Annexe 11 : Matrices de confusion calculées à partir des résultats du système à base de règles expertes

Les matrices de confusion présentées dans cette annexe sont obtenues en comparant les décisions prises à l'aide du système à base de règles expertes avec la vérité terrain lors de la détection automatique des coupes sur les exploitations de M. Barau et de M. Caruel en 2003 et 2004. Pour chaque année, une série temporelle de 9 images a été utilisée (voir Chapitre 2). Nous montrons les matrices calculées pour les différentes valeurs de μ_{conf} .

$\mu_{conf} = 0$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	123	2	3	128	96.09%	3.91%
Non Coupe	2	866	5	873	99.20%	0.80%
Total colonne	125	868	8			
Précision de l'utilisateur	98.40%	99.77%			PG =	98.80%
Erreur de commission	1.60%	0.23%			PJSP =	0.80%

$\mu_{conf} = 0,1$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	123	2	3	128	96.09%	3.91%
Non Coupe	3	866	5	874	99.08%	0.92%
Total colonne	126	868	8			
Précision de l'utilisateur	97.62%	99.77%			PG =	98.70%
Erreur de commission	2.38%	0.23%			PJSP =	0.80%

$\mu_{conf} = 0,2$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	123	2	3	128	96.09%	3.91%
Non Coupe	3	866	5	874	99.08%	0.92%
Total colonne	126	868	8			
Précision de l'utilisateur	97.62%	99.77%			PG =	98.70%
Erreur de commission	2.38%	0.23%			PJSP =	0.80%

$\mu_{conf} = 0,3$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	123	2	3	128	96.09%	3.91%
Non Coupe	3	866	5	874	99.08%	0.92%
Total colonne	126	868	8			
Précision de l'utilisateur	97.62%	99.77%			PG =	98.70%
Erreur de commission	2.38%	0.23%			PJSP =	0.80%

$\mu_{conf} = 0,4$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	123	2	3	128	96.09%	3.91%
Non Coupe	3	866	5	874	99.08%	0.92%
Total colonne	126	868	8			
Précision de l'utilisateur	97.62%	99.77%			PG =	98.70%
Erreur de commission	2.38%	0.23%			PJSP =	0.80%

$\mu_{conf} = 0,5$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	123	2	3	128	96.09%	3.91%
Non Coupe	3	866	5	874	99.08%	0.92%
Total colonne	126	868	8			
Précision de l'utilisateur	97.62%	99.77%			PG =	98.70%
Erreur de commission	2.38%	0.23%			PJSP =	0.80%

$\mu_{conf} = 0,6$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	120	2	6	128	93.75%	6.25%
Non Coupe	3	865	6	874	98.97%	1.03%
Total colonne	123	867	12			
Précision de l'utilisateur	97.56%	99.77%			PG =	98.30%
Erreur de commission	2.44%	0.23%			PJSP =	1.20%

Annexe 11 : Matrices de confusion calculées à partir des résultats du système à base de règles expertes

$\mu_{conf} = 0,7$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	118	2	8	128	92.19%	7.81%
Non Coupe	3	809	62	874	92.56%	7.44%
Total colonne	121	811	70			
Précision de l'utilisateur	97.52%	99.75%			PG =	92.51%
Erreur de commission	2.48%	0.25%			PJSP =	6.99%

$\mu_{conf} = 0,8$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	118	1	9	128	92.19%	7.81%
Non Coupe	3	799	72	874	91.42%	8.58%
Total colonne	121	800	81			
Précision de l'utilisateur	97.52%	99.88%			PG =	91.52%
Erreur de commission	2.48%	0.12%			PJSP =	8.08%

$\mu_{conf} = 0,9$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	118	1	9	128	92.19%	7.81%
Non Coupe	3	784	87	874	89.70%	10.30%
Total colonne	121	785	96			
Précision de l'utilisateur	97.52%	99.87%			PG =	90.02%
Erreur de commission	2.48%	0.13%			PJSP =	9.58%

$\mu_{conf} = 1$	Coupe	Non Coupe	Je ne sais pas	Total ligne	Précision du producteur	Erreur d'omission
Coupe	114	0	14	128	89.06%	10.94%
Non Coupe	3	645	226	874	73.80%	26.20%
Total colonne	117	645	240			
Précision de l'utilisateur	97.44%	100.00%			PG =	75.75%
Erreur de commission	2.56%	0.00%			PJSP =	23.95%

Annexe 12 : Numéros et degrés d'activation des règles expertes impliquées dans la détection des coupes sur l'exploitation de M. Caruel en utilisant la série de 2003

	17/03/2004	11/04/2004	13/05/2004	18/06/2004	09/07/2004	19/08/2004	26/10/2004	06/11/2004	07/12/2004
C39	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	-	(19;1.000)	(28;0.600)
C3	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.660)
C46	-	(1;1.000)	(1;1.000)	(1;1.000)	-	(4;1.000)	(26;0.092)(27;0.092)(28;0.908)	(28;1.000)	(28;1.000)
C8	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.632)
C10	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(19;1.000)	(28;0.216)	(6;0.220)(7;0.220)(11;0.216)(17;0.780)(18;0.216)	(4;0.220)(6;0.780)(7;0.780)
C35	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.520)
C26	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(4;0.895)	(2;0.895)(4;0.895)(6;0.105)(10;0.105)	(19;0.935)(24;0.065)(26;0.065)
C28	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.168)
C12	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(19;1.000)	(28;0.144)	(6;0.135)(7;0.135)(11;0.135)(17;0.856)(18;0.144)	(4;0.135)(6;0.750)(7;0.750)(14;0.135)(17;0.250)
C50	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(4;1.000)	(4;1.000)	(4;1.000)
C56	-	(1;1.000)	(1;1.000)	-	(29;1.000)	(2;1.000)	(28;0.016)	(6;0.370)(11;0.016)(18;0.016)	(6;0.630)
C52	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(4;0.880)	-	(19;0.880)(23;0.120)(26;0.120)
C53	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(4;1.000)	(4;0.940)	(19;0.940)(23;0.060)(26;0.053)
C58	-	(1;1.000)	(1;1.000)	-	(29;1.000)	(4;1.000)	(4;1.000)	(4;1.000)	(19;1.000)
C54	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.496)
C60	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;0.748)	(6;0.095)(7;0.095)(11;0.095)(17;0.252)(18;0.748)
C36	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(4;0.875)	(2;0.855)(4;0.855)(6;0.125)(10;0.125)	(19;0.855)(23;0.145)(26;0.048)
C47	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.936)
C17	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(4;1.000)	(4;1.000)	-
C27	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;1.000)
C16	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	-	(12;0.536)(13;0.536)(19;0.464)	(17;0.536)(18;0.464)(26;0.020)(27;0.020)(28;0.020)	(6;0.850)(7;0.850)(11;0.020)(17;0.150)(18;0.020)
C29	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.088)
C7	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(12;0.116)(13;0.116)(19;0.884)	(17;0.116)(18;0.548)(26;0.116)(27;0.116)(28;0.452)	(6;0.500)(7;0.500)(11;0.452)(17;0.500)(18;0.452)
C6	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(19;1.000)	-	-	-
C4	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(12;0.080)(13;0.080)(19;0.920)	(17;0.080)(18;0.576)(26;0.080)(27;0.080)(28;0.424)	(6;0.576)(7;0.576)(11;0.424)(17;0.380)(18;0.380)
C9	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(12;0.444)(13;0.444)(19;0.556)	(17;0.444)(18;0.556)(26;0.192)(27;0.192)(28;0.192)	(6;0.665)(7;0.665)(11;0.192)(17;0.335)(18;0.192)
C31	-	(1;1.000)	(1;1.000)	(1;1.000)	(31;0.160)(33;0.160)(37;0.712)(39;0.288)	(6;0.195)	(6;0.805)	(4;1.000)	(4;1.000)
C30	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(19;1.000)	(28;0.284)	(6;0.150)(7;0.150)(11;0.150)(17;0.716)(18;0.284)	(4;0.150)(6;0.750)(7;0.750)(14;0.150)(17;0.250)
C11	-	-	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;0.948)	(17;0.052)(18;0.888)(26;0.052)(27;0.052)(28;0.112)
C25	-	(1;1.000)	(1;1.000)	(1;1.000)	-	(4;1.000)	(18;0.405)	(4;0.595)(6;0.405)(7;0.405)(14;0.160)(17;0.160)	(4;0.840)(6;0.160)(7;0.160)
C39	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(1;1.000)	(19;1.000)	(28;0.600)
C3	-	(1;1.000)	(1;1.000)	(1;1.000)	(29;1.000)	(4;1.000)	(19;1.000)	(28;1.000)	(28;0.660)

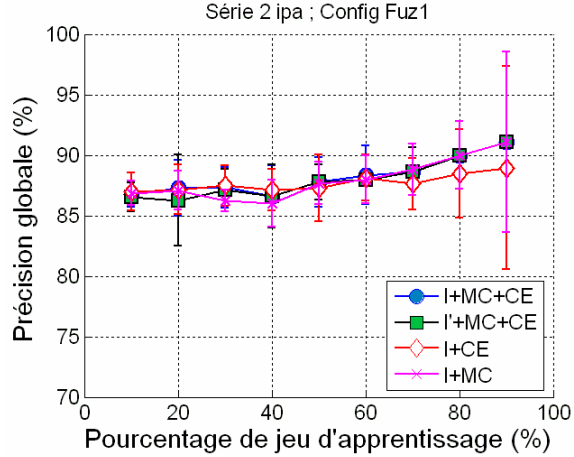
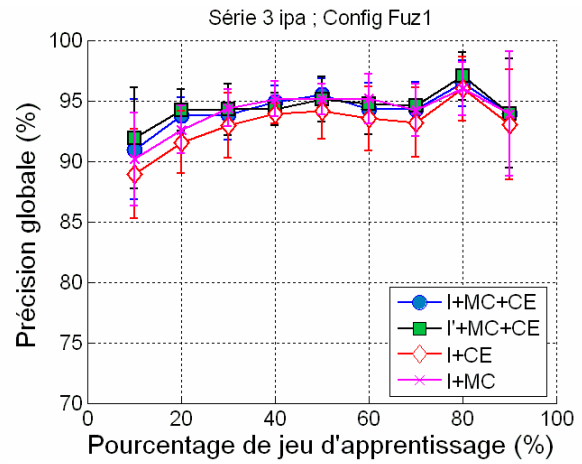
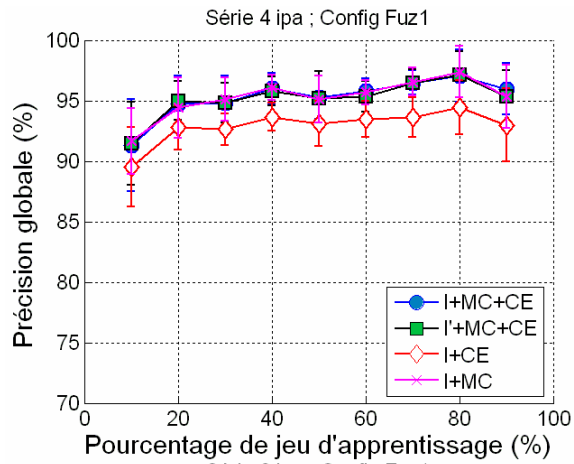
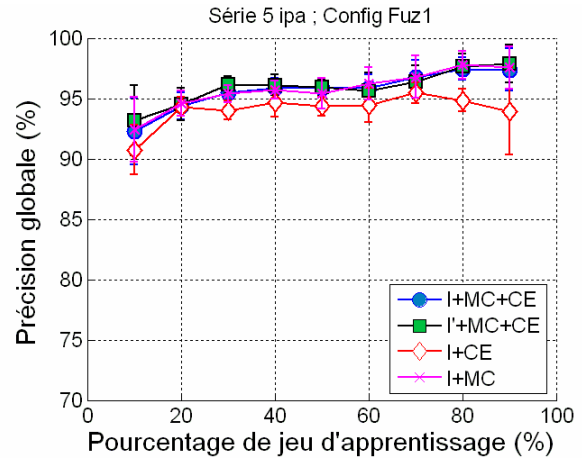
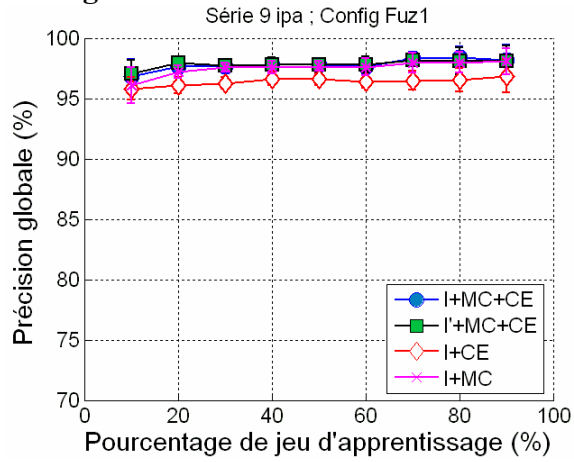
Règles activées lors de l'application du système à base de règles expertes sur l'exploitation de M. Caruel en utilisant la série temporelle d'images SPOT acquises en 2003. Chaque cellule du tableau contient à la fois le(s) numéro(s) de(s) règle(s) activée(s) pour faire l'inférence ainsi que le(s) degré(s) d'activation : (numéro de règle, degré d'activation de la règle). Les (-) correspondent aux parcelles pour lesquelles aucune décision n'est prise soit parce qu'elles sont couvertes par des nuages soit parce qu'il n'existe pas de couple d'images dégagées sur la parcelle.

Annexe 13 : Résultats du système à base de règles générées par apprentissage

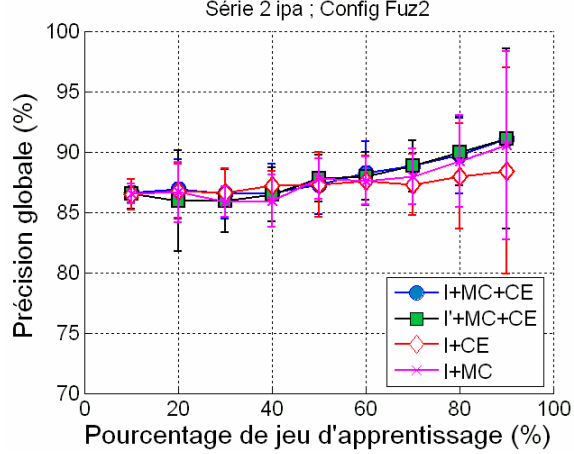
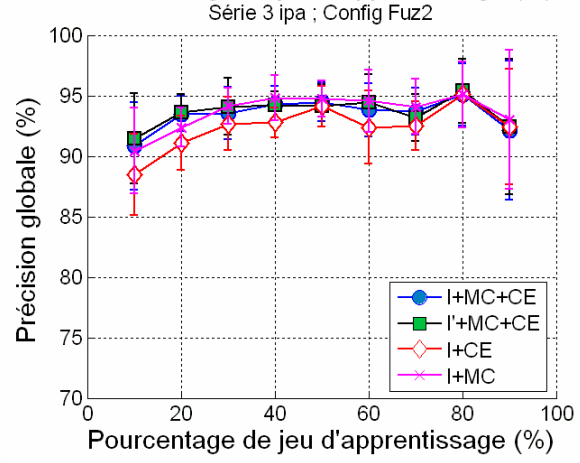
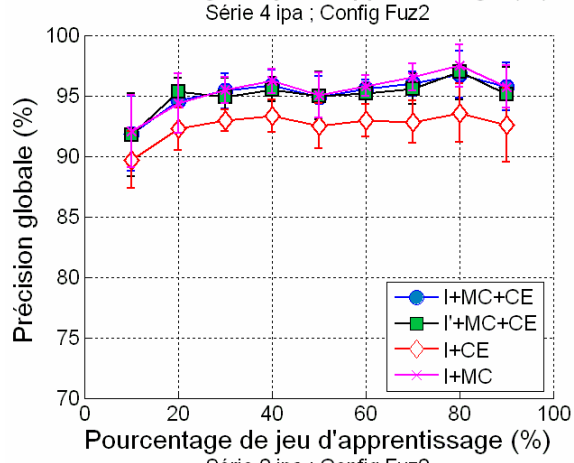
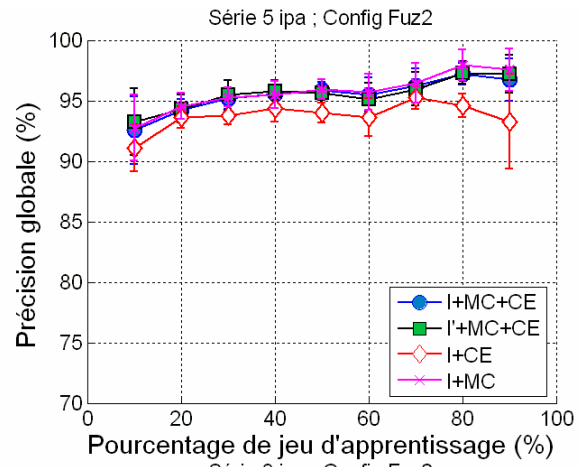
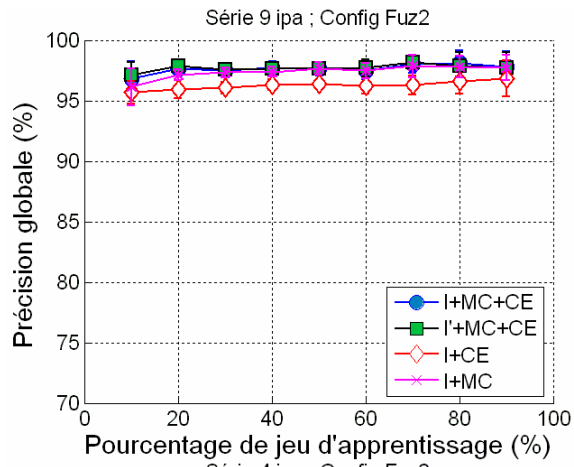
Dans cette annexe nous illustrons les précisions globales (PG) obtenus avec les différentes configurations de partition, les différentes séries d'images, les différentes sources d'informations et les différents pourcentages de jeux d'apprentissage. Nous montrons le nombre de règles générées dans chaque cas. Nous présentons également les courbes (PG vs. Nombre d'images par an) qui permettent d'analyser la robustesse du système.

Précision globale

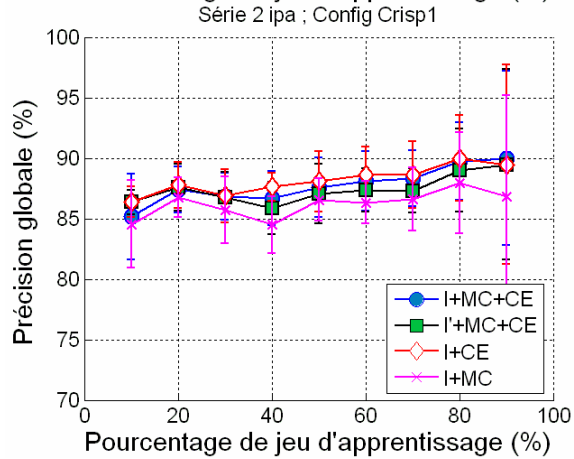
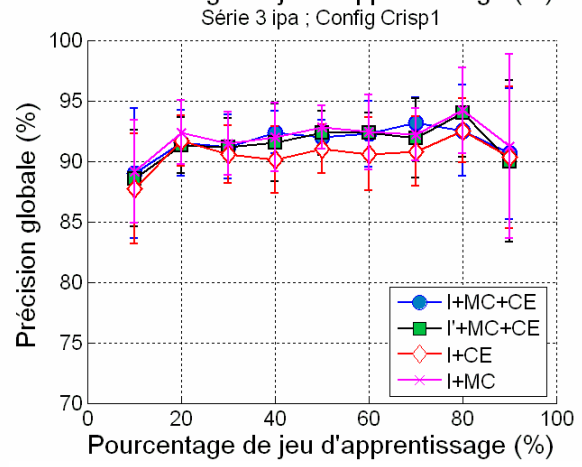
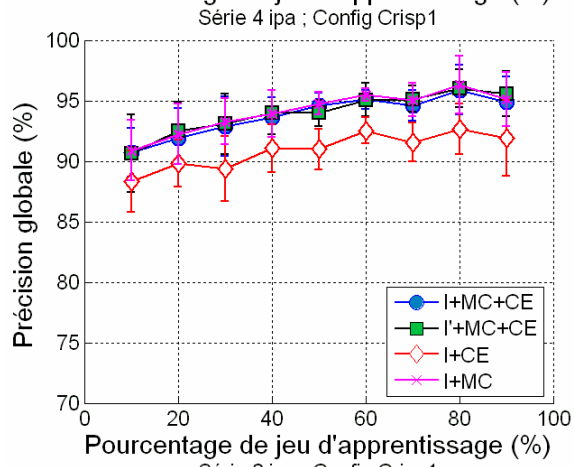
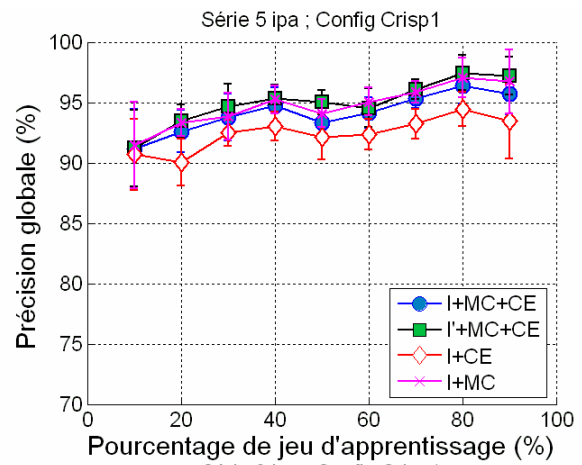
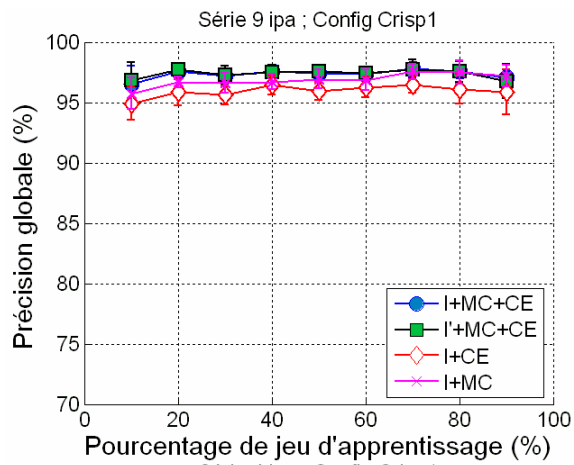
Config Fuz1



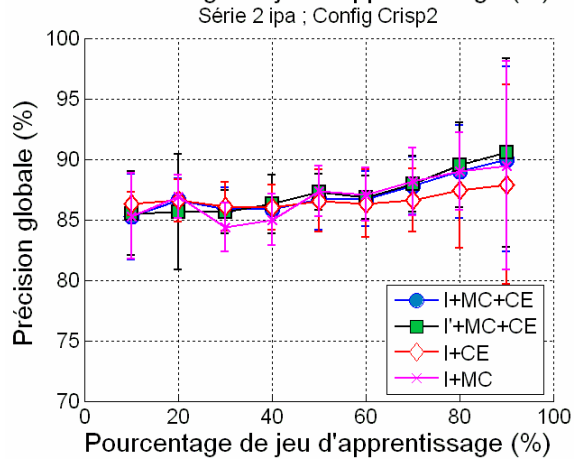
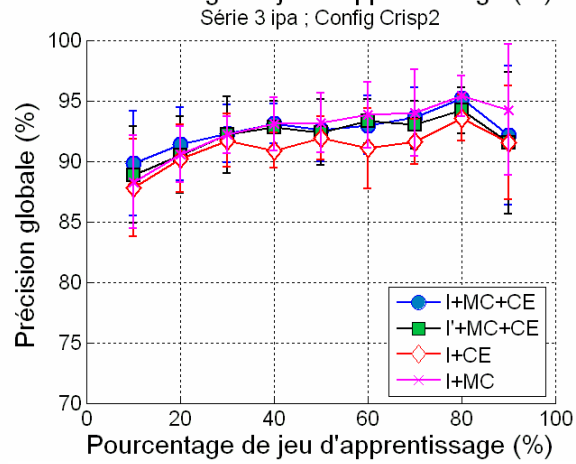
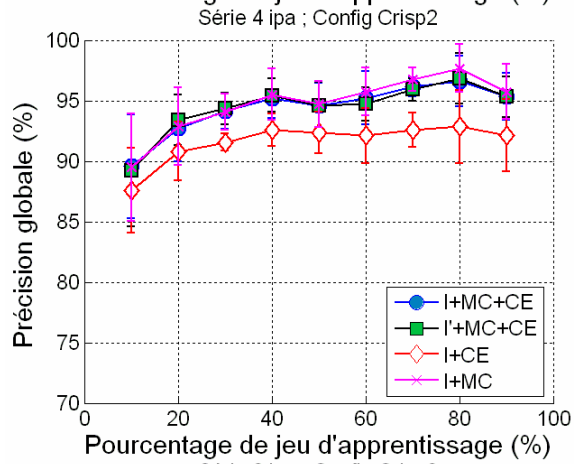
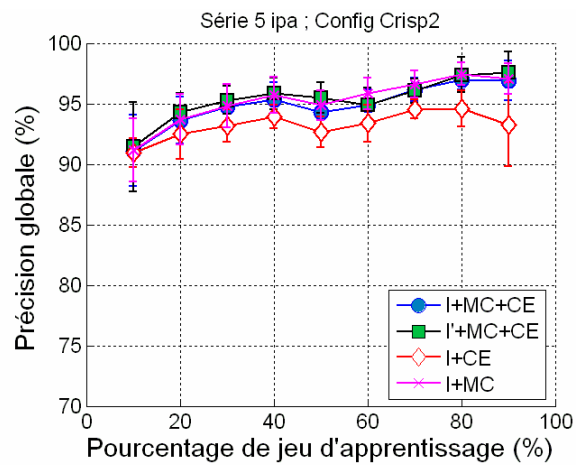
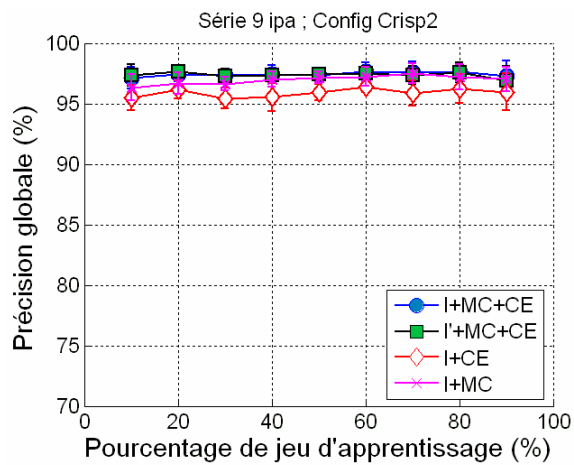
Config Fuz2



Config Crisp1

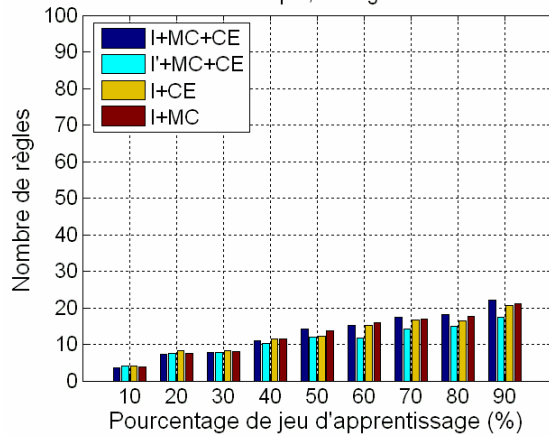
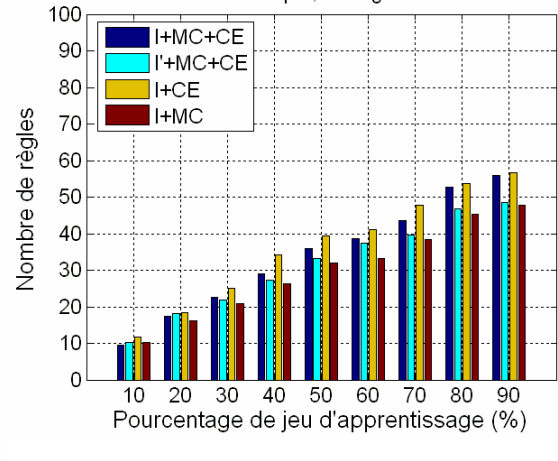
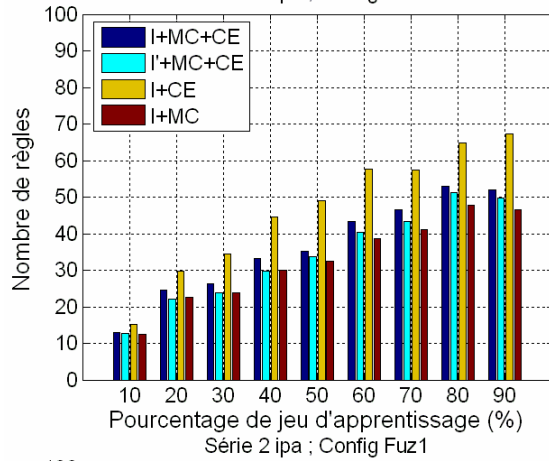
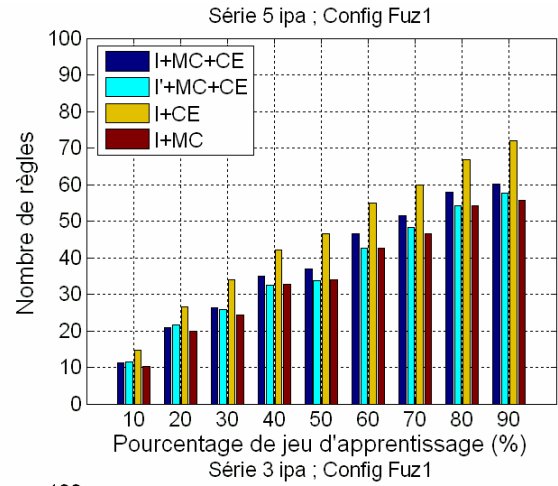
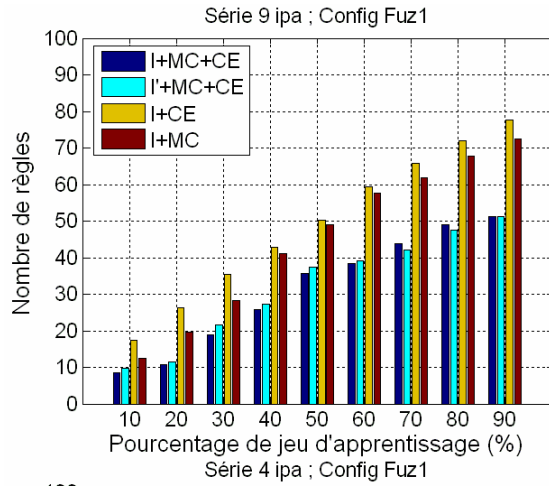


Config Crisp2

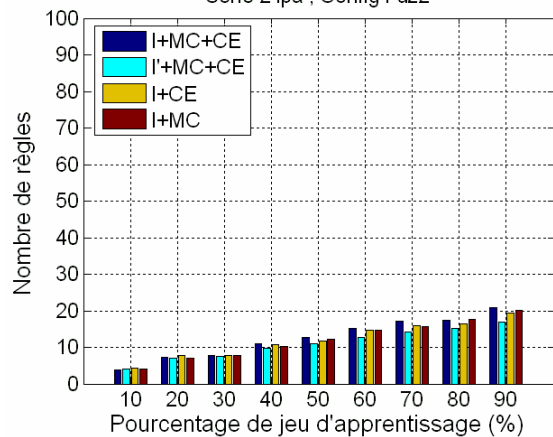
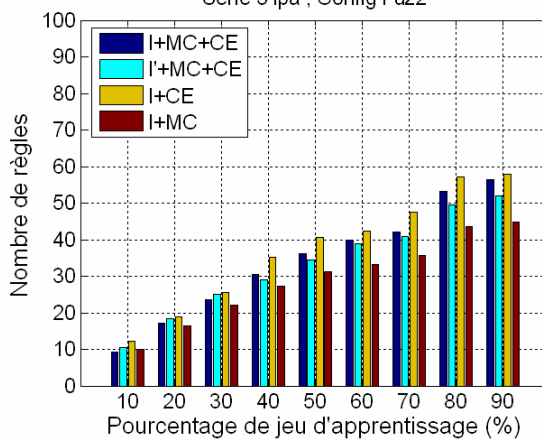
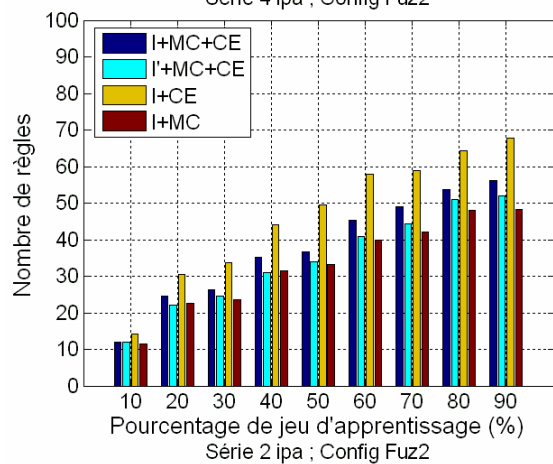
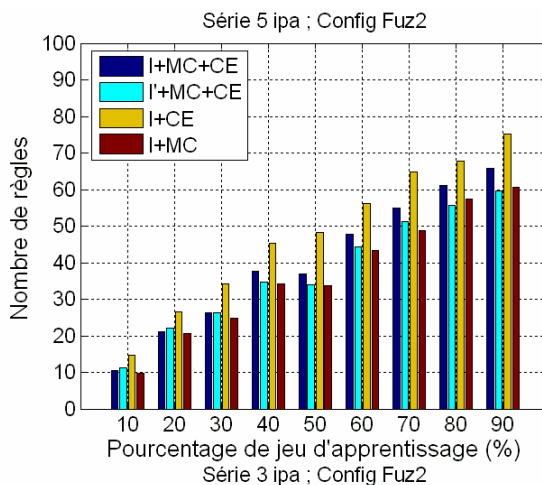
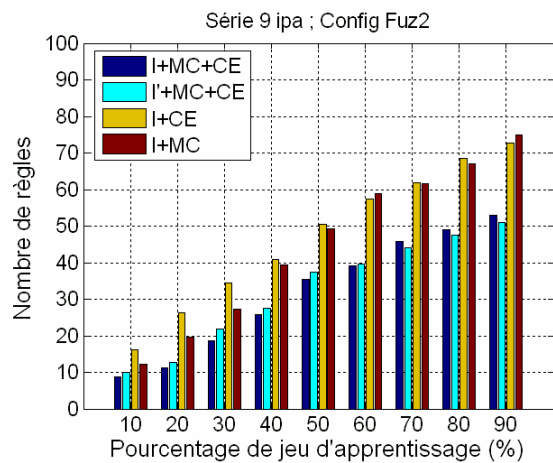


Nombre de règles

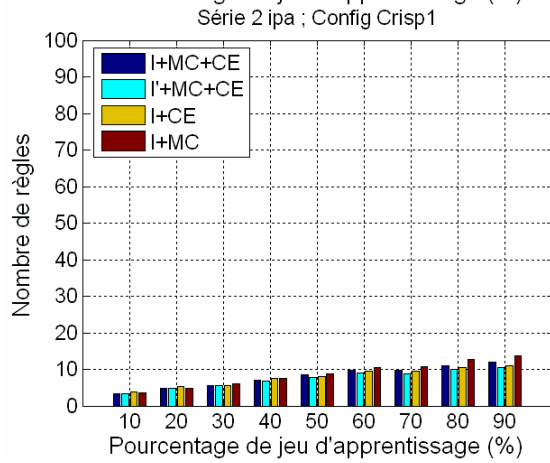
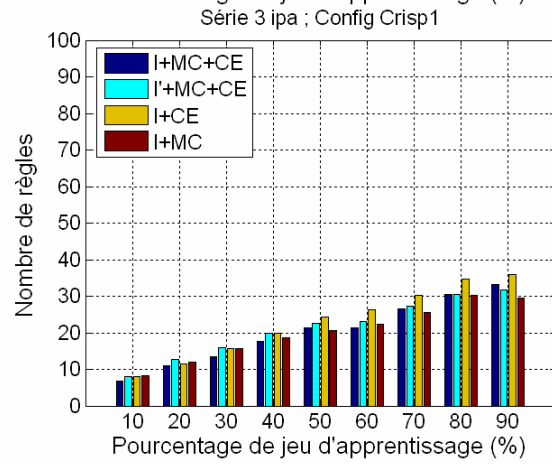
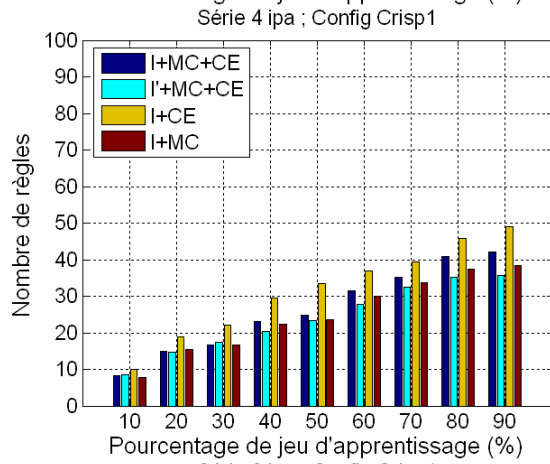
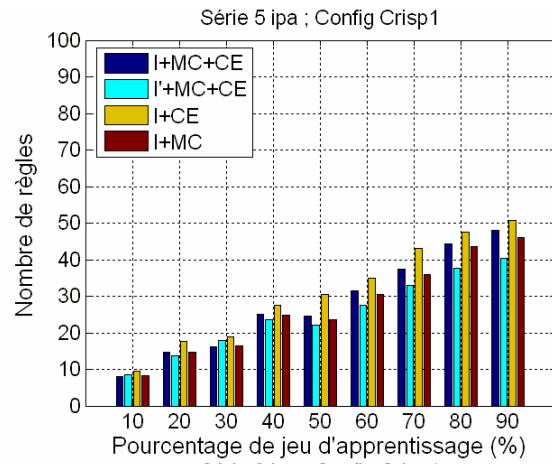
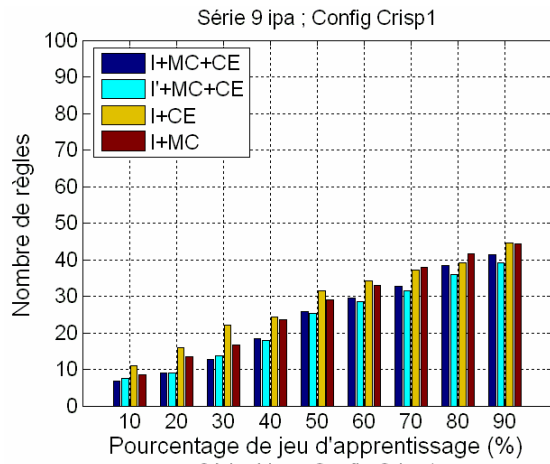
Config Fuz1



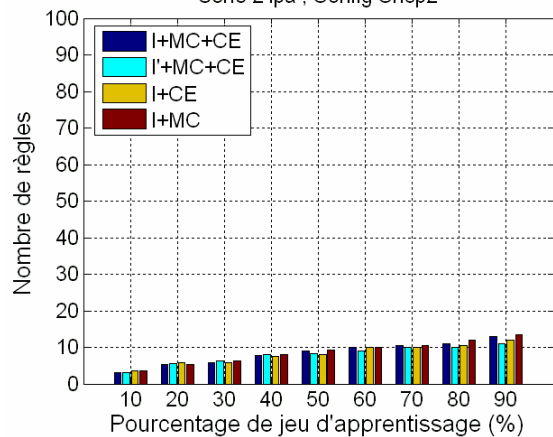
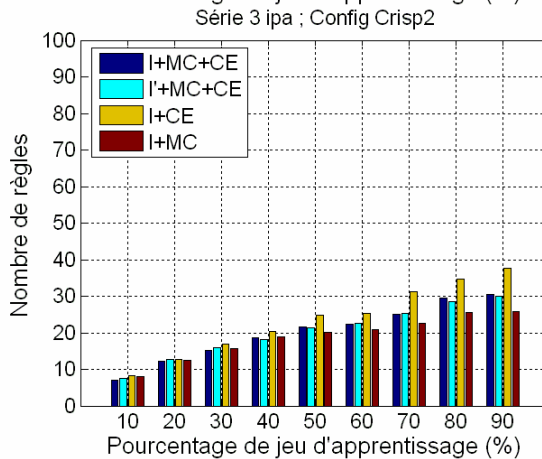
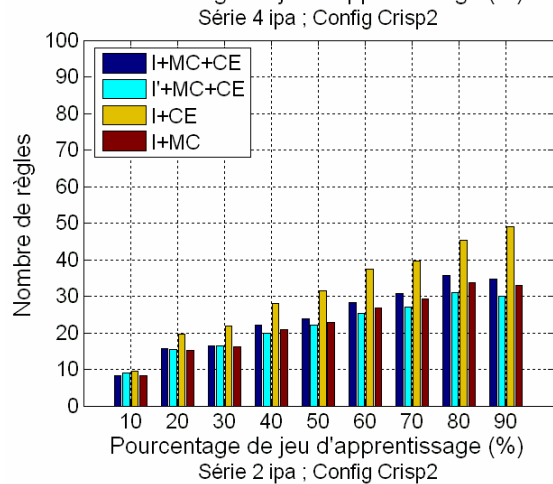
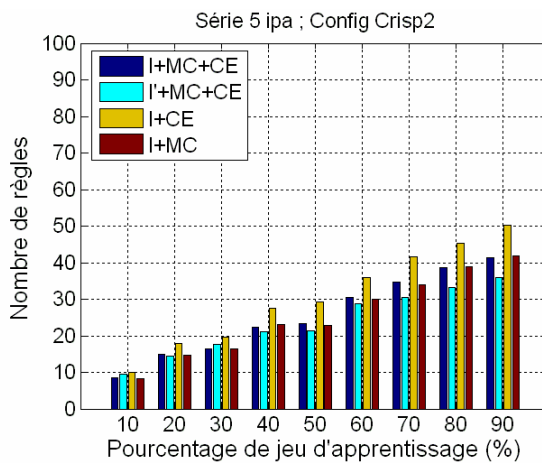
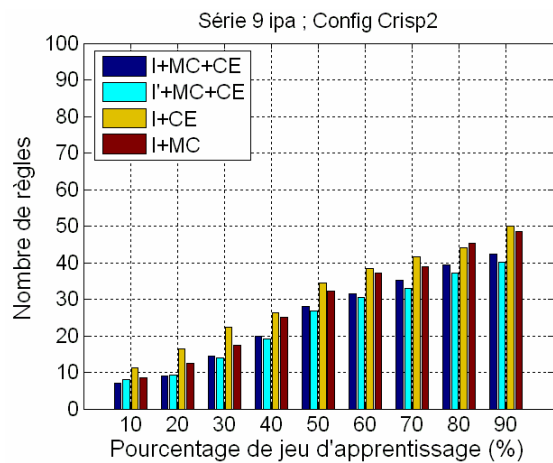
Config Fuz2



Config Crisp1

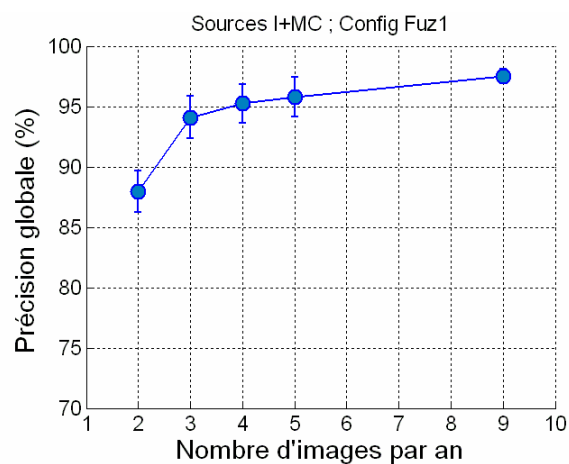
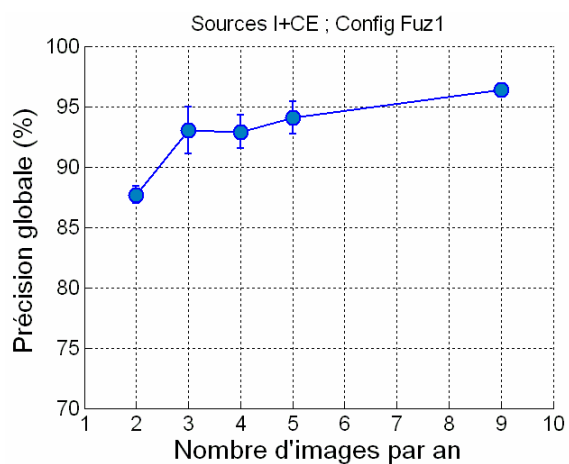
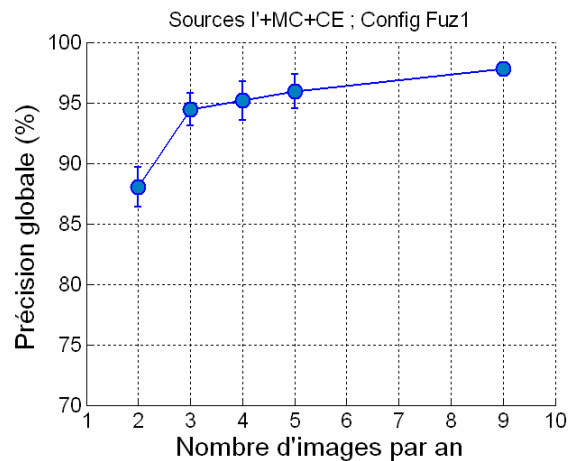
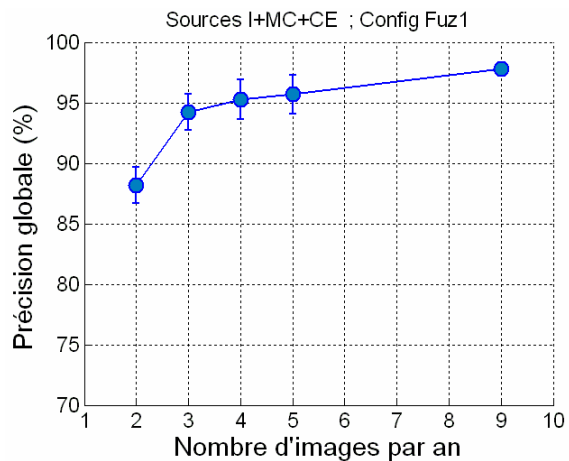


Config Crisp2

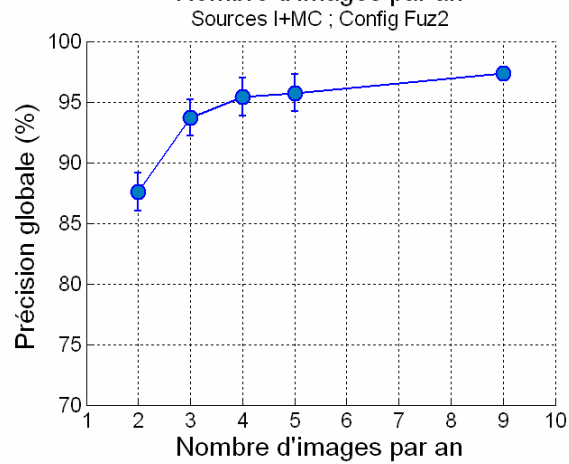
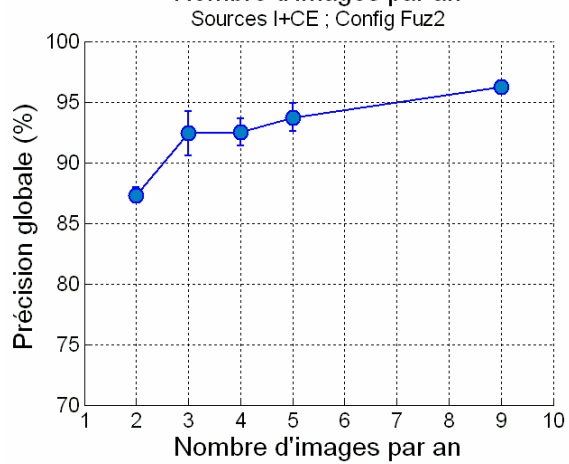
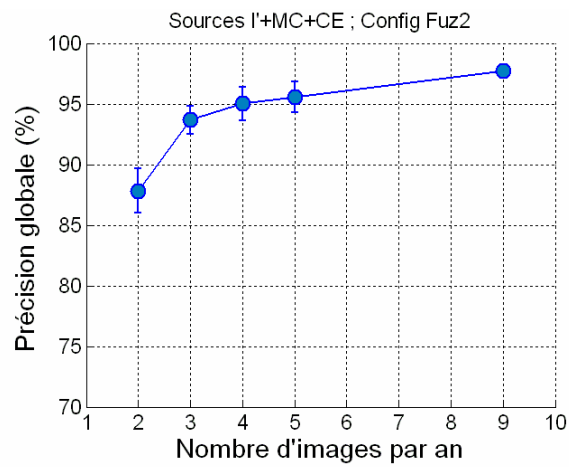
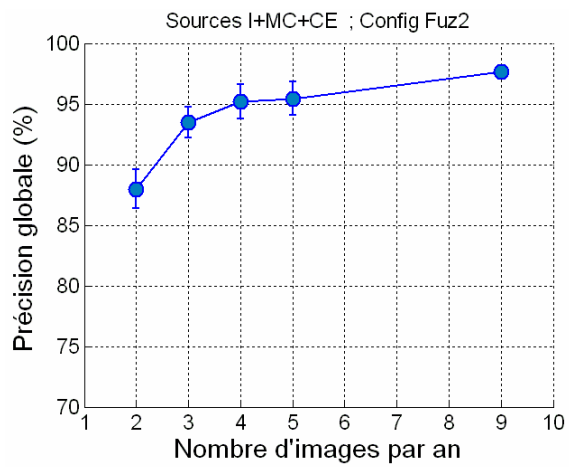


Robustesse

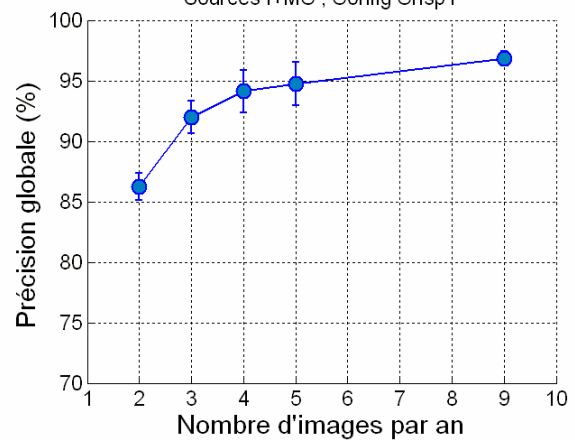
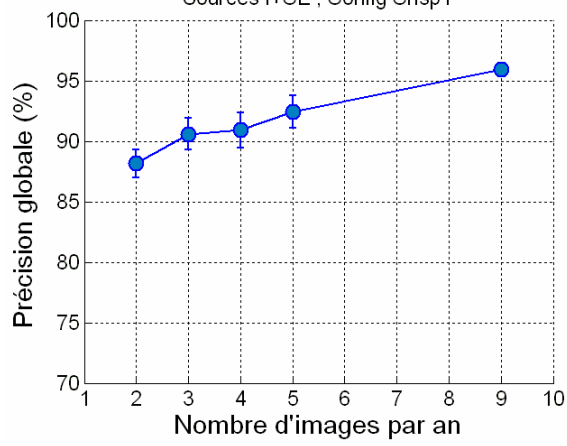
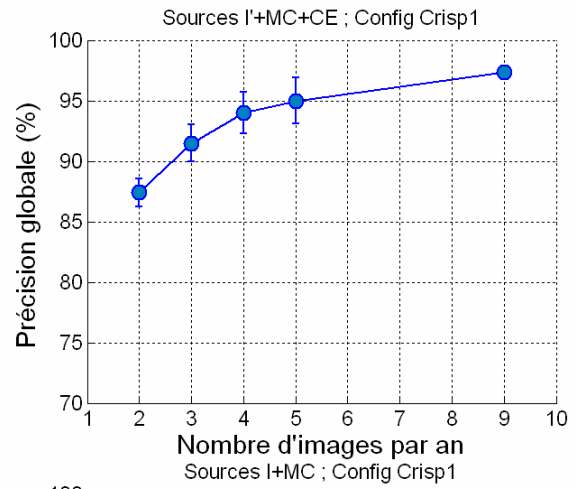
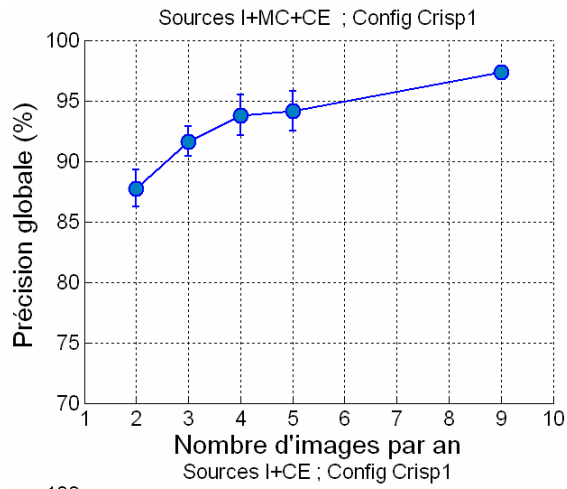
Config Fuz1



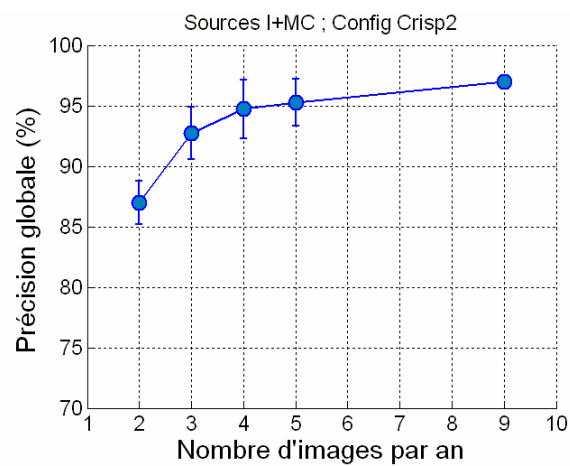
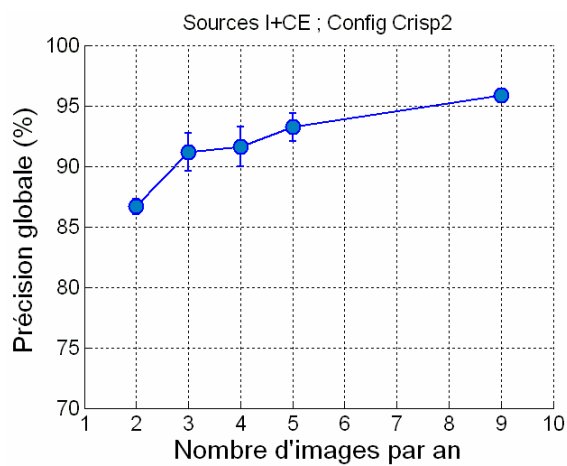
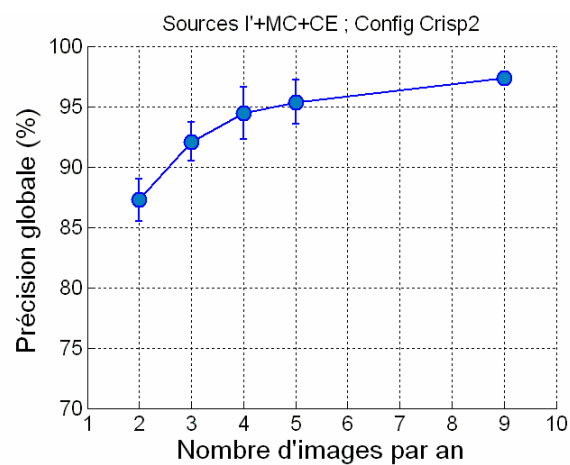
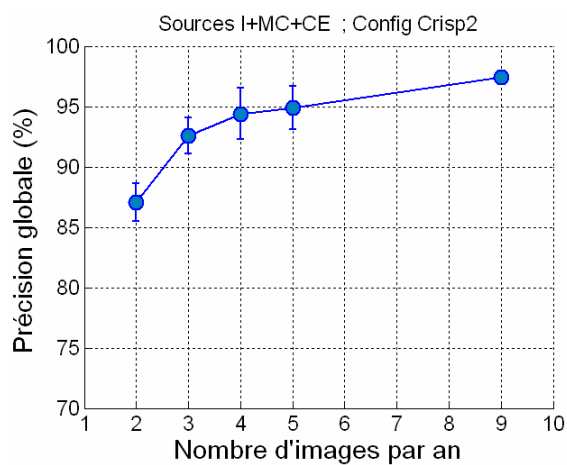
Config Fuz2



Config Crisp1



Config Crisp2



Annexes publications

Annexe P.1 :

Relative radiometric normalization and atmospheric correction of a SPOT 5 time series.....	235
Radiometric normalization of a SPOT 4 and SPOT 5 time series of images (ISLE-REUNION) for agriculture applications	255
Radiometric normalization of high spatial resolution multi-temporal imagery: A comparison between a relative method and atmospheric correction.....	263

Annexe P.2 :

Integrating SPOT-5 time series, crop growth modeling and expert knowledge for monitoring agricultural practices - The case of sugarcane harvest on Reunion Island.....	273
Multi-source information fusion: Monitoring sugarcane harvest using multi-temporal images, crop growth modelling, and expert knowledge.....	307
Combining multi-source information for crop monitoring.....	313
Fusion d'informations multi-sources pour le suivi des coupes de canne à sucre à La Réunion.....	321

