



Organisation des Nations Unies
pour l'alimentation et l'agriculture

Manuel sur l'application de la télédétection aux statistiques agricoles



Manuel sur l'application de la télédétection aux statistiques agricoles



Table des matières

Abréviations et acronymes	viii
Tableaux et figures	x
Recommandations de citations	xv
Remerciements	xvi
Synthèse	xvii
CHAPITRE	
LOGICIELS D'ACCÈS ET D'ANALYSE ET DES DONNÉES	1
1.1. Introduction	1
1.2. Capteurs et jeux de données	1
1.2.1. Facteurs de coûts liés à l'imagerie commerciale	6
1.2.2. Accès libre et gratuit	7
1.3. Traitement des images et travaux préparatoires à l'analyse des données	8
1.4. Données de référence	14
1.5. Informatique dématérialisé	15
1.6. Étude de cas	16
1.7. Conclusion	20
CHAPITRE 2	
CARTOGRAPHIE ET SUIVI DE LA COUVERTURE DU SOL	21
2.1. Introduction	21
2.2. Le concept de couverture du sol	23
2.2.1. Systèmes de classification de la couverture du sol (LCCS et LCML)	23
2.2.2. L'agriculture dans la typologie de la couverture du sol	24
2.2.3. Autres méthodes de caractérisation du sol	25
2.3. Réalisation de cartes de couverture du sol	27
2.3.1. Données de télédétection	27
2.3.2. Collecte de données <i>in situ</i>	29
2.3.3. Traitement des images et réalisation de cartes	35
2.4. Pratiques actuelles et jeux de données de couverture du sol existants	43
2.4.1. Métadonnées, politiques sur les données et production participative	43
2.4.2. Inventaire détaillé des jeux de données relatifs à la couverture du sol et aux terres cultivées existants	44
2.4.3. Détection des changements de la couverture du sol	46
2.5. Initiatives de cartographie en cours à l'échelle multinationale ou mondiale	48
2.5.1. Carte mondiale des terres cultivées	48
2.5.2. Système Sentinel-2 pour le suivi de l'agriculture	49
2.5.3. Cartographie du riz irrigué	50
2.6. Références bibliographiques	52

CHAPITRE 3

LA TÉLÉDÉTECTION AU SERVICE DE LA CONCEPTION DE BASES DE SONDAGE

3.1. Bases de sondage de type liste et bases aréolaires	62
3.1.1. Stratification	62
3.1.2. Erreurs de stratification	62
3.2. Bases de sondage aréolaires	63
3.2.1. Bases contenant des segments délimités par des frontières physiques	63
3.2.2. Bases fondées sur des grilles régulières	64
3.2.3. Stratification à partir de cartes mondiales des terres cultivées	68
3.2.4. Stratification fondée sur des données plus détaillées relatives à une année précédente	69
3.2.5. Photo-interprétation d'images du domaine public et production participative	70
3.2.6. Bases aréolaires constituées de points en grappe	72
3.2.7. Bases aréolaires constituées de points non regroupés	73
3.2.8. Étude comparative bases constituées de segments/de points	74
3.2.9. Bases aréolaires constituées de transects	75
3.3. Utilisation de la télédétection dans les bases de sondage de type liste	78
3.3.1. Localisation d'une exploitation agricole : un problème difficile à résoudre	78
3.3.2. Registres administratifs	78
3.3.3. Utilisation des districts de recensement comme unités primaires d'échantillonnage	78
3.3.4. Échantillonnage de Districts de Recensement avec une probabilité proportionnelle à la zone couverte	80
3.3.5. Stratification des districts de recensement	82
3.4. Mesure des parcelles	84
3.5. Échantillonnage d'images satellite	85
3.5.1. Expériences menées dans les années 1970	85
3.5.2. Échantillonnage d'images à résolution moyenne	85
3.5.3. Échantillonnage d'images à très haute résolution	86
3.6. Enquêtes le long des routes	87
3.7. Références bibliographiques	89

CHAPITRE 4

RÉALISATION D'UNE CARTOGRAPHIE DÉTAILLÉE DES CULTURES À PARTIR DE DONNÉES DE TÉLÉDÉTECTION (COUCHES DE DONNÉES SUR LES CULTURES)

4.1. Introduction	93
4.2. Classification de l'imagerie satellite aux fins de réalisation d'une cartographie détaillée des cultures	94
4.2.1. Classification supervisée de l'imagerie satellite	94
4.2.2. Classification par pixel ou par objet	96
4.3. Couches de données d'entrée servant à la classification des cultures	97
4.3.1. Données de terrain (<i>in situ</i>)	97
4.3.2. Données d'observation de la Terre	99
4.3.3. Données et informations auxiliaires (secondaires)	103
4.4. Réalisation d'une cartographie des cultures à l'échelle nationale	104
4.4.1. De la recherche à la pratique	104
4.4.2. Les systèmes de surveillance en milieu agricole	104
4.4.3. Étude de cas : l'inventaire annuel des cultures du Canada fondé sur des données spatiales	107
4.5. Conclusion	118
4.6. Références bibliographiques	119

CHAPITRE 5

ESTIMATION DES SURFACES CULTIVÉES À PARTIR DES DONNÉES DE TÉLÉDÉTECTION 133

5.1. Introduction sur l'estimation des surfaces cultivées	133
5.2. Méthodes d'estimation des surfaces cultivées par télédétection	135
5.2.1. Conception de Bases de Sondage Aréolaire	135
5.2.2. Analyse de données monodates	142
5.2.3. Exploitation des données SAR aux fins d'estimation des surfaces cultivées	145
5.2.4. Données sur la réalité de terrain	146
5.2.5. Estimation de la précision : matrice de confusion et écart relatif	147
5.3. Estimateur par régression	148
5.4. Estimateur par calage	150
5.5. Estimateur sur petits domaines	150
5.6. Exemples de programmes nationaux, régionaux et mondiaux d'estimation des surfaces cultivées	150
5.6.1. Programmes nationaux	150
5.6.2. Programmes régionaux: Programme MARS du JRC européen	153
5.6.3. Programmes mondiaux	153
5.7. Rentabilité de l'évaluation des surfaces au moyen de la télédétection	155
5.8. Difficultés et limites	155
5.9. Références bibliographiques	158

CHAPITRE 6

SYSTÈME D'ALERTE PRÉCOCE ET ESTIMATION DES RENDEMENTS AGRICOLES 165

6.1. Introduction	165
6.2. Système d'alerte précoce	166
6.2.1. Ravageurs et maladies des plantes	167
6.2.2. Le modèle phénologique de surveillance de la végétation	168
6.2.3. Quels délais pour une alerte précoce ?	169
6.3. Alerte et intervention précoces	170
6.4. Prévision des rendements agricoles	172
6.4.1. Le modèle agrométéorologique	173
6.4.2. Le modèle fondé sur la télédétection	176
6.5. Conclusion	181
6.6. Références bibliographiques	186

CHAPITRE 7

SUIVI DU COUVERT FORESTIER ET DE LA DÉFORESTATION 189

7.1. Introduction et objectifs principaux	189
7.2. Informations générales sur l'application de la télédétection au suivi du couvert forestier	191
7.2.1. Définition des terres forestières, de la déforestation et de la dégradation	191
7.2.2. Caractéristiques du suivi de la déforestation par télédétection	192
7.2.3. Caractéristiques du suivi de la dégradation des forêts par télédétection	194
7.2.4. Disponibilité des données Landsat	194
7.2.5. Disponibilité des données Sentinel-2	195
7.3. Évaluation des ressources forestières mondiales de la FAO : enquête mondiale par télédétection	196
7.3.1. Informations générales sur l'échantillonnage statistique conçu pour estimer la déforestation à partir de capteurs optiques de résolution spatiale moyenne	196
7.3.2. Méthode d'échantillonnage générale utilisée dans l'enquête mondiale par télédétection	197
7.3.3. Sélection et prétraitement des images satellite	198
7.3.4. Traitement et analyse des images satellite	199
7.3.5. Analyse statistique	200
7.3.6. Évaluation de la précision et de la cohérence des estimations des changements du couvert forestier	202
7.3.7. Résultats pour les tropiques	202
7.3.8. Précision des estimations pour les tropiques	202
7.3.9. Intensification du système d'échantillonnage pour l'estimation des changements du couvert forestier à l'échelle nationale	203
7.3.10. L'avenir de l'évaluation des ressources forestières mondiales : objectif FRA 2020	203
7.4. Autres exemples d'enquêtes par télédétection aux fins de statistiques forestières	204
7.4.1. Statistiques sur la déforestation du produit « couvert arboré mondial », université du Maryland	204
7.4.2. Exemple au niveau national : unités paysagères de l'inventaire forestier national du Brésil	205
7.4.3. Enquête mondiale de la FAO sur les forêts	208
7.5. Complémentarité des estimations des changements dans les domaines forestier et agricole	211
7.6. Références bibliographiques	213

CHAPITRE 8

ORGANISATIONS, RESSOURCES ET COMPÉTENCES NÉCESSAIRES À L'ADOPTION DE LA TÉLÉDÉTECTION DANS LE DOMAINE DES STATISTIQUES AGRICOLES 219

8.1. Contexte	219
8.2. Organisations	220
8.3. Ressources	221
8.3.1. Personnel qualifié	222
8.3.2. Laboratoires : configuration matérielle et logicielle requise	222
8.3.3. Données d'entrée	224
8.3.4. Planification du travail	224
8.3.5. Formation requise	225
8.3.6. Financement	226
8.4. Mise en œuvre du programme : études de cas	227
8.4.1. Exemple 1 : Éthiopie – application de la stratification d'une base de sondage aréolaire	227
8.4.2. Exemple 2 : Pakistan – Service d'information sur les cultures	228
8.4.3. Exemple 3 : Pakistan – Unité géospatiale opérationnelle SUPARCO/FAO	229
8.4.4. Exemple 4 : Bangladesh – unité géospatiale opérationnelle	232
8.4.5. Exemple 5 : Rwanda – enquêtes par sondage probabiliste à objectifs multiples	235
8.5. Conclusion	237
8.6. Références bibliographiques	238

CHAPITRE 9

RENTABILITÉ DE LA TÉLÉDÉTECTION AUX FINS DE STATISTIQUES AGRICOLES 239

9.1. La question des coûts	239
9.2. Domaines d'application et gains relatifs	240
9.2.1. Optimisation des plans d'échantillonnage	240
9.2.2. Couches de données sur les cultures (CDL)	243
9.2.3. Estimateurs améliorés	243
9.2.4. Suivi des cultures et prévision des récoltes	244
9.3. Pertinence des capteurs	245
9.4. Conclusion	247
9.5. Références bibliographiques	248

ANNEXE

ESTIMATION ET CORRECTION DES BIAIS LIÉS AU COMPTAGE DE PIXELS 251

A1. Introduction	251
A2. Biais lié au comptage de pixels	253
A2.1. Estimateurs par calage	254
A2.2. Calcul de la variance de certains estimateurs	255
A2.3. Matrices de confusion exprimées sous forme de nombres de points	256
A2.4. Quelle est la meilleure méthode pour corriger le biais ?	258
A3. Le problème de subjectivité de l'estimation par comptage de pixels	259
A4. Références bibliographiques	260

Abréviations et acronymes

AAC	Agriculture et Agroalimentaire Canada
AEE	Agence européenne pour l'environnement
AMIS	Système d'information sur les marchés agricoles (FAO)
ASC	Académie des sciences de Chine (Chine)
Asia-RiCE	Initiative pour l'estimation et le suivi des cultures rizicoles en Asie
ASIS	Système d'indice de stress agricole (FAO)
BSA	Base de sondage aréolaire
BTL	Base de sondage de type liste
CAPE	Programme d'estimation des superficies consacrées aux cultures et de la production végétale (Inde)
CCNUCC	Convention-cadre des Nations Unies sur les changements climatiques
CDE	Cadre directeur d'échantillonnage
CDL	Cropland Data Layer
CHIRPS	Climate Hazards Group Infrared Precipitation with Station data [base de données]
CLC	CORINE Land Cover
DR	District de recensement
ENSO	El Niño-Oscillation australe
EQM	Erreur quadratique moyenne
ESA	Agence spatiale européenne
FAO	Organisation des Nations Unies pour l'alimentation et l'agriculture
fAPAR	Fraction de rayonnement photosynthétiquement actif
FAS	Service agricole à l'étranger (USDA)
FASAL	Programme de prévision de la production agricole à l'aide d'observations spatiales, agrométéorologiques et terrestres (Inde)
FRA	Évaluation des ressources forestières (FAO)
FSA	Agence de services aux exploitations agricoles (USDA)
GEE	Google Earth Engine
GEOGLAM	Initiative de suivi satellitaire de l'agriculture mondiale du Group on Earth Observations
GFW	Projet Global Forest Watch
GLCN	Réseau mondial sur le couvert végétal
GNSS	Système mondial de navigation par satellite
GPS	Système de positionnement universel
GPU	Processeur graphique
GRD	SAR de niveau 1 (Ground Range Detected)
GSD	Distance d'échantillonnage au sol
IAC	Inventaire annuel des cultures (Canada)
IIASA	Institut international pour l'analyse des systèmes appliqués (Vienne)
JECAM	Expérience conjointe sur l'évaluation et le suivi des récoltes
JRC	Centre commun de recherche (UE)
KML	Format de fichier Keyhole Markup Language
LACIE	Large Area Crop Inventory Experiment
LAI	Surface foliaire
LCCS	Système de classification de la couverture du sol
LCML	Métalangage de couverture du sol
MAN	Modèle altimétrique numérique
MARS	Surveillance de l'agriculture par télédétection (UE)

MVS	Machine à vecteurs de support
NASA	Agence spatiale américaine (États-Unis)
NASS	Service national de statistiques agricoles (USDA)
NCFC	Centre national Mahalanobis de prévision des récoltes (Inde)
NDVI	Indice différentiel normalisé de végétation
NIR	Proche infrarouge
NOAA	Agence américaine d'observation océanique et atmosphérique (États-Unis)
OSM	Open Street Map
PEP	Productivité de l'eau de pluie
PG	Précision globale (des classifications)
PLC	Produits logiciels du commerce
PNUE	Programme des Nations Unies pour l'environnement
PP	Précision du producteur
PPT	Probabilité proportionnelle à la taille (échantillonnage)
PU	Précision de l'utilisateur
RADI	Institut de télédétection et Terre numérique (Chine)
REDD+	Réduction des émissions liées à la déforestation et à la dégradation des forêts
SAP	Système d'Alerte Précoce
S1tbx	Boîte à outils Sentinel-1
SAE	Estimation par petites zones
SAR	Radar à synthèse d'ouverture
SIA	Système d'information agroclimatique
SIG	Système d'information géographique
SIGMA	Stimuler l'innovation pour la surveillance mondiale de l'agriculture
SNAP	Plateforme d'application Sentinel
SRTM	Mission de cartographie radar de la Terre
SWIR	Infrarouge à ondes courtes
THR	(Imagerie) très haute résolution
TOA	Réflectance au sommet de l'atmosphère
TOC	Réflectance au sommet de la canopée
UAV	Véhicule aérien sans pilote
UCM	Unité cartographique minimale
UE	Union européenne
UPE	Unité primaire d'échantillonnage
USDA	Département de l'agriculture des États-Unis
USE	Unité secondaire d'échantillonnage
USGS	Institut d'études géologiques des États-Unis
WCA 2020	Programme mondial de recensement de l'agriculture 2020
WMS	Web Map Service
XML	Langage de balisage extensible

Tableaux et figures

TABLEAUX

CHAPITRE 1	1
Tableau 1. Classification des catégories de capteurs satellitaires inspirée de la nomenclature utilisée par l'Agence spatiale européenne dans le cadre du programme Copernicus.	3
Tableau 2. Couches de données de référence pertinentes pour la préparation d'enquêtes agricoles et d'analyses connexes faites à partir d'observations de la Terre.	14
CHAPITRE 2	21
Tableau 1. Critères applicables à la sélection des sources de données de télédétection appropriées.	29
Tableau 2. Matrice de confusion dans laquelle n_{ij} représente le nombre d'échantillons de validation correspondant sur la carte à la classe de couverture du sol i et aux informations de validation j .	34
Tableau 3. Forces et faiblesses des algorithmes utilisés en vue de la classification des données issues d'images satellite pour des superficies importantes (fondé sur Gómez et al., 2016).	38
CHAPITRE 4	93
Tableau 1. Conditions relatives à l'utilisation des données satellitaires d'observation de la Terre, conçues par une équipe spéciale du Comité sur les satellites d'observation de la Terre (CEOS, de l'anglais Committee on Earth Observation Satellites) pour GEOGLAM (CEOS, 2014 ; GEOGLAM, 2014).	100
Tableau 2. Recensement des cartes nationales, régionales et mondiales de la couverture du sol (source : Waldner <i>et al.</i> , 2015). Ce tableau ne contient que des données accessibles à tous. Les ensembles de données dont l'utilisation ou l'accès est limité par une politique de diffusion n'ont pas été recensés.	105
Tableau 3. Variabilité par province de la précision de la classification de l'inventaire annuel des cultures d'AAC, 2015	113
CHAPITRE 5	133
Tableau 1. Exploitation de la télédétection aux fins d'estimation des surfaces cultivées dans différents pays.	137
Tableau 2. Plan d'échantillonnage des zones rizicoles, coefficients de variation (CV) et efficacité de la stratification dans plusieurs États indiens couverts par le projet FASAL.	140
Tableau 3. Exemples de coûts des différentes données satellitaires optiques et SAR utilisées pour estimer les surfaces cultivées.	142
Tableau 4. Exemple de matrice de confusion utilisée pour une étude pilote au Kazakhstan.	147
Tableau 5. Exemple de matrice de confusion utilisée pour l'État du Madhya Pradesh, en Inde.	148
Tableau 6. Taille moyenne et fragmentation des exploitations agricoles, 1995-2005.	157
CHAPITRE 6	165
Tableau 1. Liste des indices de végétation courants, formules mathématiques, échelle de développement et paramètres estimés.	184

CHAPITRE 7	189
Tableau 1. Caractéristiques des capteurs Landsat 8 OLI et Sentinel-2	193
Tableau 2. Classes de couverture du sol utilisées par le JRC	199
CHAPITRE 8	219
Tableau 1. Synthèse des disciplines requises pour utiliser la télédétection dans le domaine des statistiques et de la diffusion des données agricoles.	222
Tableau 2. Synthèse des besoins d'un laboratoire de statistiques agricoles géospatiales (matériel informatique).	223
Tableau 3. Synthèse des besoins d'un laboratoire de statistiques agricoles géospatiales (logiciels).	223
Tableau 4. Exemple de calendrier des cultures en Asie destiné à identifier les périodes d'acquisition des images et à atteindre les objectifs d'échantillonnage.	225
Tableau 5. Techniques d'identification de la couverture cultivée.	229
CHAPITRE 9	239
Tableau 1. Comparaison des estimations des cultures par stratification traditionnelle (ST) et stratification automatique (SA), Dakota du Sud, 2013.	242
Tableau 2. Superficie (en million d'hectares) occupée par les différentes classes de taille des parcelles, ventilée par région.	245
Tableau 3. Résolution des satellites et pourcentage relatif de compatibilité avec les superficies cultivées.	246
Tableau 4. Classification des pays d'Afrique et d'Asie par niveau de difficulté d'utilisation de l'imagerie satellite aux fins de statistiques agricoles, en fonction du pourcentage de terres cultivées, par classe de taille de parcelle.	247
ANNEXE	251
Tableau A1. Matrice de confusion contenant une classe M. Jaune : erreur de commission pour la classe 1 (blé). Rose : erreur d'omission pour la même classe.	253

FIGURES

CHAPITRE 1	1
Figure 1. Schéma du workflow de traitement des informations SAR dans s1tbx, visant à générer des coefficients de rétrodiffusion calibrés et géocodés à partir des jeux de données GRD de Sentinel-1	12
Figure 2. Composite de coefficients de rétrodiffusion Sentinel-1 géocodés fondé sur des images ascendantes (orbite 132 relative, polarisation VV) d'une zone agricole proche de Kura, au Nigéria, prises les 27 juillet, 9 septembre et 23 octobre 2016.	13
Figure 3. Superposition d'une carte générale de la commune de Hollands Kroon (Pays-Bas) et d'une image de Sentinel-2 du 5 juillet 2016 (combinaison des bandes B8-B11-B4).	18
Figure 4. Matrice de confusion relative à une classification des forêts aléatoire à partir des images Sentinel-2 de Hollands Kroon (5 juillet 2015), fondée sur les déclarations du BRP 2015 aux fins d'essai et d'apprentissage.	19

CHAPITRE 2

21

Figure 1.	Processus de cartographie de la couverture du sol à partir de séries temporelles d'observations par satellite.	22
Figure 2.	Échantillonnage par enquête superficielle	33
Figure 3.	(a) Carte des indicateurs de priorité et (b) typologie des actualisations pour chaque pays.	45
Figure 4.	Changements de couverture du sol survenus entre 1992 et 2015 au Cambodge, détectés à 1 km et cartographiés à une résolution de 300 m.	47
Figure 5.	État d'avancement actuel (mars 2017) du projet GFSAD30 de l'USGS, qui vise à cartographier les terres cultivées à une résolution de 30 m (https://croplands.org/).	48
Figure 6.	Compositions en couleurs naturelles d'une résolution de 10 m de Sen2-Agri (à gauche) et cartes de terres cultivées correspondantes (à droite) pour la saison de végétation 2016 au Mali.	50

CHAPITRE 3

61

Figure 1.	UPE délimitée par des frontières physiques au sein de laquelle un segment a été sélectionné.	63
Figure 2.	Superposition d'une grille carrée à une région administrative (gauche) et grille carrée après suppression des cellules de petite taille (droite).	64
Figure 3.	Carte mondiale de la superficie la plus courante des champs.	65
Figure 4.	Segment réduit compte tenu du nombre trop élevé de champs qu'il englobe.	66
Figure 5.	Superposition d'une carte de couverture du sol et d'une grille d'échantillonnage.	67
Figure 6.	Superposition d'une grille d'échantillonnage à une carte de couverture du sol haute résolution (gauche) et stratification qui en découle (droite)	68
Figure 7.	Stratification à partir d'une carte de probabilité de présence de cultures et d'une grille d'échantillonnage composée de rectangles, à Sétif (Algérie).	69
Figure 8.	Grille de segments carrés superposée à une image classée.	70
Figure 9.	Interface simple de photo-interprétation participative.	71
Figure 10.	Système de photo-interprétation plus élaboré, composé de polygones préalablement délimités par un logiciel de segmentation automatique.	72
Figure 11.	Grille régulière de points appliquée à un segment carré (système d'échantillonnage à deux degrés).	73
Figure 12.	Échantillon stratifié de la première phase (Europe centrale).	74
Figure 13.	Transect généré à partir d'un point échantillonné pour enquêter sur des parcelles longues et fines.	76
Figure 14.	Système d'échantillonnage par bandes dans le cadre d'une étude de simulation (Pays-Bas).	77
Figure 15.	Exemple de DR pour lequel il existe des doutes quant à la précision géométrique.	79
Figure 16.	Modèle de répartition de communes en dents de scie (étape préalable au processus d'échantillonnage systématique).	80
Figure 17.	Échantillons avec PPT de communes de Castille-et-León (Espagne) et représentation de la zone géographique correspondante (région) servant d'outil de mesure de la taille.	81
Figure 18.	Point photo-interprété comme terre cultivée (gauche) et point photo-interprété comme terre non cultivée (droite)	82
Figure 19.	Répartition des terres arables en Castille-et-León (gauche) et échantillon systématique de communes présentant une probabilité proportionnelle à la surface de terres arables (droite).	82
Figure 20.	Grille de points superposée à un DR, pouvant être utilisée comme échantillon de première phase au cours de la deuxième étape du processus d'échantillonnage à deux degrés.	83
Figure 21.	Photo-interprétation de parcelles à l'aide d'une orthophotographie (résolution de 0,5 m) et d'une image satellite (résolution de 2,5 m).	84
Figure 22.	Échantillon stratifié de scènes Landsat et quarts de scènes utilisés pour estimer l'évolution de la forêt tropicale	86

Figure 23.	Échantillon de sites de 10 × 10 km devant être analysés à l'aide d'images THR dans le cadre du projet Geoland 2.	87
Figure 24.	Échantillon de points obtenu dans le cadre d'une enquête le long d'une route par l'échantillonnage de larges segments carrés.	88

CHAPITRE 4 93

Figure 1.	Les étapes de la classification supervisée. L'approche adoptée par Agriculture et Agroalimentaire Canada (AAC) pour cartographier les cultures au niveau des parcelles suit les flèches rouges.	96
Figure 2.	Calendriers culturaux du blé (gauche) et du riz (droite).	98
Figure 3.	Carte de l'inventaire annuel des cultures 2016 du Canada.	108
Figure 4.	Sources et volumes des données utilisées pour les inventaires annuels des cultures d'AAC, 2011-2015.	110
Figure 5.	Couverture géographique et précision globale de la cartographie dans l'inventaire annuel des cultures 2016 du Canada.	112
Figure 6.	Comparaison entre les estimations des surfaces cultivées de l'inventaire annuel des cultures 2011 d'AAC et les données du Recensement de l'agriculture réalisé la même année par Statistique Canada.	115

CHAPITRE 5 133

Figure 1.	Conception de l'échantillonnage aréolaire pour l'État du Bihar (Inde) dans le cadre du programme FASAL.	136
Figure 2.	Coefficients de variation des estimations en fonction des surfaces rizicoles cultivées dans les différents États indiens.	141
Figure 3.	Images composites de l'indice NDVI hebdomadaire/bimensuel pour l'État de l'Uttar Pradesh, en Inde.	143
Figure 4.	Profils temporels (gradués) des indices NDVI de plusieurs classes de cultures pour l'État de l'Uttar Pradesh, en Inde.	144
Figure 5.	Images de Landsat FCC (gauche) et images classifiées (blé en vert et moutarde en jaune) (droite) pour le district de Bhiwani, dans l'État d'Haryana, en Inde.	144
Figure 6.	Profil temporel des différentes classes de riz pour le district de Mirzapur, dans l'État d'Uttar Pradesh, en Inde.	145
Figure 7.	Données de terrain recueillies par smartphone et disponibles sur le géoportail Bhuvan	147
Figure 8.	Cropland Data Layer 2009.	151
Figure 9.	Stratégies de prévision de la production végétale du projet FASAL.	152

CHAPITRE 6 165

Figure 1.	Caractéristiques des Système d'alerte précoce.	166
Figure 2.	Conditions climatiques favorisant ou non le développement de la rouille du caféier au Panama.	168
Figure 3.	Moyenne en début la saison (gauche), écart-type temporel (σ_t) en début de saison (au milieu, en jours) et corrélation de Spearman en début de saison (droite), fondés sur la série temporelle NDVI calculée par l'AVHRR. La partie inférieure indique les valeurs relatives à une éventuelle deuxième saison, le cas échéant.	168
Figure 4.	Quels délais pour une alerte rapide ? Ce graphique montre le délai de réaction des SAP et les superficies d'impact (en précisant le diamètre de l'aire sphérique) pour différents risques hydrométéorologiques. Source : Golnaraghi, 2005.	169

Figure 5.	Système d'alerte précoce sur les sécheresses agricoles et intervention rapide. La carte représente le pourcentage de terres agricoles touchées par une sécheresse (Système d'information agroclimatique [SIA]) en 2006.	172
Figure 6.	Relation entre le rendement national du blé dur (kg/ha) et la pluviométrie de la campagne agricole (mm) (données de 1988 à 2011).	174
Figure 7.	Rapport entre les rendements observés et calculés pour les variétés de riz à haut rendement au Rajshahi, Bangladesh (1983-1998).	174
Figure 8.	Simulation à long terme du rendement médian du blé à l'échelle des districts (1901-2015).	176
Figure 9.	Rapport entre les estimations de rendement observés et estimés au Kenya.	177
Figure 10.	Rapport entre l'indice de stress agricole (ASI) et le rendement du blé en Syrie (1985-2012)	178
Figure 11.	Rapport entre le NDVI et la production de millet pour la période 1982-2006 (gauche) ; prix moyens et prédictions à quatre mois (droite), au Burkina Faso.	179
Figure 12.	Corrélation entre les rendements de blé observés au niveau des districts du Queensland et les prévisions <i>a posteriori</i> effectuées à partir du MCG les a) 1 ^{er} mai, b) 1 ^{er} juin, c) 1 ^{er} juillet et d) 1 ^{er} août, entre 1975 et 1993, ajustée aux tendances technologiques de 2001.	180
CHAPITRE 7		189
Figure 1.	Répartition des unités d'échantillonnage de l'enquête TREES-II	197
Figure 2.	Répartition des unités d'échantillonnage de l'évaluation des ressources forestières mondiales 2000	197
Figure 3.	Répartition des unités d'échantillonnage de l'enquête mondiale par télédétection sur les forêts tropicales du FRA 2010	198
Figure 4.	Répartition des unités d'échantillonnage paysagères de l'IFN du Brésil dans l'État du Paraná	206
Figure 5.	Illustration du taux de sondage de la première évaluation mondiale des zones arides	209
CHAPITRE 8		219
Figure 1.	Répartition des coûts d'un système de suivi agricole faisant appel à la télédétection ou à un SIG.	226

Recommandations de citations

MANUEL :

Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS). 2017. *Manuel sur l'application de la télédétection aux statistiques agricoles*. Manuel de la Stratégie mondiale : Rome.

CHAPITRES :

Lemoine, G. 2017. « Logiciels d'analyse et accès aux données ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 1). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Defourny, P. 2017. « Cartographie et suivi de la couverture du sol ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 2). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Gallego, J. 2017. « La télédétection au service de la conception de bases de sondage ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 3). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Davidson, A.M., Fisette, T., McNairn, H. et Daneshfar, B. 2017. « Réalisation d'une cartographie détaillée des cultures à partir de données de télédétection (couches de données sur les cultures) ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 4). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Ray, S. et Neetu. 2017. « Estimation des surfaces cultivées à partir des données de télédétection ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 5). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Rojas, O. 2017. « Mécanismes d'alerte précoce et estimation des rendements agricoles ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 6). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Achard, F., Malheiros de Oliveira, Y.M. et Mollicone, D. 2017. « Suivi du couvert forestier et de la déforestation ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 7). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Latham, J. 2017. « Organisations, ressources et compétences nécessaires à l'adoption de la télédétection dans le domaine des statistiques agricoles ». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 8). Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Delincé, J. 2017. « Rentabilité de la télédétection aux fins de statistiques agricoles ». In : Delincé, J. (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 9), Manuel de la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS) : Rome.

Remerciements

Ce Manuel sur l'application de la télédétection aux statistiques agricoles a été rédigé sous l'égide de Jacques Delincé, consultant pour le bureau de la Stratégie mondiale, par une équipe d'experts internationaux de haut niveau dotés de connaissances approfondies et de plusieurs dizaines d'années d'expérience dans le domaine de la télédétection appliquée à différentes régions du monde. Les auteurs sont affiliés aux institutions suivantes :

Frédéric Achard, Javier Gallego, Guido Lemoine : Centre commun de recherche, Commission européenne, Ispra, Italie

Pierre Defourny : Université Catholique de Louvain-la-Neuve, Belgique

Andrew Davidson, Thierry Fisette, Heather McNairn et Bahram Daneshfar : Agriculture et Agroalimentaire Canada, Ottawa, Ontario

Jacques Delincé, Danilo Mollicone et John Latham : FAO, Rome, Italie

Shibendu Ray et Neetu : Centre national Mahalanobis de prévision des récoltes, New Delhi, Inde

Oscar Rojas : FAO-SLM, Panama

Ce Manuel a été révisé par un groupe de travail composé de spécialistes qui s'est réuni à Rome, le 24 novembre 2016. Nous tenons à remercier tout particulièrement les participants à cette réunion – Giulia Conchedda (FAO, Italie), Carlos di Bella (INTA, Argentine), Nasreen Khan (IRRI, Philippines), Andries Potgieter (Université du Queensland, Australie) et Pierre Sibiri Traore (ICRISAT, Mali) – pour leurs commentaires constructifs et détaillés, dont il a été tenu compte dans la version finale de cette publication.

Celle-ci a été élaborée grâce à l'appui du Fonds d'affectation spéciale de la Stratégie mondiale, financé par le ministère britannique du Développement international (DfID) et la Fondation Bill & Melinda Gates ; et au soutien financier et technique de la Banque mondiale et du Centre commun de recherche de l'Union européenne.

Finalement, soulignons le travail remarquable des personnes impliquées dans la traduction et mise en page de ces lignes directrices. La traduction française de ce document a été effectuée par Angeline Hadman avec la collaboration de Gwendoline Baillet, Rita Douchamps et Jeanne Vandewattyne et révisée par Amsata Niang et Ankouvi Nayo. Mise en page et mise en texte sous la direction de Francesca Zoppi.

Synthèse

Jacques Delincé

Le présent *Manuel sur l'application de la télédétection aux statistiques agricoles* vise à fournir des directives concernant l'utilisation de la télédétection dans le cadre des statistiques agricoles. Le potentiel de la télédétection en vue de l'amélioration de ce type de statistiques est reconnu depuis le milieu des années 1970. Cette technique a été utilisée de différentes façons sur tous les continents et plusieurs de ces approches semblent aujourd'hui avoir atteint un degré de maturité suffisant pour contribuer à la pérennisation des statistiques agricoles. Selon la Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (ci-après Stratégie mondiale ou GSARS, de l'anglais *Global Strategy to improve Agricultural and Rural Statistics*; voir Banque mondiale, 2011), la télédétection contribue de manière essentielle à la localisation et au géocodage des unités d'échantillonnage, et sert de point de référence aux cadres directeurs d'échantillonnage (CDE). Elle constitue également un progrès méthodologique en matière de conception et d'estimation, un moyen d'assurer la durabilité, et un corpus de données essentiel pour les indicateurs relatifs aux différentes utilisations des terres et couvertures du sol.

La télédétection fait partie des domaines thématiques de la division de la recherche de la Stratégie mondiale et a, dans ce cadre, fait l'objet de cinq publications :

- *Developing More Efficient and Accurate Methods for the Use of Remote Sensing* (2014)
- *Technical Report on Improving the Use of GPS, GIS and Remote Sensing in Setting Up Master Sampling Frames* (2014)
- *Spatial Disaggregation and Small-Area Estimation Methods for Agricultural Surveys: Solutions and Perspectives* (2015)
- *Technical Report on Cost-Effectiveness of using Remote Sensing for Agricultural Statistics in Developing and Emerging Economies* (2015)
- *Information on Land in the Context of Agricultural Statistics* (2016)

Dans son *Manuel sur l'infrastructure géospatiale à l'appui des activités de recensement*, paru en français en 2010, la Division de la statistique des Nations Unies recommande pour la première fois l'utilisation de l'imagerie satellite et aérienne aux fins de planification, d'exécution et de diffusion des activités de recensement. En 2010, la communauté de pratique du Système mondial des systèmes d'observation de la Terre (GEOSS, de l'anglais *Global Earth Observation System of Systems*) a publié un guide des bonnes pratiques relatives à l'estimation des surfaces cultivées au moyen de la télédétection. Enfin, depuis 2014, le Système d'information sur les marchés agricoles (AMIS, de l'anglais *Agricultural Market Information System*), une initiative hébergée par l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), publie les analyses mensuelles réalisées par l'Initiative de suivi satellitaire de l'agriculture mondiale du Groupe pour l'observation terrestre (GEOGLAM, de l'anglais *Group on Earth Observations GLobal Agricultural Monitoring*), dans le but de suivre en temps réel les surfaces et la production des principales matières premières agricoles.

Les diverses publications traitant de l'application de la télédétection aux statistiques officielles (FAO, 2015) restent très largement axées sur certaines difficultés techniques spécifiques et ne proposent aucune piste pour intégrer la télédétection dans le processus de compilation des statistiques officielles.

L'objectif du présent manuel est de permettre aux lecteurs intéressés de vérifier si la télédétection correspond à leurs besoins et si son adoption les aiderait à améliorer durablement la rapidité, la couverture, la précision et/ou les coûts.

Conformément aux trois volets de la Stratégie mondiale (Banque mondiale, 2011), les services de statistique agricole devraient actuellement se fixer les priorités suivantes : 1) mise en place d'un cadre directeur pour favoriser l'intégration des statistiques agricoles dans les systèmes statistiques nationaux ; 2) amélioration de la couverture, du biais et de la précision des estimations relatives aux indicateurs clés ; 3) choix de pratiques durables en matière de rapport coût-efficacité, d'adaptabilité et d'accessibilité. Les progrès techniques réalisés en matière de gestion numérique de l'information, le positionnement universel et la liberté d'accès aux technologies de télédétection, qui sont eux-mêmes à l'origine de la préparation et de la publication du présent manuel, offrent une excellente occasion de répondre à ces priorités.

La structure du présent manuel reflète la diversité et la complexité de la statistique agricole, ainsi que les aspects techniques de la télédétection :

- Les systèmes de statistique agricole s'échelonnent sur plusieurs niveaux, correspondant chacun à différents domaines statistiques et besoins sociétaux de premier plan. La télédétection se révèle particulièrement utile pour améliorer les domaines de la Stratégie mondiale relatifs aux surfaces cultivées, aux rendements et à la production. Elle permet de répondre à des besoins très divers, entre autres l'optimisation des modèles d'échantillonnage et la rationalisation du travail de terrain des agents recenseurs, de l'assurance qualité, voire de la production des données. Les besoins sociétaux peuvent être répartis en deux composantes : 1) les prévisions de production (du début de saison jusqu'à la période de prérécolte) aux fins de suivi de la sécurité alimentaire et 2) les statistiques agricoles classiques, dont la continuité et la cohérence sur le long terme permettront aux décideurs de planifier et d'évaluer les politiques agricoles et leurs retombées positives sur la productivité totale des facteurs, le revenu des agriculteurs et le développement rural.
- Les techniques associées à la télédétection suscitent des interrogations concernant les capteurs (capteurs optiques ou radars), la résolution des images (de 30 cm à 5 m) et le temps de revisite (entre 1 heure et 16 jours), mais également l'accès libre (ou non) et les tarifs pratiqués (généralement prohibitifs), ainsi que le matériel et les logiciels d'analyse d'images disponibles (logiciels libres ou payants, systèmes informatiques locaux ou sur hébergés à distance). Par conséquent, les gestionnaires devront évaluer le temps, les moyens et les compétences humaines nécessaires pour passer de la phase d'expérimentation à la phase d'exploitation.

Le **chapitre 1** fournit une description des corpus de données utiles à l'intégration de la télédétection dans les statistiques agricoles. La première partie porte sur les sources des données de télédétection. Le rôle des niveaux des données de référence et auxiliaires, ainsi que leur utilisation aux fins de stratification et d'agrégation, sont également abordés. Par ailleurs, l'essor du libre accès est l'une des grandes questions récentes en matière d'accès aux données. C'est pourquoi ce chapitre propose, outre une liste de référence des grands programmes de télédétection commerciaux, une étude des sources en libre accès, afin de : 1) mieux faire connaître les corpus de données accessibles librement et gratuitement ; 2) fournir une analyse détaillée des méthodes d'évaluation de ces données ; 3) réfléchir aux moyens d'optimiser l'acquisition (c'est-à-dire de minimiser les coûts) des données payantes par l'utilisation de données en libre accès. Les données de télédétection, qu'elles proviennent de sources libres ou payantes, doivent généralement faire l'objet de vérifications et corrections supplémentaires avant d'être utilisées dans les analyses statistiques aux fins de suivi. L'utilisation de logiciels libres de traitement des images et d'analyse géospatiale constitue une autre évolution importante, car elle permet de généraliser plus rapidement le recours aux données de télédétection. Ce chapitre propose une analyse du recours aux logiciels libres, ainsi qu'une liste des logiciels payants. La dernière partie du chapitre analyse la tendance récente consistant à transférer l'analyse des données sur des plateformes dématérialisées.

Le **chapitre 2** porte sur la cartographie de la couverture du sol. Il introduit le concept de couverture du sol avant de passer en revue plusieurs aspects clés de la cartographie dans ce domaine. Les cartes de couverture du sol existantes sont systématiquement étudiées à partir d'un ensemble de critères ciblés. Bien que les cartes favorisant la stratification portent toujours sur les années passées, de nouvelles productions cartographiques menées pendant la

saison en cours sont également étudiées dans ce chapitre. De nos jours, les surfaces terrestres peuvent être décrites de plusieurs manières, grâce aux progrès inédits réalisés dans les domaines des technologies de l'information et des capacités d'observation, à l'instar des véhicules aériens sans pilote (UAV, de l'anglais *Unmanned Aerial Vehicle*) ou des satellites d'observation de la Terre. La télédétection par satellite est une source de données incontestée sur les sols pour un large éventail d'utilisateurs à toutes les échelles géographiques. Au vu de l'écart croissant entre les fabricants d'outils de télédétection et les utilisateurs des cartes, favorisé en grande partie par la mise à disposition généralisée d'informations géographiques par les programmes de données spatiales, il est essentiel de comprendre les différents concepts et contraintes liés à ce type de cartographie, notamment lorsque celle-ci vise à améliorer la stratification au niveau du plan d'échantillonnage, dans le cadre des statistiques agricoles. En effet, les cartes obtenues à partir des données de télédétection et indiquant, par exemple, les niveaux d'intensité des cultures, peuvent contribuer à réduire sensiblement les variances d'échantillonnage ou, simplement, alléger le travail d'échantillonnage sur le terrain et les coûts associés. Les cartes de couverture du sol peuvent mettre en évidence les strates non agricoles à exclure des échantillons ou celles pouvant être échantillonnées différemment. L'efficacité de la stratification dépend évidemment de l'adéquation de la carte de couverture du sol utilisée à cette fin.

Le **chapitre 3** porte sur l'utilisation de la télédétection au stade de la conception dans les bases de sondage de type liste et les bases aréolaires. L'imagerie satellite peut s'avérer précieuse pour définir et optimiser les solutions de conception, dans le cadre des recensements, des enquêtes ou des registres. Elle peut être essentielle en cas d'absence ou d'obsolescence des cartes de référence, car elle permet de délimiter clairement le district de recensement (DR), de dénombrer les habitations et de planifier les tâches. En ce qui concerne les enquêtes, la stratification à partir d'images classées permet de réduire la variance d'échantillonnage et de varier le taux de sondage (ou la probabilité proportionnelle à la taille [PPT]) en fonction de l'intensité agricole. Une attention particulière est accordée à la création de listes d'unités statistiques, à commencer par la base aréolaire par points. Concernant les bases aréolaires, si, dans le cadre de la stratification, les strates doivent être aussi éloignées les unes des autres que possible, les unités primaires d'échantillonnage (UPE) doivent quant à elles être aussi proches que possible dans le cadre des échantillonnages à deux degrés. L'imagerie satellite se révèle extrêmement utile dans les deux cas. Le chapitre 3 fournit plusieurs exemples pratiques tirés des pays développés et en développement, afin d'illustrer le degré d'efficacité et d'homogénéité pouvant être atteint. Il émet par ailleurs des recommandations quant à l'optimisation de la taille des segments, en fonction de la complexité de l'agencement des parcelles.

Le **chapitre 4** vise globalement à fournir une vue d'ensemble des approches en matière de télédétection aux fins de cartographie nationale annuelle détaillée (au niveau des parcelles) des cultures. La première partie présente les méthodes existantes de cartographie par télédétection des terres cultivées, avec notamment un bref aperçu de la classification supervisée des images et de la classification fondée sur les pixels et les objets. La seconde partie examine les différents types de données satellite, de données de terrain et de données secondaires servant à la cartographie détaillée des terres cultivées. La troisième partie décrit la mise en œuvre opérationnelle d'un programme national de cartographie des cultures en s'appuyant sur des exemples spécifiques tirés de l'inventaire annuel des cultures du Canada. La dernière partie présente les principales difficultés et perspectives auxquelles s'attendre à l'avenir en matière de cartographie nationale des types de cultures. Au cours de la dernière décennie, plusieurs projets ont tenté d'exposer le besoin de données de télédétection spatialement explicites pour cartographier les systèmes de culture, en particulier le lieu, le moment, la fréquence, la raie spectrale et la résolution spatiale des relevés. La sélection des meilleures données et méthodologies aux fins de cartographie des cultures reste l'une des grandes priorités de la recherche internationale. Plusieurs initiatives internationales ont d'ailleurs tenté d'harmoniser les approches et de mettre au point des protocoles de suivi et de compte rendu, ainsi que des bonnes pratiques, pour divers systèmes agricoles mondiaux (p. ex., l'initiative GEOGLAM, qui comprend l'Expérience conjointe sur l'évaluation et le suivi des récoltes [JECAM, de l'anglais *Joint Experiment of Crop Assessment and Monitoring*], l'Initiative pour l'estimation et le suivi des cultures rizicoles en Asie [Asia-RiCE, de l'anglais *Asian Rice Crop Estimation and Monitoring*], le projet Stimuler l'innovation pour la surveillance mondiale de l'agriculture [SIGMA, de l'anglais

Stimulating Innovation for Global Monitoring of Agriculture] et des contributions du projet Sentinel-2 pour le suivi de l'agriculture [Sen2-Agri]).

Le **chapitre 5** porte sur l'estimation des zones cultivées au moyen de la télédétection. Il commence par présenter l'historique de l'estimation des zones cultivées et examine l'évolution de la pratique, depuis l'utilisation des méthodes conventionnelles jusqu'à celle des données satellite, ainsi que les difficultés et complexités associées. Les tout premiers grands programmes d'estimation des zones cultivées faisant appel aux données satellite, tels que LACIE (de l'anglais *Large Area Crop Inventory Experiment*) et AgRISTARS (de l'anglais *Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing*), sont présentés. Les différentes méthodes d'estimation, telles que les bases de sondage aréolaires (BSA), le comptage de pixels et les estimateurs de régression ou de calage, sont décrites et illustrées à l'aide d'exemples. Le chapitre détaille également les grands programmes actuels de télédétection aux fins d'estimation des zones cultivées sur trois niveaux : i) national (programme Cropland Data Layer [CDL] du Service national de statistiques agricoles du Département de l'agriculture des États-Unis [USDA-NASS, de l'anglais *United States Department of Agriculture – National Agriculture Statistical Service*] et programme de prévision de la production agricole à l'aide d'observations spatiales, agrométéorologiques et terrestres [FASAL, de l'anglais *Forecasting Agricultural output using Space, Agrometeorology and Land based observations*] de l'Inde); ii) régional (programme de surveillance de l'agriculture par télédétection [MARS, de l'anglais *Monitoring Agricultural with Remote Sensing*] de la Commission européenne); iii) mondial (Service agricole à l'étranger [FAS, de l'anglais *Foreign Agricultural Service*] de l'USDA, CropWatch en Chine et GEOGLAM). La dernière partie de ce chapitre porte sur les principaux obstacles et difficultés liés aux estimations fondées sur la télédétection, et présente la marche à suivre.

Le **chapitre 6** passe en revue les concepts fondamentaux relatifs aux systèmes d'alerte précoces (SAP) et à la prévision des rendements agricoles, afin de mieux gérer les risques climatiques ayant une incidence sur la sécurité alimentaire. Le déclenchement des alertes rapides et la définition de la probabilité du risque à l'échelle du système sont des composantes fondamentales de ce type de système; des liaisons systématiques avec les solutions et les moyens d'alerte précoce permettraient de sauver bon nombre de vies et de moyens d'existence. Il est essentiel que les mécanismes d'alerte précoce visant à évaluer la disponibilité et la demande alimentaires d'une région ou d'un pays donnés puissent prévoir les rendements agricoles et la production globale. L'analyse méthodique des conditions de marché et le repérage des zones de surplus et de déficit d'un pays ou d'une région sont extrêmement utiles pour concevoir des politiques de lutte contre l'insécurité alimentaire adaptées. Des statistiques agricoles fiables et précises sont indispensables pour atteindre ces objectifs importants. Dans ce contexte, les données de télédétection sont essentielles à l'amélioration de la production de statistiques agricoles, car elles permettent d'introduire des mécanismes de vérification indépendants, notamment en cas de recours à des modèles d'échantillonnage utilisant une base aréolaire ou multiple. Les informations et données de télédétection peuvent être introduites au moment de la conception comme de l'estimation.

Le **chapitre 7** porte sur l'estimation du couvert forestier et de la déforestation à l'échelle mondiale et nationale, au moyen de technologies d'observation de la Terre. Il nous a semblé pertinent de réserver un chapitre aux ressources forestières et à la déforestation compte tenu des particularités que présentent les statistiques forestières (permanence des peuplements d'une année sur l'autre, taille des terrains très supérieure à celle des pixels, gestion à long terme, disponibilité des registres de gestion pour les forêts non naturelles). Le chapitre passe en revue les principales méthodologies d'application de la télédétection à l'évaluation et à la modification du couvert forestier, en s'attardant sur les spécificités et les résultats tels que présentés dans la littérature récente. Il analyse ensuite les données de base sur l'application de la télédétection au suivi du couvert forestier. L'enquête par télédétection de l'Évaluation des ressources forestières mondiales de la FAO est décrite, ainsi que d'autres exemples d'enquêtes de télédétection réalisées aux fins de statistiques forestières. La dernière partie analyse la complémentarité entre les estimations des changements du couvert forestier et de l'agriculture.

Le **chapitre 8** présente les exigences et critères fondamentaux visant les organisations qui débutent dans l'analyse géospatiale et, en particulier, la télédétection appliquée aux statistiques agricoles. Il décrit les ressources et les compétences nécessaires à l'utilisation des systèmes de télédétection dans le cadre de collectes de données agricoles et de formations. Les caractéristiques des équipes pluridisciplinaires (ressources humaines, qualifications, taille, budget) sont également considérées.

Des exemples de collaboration entre les services statistiques et les agences de cartographie sont proposés, ainsi que des explications sur l'importance d'interagir étroitement avec les parties prenantes.

Les budgets et les plans d'activité nécessaires sont également présentés.

Enfin, le **chapitre 9** présente la méthode d'évaluation du rapport coût-efficacité de la télédétection. En nous appuyant sur des exemples d'utilisations antérieures et actuelles, nous montrons pourquoi et à quel niveau il est possible d'améliorer nettement le rapport coût-efficacité. La tendance actuelle en faveur de l'accès libre et gratuit à l'imagerie satellite et l'enrichissement du contenu informatif (spectral, spatial ou textural) devraient bientôt permettre de maîtriser la complexité agricole.



Chapitre 1

Logiciels d'accès et d'analyse et des données

1.1. INTRODUCTION

Ce chapitre fournit une description des jeux de données utiles à l'intégration de la télédétection dans les statistiques agricoles. La première partie porte sur les sources des données de télédétection. Le rôle des données de référence et des couches de données auxiliaires, ainsi que leur utilisation aux fins de stratification et d'agrégation, sont également abordés. Par ailleurs, l'essor du libre accès est l'une des grandes questions récentes en matière d'accès aux données. C'est pourquoi ce chapitre propose, outre une liste des grands programmes de télédétection commerciaux, une étude des sources en libre accès, afin de : 1) mieux faire connaître les jeux de données accessibles librement et gratuitement ; 2) favoriser la compréhension du processus d'accès à ces données ; 3) réfléchir aux moyens d'optimiser l'acquisition (c'est-à-dire de minimiser les coûts associés) des données payantes par l'utilisation de données en libre accès. Les données de télédétection, qu'elles proviennent de sources libres ou payantes, doivent généralement faire l'objet de prétraitement avant d'être utilisées dans les analyses statistiques aux fins de suivi. L'utilisation de logiciels libres de traitement des images et d'analyse géospatiale constitue une autre évolution importante, car elle permet de généraliser plus rapidement le recours aux données de télédétection. Ce chapitre propose une analyse du recours aux logiciels libres, ainsi qu'une liste des logiciels payants. La dernière partie du chapitre analyse la tendance récente consistant à utiliser un réseau de serveurs distants hébergés sur Internet pour stocker, gérer et traiter des données, plutôt qu'un serveur local ou un ordinateur personnel.

1.2. CAPTEURS ET JEUX DE DONNÉES

Les données de télédétection utilisées dans le cadre des statistiques agricoles peuvent provenir d'un grand nombre de capteurs différents. Les sources principales d'information étaient traditionnellement les capteurs satellitaires et les instruments aéroportés. Ces derniers sont généralement utilisés pour créer des couvertures d'ortho-imagerie à très fine résolution spatiale (supérieure à 1 m) à des fins de réalisation de cartes topographiques, ou pour

établir des titres fonciers ou des cadastres ruraux. L'essor rapide de la technologie numérique n'a pas épargné le domaine de la production d'ortho-images aériennes, ce qui a permis d'améliorer la qualité et la vitesse de production d'ortho-images par les véhicules aériens dotés de caméras stéréoscopiques. Il s'agit souvent d'une alternative rentable à l'imagerie satellite haute résolution pour les grandes superficies. Aux États-Unis et dans de nombreux pays européens, l'ortho-imagerie aérienne est souvent accessible dans le cadre de licences libres, ou sert de « couche d'arrière plan » aux logiciels de géovisualisation (tels que Google Earth, Google Maps ou Bing Maps) ou de Systèmes d'Information Géographique (SIG) [tels qu'ArcMap ou Quantum GIS] populaires. Ces couches peuvent aider à la numérisation manuelle, afin, par exemple, de délimiter les parcelles agricoles ou les infrastructures. L'accès libre aux tuiles d'images à proprement parler facilite les usages plus complexes, tels que l'orthorectification automatique d'autres jeux d'images.

Les instruments aéroportés sont également utilisés à des fins de prospection spécialisée, par exemple en association avec des capteurs hyperspectraux dans le cadre d'analyses détaillées de la signature spectrale, avec des technologies LIDAR (détection et télémétrie par ondes lumineuses) pour la modélisation de terrains ou avec des radars à synthèse d'ouverture (SAR, de l'anglais *Synthetic Aperture Radar*) ou d'autres techniques de détection électromagnétique. Les plateformes aéroportées présentent l'avantage évident d'être facilement configurables et déployables sur demande. Cette constatation semble encore plus évidente avec le recours croissant aux véhicules aériens sans pilote¹ (UAV), très modulables, qui utilisent des capteurs légers à des fins de collecte de données spécifiques à très haute résolution. Le recours aux UAV s'avère toutefois plus difficile sur les superficies étendues. Les règles de sécurité (p. ex., obligation de faire appel à un opérateur certifié pour les opérations relatives aux lignes de visée ou obligations relatives à l'altitude maximale de vol) peuvent empêcher l'utilisation de ces plateformes en dehors d'applications spécialisées à forte valeur ajoutée, comme l'agriculture de précision. Les parties suivantes de ce chapitre mettent l'accent sur les options de la télédétection spatiale. Il est cependant essentiel, avant la conception d'une enquête statistique agricole, de dresser l'inventaire des outils d'ortho-imagerie aéroportés et des capacités de collecte de données par levés aériens ou par UAV pouvant être utilisés dans la région concernée.

Le nombre croissant de capteurs satellitaires disponibles destinés à l'observation de la terre rend de plus en plus difficile l'identification de la disponibilité optimale dans le cadre de la conception d'enquêtes statistiques agricoles. Ce chapitre présente les capteurs satellitaires ayant démontré leur potentiel dans le cadre d'exercices de classification de l'utilisation des terres et de la couverture des sols, de délimitation des parcelles et de caractérisation des cultures. L'Agence spatiale européenne (ESA, de l'anglais *European Space Agency*) a créé une catégorisation utile des jeux de données provenant de capteurs, qui regroupe les missions contribuant au programme Copernicus². Lesdits capteurs sont classés en groupes de mission, eux-mêmes définis par type de mission et résolution spatiale. Le tableau 1 présente les groupes de mission, ainsi qu'une typologie des résolutions utilisées et des exemples de missions actuellement menées³. La dernière colonne liste les avantages et les inconvénients de chaque catégorie de mission pour les statistiques agricoles.

En ce qui concerne les caractéristiques des capteurs, il est possible de procéder à une première distinction entre les capteurs fonctionnant 1) dans la région spectrale visible et infrarouge (multibande) et ceux fonctionnant (2) dans le spectre des hyperfréquences (SAR). Cette distinction permet de déterminer s'il est possible d'obtenir des images indépendamment du degré d'ensoleillement et de la couverture nuageuse. Les capteurs multibandes mesurent la lumière solaire réfléchie, et sont donc clairement dépendants d'un ensoleillement incident suffisant et de l'absence de nébulosité. Ce point est important, non seulement pour l'obtention instantanée d'images, mais également pour la création de séries temporelles cohérentes, par exemple pour comparer les données de plusieurs campagnes agricoles. Les SAR mesurent les rayonnements rétrodiffusés à l'aide d'une impulsion hyperfréquence qu'ils

1 Le terme équivalent « systèmes d'aéronefs pilotés à distance » est également souvent utilisé.

2 <https://spacedata.copernicus.eu/web/cscda/data-offer/mission-groups>.

3 Pour une liste exhaustive et actualisée des capteurs satellitaires opérationnels, voir : <https://directory.eoportal.org/web/eoportal/satellite-missions> et <http://database.eohandbook.com/database/missiontable.aspx>.

émettent eux-mêmes, indépendamment de tout ensoleillement. Pour les fréquences plus basses (bandes C et L), les SAR sont généralement insensibles aux conditions atmosphériques, sauf en cas de très fortes pluies. Ils présentent ainsi l'avantage d'être capables de garantir la cohérence des séries temporelles utilisées aux fins de délimitation de parcelles et de mesure des aires. Cela est particulièrement vrai pour Sentinel-1A et 1B, qui associent une résolution spatiale de 10 m et une fréquence de revisite de six jours.

La deuxième subdivision concernant les groupes de capteurs dépend de la résolution spatiale. La dénomination de l'ESA repose sur les plages de résolution pratiques < 4 m, 4-10 m, 10-30 m, 30-300 m et > 300 m, qui servent à regrouper les capteurs par familles. Outre la différence de résolution, la plage 10–30 m marque également la limite entre les données provenant de capteurs (essentiellement) accessibles librement et gratuitement, et les catégories (< 10 m) relevant du domaine commercial.

TABLEAU 1. CLASSIFICATION DES CATÉGORIES DE CAPTEURS SATELLITAIRES INSPIRÉE DE LA NOMENCLATURE UTILISÉE PAR L'AGENCE SPATIALE EUROPÉENNE DANS LE CADRE DU PROGRAMME COPERNICUS.

Groupe de mission	Résolution	Exemples	Utilité pour les statistiques agricoles
SAR	Haute, 4-30 m	Sentinel-1 Radarsat-2 ALOS-PALSAR 2 RISAT	Revisite de 6 à 24 jours, bande C (S1, R2) et bande L (ALOS) Toutes conditions météorologiques, imagerie jour et nuit, dont interférométrie Gratuit (S1) et payant (R2, ALOS, RISAT) Contribuent à la délimitation des parcelles et à l'identification des cultures (résolution < 20 m)
	Très haute, 1-4 m	TerraSAR-X CosmoSkyMed Radarsat-2 (mode fin)	Capteurs multiples, revisite variable (pouvant aller jusqu'à une fréquence quotidienne) Haute résolution mais fauchée réduite La bande X (TSX, CSM) est moins adaptée à l'identification des cultures. Tous en mesure d'obtenir exclusivement des images à haute et faible résolution Payants, chers.
Multibande	Haute, 4-30 m	Landsat, Sentinel-2 ASTER SPOT 6/7 RapidEye CBERS IRS LISS DMC	Accès libre et gratuit aux 10-30 m (Landsat, S2, ASTER), à large fauchée, comprenant des bandes infrarouges à ondes courtes (SWIR, de l'anglais <i>Short-Wave Infrared</i>) Revisite de 5 à 16 jours ou moins (p. ex., RapidEye) SPOT combine mode multibande (6 m) et panchromatique (1,5 m) Adapté à la délimitation des parcelles, à l'identification des cultures et à l'estimation des surfaces (résolution de 4-10 m), présence et situation des cultures à l'échelle régionale (4-30 m)
	Très haute, 0,3-4 m	WorldView-3 Pleiades PlanetLabs SkySAT DMC-III	Résolution la plus forte (0,3 m WV-3), nombre variable de bandes spectrales, dans les domaines visible et proche infrarouge (NIR, de l'anglais <i>near-infrared</i>) d'une part, et SWIR d'autre part L'arrivée de nouveaux prestataires (p. ex., PlanetLabs, SkySAT) fait chuter les prix (très élevés à l'heure actuelle). Adapté à la mesure des parcelles (résolution < 1 m)
	Moyenne, 30-300 m	Proba-V Sentinel-3 MODIS VIIRS	Revisite de 1 à 3 jours sur une large fauchée (> 500 km) Voie thermique (MODIS, VIIRS, S3) Composites permettant l'analyse des tendances périodiques sur de grandes superficies Perte fréquente des détails relatifs aux parcelles Utile pour la catégorisation des vastes zones de production et tendances saisonnières, aux fins de stratification Accès libre et gratuit, à l'exception de Proba-V (résolution de 100 m)
	Faible, > 300 m	MODIS AVHRR	À orbite (MODIS) et géostationnaire (METOP-2 AVHRR) Résolution principalement > 1 km Revisite entre 15 min et un jour Utilisés pour dresser des bilans radiatifs et obtenir des informations météorologiques

La résolution spatiale d'un capteur détermine, dans une large mesure, son utilité pour les statistiques agricoles. Les principaux facteurs à prendre en compte sont la taille nominale des unités de production (parcelles de terres cultivées et zones de pacage) et les particularités topographiques cartographiables (couverture arborée, cultures mixtes et éléments valables relatifs à l'utilisation des terres) qui jouent un rôle dans les estimations statistiques. En règle générale, il est plus facile d'obtenir des images à haute résolution des petites superficies, en raison à la fois des capacités techniques des capteurs à acquérir des images en temps utile et des coûts attenants moins élevés. La plupart des capteurs THR couvrent une séquence de 10 x 10 km², qui représente une « unité d'acquisition » typique ; en revanche, les instruments plus versatiles, tels que WorldView et Pleiades, sont capables d'acquérir plusieurs tuiles adjacentes lors de chaque trajectoire orbitale (en fonction de la couverture nuageuse). Les capteurs à faible résolution englobent des superficies (ou fauchées) beaucoup plus vastes à chaque trajectoire. La fauchée de Sentinel-2A, par exemple, couvre une zone de 285 km de large ; le satellite collecte quatre bandes spectrales avec une résolution de 10 m, six avec une résolution de 20 m et trois avec une résolution de 60 m (à des fins de correction atmosphérique). Ainsi, les séquences ultérieures prises dans le cadre d'une trajectoire orbitale donnée peuvent couvrir une part importante d'un territoire national. La fauchée des capteurs à faible résolution est encore plus importante. Celle de MODIS peut couvrir jusqu'à 2 100 km, par exemple. Toutefois, ces images ne peuvent évidemment pas capter les détails topographiques pour délimiter les zones de production.

La résolution radiométrique donne lieu à une autre subdivision des catégories de capteurs. En ce qui concerne les instruments multispectraux, le nombre et la largeur des bandes spectrales constituent d'importants paramètres. Le nombre de bandes spectrales détermine les parties du spectre radiatif sont photographiées, tandis que leur largeur rend compte de la spécificité de la bande spectrale pour laquelle la luminance énergétique est mesurée. Pour une caractérisation détaillée des cultures, il est généralement préférable d'avoir un grand nombre de bandes (et donc de capteurs hyperspectraux) étroites. Toutefois, la conception des capteurs et les taux de transfert des informations satellite limitent considérablement les possibilités réelles d'acquisition de données. Un grand nombre d'instruments à (très) haute résolution (SPOT 6/7, Pleiades et SkySAT) ont tendance à associer bandes multispectrales à faible résolution dans le visible et proche infrarouge et bandes panchromatiques à haute résolution, afin de favoriser le suréchantillonnage lors de l'étape d'« affinage panchromatique » au cours de l'étape prétraitement. L'imagerie proche infrarouge est essentielle pour faire la distinction entre couverture végétale et sols nus. La réflectance proche infrarouge est d'autant plus élevée que les couverts végétaux sont fermés.

WorldView-3 produit actuellement les images satellites les plus détaillées, grâce à une bande panchromatique dotée d'une résolution de 31 cm, quatre bandes spectrales dans la région visible et proche infrarouge de 1,24 m, huit bandes spectrales dans la région Infrarouge à courte longueur d'ondes (SWIR, de l'anglais Short-Wave Infrared) (3,7 m) et 12 bandes supplémentaires (30 m) pour la correction atmosphérique et la détection de nuages, de neige et de glace. Dans le domaine à haute résolution, une attention particulière est accordée à la compatibilité spectrale entre les capteurs (tels que Landsat et Sentinel-2), afin de faciliter l'étalonnage et la comparaison des images. Sentinel-2 vient toutefois compléter Landsat-8 à l'aide de bandes supplémentaires dans le domaine red-edge (raie spectrale de 570-720 nm) et une bande SWIR supplémentaire. Le red-edge renvoie à une région spectrale très sensible à la réflectance de la chlorophylle dans les tissus végétaux. Elle peut donc être utilisée pour caractériser les canopées partielles, et contribuer à la distinction des types de culture. L'Indice différentiel normalisé de végétation (NDVI, de l'anglais *Normalized Difference Vegetation Index*), souvent mentionné dans la littérature, se définit comme la différence normalisée entre la réflectance de la bande du proche infrarouge (au sommet de la gamme red-edge) et celle des bandes spectrales rouges (en bas de la gamme red-edge). Sentinel-2 est le premier capteur à permettre la réalisation d'une analyse complète de la bande spectrale du red-edge.

Les bandes SWIR sont particulièrement utiles pour caractériser les sols nus et les sols partiellement couverts, la sénescence du couvert végétal et les situations de stress. En association avec les bandes comprises dans le visible, elles peuvent faire la distinction entre les différents types de sols en raison des différences de réflectance

caractérisant la composition minérale et la teneur en carbone. Landsat-8 est le seul à être en mesure d'obtenir des images à haute résolution (90 m, rééchantillonnées à 30 m) de la bande infrarouge thermique.

Pour les SAR, la subdivision des différentes résolutions spatiales équivaut à celle effectuée dans le domaine multispectral. Le terme «résolution spatiale» a un sens différent lorsqu'il s'applique aux SAR; l'étude approfondie de cette distinction sort toutefois du champ de ce manuel. Sur le plan radiométrique, les SAR se caractérisent par la fréquence à laquelle ils opèrent et les polarisations auxquelles ils peuvent transmettre et recevoir des impulsions hyperfréquence. Les SAR satellites peuvent opérer sur la bande X (9,6 GHz, TerraSAR-X, CosmoSkyMed), la bande C (5,6 GHz, Sentinel-1, Radarsat-2) ou la bande L (3,0 GHz, ALOS-PALSAR-2). Ces fréquences correspondent respectivement aux longueurs d'onde de 3,2, 5,8 et 10,2 cm. En principe, plus la longueur d'onde du SAR sera longue, plus le signal pénétrera la couverture végétale (ou la surface de sol sec), et plus le nombre d'informations déduites des signaux rétrodiffusés sera important. C'est la raison pour laquelle la bande L est généralement privilégiée dans l'étude des cultures. Il convient de noter que l'ensemble de données SAR issues de Sentinel-1 est aujourd'hui le plus populaire, en raison de la grande disponibilité des données fréquentes (bande C) dans le cadre de la licence ouverte et gratuite Copernicus. Les SAR émettent des impulsions hyperfréquence cohérentes, dans le cadre desquelles il est possible à la fois de mesurer l'intensité et de déterminer la phase en cours des plans de polarisation transmis et reçus. Les mesures polarimétriques complètes fournissent davantage d'informations sur les signaux rétrodiffusés (des cultures, par exemple); mais leur disponibilité suppose souvent de faire des compromis en matière de résolution. Sentinel-1 mesure généralement l'intensité polarisée des surfaces terrestres sous forme de bandes verticalement transmises-verticalement reçues (VV) à polarisation parallèle, et verticales-horizontales (VH) à polarisation croisée. L'interférométrie se définit comme une technique consistant à combiner les mesures cohérentes effectuées dans le cadre de deux acquisitions de données distinctes par SAR à partir d'un même capteur. La cohérence interférométrique pourrait contribuer à la cartographie des utilisations de la terre; le recours à cette technique dans le cadre de la cartographie agricole est toutefois encore limité. Le traitement des données SAR interférométriques (InSAR) était, jusqu'à récemment, réservé aux spécialistes. Il pourrait toutefois gagner en importance avec la disponibilité croissante des données Sentinel-1 — notamment avec la mise en orbite des satellites Sentinel-1A et Sentinel-1B — et des logiciels en libre accès s'y rapportant (voir la partie 2.2).

Pour tous les capteurs, la résolution temporelle concerne la densité du temps à laquelle il est possible d'associer (de superposer) les couvertures répétitives pour former une composition diachronique. Cette résolution est étroitement liée à la configuration orbitale des satellites et à la possibilité de diriger et de pointer le capteur le long de et à travers la trajectoire orbitale. Les capteurs à très forte résolution spatiale sont en mesure d'acquérir des images à partir d'orbites voisines, d'atteindre une fréquence de revisite quotidienne, voire d'acquérir de multiples données le long de l'orbite (c'est le cas de WorldView et de SkySAT), formant ainsi une séquence semblable à une vidéo (ceci n'est toutefois pas particulièrement utile dans le cadre des statistiques agricoles). La fréquence de revisite des capteurs à vue fixe (p. ex., Landsat et Sentinel-2) est entièrement déterminée par leur orbite (p. ex., respectivement 16 et 12 jours). Les orbites sont généralement déterminées pour assurer une couverture continue de la fauchée du capteur à l'équateur. Par conséquent, les fauchées se chevauchent latéralement, notamment aux latitudes élevées de l'hémisphère nord et méridionales de l'hémisphère sud, ce qui accroît la fréquence de revisite à ces endroits. La manipulation de capteurs identiques dans une constellation d'orbites phasées (Sentinel-1, Sentinel-2, RapidEye et CosmoSkyMed) permet également d'augmenter la résolution temporelle. La nature active des capteurs SAR permet d'augmenter la résolution temporelle en combinant les acquisitions de données des passages ascendants et descendants.

Les catégories d'instruments à faible résolution revêtent un intérêt moindre pour le domaine des statistiques agricoles. Ils contribuent toutefois souvent à l'analyse des tendances phénologiques pluriannuelles sur une superficie étendue (voir le chapitre 6). Parmi les autres catégories de capteurs non recensés dans le tableau 1, citons les missions de radiométrie passive, les capteurs océaniques et les sondeurs atmosphériques. L'apport de ces capteurs aux statistiques agricoles est plus indirect. Ils contribuent par exemple à l'élaboration de modèles climatiques mondiaux qui alimentent les modèles de prévisions météorologiques employés à des fins de modélisation de

l'évaluation agrométéorologique des rendements. Ils dépassent donc le champ de ce chapitre, mais pourront être mentionnés dans d'autres parties de ce manuel.

Il convient de remarquer que la présente analyse se limite aux capteurs de pointe, c'est-à-dire ceux qui sont actuellement opérationnels dans le cadre de licences libres et gratuites (Landsat, Sentinel et ASTER) ou payantes. Il existe de nombreux autres systèmes, qui ne sont pas opérationnels (p. ex., missions de recherche et développement) ou largement et facilement accessibles (p. ex., capteurs de programmes nationaux). En outre, on risque de surestimer le potentiel des technologies « d'avenir » en matière d'observation de la Terre, au détriment souvent des capacités existantes. La tendance générale à faire « plus et mieux » dans ce domaine est toutefois évidente. La miniaturisation et la baisse du coût des composants de l'espace numérique favorisent déjà l'arrivée sur le marché de nouveaux protagonistes industriels (Planet, TerraBella et UrTheCast) et poussent les acteurs existants à envisager l'adoption d'importants parcs de capteurs satellites THR de petite taille capables de procéder à une revisite quotidienne. En ce qui concerne les résolutions haute et moyenne, on espère que les programmes nationaux, de plus en plus nombreux (tels que le programme spatial chinois, la nouvelle constellation Radarsat du Canada ou le programme argentin SAOCOM), adopteront des licences accessibles librement et gratuitement, contribuant ainsi aux capacités mondiales de suivi.

1.2.1. Facteurs de coûts liés à l'imagerie commerciale

L'utilisation de l'imagerie satellite commerciale se heurte à une limite pratique, celle des coûts. Plusieurs facteurs déterminent le coût total de l'exploitation des données. À noter que, pour la plupart des capteurs, cette dernière est sous licence : le client ne devient pas propriétaire des données, mais paie pour avoir le droit de les exploiter et d'en tirer des informations. Le type de licence utilisée peut largement influencer sur les coûts engendrés. Par exemple, les licences à usage unique sont généralement les moins chères, mais le partage de données peut être limité à un seul utilisateur ou un groupe d'utilisateurs préalablement identifiés. Les licences à utilisateurs multiples peuvent faire la distinction entre usage au sein d'une organisation et partage des informations avec des (groupes d') usagers externes, les prix augmentant en conséquence.

Il est également essentiel de déterminer si les travaux issus d'analyses de données — tels que les résultats de classements et les contenus numérisés — sont assujettis aux conditions des licences. Jusqu'à récemment, c'était le cas pour la plupart des capteurs. Une tendance assez alarmante consiste cependant depuis peu à introduire des licences « virales » couvrant les travaux dérivés. Elle concerne notamment les capteurs technologiques les plus avancés (WorldView et Pleiades).

D'autres facteurs déterminent le prix par unité de surface : le fait d'obtenir les données à partir d'archives existantes ne coûte pas le même prix que lorsque celles-ci sont obtenues sur demande. Le temps (urgent et dans une période limitée) de traitement de l'acquisition est également un facteur qui détermine le prix. Les images d'archive sont évidemment les plus abordables, surtout lorsqu'elles proviennent de capteurs dont la « durée de vie » est la plus courte en raison du nombre de prestataires actifs dans cette résolution. Une bonne planification permet d'éviter les demandes de données urgentes, qui coûtent généralement le plus cher. Parmi les autres facteurs de coûts, citons : 1) la qualité technique requise des images en termes de résolution spatiale et spectrale ; 2) les limites en matière d'angle de vue ; 3) la demande de traitement préliminaire ; 4) la couverture totale requise. En ce qui concerne ce dernier point, il est généralement intéressant de déterminer si les images requises doivent faire l'objet d'une couverture contiguë ou peuvent être présentées sous forme unitaire (p. ex., pour obtenir un ensemble d'images indépendantes). À noter également que pour certains capteurs basés sur des scènes, un critère de taille minimale de la surface pourrait s'appliquer.

E-geos⁴ propose une liste de prix catalogue pour un ensemble hétérogène de capteurs THR, ce qui offre un bon aperçu des facteurs susmentionnés et de leurs effets sur les prix estimés. Si les prix au km² varient entre 0,01 (Rapideye NextMap World 30) et 145 dollars US (WorldView-3 «Select Plus Tasking») (soit quatre ordres de grandeur!), l'estimation plus réaliste du prix d'une activité donnée ne sera possible qu'avec une bonne compréhension de toutes les exigences techniques et conditions applicables.

En général, pour les projets nécessitant une imagerie satellite THR (< 4 m), il est conseillé de présenter le scénario souhaité en matière d'acquisition ainsi que les exigences techniques requises à plusieurs prestataires, puis de comparer les prix et de négocier pour obtenir la meilleure offre possible. Cela permettra en effet de bénéficier d'importantes réductions sur les prix catalogue et de conditions spéciales favorisant le partage de données. Dans ces cas, les études de faisabilité proposées dans le cadre des services à la clientèle des opérateurs de satellites sont généralement gratuites. La plupart des opérateurs de capteurs publient leurs fonds d'archives dans des interfaces Web de recherche. Des efforts ont été déployés pour harmoniser l'accès aux informations communes⁵. Les courtiers et revendeurs indépendants de données proposent des services (payants) d'aide à la sélection des offres d'acquisition de données les «plus économiques». L'offre relative au segment de la résolution 1-4 m est de plus en plus saturée, ce qui devrait entraîner une amélioration des services à la clientèle et une multiplication des réductions. Pour les couvertures contiguës sur de grandes surfaces, il convient d'envisager également de faire appel à un opérateur de capteurs numériques aériens, notamment parce que l'accès à la propriété des données est généralement implicite et que les avantages en matière de coûts/de qualité peuvent être plus intéressants que ceux des opérateurs de satellites. ArcGIS World Imagery⁶ présente les technologies d'imagerie THR disponibles, parmi lesquelles se trouvent des ensembles de données satellites et aériennes mises à jour régulièrement portant sur différentes régions du monde.

1.2.2. Accès libre et gratuit

L'introduction de la licence gratuite et en libre accès couvrant toutes les données du capteur Sentinel du programme Copernicus de l'Union Européenne (UE) représente un complément indispensable aux efforts déployés par les États-Unis pour assurer l'accès libre aux jeux de données d'observation de la Terre fournies par les capteurs du gouvernement tels que MODIS et, surtout, Landsat. Cette transition vers un accès libre aux ressources publiques de données satellite a récemment été encouragée avec l'ouverture de l'accès aux données du capteur nippo-américain ASTER. Si la motivation principale sous-tendant cette évolution est d'établir une base de connaissances communes dans le but de procéder à un suivi environnemental à l'échelle mondiale, l'accès libre a des répercussions considérables sur de nombreuses applications potentielles en matière de statistiques agricoles. Comme nous l'avons vu précédemment, les données fournies par Sentinel-1 et Sentinel-2 permettent de tirer des informations spécifiques sur les cultures à des échelles cartographiques allant de 1:25 000 à 1:50 000, lesquelles servent de fondement aux études de conception d'enquêtes statistiques. Les activités menées dans ce domaine devraient donc envisager de prendre comme socle commun les données Sentinel.

À l'instar des prestataires de services commerciaux, les opérateurs de capteurs en libre accès mettent à disposition des usagers des interfaces Web de recherche⁷ qui permettent de rechercher dans les archives les images acquises et les caractéristiques orbitales pour comprendre les capacités de revisite. L'une des principales différences opposant les opérateurs de données en libre accès et leurs homologues commerciaux tient dans le fait que les archives des premiers peuvent être téléchargées à pleine résolution.

4 <http://www.e-geos.it/products/pdf/prices.pdf>.

5 <https://earth.esa.int/hma/>.

6 <http://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08feb2a9>.

7 <https://scihub.copernicus.eu/dhus/#/home> (pour Sentinel-1 et Sentinel-2 du programme Copernicus) et <http://earthexplorer.usgs.gov/> (pour Landsat et ASTER, entre autres).

Les interfaces Web graphiques de recherche se prêtent bien aux recherches et téléchargements occasionnels. Pour les demandes plus importantes ou récurrentes, il est souvent plus facile de procéder aux sélections et aux téléchargements en s'aidant de scripts. Les interfaces de recherche bien conçues présentent les protocoles sous-jacents sous forme d'ensembles d'interfaces de programmation d'application (API, de l'anglais *Application Programming Interface*)⁸ pouvant être configurés pour effectuer des sélections spécifiques dans le temps et dans l'espace et des choix de métadonnées propres à certaines missions, puis exploités sous forme de scripts automatiques de commande. Ce point est particulièrement utile pour les jeux de données volumineux, tels que les SAR de niveau 1 (Ground Range Detected) Landsat, Sentinel-2 et Sentinel-1, qui peuvent aller jusqu'à 1 Gb par scène, et le produit singulier complexe Sentinel-1 (jusqu'à 8 Gb/scène). Des exemples de scripts en langage Python sont disponibles sur Internet⁹ ou sur simple demande à l'auteur du présent chapitre.

L'accès entièrement libre et gratuit à ces données pousse les tiers à l'intégrer à leurs infrastructures de stockage et de traitement sur un serveur hébergé (cloud server).

Par exemple, Google¹⁰ et Amazon¹¹ téléchargent les images de Sentinel-2 à partir de points d'accès aux données hébergés par l'ESA, puis les mettent gratuitement à disposition de l'ensemble des internautes, notamment sur leurs propres infrastructures d'informatique à distance (en cloud). Le programme Copernicus de l'UE prévoit, en partenariat avec l'ESA, de créer prochainement des infrastructures européennes d'hébergement et de traitement des données. Nous traitons plus en détail les aspects relatifs à l'informatique dématérialisée dans la dernière partie de ce chapitre.

1.3. TRAITEMENT DES IMAGES ET TRAVAUX PRÉPARATOIRES À L'ANALYSE DES DONNÉES

En fonction du type de capteur utilisé et des niveaux de traitement des données, les images satellite (ou aériennes) fournies pourront nécessiter prétraitement à l'analyse des données effectuée ultérieurement, un processus qui a récemment été baptisé « création de « données prêtes à être analysées » ». Ce manuel présente les exigences maximales, puis les associe aux solutions logicielles existantes permettant la mise en œuvre des fonctionnalités requises. Pour mieux comprendre les questions techniques propres à chaque étape, le lecteur est invité à consulter la littérature pertinente.

Au minimum, les images satellite fournies devraient répondre à plusieurs prescriptions techniques de base, telles que la superposition cohérente des différentes bandes pour les données multibandes, la mise à l'échelle et la quantification radiométriques aux fins d'optimisation de la dynamique instrumentale du capteur pour la scène concernée (en d'autres termes, absence de saturation excessive ou de compression de portée) et l'adaptation de la profondeur de bits au type de données (octet, entier [non-] signé et float). Par ailleurs, les données tirées des images devraient s'accompagner de métadonnées fournissant des informations détaillées sur la méthode de projection des images, la géolocalisation (approximative), les niveaux de traitement et les paramètres pouvant servir à traduire les valeurs de la bande image dans les mesures physiques auxquelles elles correspondent, telles que la réflectance ou la radiance dans le cas de l'imagerie multibande, ou l'intensité de la rétrodiffusion dans le cas des SAR. Des métadonnées supplémentaires renseignent généralement sur l'attitude de la plateforme de capteurs (position orbitale et configuration de la visualisation au moment de l'acquisition des données), les paramètres environnementaux extérieurs (degré d'ensoleillement), les paramètres d'étalonnage, et, occasionnellement, les estimations statistiques calculées à partir de l'analyse des images au moment de leur traitement (p. ex., couverture nuageuse et estimation des données manquantes). Il est recommandé de vérifier l'existence de ces conditions préalables au moment de la réception des données, puisque les prestataires proposent généralement de combler les lacunes gratuitement.

⁸ Voir, par exemple, <https://scihub.copernicus.eu/userguide/5APIsAndBatchScripting>.

⁹ <http://www.cesbio.ups-tlse.fr/multitemp/?p=2782> (Landsat) et <http://www.cesbio.ups-tlse.fr/multitemp/?p=6442> (Sentinel-2).

¹⁰ <https://cloud.google.com/storage/docs/public-datasets/sentinel-2>.

¹¹ <http://sentinel-pds.s3-website.eu-central-1.amazonaws.com/>.

Les étapes de prétraitement les plus courantes en matière d'imagerie multibande sont l'orthorectification et la correction atmosphérique. Le recours à l'orthorectification dépend du niveau de traitement des images fournies, et peut englober le réglage des capteurs et la correction topographique. Cette dernière suppose d'utiliser un Modèle Altimétrique Numérique (MAN) externe visant à géoréférencer l'imagerie dans une carte de référence cohérente. La correction topographique est fondamentale pour corriger les images aux angles de vue en périphérie (obliques). La correction atmosphérique sert quant à elle à corriger les effets atmosphériques afin de calculer la réflectance au bas de l'atmosphère (BOA, de l'anglais *bottom of atmosphere* — ou son équivalent TOC, de l'anglais *top of canopy* — réflectance au sommet de la canopée) à partir de la réflectance au sommet de l'atmosphère (TOA, de l'anglais *top of atmosphere*) caractérisant généralement les images fournies. La réflectance BOA (ou TOC) fournit des valeurs cohérentes dans le temps correspondant à la réflectance « véritable » des éléments de pixel composant l'image matricielle. Pour les ensembles de données associant bandes multispectrales et bande panchromatique à haute résolution, il est possible de procéder à un affinage panchromatique afin de créer une « pseudo »-image multibande suréchantillonnée à haute résolution.

Dans le cadre de l'imagerie SAR, les étapes équivalentes sont le géocodage et l'étalonnage. Ce dernier convertit les valeurs de l'intensité des bandes des images SAR en coefficients de rétrodiffusion. Compte tenu de la visée latérale des SAR et de la spécificité de la formation de leurs images dans la direction de l'azimut et de la portée, il convient de procéder à un géocodage spécifique afin de compenser les effets du relief (en ayant recours, ici aussi, à un MAN externe). En fonction de l'usage prévu, un traitement plus poussé impliquant la multivisée et l'application d'un filtre de déchantonnement pourra s'avérer nécessaire.

Le processus présenté ci-dessus donne lieu à des images calibrées et géoréférencées prêtes pour la suite du traitement. En termes techniques, le prétraitement convertit les images fournies des niveaux 1A (« brut ») ou 1B (« 1A corrigé par le système ») aux niveaux géoréférencé (1C) et calibré (2A). Il sera parfois demandé aux prestataires de fournir un niveau 2A, généralement moyennant finance. Autrement, il conviendra de saisir les données auxiliaires applicables dans un logiciel spécialisé (p. ex., un MAN adapté) pour prendre ces mesures. En fonction de l'usage prévu, d'autres étapes de traitement pourront s'avérer nécessaires, comme le mosaïquage des scènes uniques pour obtenir des couvertures plus vastes, la superposition des images prises au fil du temps, la classification, la segmentation, etc. Ces différentes étapes supposent elles aussi de disposer de logiciels appropriés.

Une grande tendance en matière de traitement des images s'est dégagée ces dernières années, à savoir la disponibilité croissante des logiciels libres. Citons les bibliothèques Python, Java et C++ pour la manipulation d'images génériques, les boîtes à outils spécifiques à la télédétection (p. ex., Orfeo ToolBox¹²) et, plus récemment, les boîtes à outils relatives à Sentinel. Il existe par ailleurs des logiciels interactifs d'analyse et de visualisation géospatiale (p. ex., GRASS¹³ et QGIS), des bases de données rastrées (PostgreSQL/Postgis¹⁴) et des serveurs Web publiant des données spatiales (GeoServer, MapServer¹⁵). La Geospatial Data Abstraction Library (GDAL¹⁶), qui s'appuie sur des bibliothèques inférieures telles que PROJ.4 et GEOS, constitue une bibliothèque basique, mais très polyvalente, qui prend en charge la plupart de ces solutions. Citons, parmi leurs équivalents PLC (Produits Logiciels du Commerce), Excelis ENVI/IDL, ERDAS Imagine, PCI, MapInfo et ESRI-ArcGIS, entre autres¹⁷.

Le choix entre logiciel en accès libre et PLC peut dépendre de plusieurs facteurs qui déterminent le « coût de possession », lequel dépasse les simples prix d'achat et coûts d'entretien, et peut englober l'évaluation de la base de connaissances existante, les besoins de formation et d'autres éléments. La balance penche cependant radicalement vers les solutions en libre accès, notamment parce qu'elles intègrent bien plus efficacement les dernières avancées

12 <https://www.orfeo-toolbox.org/>.

13 <http://grass.osgeo.org/>.

14 <http://www.postgis.net/>.

15 <http://www.mapserver.org/>, <http://geoserver.org/>.

16 <http://gdal.globe.org/>.

17 Voir <http://www.un-spider.org/fr/liens-et-ressources/sig-logiciel-teledetection>.

(intégration dans des cadres informatiques parallèles et utilisation de composants matériels spécialisés tels que processeurs graphiques), y compris celles du champ des jeux de données de référence et de données satellite en accès libre. Les logiciels en libre accès les plus populaires bénéficient de la croissance exponentielle d'une base d'utilisateurs qui s'illustrent par leur travail collaboratif de conception et de mise en œuvre de solutions de pointe. De fait, les fabricants de PLC n'aspirent aujourd'hui qu'à rattraper les solutions en libre accès. À cet égard, il convient de noter qu'il existe bien plus de 50 formats d'images et de données vectorielles différents, dont la plupart tirent leurs origines de brevets (plutôt que de la logique) — fait d'autant plus remarquable qu'il n'existe que trois formats dans le domaine pourtant bien plus vaste du traitement de texte (DOC, PDF et ODT).

En ce qui concerne les satellites Sentinel, l'ESA appuie le développement de la Plateforme d'application Sentinel (SNAP, de l'anglais *Sentinel Application Platform*¹⁸), qui fournit une structure commune à toutes les fonctionnalités de traitement spécifiques à Sentinel. Le logiciel SNAP, basé sur la technologie Java, fonctionne avec tous les systèmes d'exploitation. Il supporte également un large éventail de fonctions de traitement standard (arithmétique de canal, analyse d'images vectorielles, reprojection, importation/exportation de nombreux formats), qui peuvent être utilisées en mode interactif ou appliquées à des graphiques de traitement prédéfinis sur la ligne de commande et par lots. La boîte à outils Sentinel-1 (s1tbx) a joué un rôle fondamental dans l'intérêt croissant porté par la communauté d'utilisateurs à l'imagerie SAR fournie par le satellite du même nom ; les SAR ont toujours été considérés comme étant techniquement difficiles à utiliser, notamment parce que les méthodes essentielles de traitement n'étaient disponibles que dans le cadre de solutions propriétaires.

Les boîtes à outils constituent également d'excellents outils pédagogiques. Si l'ESA présente SNAP comme un logiciel en libre accès, certaines fonctionnalités (telles que le module de correction atmosphérique *sen2cor* de Sentinel-2) sont en réalité encore protégées par un code source fermé. Lorsqu'associées à des fonctionnalités de sélection et de téléchargement automatisés des données, les boîtes à outils peuvent être reliées de façon à traiter les données de Sentinel-1 et de Sentinel-2 (ainsi que celles de Sentinel-3) sur demande, de manière autonome (figure 1). Ce modèle fonctionnel peut être développé en domaines d'application thématiques, comme dans le cadre du projet *Sen2Agri*¹⁹, qui vise à publier une chaîne complète de traitement en vue de l'utilisation de Sentinel-2 dans la production de masques des cultures, la classification des types de cultures et l'extraction de l'indice de surface foliaire dans les régions étendues. Son utilité a déjà été démontrée dans le cadre d'essais menés à l'échelle nationale en République tchèque et en Ukraine.

Le petit exemple²⁰ suivant présente un workflow qui recherche les tuiles de données GRD Sentinel-1 se recoupant près de Kura, au Nigéria, au cours des quatre mois précédents. Cette requête donne lieu à un jeu de données sous format XML, qui sont sauvegardées dans le fichier *s1.xml*.

```
wget "https://{USER}:{PASSWORD}@scihub.copernicus.eu/apihub/search?q=productType:GRD \
AND footprint:\\"Intersects(11.769, 8.422)\\" \
AND beginPosition:[NOW-4MONTHS TO NOW]&rows=100" -O s1.xml
```

L'accès à scihub nécessite un nom d'utilisateur et un mot de passe valides, qui peuvent être obtenus après inscription à la plateforme. Les données sont regroupées sous forme de fil d'actualité, dont le nombre d'entrées correspond aux données individuelles propres à chaque image satisfaisant aux critères de sélection. Chaque donnée s'accompagne d'un jeu de métadonnées précisant le format, la date d'acquisition et d'ingestion des données, la taille du fichier, l'empreinte, etc. Les deux éléments de métadonnées les plus pertinents sont l'élément *link*, qui permet d'obtenir l'attribut *href* qui donne accès au lien de téléchargement, et l'élément *str* accompagnant l'attribut *filename*,

18 <http://step.esa.int/main/toolboxes/snap/>.

19 <http://www.esa-sen2agri.org/SitePages/Home.aspx>.

20 Cet exemple part de l'hypothèse que l'auteur de la requête travaille sur Linux. Il existe des commandes équivalentes sur les autres systèmes d'exploitation.

qui produit un nom de fichier unique applicable au fichier téléchargé. Il est recommandé d'utiliser un analyseur²¹ pour lire le format XML, cet outil pouvant extraire les éléments et les attributs pertinents, afin, par exemple, de créer un script de téléchargement.

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/" xmlns="http://www.
w3.org/2005/Atom">
<title>Sentinels Scientific Data Hub search results for: productType:GRD AND
footprint:"Intersects(11.769, 8.422)" AND beginPosition:[NOW-4MONTHS TO NOW]</title>
<subtitle>Displaying 9 results. Request done in 0.177 seconds.</subtitle>
<updated>2016-11-08T14:22:56.407Z</updated>
<author>
<name>Sentinels Scientific Data Hub</name>
</author>
<id>https://scihub.copernicus.eu/apihub/search?q=productType:GRD AND
footprint:"Intersects(11.769, 8.422)" AND beginPosition:[NOW-4MONTHS TO NOW]</id>
<opensearch:totalResults>9</opensearch:totalResults>
<opensearch:startIndex>0</opensearch:startIndex>
<opensearch:itemsPerPage>100</opensearch:itemsPerPage>
...
<entry>
<title>S1A_IW_GRDH_1SDV_20161027T173813_20161027T173838_013679_015EF5_8838</title>
<link href="https://scihub.copernicus.eu/apihub/odata/v1/Products('aa979434-a766-
480d-a269-bala6b2b708a')/$value"/>
...
<summary>Date: 2016-10-27T17:38:13.561Z, Instrument: SAR-C SAR, Mode: VV VH,
Satellite: Sentinel-1, Size: 1.58 GB</summary>
...
<str name="filename">S1A_IW_
GRDH_1SDV_20161027T173813_20161027T173838_013679_015EF5_8838.SAFE</str>
<str name="gmlfootprint"><gml:Polygon srsName="http://www.opengis.net/gml/srs/epsg.
xml#4326" xmlns:gml="http://www.opengis.net/gml">
  <gml:outerBoundaryIs>
    <gml:LinearRing>
      <gml:coordinates>11.997962,7.224397 12.441014,9.497738 10.932872,9.795840
10.485979,7.535322 11.997962,7.224397</gml:coordinates>
    </gml:LinearRing>
  </gml:outerBoundaryIs>
</gml:Polygon></str>
...
</entry>
<entry>
</entry>
...
<entry>
...
</feed>
```

²¹ Par exemple, l'analyseur lxml pour Python, disponible à l'adresse <http://lxml.de>.

Il est possible de télécharger des images de la façon suivante grâce aux attributs href et filename :

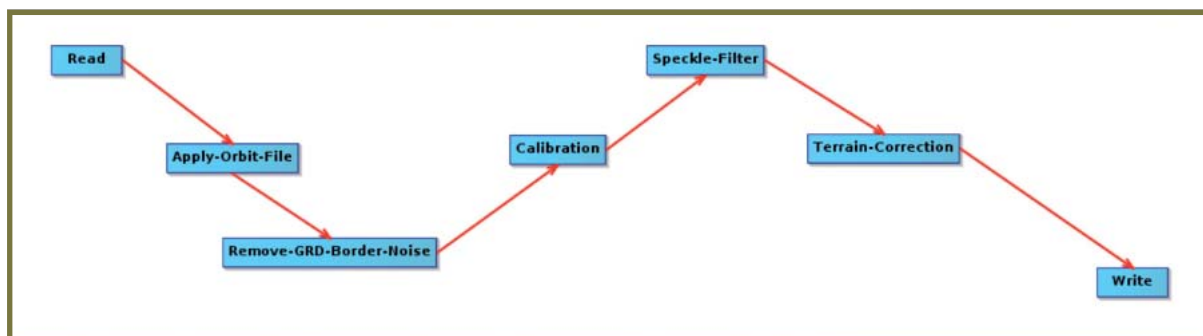
```
wget «https://{USER}:{PASSWORD}@scihub.copernicus.eu/apihub/odata/v1/  
Products('aa979434-a766-480d-a269-bala6b2b708a')/\$value» -O S1A_IW_  
GRDH_1SDV_20161027T173813_20161027T173838_013679_015EF5_8838.zip
```

Cette opération peut être effectuée pour chaque fichier compris dans la série temporelle.

Le fichier GRD Sentinel-1 contient des intensités rétrodiffusées non calibrées organisées selon la géométrie portée-azimut du SAR. Il est possible de recourir au logiciel s1tbx pour créer une chaîne de traitement et ainsi convertir ces informations en coefficients de rétrodiffusion géoréférencés et calibrés. L'une des fonctions très pratiques proposées par s1tbx, qui repose sur la notion de graphiques, permet de dessiner des schémas relatifs aux différentes étapes de traitement sous forme de workflow (voir l'exemple en figure 1). Le graphique peut alors être sauvegardé sous format XML et appliqué, dans le cadre d'un traitement par lot, à toute autre image de la même série temporelle, en remplaçant les paramètres utilisés.

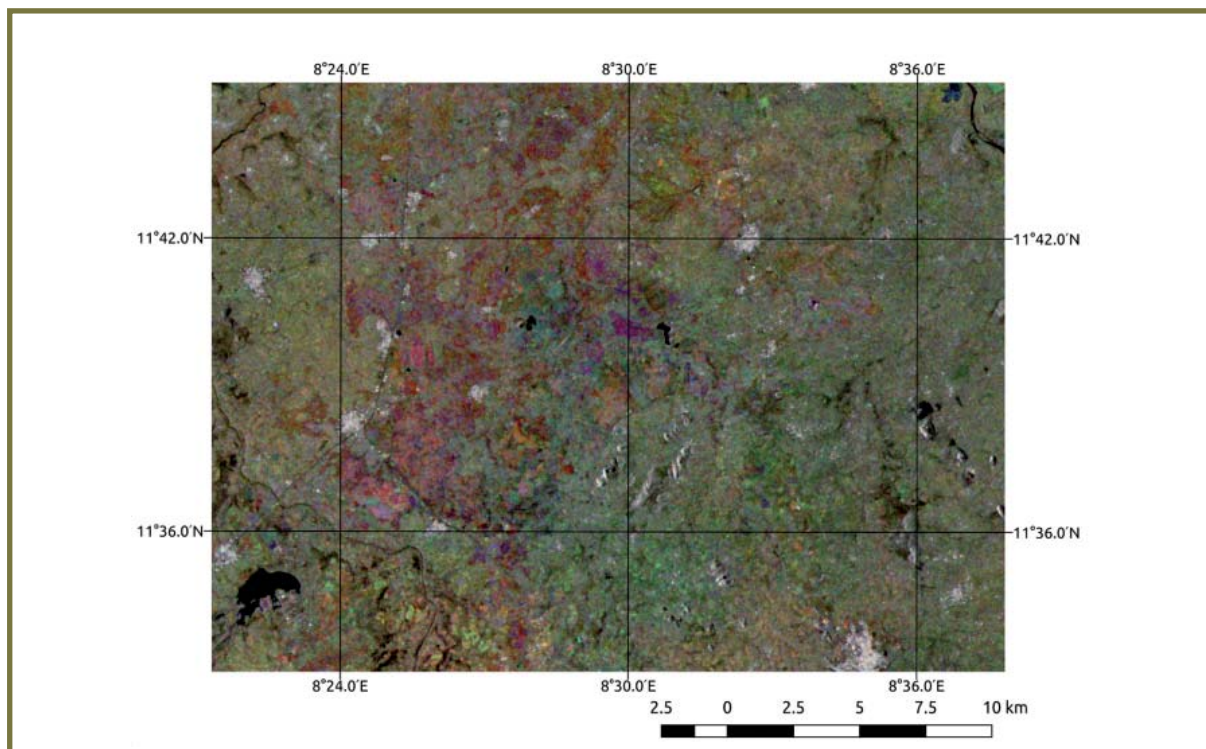
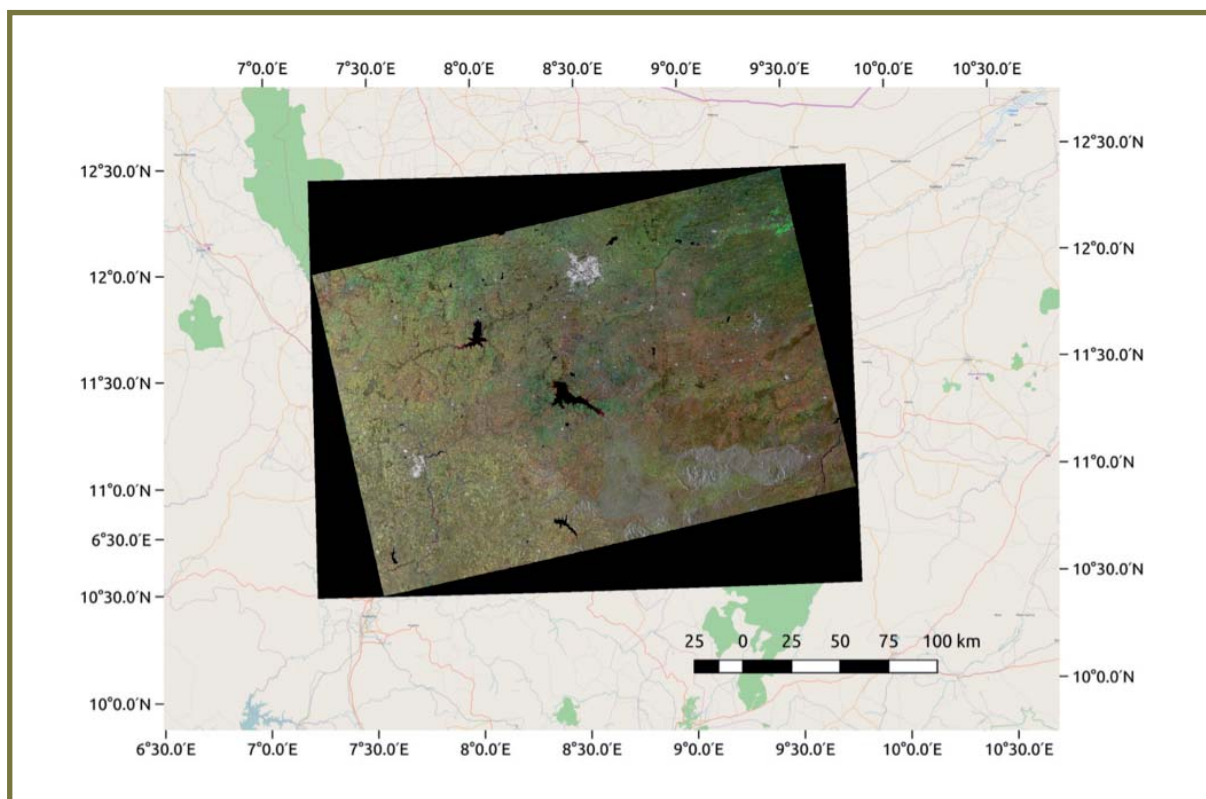
```
/home/user/snap/bin/gpt /home/user/S1A/10m_Calibrated_Geocoded.xml \  
-Pin_file=S1A_IW_GRDH_1SDV_20161027T173813_20161027T173838_013679_015EF5_8838.zip \  
-Pout_file=S1A_IW_GRDH_1SDV_20161027T173813_20161027T173838_013679_015EF5_8838.tif
```

FIGURE 1. SCHÉMA DU WORKFLOW DE TRAITEMENT DES INFORMATIONS SAR DANS S1TBX, VISANT À GÉNÉRER DES COEFFICIENTS DE RÉTRODIFFUSION CALIBRÉS ET GÉOCODÉS À PARTIR DES JEUX DE DONNÉES GRD DE SENTINEL-1



En associant recherche sur le catalogue Sentinel, analyse XML et scriptage d'un workflow automatique en un workflow unique, il est facile de créer des procédures par lot, telles que les tâches crontab, qui téléchargent et préparent des jeux de données propres à un domaine d'intérêt du type cycle de revisite. La figure 2 illustre un ensemble d'images géocodées pour notre sélection, intégrées dans une session QGIS.

FIGURE 2. COMPOSITE DE COEFFICIENTS DE RÉTRODIFFUSION SENTINEL-1 GÉOCODÉS FONDÉ SUR DES IMAGES ASCENDANTES (ORBITE 132 RELATIVE, POLARISATION VV) D'UNE ZONE AGRICOLE PROCHE DE KURA, AU NIGÉRIA, PRISES LES 27 JUILLET, 9 SEPTEMBRE ET 23 OCTOBRE 2016.



Les tons rouges et orange représentent les zones agricoles sujettes à des fluctuations saisonnières en matière de signature de rétrodiffusion. Les données Sentinel-1 sont protégées par les droits d'auteur du programme Copernicus, 2016.

1.4. DONNÉES DE RÉFÉRENCE

L'intégration de l'imagerie satellite dans l'analyse des statistiques agricoles nécessite d'avoir accès à plusieurs couches essentielles de données de référence. Celles-ci jouent un rôle fondamental dès le début, lors de la phase de réflexion sur le plan d'échantillonnage et la sélection des données (par exemple, lors de la prise de décisions sur la stratification de l'échantillon, de la planification logistique de l'exécution de l'enquête sur le terrain, de l'achat de matériel d'imagerie de qualité et de résolution requise, de la vérification et de la validation des indicateurs dérivés des images, de l'agrégation des données au niveau des unités administratives, etc.). Le tableau 2²² dresse une liste non exhaustive des premières couches pertinentes.

TABEAU 2. COUCHES DE DONNÉES DE RÉFÉRENCE PERTINENTES POUR LA PRÉPARATION D'ENQUÊTES AGRICOLES ET D'ANALYSES CONNEXES FAITES À PARTIR D'OBSERVATIONS DE LA TERRE.

Ensemble de référence	Échelle minimale	Remarques
MAN	1:100 000 à l'échelle des pays	Pour stratification, rectification d'image Libre accès aux systèmes mondiaux à 30 m (1 arcsec) fournis par SRTM, ALOS-PRISM, ASTER Pour une utilisation avec des images THR, il est recommandé de recourir à des échelles plus grandes (> 1:25 000).
Topographie numérique	1:100 000 à l'échelle des pays, 1:25 000 pour les zones étudiées	Pour conception d'enquêtes, stratification et enregistrement d'images Liens avec la logistique des denrées alimentaires (transport et stockage) et avec des facteurs de production agricole pertinents (accès à l'eau, structures d'irrigation)
Utilisation des terres/ couverture du sol	1:100 000	Stratification et conception d'enquêtes Facteurs environnementaux (conversion des terres, biodiversité, prairies ou terres de parcours)
Carte des sols	1:100 000	Adéquation des terres, facteurs de production, modes de production régionaux, risques d'érosion et de salinité
Cadastre	1:10 000	Identification des parcelles, régimes fonciers, stratification, aide à la classification des images, conception d'enquêtes, extrapolation et agrégation statistiques
Frontières administratives	1:50 000	Conception d'enquêtes et agrégation statistique

Hormis les jeux de données tels que les MAN internationaux, la disponibilité de ces couches est extrêmement hétérogène, voire inexistante en fonction du pays ou de la région concerné(e), à la fois en matière de qualité et de conditions applicables conformément aux licences. Lorsque les échelles requises sont indisponibles, des extraits provenant de jeux de données régionaux ou mondiaux à plus petite échelle peuvent fournir des premières estimations ; toutefois, des efforts supplémentaires pourront s'avérer nécessaires pour générer des cartes à l'échelle la plus pertinente, notamment pour les grandes zones de production agricole. L'imagerie satellite et aérienne peut servir de source alternative au calcul des couches de référence, comme la topographie numérique, l'utilisation des terres/la couverture du sol ; elle nécessite en revanche un grand nombre de ressources dès la phase de préparation.

OpenStreetMap²³ (OSM) est une initiative en libre accès qui vise à produire une carte mondiale numérique grâce aux contributions volontaires, à l'aide des connaissances locales, de tracés GPS et de données sources fournies gracieusement. Les données numériques attribuées peuvent être extraites d'OSM par région ou par caractéristique, puis intégrées dans un logiciel d'analyse géospatiale. Le logiciel en libre accès QGIS²⁴ est particulièrement utile

22 Pour plus de détails sur les applications de surveillance terrestre, voir Agence européenne pour l'environnement, 2013, *Note on in-situ data requirements: Update of D2.1 – Report on In-situ data requirements*, Publication GMES in-situ coordination (GISC). Disponible à l'adresse : <http://gisc.pbe.eea.europa.eu/deliverables/d2.1.pdf/download/en/1/D2.1.pdf?action=view> (p. 121 ff).

23 <https://www.openstreetmap.org>.

24 <https://www.qgis.org/fr/site/>.

pour l'analyse des données extraites d'OSM, puisqu'il fournit des interfaces simples pour extraire, modifier et télécharger les caractéristiques numérisées.

QGIS présente un autre avantage, celui d'être un outil polyvalent facilitant l'intégration de l'imagerie et des jeux de classes d'entités, ainsi que la création de processus d'analyse géospatiale sophistiqués.

1.5. INFORMATIQUE DÉMATÉRIALISÉ

Une tendance récente, souvent baptisée le « nouveau paradigme du Big Data », consiste à faire migrer le stockage et le traitement d'un volume important de données vers des solutions informatiques dématérialisées (hébergées). Les prestataires de ce type de solutions fournissent les infrastructures matérielles nécessaires à l'exécution simultanée d'algorithmes de traitement des données en présence de jeux de données volumineux. L'un des grands avantages de ce système réside dans le fait que, puisque les archives de données et les unités de traitement se trouvent logiquement proches les unes des autres, et puisque les structures et la parallélisation des données peuvent être largement optimisées dans le cadre des missions à accomplir, l'analyse générale des données peut être beaucoup plus rapide et donc être adaptée à un ensemble de données beaucoup plus vaste ou se prêter à une utilisation plus complexe des ressources informatiques. Autre avantage non négligeable, les usagers de solutions hébergées n'ont pas besoin de posséder les équipements matériels nécessaires au traitement de larges volumes de données. En réalité, les plateformes clients accédant à ce type de solution peuvent rester simples (de simples ordinateurs portables peuvent faire l'affaire). En outre, le besoin de télécharger de larges volumes de données étant fortement réduit, il n'est pas nécessaire d'augmenter la bande passante de la connexion Internet utilisée. L'informatique dématérialisée est une alternative viable dans les pays moins développés, où l'entretien des équipements et des logiciels spécialisés peut s'avérer difficile.

Les solutions hébergées présentent quelques inconvénients : d'une part, les usagers n'ont plus la maîtrise totale du workflow ; d'autre part, les données et algorithmes de référence et auxiliaires doivent être cohébergés sur l'infrastructure à distance (serveur hébergé), ce qui peut entraîner des problèmes en matière d'autorisation, de droits de propriété intellectuelle et de coûts (liés au stockage sur les plateformes hébergées et à la capacité de calcul, par exemple). Ces facteurs doivent être pris en compte à la lumière des activités prévues, des coûts d'entretien des infrastructures existantes et d'investissement dans les mises à jour, le recrutement d'experts et les formations nécessaires, etc.

Pour apprécier les capacités des services d'informatique dématérialisés et les comparer aux solutions de traitement autonomes, prenons un pays de taille médiane, comme la République de Corée (dont le territoire couvre une surface d'environ 95 000 km²). Une couverture totale nécessite 20 tuiles Sentinel-2 (10 Gb par orbite), 13 scènes Landsat-8 (13 Gb par orbite) et 16 tuiles Sentinel-1 (8 descendantes, 8 ascendantes, 16 Gb par orbite²⁵). En partant du principe que la couverture fournie par Sentinel-1A et Sentinel-1B d'une part, et par Sentinel-2A et Sentinel-2B d'autre part, soit pleinement opérationnelle (respectivement 61 et 73 orbites par an), le volume total de données saisies de niveau 1A/1C sera de $10 \times 73 \text{ (S2)} + 13 \times 22 \text{ (L8)} + 16 \times 61 \text{ (S1)} = 1\,992 \text{ Gb}$ (soit près de 2 Tb) par an. Après traitement préalable (géocodage, étalonnage, retrait de la couverture nuageuse des données optiques, dissimulation de la surface de la mer), le volume de données alimentant l'analyse statistique nationale serait de l'ordre de 1 Tb. Si ce volume demeure encore gérable dans le cadre d'un workflow opérationnel autonome, le téléchargement, le stockage et le traitement systématiques et opportuns des informations concernées nécessiteraient des efforts non négligeables et une structure informatique relativement solide.

25 Environ 96 Gb par orbite, au cas où un traitement interférométrique s'avérerait nécessaire.

Google Earth Engine²⁶ (GEE) est une solution d'informatique en dématérialisée réputée dans le domaine géospatial. Ce logiciel fournit aux usagers inscrits²⁷ un accès à la plupart des catalogues de données gratuits et en libre accès, notamment les données à pleine résolution de MODIS, de Landsat, de Sentinel-1 et de Sentinel-2, conservées sur l'infrastructure dématérialisée de stockage de Google. Il propose également d'autres catalogues de données rastrées en libre accès, tels que le MAN SRTM d'une résolution de 1 et 3 arcsecs, les grilles d'estimation des précipitations (fondées sur la base de données CHIRPS [Climate Hazards Group Infrared Precipitation with Station data]), les produits de modélisation atmosphérique du Système mondial de prévision, et les classifications mondiales de la couverture du sol ou thématiques. Les données rastrées se présentent généralement sous le même format que celui proposé par la source (par exemple, l'Agence américaine d'observation océanique et atmosphérique — NOAA, de l'anglais *National Oceanic and Atmospheric Administration* — ou l'Institut d'études géologiques des États-Unis — USGS, de l'anglais *United States Geological Survey* — ou encore le programme Copernicus de l'Union européenne). Ces données peuvent comprendre des versions ayant fait l'objet d'un traitement plus poussé, comme la réflectance à la surface de Landsat, les indices de MODIS, et les composites spatiaux et temporels. Les données de Sentinel-1 sont fournies sous forme de coefficients de rétrodiffusion calibrés et géocodés après traitement des originaux GRD formatés à l'aide de la boîte à outils Sentinel-1, qui prépare les données en vue de leur exploitation aux fins d'analyse. GEE facilite également l'intégration des jeux de classes d'entités ; il ne compte toutefois qu'un nombre limité de classes d'entités mondiales aux données plutôt grossières (p. ex., sur les frontières nationales). Les utilisateurs de GEE peuvent cependant mettre en ligne (et télécharger) des données raster et des jeux de classes d'entités sous forme de ressources privées ou publiques.

L'une des caractéristiques principales de GEE est la possibilité d'intégrer l'accès à des jeux de données mondiaux dans une logique de traitement programmable mise à disposition des usagers sous forme de bibliothèque de routines. Les bibliothèques de base englobent les tâches de traitement des données raster et des classes d'entités (filtres, opérateurs géométriques, classificateurs et agrégation statistique) qui couvrent des nombres arbitraires de bandes images, de combinaisons de séries temporelles, de compositions multicapteurs, etc. Les bibliothèques matricielles présentent quant à elles des fonctionnalités permettant l'analyse de régression, l'analyse des valeurs propres et l'(auto-)corrélation. Les usagers peuvent combiner sélection de données et fonctions des bibliothèques dans des scripts définissant des workflows pertinents, qui peuvent donner lieu à un produit rastré comme un résultat de classification, ou être réduits sous forme de tableau, comme les agrégats statistiques spatiaux. Le script peut être obtenu dans une interface de navigateur (à l'aide de JavaScript) ou via l'API Python. L'un des principaux avantages de cette méthode réside dans le fait que les usagers peuvent se partager les scripts, contrairement aux téléchargements, puisqu'il est généralement plus rentable de relancer le script que d'échanger des volumes importants de données. GEE peut également servir en arrière-plan dans le cadre de solutions (Web) plus élaborées, comme le système Collect Earth²⁸ de la FAO, qui vise à assurer la collecte et l'analyse systématiques des données de validation dans le cadre de l'utilisation des terres, y compris agricoles, et de la modification de ladite utilisation.

1.6. ÉTUDE DE CAS

Pour illustrer la notion générale d'intégration des données de télédétection dans les données de référence en vue de la réalisation de statistiques agricoles, nous allons désormais décrire le scénario idéal, dans le cadre duquel des jeux de données essentielles de qualité équivalente, voire supérieure, aux exigences générales, sont en libre accès.

Le gouvernement des Pays-Bas soutient sans réserve le libre accès aux infrastructures publiques de données, à la fois dans les lois qu'il adopte et dans les plateformes technologiques adaptées qu'il met en place. Le service public

²⁶ <https://earthengine.google.com/>.

²⁷ GEE est gratuit pour les usages non commerciaux.

²⁸ <http://www.openforis.org/tools/collect-earth.html>.

d'aide à la cartographie²⁹ aide les autorités publiques à publier leurs archives de données sous différents formats, tous en libre accès. Ces jeux de données alimentent le registre géographique national, qui fournit lui-même un accès à ces données, mais appuie également un certain nombre de solutions de recherche, de visualisation et de traitement des données, par le biais par exemple de protocoles de services Web.

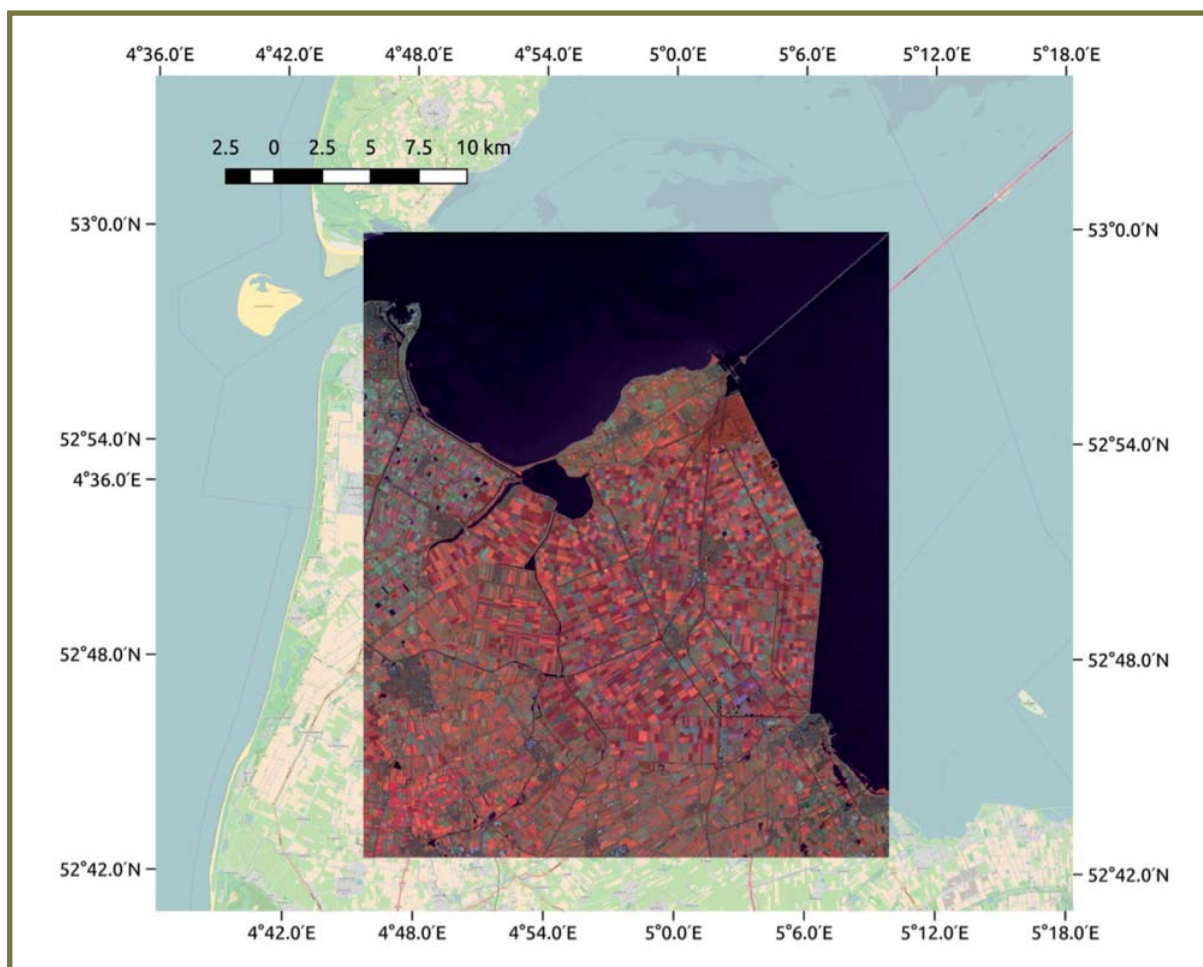
Les jeux de données thématiques englobent, entre autres, la topographie numérique à une échelle de 1:10 000, un MAN dérivé de lidar d'un espacement de 0,5 m, de l'ortho-imagerie (annuelle) de résolution 0,25 m, des couches d'informations sur les sols, et des volumes importants de couches thématiques, portant notamment sur la production agricole. L'utilisation des terres agricoles est définie dans une couche de données de référence des parcelles, qui fait partie du système intégré de gestion et de contrôle soutenant les mesures de la Politique agricole commune de l'Union européenne. Chaque année, les agriculteurs enregistrent l'usage qu'ils font de leurs terres en se fondant sur le système de référence des parcelles. Le registre géographique national³⁰ contient à la fois le système de référence des parcelles et les couches de données annuelles sur l'utilisation des terres (échelle de 1:10 000). Les déclarations annuelles anonymisées d'utilisation des terres sont généralement disponibles dans leur version finale vers la fin de l'année (décembre), bien après la période de végétation (octobre AAAA-1 à octobre AAAA).

Ainsi, bien que des statistiques détaillées concernant les cultures puissent être calculées directement à partir des données tirées des déclarations, les jeux de données constituent quant à eux d'excellents éléments de référence pour valider *a posteriori* le potentiel des données de télédétection pour le calcul d'estimations précises sur les cultures aux différents stades de la période de végétation (précédente). De plus, l'excellente qualité des données de référence permet de tester différentes méthodes d'échantillonnage pour la collecte d'informations de référence sur les cultures de saison (points échantillons ou échantillons de tuiles, dans le cadre d'enquêtes menées «en chemin», avec ou sans l'appui de données de télédétection à proprement parler. En effet, l'exemple des Pays-Bas peut servir de suite d'entraînement adaptée pour bon nombre d'expériences géostatistiques en cours de déploiement.

²⁹ <https://www.pdok.nl/>.

³⁰ Rechercher l'AAN (registre des parcelles) et le BRP (déclarations annuelles sur l'utilisation des terres), respectivement.

FIGURE 3. SUPERPOSITION D'UNE CARTE GÉNÉRALE DE LA COMMUNE DE HOLLANDS KROON (PAYS-BAS) ET D'UNE IMAGE DE SENTINEL-2 DU 5 JUILLET 2016 (COMBINAISON DES BANDES B8-B11-B4).



L'image est projetée dans le système national néerlandais de coordonnées de référence (RDS, EPSG:28992). Les données de la carte en arrière-plan sont protégées par le copyright des contributeurs d'OpenStreetMap 2017. Les données de Sentinel-2 sont protégées par le copyright du programme Copernicus, 2015.

L'exemple traité ici porte sur une classification simple fondée sur GEE. Tout d'abord, toutes les parcelles sélectionnées concernent l'année 2015 et la commune de Hollands Kroon, dans la province de Hollande-Septentrionale, aux Pays-Bas (voir la figure 3). Cette commune comprend Wieringermeer, situé sur un polder, et quelques zones environnantes. Compte tenu de son sol alluvial d'excellente qualité, la région est largement exploitée pour la production de cultures, notamment horticoles. Sur un territoire municipal total de 37 108 ha, les données issues des déclarations 2015 portent sur un total de 7 086 parcelles recouvrant 27 723 ha. La taille moyenne des parcelles déclarées est de 3,91 ha (l'ensemble des observations étant dans un intervalle de 0,1 à 47,9 ha).

Les parcelles concernées sont converties en format KML, puis saisies dans un tableau Fusion Table sur Google Drive, lequel peut être intégré au script GEE à l'aide du numéro d'identification dudit tableau. Les frontières de la commune sont saisies selon le même processus. Ces dernières sont utilisées pour sélectionner des images à partir des catalogues GEE pour une période et un capteur donnés, avec la possibilité de limiter la recherche à certains critères. Par exemple, il est possible de limiter la recherche aux données optiques présentant un pourcentage minimum de couverture nuageuse, de limiter la sélection de données SAR à une orbite donnée, etc. Pour les résolutions élevées, il est possible de choisir entre Landsat-8, ASTER, Sentinel-1 et Sentinel-2. Cet exemple portera sur Sentinel-2, pour lequel la phase de mise en service a été lancée le 5 juillet 2015.

Ensuite, un échantillon d'apprentissage provenant du jeu de données BRP 2015 est constitué. Aux fins de cet exemple, la taille de la parcelle d'entraînement est limitée à un minimum de 2 ha, afin d'éviter d'ajouter des pixels mélangés dans les limites de la parcelle. Par ailleurs, les parcelles de forme oblongue (c'est-à-dire celles pour lesquelles un côté est quatre fois plus long que l'autre) sont exclues. Il est important de procéder ainsi pour Hollands Kroon, puisque la commune englobe plusieurs prairies, très fines mais très allongées, bordant les digues intérieures et extérieures du polder. La sélection est de surcroît limitée aux types de produits qui sont cultivés dans au moins 50 parcelles du lot. Cette étape vise à éviter d'intégrer des cultures relativement rares. La sélection finale comprend 4 473 parcelles (sur 24 624 ha).

Enfin, 20 % de l'échantillon (soit 916 parcelles) sont sélectionnés de façon aléatoire aux fins d'apprentissage, et les 80 % restants (3 557 parcelles) sont utilisés pour tester la classification. Tous les paramètres utilisés dans la sélection sont des variables, qui peuvent être aisément modifiées pour procéder à une nouvelle classification et étudier l'influence de différents paramétrages.

Il est désormais possible de mettre en place le classifieur. À cette fin, les bandes suivantes sont sélectionnées à partir de l'image de Sentinel-2 : 2-4 (visible), 5-7 (red-edge), 8 et 8A (NIR) et 11 et 12 (SWIR). À noter également que certaines bandes présentent une résolution de 10 m (2-4, 8), tandis que d'autres présentent une résolution de 20 m (5-7, 8A, 11 et 12); il n'est cependant pas nécessaire de procéder à un rééchantillonnage explicite pour préparer l'image. En ce qui concerne l'outil de classement, le classifieur des forêts aléatoire, fondé sur un seul paramètre, à savoir 10 arbres initiaux par classe, a été choisi. Le classifieur est testé avec les échantillons choisis, et le résultat est ensuite appliqué aux bandes d'images. Pour chaque parcelle, l'histogramme de fréquence de l'étiquette de classe contenu dans chaque parcelle est extrait, et la parcelle est affectée à la classe majoritaire. Il est alors possible de générer une matrice de confusion fondée sur les parcelles (et non sur les pixels), comme le montre la figure 4.

FIGURE 4. MATRICE DE CONFUSION RELATIVE À UNE CLASSIFICATION E DES FORÊTS ALÉATOIR À PARTIR DES IMAGES SENTINEL-2 DE HOLLANDS KROON (5 JUILLET 2015), FONDÉE SUR LES DÉCLARATIONS DU BRP 2015 AUX FINS D'ESSAI ET D'APPRENTISSAGE.

	ALF	CRT	GRA	HOR	MAI	ONI	POT	SBT	SWH	VEG	WWH	Totaux	PU
ALF	14	0	1	0	1	0	3	0	0	1	0	20	0,70
CRT	0	9	0	0	1	9	0	0	0	2	0	21	0,43
GRA	15	3	942	3	8	0	6	1	2	3	0	983	0,96
HOR	8	28	17	586	12	31	10	3	1	95	3	794	0,74
MAI	0	3	2	2	163	4	5	0	4	5	1	189	0,86
ONI	1	2	0	6	3	99	1	4	0	11	0	127	0,78
POT	31	4	5	9	1	0	571	3	0	6	0	630	0,91
SBT	1	1	0	12	0	10	2	232	0	7	1	266	0,87
SWH	0	0	1	0	2	0	0	0	50	0	9	62	0,81
VEG	2	1	0	8	1	0	1	0	0	12	0	25	0,48
WWH	0	0	1	0	1	10	0	0	14	0	403	429	0,94
Totaux	72	51	969	626	193	163	599	243	71	142	417	3546	PG=0,87
PP	0,19	0,18	0,97	0,94	0,84	0,61	0,95	0,95	0,70	0,08	0,97		

Étiquettes des cultures : ALF = luzerne; CRT = carottes; GRA = graminées; HOR = horticulture; MAI = maïs; ONI = oignons; POT = pommes de terre; SBT = betterave à sucre; SWH = blé de printemps; VEG = légumes; WWH = blé d'hiver.
Étiquettes de classification : PP = précision du producteur; PU = précision de l'utilisateur; PG = précision globale.

Il convient de remarquer qu'il est possible d'obtenir des résultats relativement bons, avec une précision globale de 87 % pour 11 types de cultures distinctes, même en cas d'utilisation d'une seule image Sentinel-2. Début juillet est manifestement une période quasi optimale pour définir plusieurs types de cultures. À cette période de l'année, les céréales d'hiver, sénescences, sont prêtes à être récoltées, en contraste total avec les cultures d'été en pleine maturation, telles que les betteraves à sucre et les pommes de terre. Les prairies sont également bien définies, même si l'on peut s'attendre à des variations très importantes dues aux différentes pratiques de fauche. Les cultures horticoles constituent l'exception, puisque ce groupe se compose en majorité de tulipes et d'autres bulbes de fleurs de printemps qui sont généralement récoltés à la fin juin. La confusion est manifeste pour les cultures comme les oignons, les carottes et les autres légumes (mélangés aux cultures horticoles) et le blé de printemps (mélangé au blé d'hiver). Des images optiques supplémentaires prises avant et après juillet aideront à définir plus précisément ces groupes.

Si les résultats prouvent le potentiel des images fournies par Sentinel-2 pour la production de cartes de cultures, il ressort surtout de l'exemple ci-dessus qu'il est relativement simple de produire des résultats cohérents à l'aide de GEE. Dans cet exemple, il aurait été parfaitement possible de générer des résultats similaires avec d'autres logiciels en libre accès, en associant par exemple des scripts Python et les modules mentionnés précédemment. GEE présente l'avantage de permettre à ses usagers de passer à d'autres sélections d'images, par exemple les corpus de Landsat-8 et de Sentinel-1 (et des mélanges hybrides), de choisir d'autres régions et de tester d'autres configurations en adaptant légèrement le script. Les jeux d'images Sentinel-1 sont généralement privilégiés aux fins de classification, puisque la disponibilité des données est systématique (pas de problèmes liés à la couverture nuageuse) et la précision globale de la classification dépasse les 90 % pour les régions de production de cultures présentant des caractéristiques comparables à celles de Hollands Kroon.

1.7. CONCLUSION

Nous avons cherché dans ce chapitre à analyser la disponibilité des outils d'imagerie et à déterminer la meilleure façon de les exploiter pour délimiter les parcelles et caractériser les cultures, afin de faciliter le calcul de la superficie des terres de cultures, en mettant l'accent sur l'imagerie à haute résolution en libre accès produite principalement par les nouveaux capteurs Sentinel-1 et Sentinel-2. Ces capteurs, associés aux données produites par d'autres satellites en libre accès tels que Landsat et ASTER, permettent d'adopter de nouvelles approches en matière d'estimation des surfaces cultivées associant couverture large à une résolution de 10 m et images THR d'autres opérateurs aériens et vendeurs de satellites commerciaux en vue d'un échantillonnage statistiquement valide. En règle générale, les logiciels visant à appuyer l'analyse des images en vue de la création de classifications des cultures et d'estimations des zones sont aujourd'hui accessibles gratuitement. Toutefois, compte tenu des volumes importants d'images en accès libre générées à l'heure actuelle, il convient de tenir compte de plusieurs aspects pratiques relatifs au traitement des Big Data (mégadonnées), en fonction de l'échelle de l'analyse à mener. Google Earth Engine (GEE) est une plateforme dématérialisée (cloud computing) qui présente un certain intérêt pour l'étude et l'analyse rapides des données, puisqu'elle héberge déjà des images en accès libre et propose un ensemble fourni de bibliothèques permettant de réaliser des analyses géospatiales approfondies. Bien que cet outil ne soit pas en libre accès, il présente une utilité pour une vaste communauté de développeurs et d'utilisateurs qui partagent une analyse scriptée au même titre que les communautés de logiciels libres. De nouvelles évolutions, y compris le développement de solutions en libre accès total, devraient voir le jour pour stimuler davantage l'analyse à partir des solutions informatiques dématérialisées.

Chapitre 2

Cartographie et suivi de la couverture du sol

Pierre Defourny

2.1. INTRODUCTION

Il existe aujourd'hui plusieurs moyens de décrire la surface terrestre grâce aux avancées sans précédent réalisées en matière de technologies d'information et d'observation, depuis l'apparition des UAV jusqu'à la création de satellites d'observation de la Terre. La télédétection par satellite constitue une source indiscutable d'informations sur les terres pour un large éventail d'utilisateurs à toutes les échelles géographiques. Toutefois, l'écart se creuse entre les producteurs de données de télédétection et les utilisateurs des cartes, un phénomène accentué par le fait que les infrastructures de données spatiales génèrent un important volume d'informations géographiques largement accessibles. Il est donc important de comprendre les divers concepts et contraintes qui régissent la cartographie de la couverture du sol dans le contexte des statistiques agricoles, d'autant plus que les cartes de couverture du sol sont souvent utilisées dans le cadre des enquêtes agricoles pour appuyer la stratification au niveau du plan d'échantillonnage. De simples cartes des terres cultivées ou des cartes plus précises illustrant l'intensité de culture peuvent en effet réduire considérablement la variance d'échantillonnage ou le fardeau lié au dispositif des opérations de terrain ainsi que leur coût.

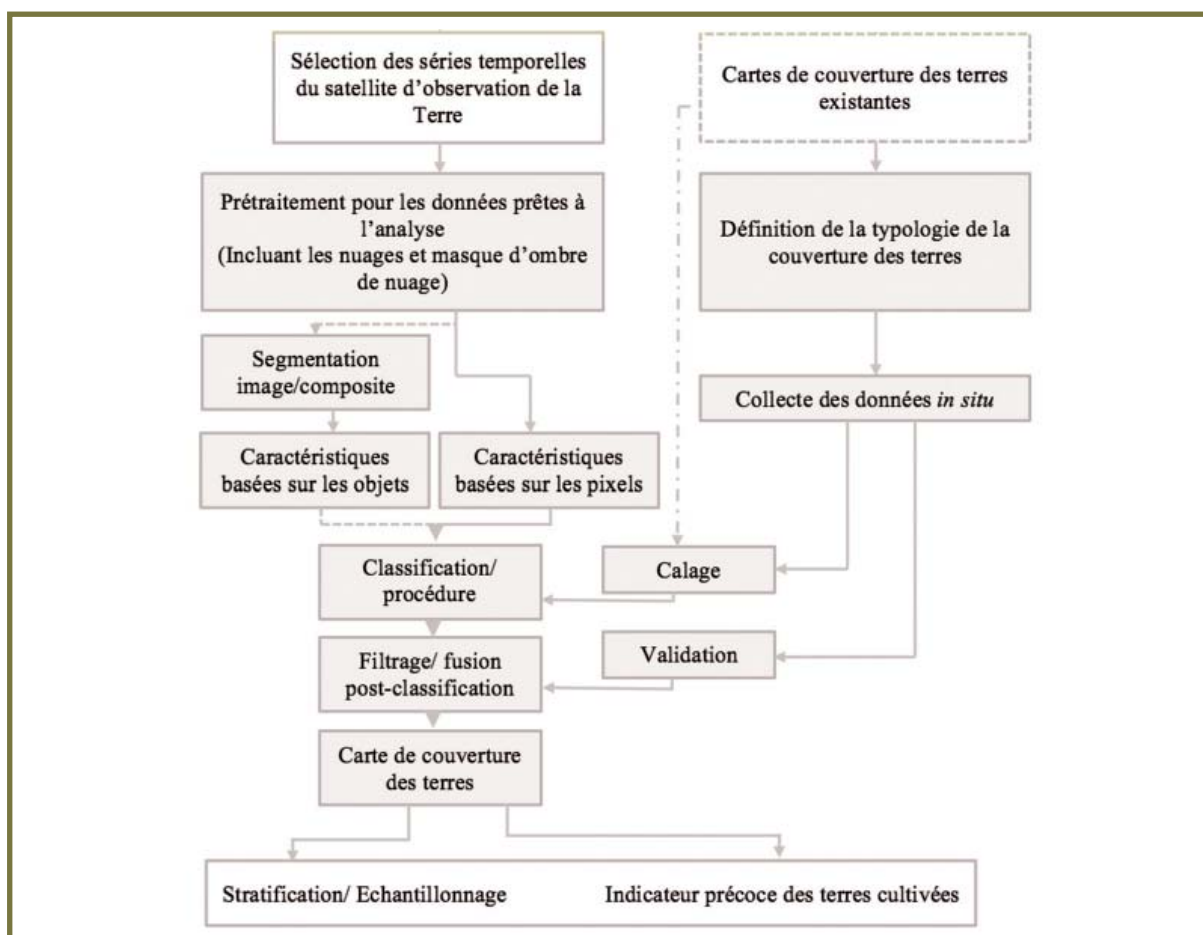
Les cartes de couverture du sol peuvent mettre en évidence les strates non agricoles qui ne doivent pas faire l'objet d'un échantillonnage ou les strates qui pourraient être échantillonnées différemment. Comme le montre Delincé (2015)¹, si un tiers de la zone administrative étudiée correspond à une strate non agricole, la réaffectation de l'ensemble de l'échantillon aux strates restantes (terres cultivées, notamment) permettra d'atteindre une efficacité

¹ GSARS, 2015, *Technical Report on Cost-Effectiveness of Remote Sensing for Agricultural Statistics in Developing and Emerging Economies*, Rapport technique de la GSARS, Rome, disponible à l'adresse <http://gsars.org/en/technical-report-on-cost-effectiveness-of-remote-sensing-for-agricultural-statistics-in-developing-and-emerging-economies/>, consulté le 9 août 2017.

relative de la stratification équivalant à 1,51 pour un coût quasi nul². L'efficacité de la stratification dépend clairement de la pertinence de la carte (de couverture du sol) retenue pour la stratification.

Ce chapitre présente le concept de couverture du sol et passe en revue certains des principaux éléments du processus de cartographie de la couverture du sol, dans l'ordre dans lequel ils sont habituellement réalisés (figure 1). Les premières étapes de ce processus consistent à sélectionner la typologie de couverture du sol appropriée, à collecter des données *in situ* et à acquérir les images de télédétection. L'exploitation numérique de ces images satellites nécessite l'exécution rigoureuse d'une série d'opérations standard, suivies de la réalisation d'une carte précise de la couverture du sol. Les cartes de couverture du sol étant facilement accessibles pour certaines régions, la pertinence des cartes existantes pour le secteur de l'agriculture sera systématiquement étudiée en fonction d'un ensemble de critères bien définis. Si les cartes de couverture du sol utilisées pour la stratification sont généralement produites sur la base des informations issues des campagnes agricoles précédentes, ce chapitre présente également d'expériences plus récentes d'élaboration de carte sur la base des observations réalisées au cours de la campagne agricole courante.

FIGURE 1. PROCESSUS DE CARTOGRAPHIE DE LA COUVERTURE DU SOL À PARTIR DE SÉRIES TEMPORELLES D'OBSERVATIONS PAR SATELLITE.



Les lignes discontinues correspondent à des possibilités alternatives.

² Sans stratification, la variance de la moyenne serait égale à $S_y/(2/3) \times n$, c'est-à-dire que 33 % de l'échantillon serait gaspillé dans des régions non cultivées. Avec la stratification, la variance devient égale à S_y/n , de sorte que l'efficacité relative est la suivante : $Er = \{ S_y / (2/3) \times n \} / \{ S_y / n \} = 3/2 = 1,5$.

2.2. LE CONCEPT DE COUVERTURE DU SOL

Historiquement, le concept d'utilisation des terres est jugé plus pertinent pour de nombreuses applications que celui de couverture du sol. Ce dernier a pourtant été choisi pour devenir une sorte de panacée dans le domaine des inventaires fonciers, étant donné qu'il se fonde sur l'observation directe et compte tenu du nombre croissant d'images satellite dont on dispose aujourd'hui. Il est utilisé comme substitut pour décrire la structure et les caractéristiques du paysage par un nombre grandissant d'utilisateurs ne connaissant pas nécessairement l'origine et la sémantique des informations relatives à la couverture du sol. Comber *et al.* (2005a) montrent que le concept de couverture du sol est perçu différemment selon les disciplines. Si les utilisateurs ne comprennent pas parfaitement sa signification et les hypothèses sur lesquelles il repose, ils imposent leur propre interprétation de ce qu'il englobe à la lumière de leurs propres contraintes, centres d'intérêt et objectifs, au risque de fausser leur évaluation des données et leurs analyses ultérieures.

Le domaine de la surface terrestre est en réalité infiniment complexe, et toute interprétation des données d'observation de la Terre implique des processus tels que l'abstraction, la classification, l'agrégation et la simplification. Depuis maintenant plusieurs décennies, il existe des divergences d'opinions quant à la signification du concept de couverture du sol et à ce qui le distingue de celui d'utilisation des terres. Aucune unité fondamentale n'ayant été arrêtée en matière d'observation terrestre, la cartographie de la couverture du sol doit être comprise comme un processus d'extraction d'informations, régi par des règles fondées sur des objectifs individuels ou institutionnels. La plupart des grandes initiatives de cartographie de la couverture du sol ont créé et décrit de façon détaillée leur propre système de classification. Aux tous débuts de l'observation par satellite, l'USGS a mis au point un système normalisé de classification de l'utilisation des terres et de la couverture du sol en s'appuyant sur 40 ans d'expérience de la cartographie à partir de photographies aériennes (Anderson *et al.*, 1976). Ces travaux sont considérés comme faisant partie des plus influents en matière d'élaboration de normes nationales destinées à divers organismes. Compte tenu des attentes de plus en plus importantes des utilisateurs, de la disponibilité toujours croissante de données et de la grande diversité des objectifs et des contextes, les efforts visant à répertorier les différentes typologies de couverture du sol s'avèrent encore vraiment nécessaires partout dans le monde.

2.2.1. Systèmes de classification de la couverture du sol (LCCS et LCML)

Dans le but de garantir une interopérabilité totale entre les typologies et de fournir des bases communes pour décrire la couverture du sol, la FAO a mis au point le Système de classification de la couverture du sol (LCCS, de l'anglais *Land Cover Classification System*), un cadre conceptuel destiné à la définition des légendes. Cet outil, grâce à une approche dichotomique et à une approche modulaire hiérarchique fondées sur plusieurs ensembles de descripteurs (appelés « classifieurs »), vise à définir clairement chaque classe de couverture du sol, et permet par conséquent d'établir des correspondances entre différentes typologies (Di Gregorio et Jansen, 2000). Ce système repose sur des critères indépendants et universellement valides de d'analyse de la couverture du sol, plutôt que sur un ensemble prédéfini de classes de couverture du sol. Ce processus permet de mettre à disposition une définition exhaustive de la couverture du sol, quels que soient l'échelle cartographique, le type de couverture du sol, la méthode d'interprétation des données ou la situation géographique (Di Gregorio, 2005).

Plus récemment, le cadre LCCS a été modifié pour devenir le Métalangage de couverture du sol (LCML, de l'anglais *Land Cover Meta-Language*), un outil plus flexible doté de classifieurs non limités, qui propose une description des classes plus détaillée. En 2012, le cadre LCML proposé par la FAO a été adopté en tant que norme internationale par le Comité technique de l'Organisation internationale de normalisation (ISO, de l'anglais *International Standardization Organization*)³. Le LCML est un système de classification orienté objet dans lequel chaque particularité de la couverture du sol est caractérisée par une série d'éléments pouvant être décrits plus en

³ Le LCML a été adopté en tant que norme ISO (ISO 19144-2:2012) : <http://www.din.de/en/getting-involved/standards-committees/nabau/standards/wdc-beuth:din21:155452459/toc-1911111/download>.

détail par un ensemble d'attributs. La signification de la classe n'est plus liée à un simple nom, mais plutôt à un modèle plus exhaustif et plus moderne constitué d'éléments et d'attributs qui caractérisent les particularités de la couverture du sol. Latham et Rosati (2016) fournissent de plus amples informations à ce sujet⁴.

Par souci de clarté, de transparence et de comparabilité, il est recommandé au niveau international d'utiliser le cadre LCML pour définir une typologie de couverture du sol donnée avant tout travail cartographique. Le projet de cartographie de la couverture du sol Globland30, achevé en 2014 au prix d'efforts considérables et d'un travail exhaustif, a ainsi mal défini les classes de couverture du sol relatives à l'agriculture, ce qui a sérieusement limité son utilisation pour de nombreuses applications dans le domaine de l'agriculture et de l'élevage.

2.2.2. L'agriculture dans la typologie de la couverture du sol

Dans le cadre des statistiques agricoles, la stratification définie en vue du plan d'échantillonnage dépend principalement des classes de couverture du sol relatives à l'agriculture. Il convient de signaler que la notion de terres cultivées ne constitue pas à proprement parler une classe de couverture du sol, mais plutôt une classe d'utilisation des terres⁵. Un champ de céréales, par exemple, correspond plus précisément à une « végétation herbacée dense » dans la nomenclature de couverture du sol, alors que dans la nomenclature de l'utilisation des terres il devrait correspondre à « agriculture » ou « activité agricole ». Toutefois, toutes les typologies de couverture du sol existantes intègrent des classes relatives à l'agriculture en raison de leur importance au niveau de la structure du paysage et pour les utilisateurs de cartes.

Bien que l'agriculture semble de prime abord être la classe de « couverture du sol » la plus facile à cartographier, il s'agit d'une source majeure de malentendus et de divergences entre les cartes de couverture du sol existantes, même lorsque ces dernières se contentent de distinguer les terres cultivées et non cultivées. Cette situation est exacerbée dès lors que l'on tient compte de l'extrême diversité des terres agricoles à travers le monde, des rizières dans lesquelles on pratique la double culture en Asie au système mésoaméricain traditionnel d'association de cultures, ou *milpa*, en passant par les jachères européennes et les plantations pérennes telles que le cacao cultivé sous la canopée forestière en Afrique.

Le Programme mondial de recensement de l'agriculture 2020 (vol. 1, p. 79) propose les définitions suivantes⁶, qui sont le résultat d'une agrégation des classes LCML :

- Les **terres arables** désignent les terres utilisées la plupart du temps pour les cultures temporaires. Elles comprennent les terres consacrées à des cultures temporaires pendant une période de référence de douze mois, ainsi que les terres qui seraient normalement utilisées à cette fin mais qui sont restées en jachère ou qui n'ont pas été semées en raison de circonstances imprévues. Les terres arables ne comprennent pas les terres consacrées aux cultures permanentes ni les terres potentiellement cultivables mais qui ne sont normalement pas cultivées. Ces dernières doivent être classées comme des « prairies et pâturages permanents » si elles sont utilisées pour le pâturage ou la fenaïson, comme des « forêts et autres terres boisées » si elles sont envahies d'arbres et non utilisées pour le pâturage ou la fenaïson, ou comme des « autres zones non classées ailleurs » si elles deviennent des friches.
- Les **terres cultivées** correspondent à la somme des terres arables et des terres consacrées aux cultures permanentes.
- Les **terres agricoles** correspondent à la somme des terres cultivées et des prairies et pâturages permanents.

4 Latham, J. et Rosati, I., 2016, *Land Information in the Context of Agricultural Statistics*, Rapport technique GO-15-2016 de la GSARS, disponible à l'adresse http://gsars.org/wp-content/uploads/2016/08/TR_Information-on-Land-in-the-Context-of-Ag-Statistics-180816.pdf, consulté le 10 juin 2017.

5 Voir <http://www.fao.org/faostat/fr/#data/RL> pour consulter les statistiques communiquées par les pays concernant l'utilisation agricole des terres.

6 Programme mondial de recensement de l'agriculture 2020, vol. 1, p. 90.

- Les **terres consacrées à l'agriculture** correspondent à la somme des « terres agricoles » et des « terres consacrées aux bâtiments et cours de ferme ».

À partir du cadre LCML, Di Gregorio (2013) a établi une nomenclature précise et exhaustive des terres cultivées afin de définir ce concept. Cependant, dans le cadre des statistiques agricoles, cette définition peut soulever de nouvelles questions, notamment quant au fait que la surface cultivée qui nous intéresse n'est ni la surface semée ni la surface exploitable, mais plutôt la surface réellement récoltée. Cette distinction ne se réduit pas à une discussion sémantique réservée aux chercheurs, car les différences peuvent être importantes en cas de sécheresse ou d'inondations.

En outre, la typologie de la couverture du sol doit être exploitable et compatible avec la source des données. En ce qui concerne la télédétection par satellite, la JECAM⁷ a adopté une définition des terres cultivées annuelles en raison de la nature annuelle des séries temporelles de données d'observation de la Terre :

« du point de vue de la télédétection, les terres emblavées en cultures annuelles désignent les parcelles d'une surface minimale de 0,25 ha (et d'une largeur minimale de 30 m) qui ont été semées/plantées et qui pourront faire l'objet d'au moins une récolte dans les douze mois suivant la date de semis. Ces terres produisent une couverture herbacée⁸ et sont parfois associées à une végétation arborée ou ligneuse⁹. »

La notion de terres cultivées annuelles est plus précise d'un point de vue cartographique et permet de tenir compte de l'évolution de la couverture du sol d'une année sur l'autre (due par exemple à l'extension ou à l'abandon de certaines terres cultivées).

Il est important de signaler que la définition adoptée par la JECAM fait également figurer le concept d'unité cartographique minimale (UCM), qui désigne la plus petite unité à prendre en compte dans le processus de cartographie. Le processus de cartographie utilisé pour le projet européen CORINE Land Cover, par exemple, avait initialement fixé une UCM de 25 ha : de ce fait, la base de données ne tenait compte que des éléments paysagers d'une surface supérieure à 25 ha. Une telle spécification risque d'écarter les petits champs éparpillés dans un paysage urbain ou forestier, et donc d'introduire un biais important dans la carte des terres agricoles ainsi créée.

2.2.3. Autres méthodes de caractérisation du sol

D'autres initiatives, axées sur des objectifs bien ciblés, s'attachent à réaliser des produits portant sur une seule classe de couverture du sol ou des masques binaires. La superficie mondiale des terres cultivées a ainsi été calculée à partir d'une série temporelle pluriannuelle d'images MODIS d'une résolution de 250 m utilisant un ensemble de 39 indicateurs afin de rendre compte de la phénologie des terres cultivées et d'obtenir une couche de probabilité de la présence de terres cultivées pixel par pixel au niveau mondial à l'aide d'un algorithme de classification mondiale par arbre de décision (Pittman *et al.*, 2010). Hansen *et al.* (2013) ont réalisé une carte des sols nus et non nus à l'échelle mondiale en traitant l'ensemble des archives des données enregistrées par les satellites Landsat depuis 2000 pour leur produit consacré à la couverture arborée. Toutes ces initiatives présentent l'avantage de

⁷ www.jecam.org.

⁸ La végétation herbacée exprimée sous la forme fCouverture (fraction du sol couverte par de la végétation vivante) est censée atteindre au moins 30 %, tandis que la couverture arborée ou ligneuse (d'une hauteur supérieure à 2 m) ne doit normalement pas dépasser une fCouverture de 20 %.

⁹ Il existe trois exceptions notoires à cette définition. La première concerne la canne à sucre et le manioc, qui figurent dans la classe des terres cultivées bien que leur cycle de végétation soit plus long et qu'ils ne soient pas plantés annuellement. Ensuite, les petites parcelles individuelles de légumineuses par exemple ne satisfont pas au critère de taille minimale de la définition des terres cultivées. Elles doivent toutefois être comprises dans les terres cultivées lorsqu'elles sont prises collectivement sous forme de champ hétérogène continu. Le troisième cas porte sur les cultures de serre, qui ne peuvent être contrôlées par télédétection et qui doivent donc être exclues de la définition.

proposer un produit cartographique axé sur la classe de couverture du sol qui nous intéresse. Leur inconvénient majeur, en revanche, est que rien n'a été fait pour assurer une complémentarité entre les différents produits, ce qui peut entraîner d'importantes incohérences spatiales ou incompatibilités sémantiques.

Le programme européen Copernicus a produit cinq couches d'information distinctes (appelées « couches haute résolution »), consacrées respectivement aux forêts, aux prairies, aux surfaces en eau permanentes, aux sols imperméabilisés (essentiellement des zones construites) et aux zones humides. Cette approche repose spécifiquement sur un élément paysager donné, ce qui simplifie le processus d'interprétation des données. Toutefois, la complémentarité spatiale entre ces couches produites séparément ne peut être assurée que dans un deuxième temps moyennant une étape supplémentaire, et il n'est pas encore possible d'obtenir une carte complète de la couverture du sol.

Réduire la diversité des caractéristiques terrestres à un nombre limité de classes prédéfinies par la typologie de la couverture du sol est toujours un exercice difficile dans les paysages complexes. Comme l'indiquent Defourny et Bontemps (2013), une stratégie alternative a été proposée afin de décrire la végétation en termes de champs continus (Smith *et al.*, 1990 ; DeFries *et al.*, 1995). Cette approche, également appelée « approche des champs continus », consiste à cartographier la fraction respective des éléments essentiels de la surface terrestre, afin par exemple de représenter le pourcentage de sol nu et de couverture herbacée et arborée correspondant à chaque pixel (Hansen *et al.*, 2005). Les champs continus sont généralement obtenus à l'aide d'un algorithme d'arbre de régression qui établit une correspondance entre un jeu de données d'apprentissage dense et continu portant sur l'ensemble des couvertures végétales existantes et un vaste ensemble d'indicateurs multitemporels fondés sur un jeu complet de séries temporelles de données. Les propriétés des champs continus de végétation présentent un certain avantage par rapport aux classifications discontinues traditionnelles : ils permettent en effet une représentation plus précise des zones hétérogènes, car chaque pixel illustre la proportion des différents types de couvertures qu'il représente. À cet égard, cette approche est attrayante et pertinente pour de nombreux paysages naturels et semi-naturels. En revanche, ces produits se sont avérés plutôt difficiles à valider, parce qu'il n'existe pas de jeu de données de référence approprié et que c'est aux utilisateurs qu'il appartient de définir des seuils spécifiques en vue de convertir ces produits continus en cartes exploitables.

De même, l'extraction de variables biophysiques à partir des séries temporelles de données satellites aboutit également à une description quantitative de la surface terrestre grâce à la régression empirique ou à la méthode d'inversion des modèles physiques. En effet, les produits de télédétection correspondant à la surface foliaire, à la fraction de rayonnement photosynthétiquement actif absorbé par la végétation (FAPAR, de l'anglais *Fraction of Absorbed Photosynthetically Active Radiation*), à l'albédo, etc. fournissent des estimations directes de variables indiscutables pouvant également être mesurées sur le terrain. L'évolution saisonnière de ces variables biophysiques permet de caractériser la surface terrestre, et peut parfois même être interprétée pour déterminer les classes agricoles de couverture du sol qui nous intéressent, ou directement utilisée à des fins de stratification. Toutefois, la possibilité d'identifier ces variables biophysiques à partir d'images satellites à haute résolution, gratuites et librement accessibles, telles que celles fournies par Sentinel-1 et Sentinel-2, ne s'est développée que très récemment. Les séries temporelles d'une résolution spatiale relativement grossière (250 m à 1 km) disponibles depuis plusieurs années ne sont utiles à des fins de stratification que dans certains paysages agricoles, constitués soit de champs très vastes (comme c'est généralement le cas en Argentine, aux États-Unis, en Russie, en Ukraine, etc.) soit de paysages uniformes et non fragmentés composés de nombreux champs exigus mais similaires, cultivés suivant un même calendrier des cultures (grande plaine du nord de la Chine ou rizières irriguées, par exemple).

2.3. RÉALISATION DE CARTES DE COUVERTURE DU SOL

Compte tenu de la disponibilité croissante de données issues des satellites d'observation de la Terre, la cartographie de la couverture du sol à partir de signatures spectrales et temporelles est progressivement devenue l'une des méthodes les plus populaires pour décrire la surface terrestre. Différentes régions du monde ont été cartographiées et catégorisées à plusieurs reprises, soit de façon régulière par des organismes nationaux, soit par des programmes internationaux (Africover, SERVIR, CORINE Land Cover, etc.), et plusieurs cartes mondiales de couverture du sol ont été publiées avec des résolutions allant de 30 m à 500 m.

Le chapitre 1 du présent manuel examine en détail les différentes sources de données, leur évolution et l'apparition de nouveaux environnements de traitement. La possibilité d'accéder à des outils de calcul informatique ultrarapide en ligne et l'accès libre et gratuit à des séries temporelles de données prêtes à être analysées facilitent grandement la réalisation de cartes de couverture du sol à l'échelle nationale, devenue par conséquent bien plus accessible qu'il y a seulement quelques années. Il est toutefois important de souligner les divergences conceptuelles et méthodologiques entre les méthodes de classification existantes, conçues pour examiner les scènes une par une ou pour traiter de façon interactive un certain nombre d'images sur une zone d'étude limitée (échelle locale à nationale), et les chaînes de traitement automatisé capables d'exploiter l'ensemble des images acquises sur une importante zone d'étude (échelle nationale à régionale, voire continentale).

La présente section examine systématiquement les principaux éléments qui interviennent dans la cartographie de la couverture du sol, de façon à évaluer la qualité des cartes de couverture du sol existantes et à favoriser une conception appropriée des initiatives de cartographie de la couverture du sol à l'échelle nationale. L'objectif est de proposer une vue d'ensemble de tous les aspects à prendre en compte, plutôt que d'expliquer les différentes méthodes de télédétection. Les chapitres suivants décriront plus en détail ces éléments, en se concentrant sur les applications particulières, telles que la cartographie des types de cultures.

2.3.1. Données de télédétection

Le choix de la source de données limite fortement la qualité et le niveau de détail du produit cartographique. Outre les questions de coût et d'accessibilité des données, étudiées dans le chapitre 1, le compromis optimal en matière de cartographie de la couverture du sol doit se fonder sur quatre grands critères, présentés ci-dessous et synthétisés dans le tableau 1 :

1. La résolution spatiale, ou plus précisément la distance d'échantillonnage au sol (GSD, de l'anglais *Ground Sampling Distance*) de l'instrument caractérisé par sa fonction d'étalement du point, définit la plus petite caractéristique du terrain pouvant être détectée puis, le cas échéant, cartographiée. Il est primordial que la résolution spatiale soit inférieure à la taille de la plupart des parcelles agricoles afin d'obtenir un nombre suffisant de pixels couvrant uniquement des terres cultivées. Toutefois, la résolution spatiale la plus adaptée pour cartographier des terres cultivées dépend également de la fragmentation du paysage, de la diversité des types de cultures et de leur répartition dans l'espace ; ainsi, des champs de faible surface mais adjacents et similaires (consacrés à un même type de culture) peuvent être considérés comme un seul champ très vaste si leur développement végétatif est relativement synchronisé. Bien qu'il puisse être intéressant de discerner des éléments paysagers linéaires tels que des haies, des alignements d'arbres ou des routes rurales, ces informations ne sont pas strictement nécessaires du point de vue de la stratification des paysages. Il est important de signaler qu'une meilleure résolution spatiale n'est pas toujours une solution plus adaptée lorsque l'on procède au traitement d'images numériques. En effet, l'amélioration de la résolution spatiale accroît de façon exponentielle le volume de données et les capacités informatiques nécessaires pour les traiter. Passer d'une résolution de 30 m à une résolution de 10 m, par exemple, multiplie le stockage de données par 9. Par ailleurs, s'il est plus facile d'interpréter visuellement une image lorsque sa résolution spatiale permet de saisir des éléments paysagers tels que des arbres ou des maisons, cela ne vaut pas pour les chaînes de traitement des images numériques plus

automatisées. Duveiller et Defourny (2010) approfondissent cette question et proposent un cadre conceptuel visant à ajuster la résolution spatiale en fonction d'un élément du paysage agricole donné.

2. La fréquence des observations valides et sans nuages constitue le deuxième critère le plus important pour cartographier précisément les paysages agricoles. En effet, la dynamique saisonnière des terres agricoles ne peut être appréhendée qu'à l'aide d'une série temporelle dense d'observations sans nuages. Les données acquises au cours d'un même cycle de revisite par un système d'observation de la Terre peuvent ainsi présenter des densités temporelles d'observations exploitables très différentes en fonction de la présence de nuages. La discrimination des différents types de terres agricoles doit donc prendre en compte la fréquence effective d'images valides pour la saison de végétation, et non uniquement la capacité de revisite (Whitcraft *et al.*, 2015). Dans les paysages hétérogènes, obtenir l'ensemble du profil saisonnier de la signature des champs permet de faire la distinction entre les différentes terres cultivées, notamment entre les prairies permanentes, les pelouses naturelles et les champs de céréales, qui peuvent avoir une apparence très similaire pendant une longue période de la saison de végétation. Il existe des constellations de satellites équipés d'instruments interférométriques, qui peuvent constituer une solution compétitive pour accroître la fréquence d'observation en utilisant une résolution spatiale élevée. Dans les régions nuageuses, la meilleure solution pour obtenir une série temporelle d'observations dense consiste peut-être à utiliser des capteurs SAR tels que Sentinel-1 ou Radarsat-2, car les bandes hyperfréquences ne sont pas affectées par la plupart des perturbations atmosphériques ni par les nuages.
3. Le nombre de bandes spectrales et leur position sur le spectre électromagnétique sont un autre critère important, non seulement pour la discrimination de la couverture du sol, mais également pour la détection et la correction de la brume atmosphérique et le filtrage efficace des nuages et de leurs ombres. À l'heure actuelle, la plupart des satellites d'observation de la Terre disposent de bandes visibles et proches infrarouge étroites ; celles-ci sont idéalement complétées par des bandes SWIR, également très utiles pour discriminer les types de couvertures du sol. La détection des nuages, en revanche, reste problématique. À l'exception de MODIS, Worldview-3 et Sentinel-3 OLCI, aucun des capteurs existants ne possède, sur sa propre plateforme, de bandes spectrales adéquates pour gérer l'ensemble des aérosols et des types de nuages, afin d'éviter de les confondre avec de la neige ou de la glace. Certains satellites extrêmement populaires tels que Landsat-8 et Sentinel-2 disposent d'un ingénieux jeu de bandes (et notamment d'une bande cirrus efficace) permettant de filtrer les perturbations atmosphériques. Leur utilisation opérationnelle reste toutefois limitée dans certaines régions. De nombreux autres capteurs s'avèrent problématiques dès lors que l'on souhaite obtenir des séries temporelles cohérentes sur des superficies importantes, car les perturbations atmosphériques des signaux sont difficiles à détecter et à corriger. Ces corrections sont pourtant nécessaires si l'on cherche à combiner des valeurs de pixels enregistrées à différentes dates d'observation ou par différents capteurs de façon à obtenir une seule image (image composée ou mosaïque lorsque l'on ne dispose que d'une seule observation par pixel) qui soit ininterrompue et sans nuages. Cependant, certaines observations sans nuages réalisées sur des superficies importantes et bien réparties au cours de la saison peuvent faire l'objet d'une classification sans qu'il soit nécessaire de procéder à une correction atmosphérique, et pourraient s'avérer suffisantes pour cartographier la couverture du sol.
4. Les capteurs à large fauchée couvrent des zones très importantes en un seul passage (290 km pour Sentinel-2, 650 km pour les satellites DMC, et plus de 1 200 km pour les instruments plus grossiers tels que MODIS, Sentinel-3 OLCI et PROBA-V). La largeur de la fauchée est donc également un critère à prendre en compte pour une application à grande échelle. Même si les images de taille plus réduite peuvent être assemblées de façon à former une mosaïque, il convient de procéder à une correction atmosphérique si l'on souhaite en faire une image ininterrompue ou un produit cartographique cohérent sur une superficie importante. Les capteurs ayant une fauchée étroite (environ 50 km en règle générale, voire moins de 20 km pour les satellites à très haute résolution), couvrent généralement des zones étendues en réalisant une mosaïque à partir de petites images enregistrées à partir de différents angles de vue. Compte tenu de ce que l'on appelle la fonction de distribution de la réflectance bidirectionnelle, cette variabilité de l'angle de vue peut induire des différences de signature spectrale pour une surface de terrain donnée, ce qui rend le processus de classification encore plus délicat.

TABEAU 1. CRITÈRES APPLICABLES À LA SÉLECTION DES SOURCES DE DONNÉES DE TÉLÉDÉTECTION APPROPRIÉES.

Critères	Résolution spatiale	Fréquence des obs. valides	Résolution spectrale	Largeur de la fauchée
Raison	Mise en correspondance avec la taille des éléments paysagers	Profil temporel tout au long de la saison de végétation	Discrimination de la couverture du sol Filtrage des nuages	Image ininterrompue sur une superficie étendue
Variable en fonction de	Taille, distribution et fragmentation des parcelles	Persistance de la couverture nuageuse	Diversité de la couverture du sol Fréquence de brume et de couverture nuageuse	Taille du pays

La sélection de la source de données s'appuie également sur un autre critère, à savoir le niveau et la qualité du prétraitement des données. Les fournisseurs de données livrent de plus en plus souvent des images prêtes à l'emploi ou prêtes à être analysées, en se chargeant notamment de l'étalonnage radiométrique de pointe, de l'orthorectification, de la correction atmosphérique et du filtrage des nuages et de leur ombre. Dans le cas contraire, une chaîne de traitement préalable doit être mise en place afin de convertir l'énergie radiative enregistrée dans le cadre de la réflectance multibande au sommet et au bas de l'atmosphère. Les seuls avantages d'un traitement préalable en interne sont que certains algorithmes peuvent être affinés en fonction de la zone d'étude (disponibilité locale d'informations atmosphériques plus pertinentes, meilleur modèle altimétrique numérique, etc.) et qu'il est possible d'intégrer certaines étapes de traitement plus complexes, telles que la correction topographique, utile dans les régions où d'importantes superficies sont cultivées dans des zones vallonnées ou montagneuses.

2.3.2. Collecte de données *in situ*

Outre les images des satellites d'observation de la Terre ou les séries temporelles de données, la cartographie de la couverture du sol utilise toujours au moins une autre source d'informations pour faciliter le processus de classification (soit pour un entraînement préalable de l'algorithme de classification, soit pour labéliser, *a posteriori*, les classes résultantes de l'opération de classification) et pour évaluer la qualité des cartes.

La collecte de données *in situ* répond essentiellement à deux grands objectifs : l'étalonnage de l'algorithme de classification et la validation du résultat de la classification (figure 1). Le jeu de données d'étalonnage permet d'entraîner les algorithmes de classification des données d'observation de la Terre afin de générer par exemple une carte de couverture du sol, un masque des terres cultivées ou même une carte des types de cultures. Les exigences à respecter, pour collecter un jeu de données approprié pour l'étalonnage, sont examinées en détail dans la section 2.2.3.

Au lieu de se servir de données *in situ*, Matton *et al.* (2015) et Desclée *et al.* (2006) ont élaboré des méthodes spécifiques visant à utiliser des cartes de couverture du sol obsolètes en tant qu'informations préalables afin d'entraîner les algorithmes de classification et produire respectivement un masque des terres cultivées et un masque des terres forestières. Les algorithmes de classification non supervisée, quant à eux, ne nécessitent pas de jeu de données d'étalonnage, mais un jeu de données d'étiquetage, pouvant se composer soit de données *in situ*, soit de cartes obsolètes, une méthode d'abord mise en œuvre dans le cadre du projet GlobCover (Defourny *et al.*, 2006).

En revanche, le jeu de données de validation doit nécessairement être constitué de données de référence de qualité, qui seront utilisées en tant que « réalité de terrain » afin d'évaluer la précision d'une carte de couverture du sol. La validation ne doit en aucun cas être confondue avec la comparaison croisée des cartes (un processus qui quantifie les divergences entre les cartes, mais qui ne permet pas d'attribuer rigoureusement les éventuelles erreurs à une carte plutôt qu'à une autre).

La collecte de données sur le terrain est une activité qui mobilise des ressources conséquentes et exige un temps considérable, en particulier si l'on cherche à couvrir des superficies importantes. Bien que la conception des opérations de terrain soit censée être définie par la théorie des sondages, il arrive le plus souvent que les contraintes logistiques et celles liées aux ressources conduisent à considérer un compromis optimal entre performance et rentabilité. Quelques figures et principes de base visant à orienter les éventuels producteurs de cartes et à évaluer la qualité des cartes réalisées sont proposés plus loin dans cette section. Dans certains pays, des systèmes actualisés d'information sur les parcelles agricoles constituent un jeu de données de référence complet et prêt à l'emploi. Dans d'autres, les enquêtes sur le terrain peuvent s'avérer impossibles pour des raisons de sécurité ou faute de pouvoir accéder physiquement au territoire.

Les jeux de données d'apprentissage et de validation peuvent également être collectés à l'aide de délimitations à l'écran fondées sur l'interprétation visuelle et interactive de compositions colorées multispectrales. Les images à très haute résolution spatiale mises en ligne par Google Earth, Bing et autres portails géographiques similaires sont souvent utilisées pour numériser des polygones plutôt que pour recueillir des données de terrain. Les dates d'acquisition des images publiées sur ces portails géographiques peuvent être très variables et ne correspondent pas nécessairement à une année d'observation donnée. Cette solution est donc relativement efficace pour les types de couvertures du sol les plus stables, mais elle ne doit pas être utilisée pour déterminer les types de cultures, par exemple, ceux-ci étant susceptibles de changer chaque année. Comme l'a montré la validation des cartes réalisées dans le cadre du projet GlobCover (Defourny *et al.*, 2012b), la fiabilité d'une interprétation fondée uniquement sur des images THR dépend en grande partie de leur date d'observation, un facteur qui échappe au contrôle de la personne chargée de ce processus. L'efficacité de l'interprétation des images THR peut être accrue par les profils temporels du NDVI fournis pour chaque pixel correspondant : c'est ce que propose l'interface de l'initiative CCI Land Cover de l'ESA, qui couvre le monde entier¹⁰. Plus récemment, plusieurs initiatives de production participative ont démontré qu'il était possible de recourir à la mobilisation citoyenne pour collecter d'importants volumes de données sur l'agriculture ou la couverture du sol, comme on l'a vu sur des programmes tels que Google Earth. Geowiki (Fritz *et al.*, 2015) et Collect Earth¹¹ sont des applications en ligne qui permettent à un nombre incroyable de particuliers non spécialisés de se livrer à des interprétations visuelles ; elles utilisent parfois la ludification et partent du principe qu'une grande quantité d'informations finit en quelque sorte par garantir des informations de qualité. Ce domaine de recherche est très actif, et un certain nombre d'applications expérimentales menées à l'échelle internationale se sont déjà avérées plus ou moins fructueuses.

Étalonnage

L'objectif de l'échantillonnage destiné au jeu de données d'étalonnage consiste à tenir compte de toutes les situations existant dans la zone d'étude (territoire national, par exemple) afin de représenter l'éventail des différentes signatures spectrales et temporelles possibles. Si l'on souhaite produire un masque des terres agricoles délimitant les terres cultivées et non cultivées, tous les types de couvertures du sol doivent figurer dans le jeu de données d'étalonnage. Dans le cas contraire, les types de couvertures du sol non répertoriés seraient affectés de manière anarchique à la classe des terres cultivées ou à celle des terres non cultivées.

Les échantillons sont ciblés de façon à établir les liens nécessaires entre un type de paysage donné (zone humide, zone urbaine, surface en eau, etc.) et l'éventail existant de signatures spectrales et temporelles correspondantes. Contrairement au processus de validation, qui doit respecter un plan d'échantillonnage rigoureux, différents types de stratégies d'échantillonnage sont acceptables pour l'étalonnage, du moment qu'elles sont réparties de manière équitable sur le territoire afin de mettre en évidence la variabilité de chaque type de couverture du sol.

Pour rendre compte de cette diversité et des variations graduelles des conditions agroclimatiques souvent observées sur des superficies importantes (territoire national, par exemple), il est recommandé de stratifier le

¹⁰ <http://maps.elie.ucl.ac.be/CCI/viewer/>.

¹¹ <http://www.mdpi.com/2072-4292/8/10/807/html>.

paysage en fonction des interactions entre les types d'agro-écosystèmes, de la répartition des types de cultures, des calendriers de culture et du contexte biophysique général (plaine inondable, zones montagneuses, etc.). Cette forme de stratification, que nous appellerons ici « zonage agro-écologique », existe déjà dans de nombreux pays. Il convient en principe d'étudier un nombre limité de zones (généralement 3 à 12 zones pour un pays d'une superficie d'environ 500 000 km²) afin qu'il y ait suffisamment d'échantillons d'apprentissage pour toutes les classes de chaque strate et que l'échantillonnage reste raisonnable.

Comme l'expliquent les directives de la JECAM (2014) au sujet des classes non agricoles de couverture du sol, une interprétation visuelle des images publiées sur les portails géographiques peut être très efficace pour définir au moins 15 échantillons d'apprentissage par classe et par strate, en particulier pour les régions reculées et les zones présentant une faible variation de la couverture du sol. Comme nous l'avons expliqué ci-dessus, l'interprétation visuelle des photographies aériennes ou des images THR à jour et disponibles est avantageusement remplacée par des informations temporelles telles que les profils du NDVI pour rendre compte de la diversité des types non agricoles de couvertures du sol.

Comme indiqué plus haut, l'ensemble des différentes classes agricoles (différents types de cultures, prairies, pâturages, cultures permanentes, etc.) doit également être décrit dans le jeu de données d'étalonnage, bien que ces informations soient susceptibles de varier d'une année sur l'autre. Une enquête sur le terrain correspondant à la saison agricole du jeu de données de télédétection à traiter est donc le meilleur moyen de délimiter un ensemble d'échantillons d'apprentissage pour chaque strate (environ 30 échantillons par culture présentant une superficie supérieure à 5 pixels × 5 pixels, par exemple), de façon à rendre compte de la diversité des signatures d'une culture donnée. La qualité de la classification s'améliore souvent avec celle des échantillons. Si l'objectif consiste à cartographier uniquement les terres cultivées en vue d'obtenir un masque pour cette seule classe, il peut être suffisant de délimiter toutes les différentes signatures de terres cultivées directement sur les compositions colorées afin de définir différentes classes spectrales, que l'on regroupera ensuite en une seule classe de terres cultivées. Dans les pays qui connaissent une forte saisonnalité agricole, la sélection d'une période d'observation appropriée pourra considérablement simplifier l'échantillonnage des terres agricoles : tous les sols nus à une période donnée de l'année pourraient par exemple correspondre à l'ensemble des types de cultures annuelles.

Il convient de signaler que l'échantillonnage exhaustif des types de cultures, souvent nécessaire pour cartographier la couverture du sol, peut également servir à réaliser une carte des types de cultures. Dans ce cas, il peut s'avérer plus efficace de produire en premier lieu un masque des terres agricoles, puis de déterminer les types de cultures uniquement pour les pixels appartenant à ce masque, plutôt que d'opérer une nouvelle classification de l'ensemble de la zone d'étude.

L'une des différences majeures entre un masque des terres cultivées et une carte des types de cultures réside dans le caractère actuel de la carte. Si la superficie des terres cultivées varie relativement lentement, ce qui rend un masque des terres cultivées susceptible de conserver sa pertinence pendant plusieurs années, la répartition des types de culture peut changer considérablement chaque année et n'est donc valable que pour l'année d'observation. Cela vaut également pour le jeu de données utilisé pour l'étalonnage.

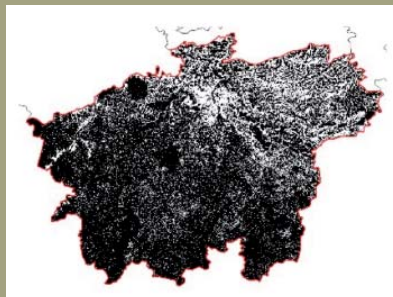
ENCADRÉ 1.MALI : STRATIFICATION ET ÉCHANTILLONNAGE EN VUE DU PROJET SENTINEL-2 POUR LE SUIVI DE L'AGRICULTURE, SYSTÈME SOUTENU PAR L'ESA.

Au Mali, le zonage agro-écologique fondé sur l'atlas du Projet Inventaire des ressources terrestres (PIRT), qui couvre plus de 447 000 km², délimite cinq strates distinctes, notamment le bassin cotonnier et la zone de production du riz. Pour les terres non cultivées, 12 types de couvertures du sol doivent être retenus séparément pour les cinq strates, soit un total d'environ 1 125 échantillons d'apprentissage pour les terres non cultivées (15 échantillons × 15 types de couvertures du sol × 5 strates).

La répartition des échantillons au sein d'une strate est censée couvrir tout l'éventail attendu des différentes typologies de couverture du sol ; ce travail a été réalisé à l'aide d'une interprétation Google Earth d'images de l'année en cours ou de l'année précédente, ainsi que du profil du NDVI correspondant si nécessaire.

Afin de rendre compte de leur diversité de façon à pouvoir les cartographier, chacun des principaux types de cultures doit faire l'objet d'un échantillonnage de grande ampleur au sein de chaque strate. En effet, la discrimination entre des types de cultures similaires nécessite en règle générale un plus grand nombre de données d'apprentissage que lorsque l'on travaille sur des classes de couverture du sol contrastées. Par conséquent, l'objectif de 1 800 échantillons collectés, soit 30 échantillons × 12 types de cultures × 5 strates, est probablement un minimum. Notons que ces 12 grands types de culture constituent une simplification des 40 types de cultures utilisés par la Cellule de planification et de statistiques (CPS) malienne.

Il convient de mentionner que pour leurs statistiques sur les cultures, les 275 agents de la CPS collectent chaque année 30 000 échantillons géoréférencés, correspondant aux champs cultivés répartis sur l'ensemble du territoire national. Une fois soumises à un contrôle qualité, ces informations ont pu servir de données d'apprentissage en vue de la cartographie des terres cultivées réalisée par le système Sentinel-2 pour le suivi de l'agriculture. Pour la strate située au sud du pays, la combinaison des échantillons de terres non cultivées délimités sur Google Earth et des données fournies par la CPS ont permis de cartographier les terres cultivées (en blanc) et non cultivées (en noir) avec une résolution de 10 m, en utilisant Sentinel-2 avec une précision globale de 94 %. Cette carte permettra sans nul doute d'optimiser le plan d'échantillonnage de la CPS à l'avenir



Validation

L'objectif d'un jeu de données de validation est de fournir une estimation statistiquement rigoureuse de la précision de la carte produite en s'appuyant sur une source indépendante d'informations de référence. La précision est évaluée en mesurant le degré de concordance entre la carte produite et le jeu de données de validation.

Plusieurs indicateurs de la précision de la classification peuvent être calculés à l'aide de la matrice de confusion, qui est un tableau de contingence à deux dimensions dans lequel on met en correspondance la classe générée (cartographiée) et la classe observée pour chaque échantillon.

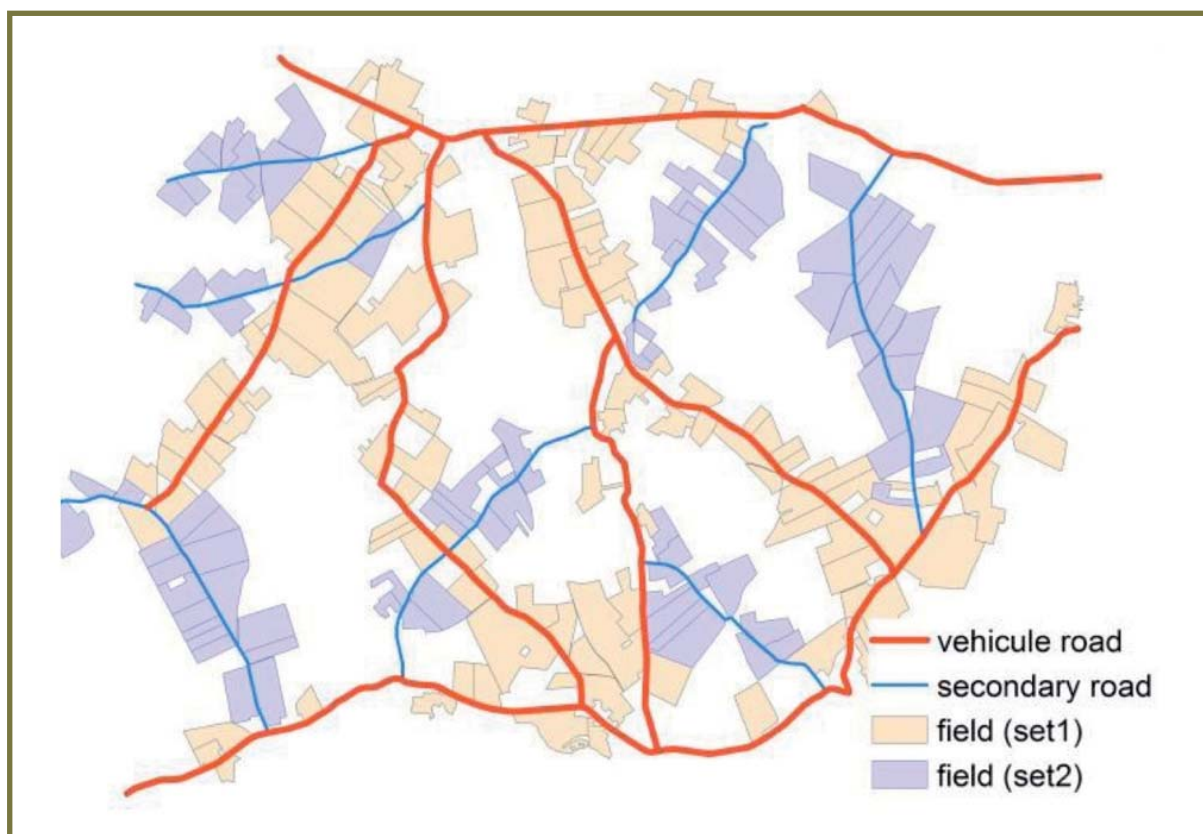
Les jeux de données d'apprentissage et de validation doivent être totalement différents et spatialement indépendants. Lorsque les données *in situ* sont réparties en deux sous-ensembles, le premier est généralement utilisé pour entraîner l'algorithme, tandis que le second sert souvent à évaluer la qualité de la carte produite. En réalité, cette méthode de validation ne doit pas être considérée comme une solution appropriée, car le deuxième jeu de données

est en général spatialement corrélé au jeu de données d'apprentissage. Une telle approche permet d'évaluer les performances en matière d'étalonnage de l'algorithme, mais pas la qualité de la carte produite dans son ensemble. Le jeu de données de validation ne peut être jugé adéquat que si les échantillons ont été collectés sur la totalité de la zone d'étude puis répartis aléatoirement dans les sous-ensembles destinés à l'étalonnage et à la validation.

La qualité de l'estimation de la précision est déterminée par le nombre d'observations, le plan d'échantillonnage, la méthodologie d'évaluation sur le terrain et la fiabilité de la source des données de validation. Selon la théorie des sondages, les méthodes les plus à même de fournir des informations indépendantes et d'éviter les biais spatiaux dus à la localisation des observations de terrain sont l'échantillonnage aléatoire simple et l'échantillonnage aléatoire systématique. De même, un plan d'échantillonnage fondé sur des tirages dans des grappes d'échantillons, sélectionnées suivant une approche systématique à partir d'un point de départ aléatoire garantit la représentativité et la bonne répartition spatiale du jeu de données sélectionné. Ces plans d'échantillonnage, relativement théoriques, sont difficiles à appliquer sur le terrain dans un contexte opérationnel, en raison de contraintes pratiques d'accessibilité des zones échantillonnées et de ressources limitées.

Une stratégie d'échantillonnage à deux degrés est souvent adoptée afin d'assurer une bonne répartition et une sélection relativement aléatoire des échantillons. Des unités primaires d'échantillonnage (UPE) définies à partir des données auxiliaires (régions administratives, en règle générale) sont tout d'abord sélectionnées de façon aléatoire dans les différentes strates en fonction de leur superficie de terres cultivées (probabilité d'inclusion pondérée en fonction de la superficie des terres cultivées). La constitution d'une base de sondage au second degré peut être effectuée dans chaque UPE au moyen d'un « dénombrement » visant à identifier, depuis un véhicule motorisé, des unités secondaires d'échantillonnage (ou USE, d'une dimension minimale de 3 pixels \times 3 pixels) situées le long des routes.

FIGURE 2. ÉCHANTILLONNAGE PAR ENQUÊTE SUPERFICIELLE



Source : JECAM, 2014.

En termes de méthodologie d'évaluation sur le terrain, cette stratégie permet à un agent de collecte de données d'appréhender facilement et rapidement la diversité des cultures dans l'ensemble des champs visibles (série 1 de la figure 2). Cette méthode vise essentiellement à identifier de longs transects dans l'UPE en sélectionnant un ensemble de routes appropriées. Cependant, il est recommandé de compléter cette stratégie d'échantillonnage par des transects supplémentaires (série 2) en utilisant des routes secondaires et des sentiers de façon à réduire le biais spatial introduit par un échantillonnage réalisé uniquement le long des routes. Cette recommandation est primordiale lorsque la répartition des terres cultivées ou des types de cultures dépend de leur distance par rapport aux routes principales. Plusieurs transects orientés de différentes façons (nord-sud et est-ouest, par exemple) garantissent une couverture complète de la zone d'étude, tandis qu'un échantillonnage réalisé le long des routes secondaires permet de compléter le jeu de validation par des données moins biaisées.

Pour la validation des cartes de couverture du sol, la taille minimale de l'échantillonnage doit être de 30 pour la classe la moins représentée, et environ une observation pour 100 à 150 km² pour les superficies importantes. Une taille d'échantillonnage plus élevée permettrait d'améliorer la qualité de l'estimation de la précision. Étant donné que les classes de couverture du sol moins fréquentes ou marginales ne sont pas suffisamment représentées dans un plan d'échantillonnage global, il convient d'adopter une stratégie d'échantillonnage spécifique afin de cibler de ces types spécifiques de couvertures du sol tout en veillant à constituer l'ensemble de l'échantillon de manière aléatoire. Lorsque l'on dispose d'images THR à jour, un échantillonnage aléatoire systématique de taille importante peut également permettre, grâce à l'interprétation visuelle, de collecter suffisamment d'unités correspondant aux types de couvertures du sol les moins répandus. Dans le cas contraire, il serait nécessaire de procéder à un échantillonnage sur le terrain en favorisant les classes concernées.

Dans certains cas, les bases de données existantes peuvent fournir des données réelles de terrain pour la validation des cartes de couverture du sol. Ainsi, la base de données indépendante LUCAS, mise en place par l'Union européenne (UE) dans le but de fournir des statistiques sur l'occupation et l'utilisation des sols, peut être utilisée pour réaliser une évaluation de la précision des cartes de couverture du sol et des terres cultivées pour l'année correspondante.

Les données de validation sont alors utilisées pour évaluer la précision des cartes en fonction de la concordance observée entre la cartographique issue de la classification et les données de validation supposées refléter la réalité de terrain. La manière la plus courante de déterminer la précision d'une carte consiste à analyser la matrice de confusion, qui est une matrice de cooccurrence carrée qui associe le nombre d'échantillons correspondant à une classe de couverture du sol donnée avec des données de validation. Les valeurs diagonales représentent la fréquence de concordance entre les données de validation et le produit cartographique, tandis que les valeurs non diagonales représentent les erreurs.

TABLEAU 2. MATRICE DE CONFUSION DANS LAQUELLE n_{ij} REPRÉSENTE LE NOMBRE D'ÉCHANTILLONS DE VALIDATION CORRESPONDANT SUR LA CARTE À LA CLASSE DE COUVERTURE DU SOL / ET AUX INFORMATIONS DE VALIDATION j .

Données de validation						
Carte produite	Terres cultivées	Terres Forestières	Prairies	Zones urbaines	Terres humides	Total
Terres cultivées	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	$n_{1.}$
Terres Forestières	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}	$n_{2.}$
Prairies	n_{31}	n_{32}	n_{33}	n_{34}	n_{35}	$n_{3.}$
Zones urbaines	n_{41}	n_{42}	n_{43}	n_{44}	n_{45}	$n_{4.}$
Terres humides	n_{51}	n_{52}	n_{53}	n_{54}	n_{55}	$n_{5.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	$n_{.5}$	$N_{..}$

Parmi 14 indicateurs de précision portant sur les classes et 20 portant sur les cartes, Liu *et al.* (2007) recommandent d'utiliser en priorité la précision de l'utilisateur, la précision du producteur et la précision globale. Pour les cartes binaires (masque des terres cultivées, par exemple), cette dernière dépend dans une large mesure de la proportion respective occupée par chaque classe dans le jeu de données de validation. Dans ce cas, il sera plus instructif d'utiliser le F-score, un indicateur de précision adopté récemment.

La précision globale (PG) correspond au rapport entre le nombre d'échantillons correctement classés et le nombre total (N) d'échantillons de validation. Pour une carte de couverture du sol, l'objectif standard en matière de précision globale est généralement de 85 %. Dans certains cas, cette précision peut atteindre 90 % pour des cartes de couverture du sol simples comptant un nombre de classes très limité.

$$PG (\%) = (100 \times \sum_{k=1}^q n_{kk}) / N \quad (\text{équation 1})$$

La précision de l'utilisateur (PU) pour une classe de couverture du sol i correspond au rapport entre le nombre d'échantillons correctement affectés à cette classe et l'ensemble des échantillons affectés à cette classe.

$$PU_i (\%) = 100 \times \frac{n_{ii}}{n_{i+}} \quad (\text{équation 2})$$

La précision du producteur (PP) pour une classe i correspond au rapport entre le nombre d'échantillons correctement classés et l'ensemble des échantillons appartenant effectivement à cette classe, selon les données de validation.

$$PP_i (\%) = 100 \times \frac{n_{ii}}{n_{+i}} \quad (\text{équation 3})$$

Le F-score est calculé en combinant la précision du producteur et celle de l'utilisateur pour une classe de couverture du sol i :

$$F\text{-score}_i = (2 \times PP_i \times PU_i) / (PP_i + PU_i) \quad (\text{équation 4})$$

2.3.3. Traitement des images et réalisation de cartes

L'extraction d'informations relatives à la surface terrestre à partir des données de télédétection dépend d'une série de processus complexes, car la radiance mesurée par les capteurs en watt/m².str ne permet pas de déduire directement la couverture du sol. Plusieurs systèmes de cartographie opérationnelle employés par le passé se fondaient sur l'interprétation visuelle interactive sur écran d'une ou deux images acquises à certaines périodes de l'année, et s'appuyaient essentiellement sur l'interprétation des experts. Cette approche a été progressivement facilitée par des outils de traitement d'images utilisés de façon interactive ou appliqués une fois pour toutes. En passant d'une sélection très rigoureuse des meilleures images au traitement de séries temporelles, le traitement numérique tend à réduire la main-d'œuvre nécessaire au processus de gestion des données afin de concentrer l'intervention humaine interactive sur les étapes les plus critiques.

La réalisation de cartes de couverture du sol comprend systématiquement une séquence d'étapes principales de traitement. Pour chacune de ces étapes, plusieurs choix conceptuels et algorithmiques sont possibles. Comme l'ont montré Waldner *et al.* (2016), la précision des masques de culture varie davantage entre deux régions agricoles

qu'entre une méthode de pointe à une autre. Évidemment, certains choix méthodologiques peuvent être plus appropriés que d'autres ; toutefois, en définitive, la qualité et la quantité des données de télédétection utilisées comme données d'entrée et du jeu de données d'étalonnage jouent dans la plupart des cas un rôle encore plus important. La clé de la réussite réside probablement dans l'adéquation des choix méthodologiques adoptés pour une certaine quantité et qualité d'images d'observation de la Terre et de données d'étalonnage *in situ*, suivant les caractéristiques paysagères à cartographier.

Comme l'a illustré la figure 1, quatre grandes étapes se dégagent clairement de la chaîne de production de cartes de couverture du sol : 1) segmentation d'image ; 2) extraction de caractéristiques ; 3) classification ; 4) post-traitement, notamment filtrage et/ou fusion.

Segmentation d'image

Le sol est découpé en pixels par les images satellite, et l'interprétation visuelle sur écran permet de délimiter des formes homogènes. Les deux principaux modèles conceptuels utilisés pour représenter la dimension spatiale du monde sont les images matricielles, composées de pixels, et les images vectorielles, composées d'objets. Lorsque la résolution spatiale est équivalente ou supérieure à la taille des éléments à cartographier, les informations relatives à la couverture du sol sont généralement extraites au niveau des pixels et l'étape de segmentation n'est pas nécessaire. En ce qui concerne les images à haute ou très haute résolution, constituées de pixels bien plus petits que les éléments de couverture du sol, le modèle vectoriel est couramment privilégié et l'image doit être segmentée en objets par des algorithmes de segmentation d'image.

La segmentation d'image regroupe des pixels adjacents en objets spatialement continus en fonction de leurs caractéristiques spectrales et de leur contexte spatial, dans le but d'appréhender les objets terrestres spatialement distincts dignes d'intérêt. L'approche par objets est bien adaptée à l'extraction de textures d'images, dispose d'informations contextuelles intrinsèques qui lui permettent d'éviter l'effet « poivre et sel ¹² » dans le résultat de classification, et favorise une interprétation multiéchelles grâce à une segmentation hiérarchique ou multiniveaux (Radoux et Defourny, 2008). Toutefois, cette étape représente également une source d'erreur supplémentaire par rapport à l'approche par pixels. Comme nous l'avons expliqué plus haut, il est le plus souvent recommandé de procéder à une classification par objets lorsque la taille des pixels est largement inférieure à celle des éléments paysagers. En règle générale, les images qui présentent une résolution métrique ou décamétrique sont souvent segmentées en objets, contrairement des images ayant une résolution hectométrique qui ne subissent pas ce type de traitement. Il existe de très rares chaînes de production fondées à la fois sur une approche par pixels et par objets ; c'est le cas par exemple de la production interactive de la carte de couverture du sol réalisée dans le cadre du projet GlobeLand30 (Jun Chen *et al.*, 2015).

La segmentation d'images peut être effectuée suivant deux approches différentes : les méthodes fondées sur les gradients d'image, qui reposent sur une détection locale des contours (délimitation des bassins versants en fonction du gradient d'intensité, par exemple), et les méthodes par croissance de régions, qui identifient des grappes spatiales de pixels cohérents. L'un des algorithmes de croissance de régions les plus populaires dans le domaine de la télédétection regroupe des objets tant que la variance normalisée des valeurs de pixels au sein de l'objet aggloméré reste inférieure à un seuil donné (Baatz et Schäape, 2000). Outre ce critère d'homogénéité spectrale, l'agglomération des objets peut également être limitée par leur forme, l'objectif étant d'améliorer la correspondance avec les objets spatiaux de couverture du sol. Cet algorithme a été mis en œuvre dans le logiciel commercial eCognition, tandis que les algorithmes « mean-shift » et la délimitation des bassins versants sont utilisés dans le logiciel libre ORFEO Toolbox (<https://www.orfeo-toolbox.org/>) et disponibles dans le logiciel libre Quantum GIS (QGIS) via l'extension SEXTANTE.

12 Tâche ou point de nuances sombres et claires entremêlés

Extraction de caractéristiques

L'étape d'extraction de caractéristiques consiste à calculer, à partir des images ou des séries temporelles d'images de télédétection, les variables les plus discriminantes, qui seront utilisées comme données d'entrée pour l'algorithme de classification. Ces caractéristiques peuvent être de différentes natures : 1) spectrales, telles que la réflectance multibande ou les indices dérivés, notamment l'indice différentiel normalisé de végétation (NDVI) ou tout autre indice relatif à la végétation, à la chlorophylle ou au sol ; 2) temporelles, telles que la valeur minimale, maximale ou l'amplitude d'une variable au cours d'une période donnée ; 3) texturales, telles que le contraste local, l'entropie ou toute autre variable obtenue à partir de la matrice de cooccurrence ; et 4) une variable spatiale ou contextuelle, particulièrement adaptée à l'approche par objets.

À l'heure actuelle, on peut observer trois grandes stratégies dans le domaine de la cartographie de la couverture du sol. Premièrement, les stratégies classiques reposent essentiellement sur les caractéristiques spectrales et, éventuellement, sur certaines caractéristiques temporelles simples fondées sur des séries temporelles du NDVI, vu que ces deux types de caractéristiques sont dans tous les cas la source de toutes les autres. Compte tenu de l'important développement des performances de calcul informatique et de la diffusion des algorithmes d'apprentissage automatique, de nombreux spécialistes de la télédétection considèrent aujourd'hui que « plus on dispose de caractéristiques, mieux c'est » et utilisent des algorithmes de classification pour sélectionner les caractéristiques les plus discriminantes. Troisièmement, les stratégies fondées sur le savoir cherchent à intégrer les connaissances d'experts externes en créant des caractéristiques ad hoc en fonction de l'objectif de classification et en ne conservant que celles jugées significatives par les experts (Lambert *et al.*, 2016).

Classification

L'étape de classification comprend un ou plusieurs processus numériques visant enfin à affecter chaque pixel ou objet à l'une des classes de couverture du sol. Les algorithmes de classification sont très divers mais peuvent être répartis en deux grandes catégories : la classification supervisée, qui utilise un jeu de données d'apprentissage afin d'étalonner l'algorithme au préalable, et la classification non supervisée, qui produit des grappes de pixels qui seront étiquetées *a posteriori* en tant que classes de couverture du sol en fonction des données *in situ* ou des informations auxiliaires. Plus récemment, la classification supervisée s'est enrichie d'étapes préalables très utiles consistant à soumettre les jeux de données d'apprentissage *in situ* à un nettoyage automatique ou à procéder à un apprentissage actif visant à construire un jeu de données d'apprentissage plus efficace, en améliorant de façon itérative les performances du modèle de classification.

Le nombre de méthodes employées pour la classification des images en classes de couverture du sol ne cesse de s'accroître ; leurs forces et leurs faiblesses respectives sont synthétisées dans le tableau 3. Un examen de ces différentes méthodes récemment réalisé par Davidson (2016) est également proposé ci-dessous.

TABLEAU 3. FORCES ET FAIBLESSES DES ALGORITHMES UTILISÉS EN VUE DE LA CLASSIFICATION DES DONNÉES ISSUES D'IMAGES SATELLITE POUR DES SUPERFICIES IMPORTANTES (FONDÉ SUR GÓMEZ ET AL., 2016).

Algorithme	Forces/caractéristiques	Faiblesses
Maximum de vraisemblance (Paramétrique)	<ul style="list-style-type: none"> • Application simple • Facile à comprendre et à interpréter • Prédit la probabilité d'appartenance à une classe 	<ul style="list-style-type: none"> • Paramétrique • Part du principe que les données présentent une distribution normale • Nécessite un important échantillon d'apprentissage
Réseaux de neurones artificiels (Non paramétriques)	<ul style="list-style-type: none"> • Capables de gérer correctement un espace de caractéristiques de grande dimension • Indiquent la probabilité de l'appartenance à une classe • Classification généralement très précise • S'adaptent aux manques de données d'apprentissage : nécessitent moins de données d'apprentissage que les arbres de décision 	<ul style="list-style-type: none"> • Besoin de paramètres pour la conception du réseau • Tendance à surclasser les données • Boîte noire (règles inconnues) • Nécessitent une importante puissance de calcul informatique • Apprentissage lent
Machines à vecteurs de support (Non paramétriques)	<ul style="list-style-type: none"> • Capables de gérer correctement un espace de caractéristiques de grande dimension • Insensibles au phénomène de Hughes • Fonctionnent bien avec des jeux de données d'apprentissage limités • Pas de surclassement 	<ul style="list-style-type: none"> • Besoin de paramètres : régularisation et noyau • Performances médiocres avec un espace de caractéristiques limité • Nécessitent une importante puissance de calcul informatique • Conçus comme un outil binaire, bien qu'il existe des variations
Arbres de décision (Non paramétriques)	<ul style="list-style-type: none"> • Pas besoin de paramètres • Faciles à appliquer et à interpréter • Gèrent les données manquantes • Gèrent des données de différents types (continues, catégoriques) et différentes échelles • Gèrent les relations non linéaires 	<ul style="list-style-type: none"> • Sensibles au bruit • Tendance au surclassement • Pas aussi performants que d'autres outils dans les espaces de caractéristiques de grande dimension • Nécessitent un important échantillon d'apprentissage
Forêts aléatoires (Non paramétriques)	<ul style="list-style-type: none"> • Capables de déterminer l'importance des variables • Fiables en ce qui concerne la réduction de données • Pas de surclassement • Permettent une estimation non biaisée de la précision • Plus précises que les arbres de décision 	<ul style="list-style-type: none"> • Règles de décision inconnues (« boîte noire ») • Nécessitent une importante puissance de calcul informatique • Nécessitent des paramètres d'entrée (#arbres et #variables pour chaque nœud)

Classification par maximum de vraisemblance

Jusqu'à récemment, la méthode de classification par maximum de vraisemblance était la plus couramment utilisée en matière de classification supervisée des données de télédétection (Lu et Weng, 2007 ; Bhatta, 2008 ; Kumar *et al.*, 2016). La règle de décision de cette méthode repose sur la probabilité. Les données d'apprentissage sont ici utilisées pour décrire statistiquement les classes cibles selon leur fonction de densité de probabilité multidimensionnelle. Chaque fonction de densité représente la probabilité selon laquelle la structure spectrale d'une classe est présente dans une région donnée de l'espace spectral multidimensionnel. La signature spectrale de chaque pixel est ensuite affectée à la classe à laquelle il a le plus de chances d'appartenir (Jensen, 1986).

Si le principal avantage de cette méthode est qu'elle laisse à l'analyste une totale maîtrise des classes de couverture du sol à utiliser dans la classification finale, son application est limitée par le fait qu'elle repose sur une distribution gaussienne des données d'entrée, une supposition souvent contredite lorsque l'on utilise des

données multitemporelles aux nombreuses caractéristiques spectrales et distributions multimodales (Gislason *et al.*, 2006 ; Glanz *et al.*, 2014). En outre, la classification par maximum de vraisemblance utilise le même ensemble de caractéristiques pour toutes les classes et nécessite un grand nombre de calculs pour classer complètement les données d'images, en particulier lorsque l'on utilise un grand nombre de caractéristiques comme données d'entrée dans le processus de classification ou qu'il faut différencier un grand nombre de classes spectrales. Dans ce cas, l'application de ce type de classifieur peut s'avérer nettement plus longue que l'utilisation d'autres techniques de classification supervisée. Les différentes limites liées à la classification par maximum de vraisemblance ont entraîné le développement rapide de nouveaux algorithmes de classification destinés au domaine de la télédétection. Parmi ces nouvelles méthodes, plusieurs s'annoncent très prometteuses, notamment les réseaux de neurones artificiels (Rumelhart *et al.*, 1986 ; Rigol-Sanchez *et al.*, 2003), les machines à vecteurs de support (Abedi *et al.*, 2013 ; Al-Anazi et Gates, 2010 ; Cortes et Vapnik, 1995 ; Ghimire *et al.*, 2012 ; Zuo et Carranza, 2011), les arbres de décision (Breiman, 1984) et les ensembles d'arbres de classification tels que les forêts aléatoires (Breiman, 2001 ; Vincenzi *et al.*, 2011 ; Waske et Braun, 2009 ; Ghimire *et al.*, 2012 ; Rodriguez-Galiano et Chica-Olmo, 2012).

Réseaux de neurones artificiels

L'utilisation de réseaux de neurones artificiels à des fins de classification dans le domaine de la télédétection découle de la prise de conscience que le cerveau humain est capable de traiter efficacement d'importantes quantités de données issues de sources très diverses, et que la traduction mathématique d'un tel processus pourrait s'avérer utile pour traiter et interpréter des données d'images. Un réseau de neurones artificiel appliqué à la classification d'images prend la forme d'un processeur massivement parallèle constitué de simples unités de traitement qui acquièrent des connaissances à partir de leur environnement au moyen d'un processus d'autoapprentissage, afin de s'y adapter et d'établir des liens entre les données d'entrée (caractéristiques des images satellite, notamment), et les données de sortie (classes de couverture cibles, par exemple) (Rumelhart *et al.*, 1986 ; Rigol-Sanchez *et al.*, 2003). Plusieurs réseaux de neurones artificiels méritent d'être mentionnés, notamment le perceptron multicouche à rétropropagation (MLP, de l'anglais *Multi-Layer Perceptron*) (Wilkinson, 1997), les cartes auto-organisantes de Kohonen (KSOFM, de l'anglais *Kohonen's Self-Organizing Feature Maps*) (Ji, 2000 ; Pal *et al.*, 2005) et le réseau Fuzzy ARTMAP (Carpenter *et al.*, 1992 ; Mannan *et al.*, 1998). Si ces méthodes ne s'appliquent pas exactement de la même manière, toutes nécessitent une phase d'apprentissage et de classification afin de pouvoir extraire des informations utiles à partir de données d'images obtenues par télédétection (Jensen, 2016). Durant la phase d'apprentissage, des données d'images provenant de lieux dont les attributs (classes) sont déjà connus sont transmises au réseau en tant que données d'entrée. Le réseau utilise ces informations dans le cadre d'une procédure itérative visant à définir les règles qui permettront de produire les meilleurs résultats possibles en termes de classification. Ces règles sont ensuite utilisées lors de la phase de classification afin d'affecter les données aux classes d'apprentissage auxquelles elles sont le plus susceptibles d'appartenir (degré d'appartenance flou).

Les avantages des réseaux de neurones artificiels résident notamment dans leur capacité à : 1) obtenir des résultats plus précis lorsque l'on utilise comme variables d'entrée de nombreux jeux de données volumineux, mesurés à différentes échelles et présentant une distribution statistique non normale ; 2) apprendre et mettre à jour continuellement des modèles complexes, notamment des relations non linéaires entre les données d'entrée et les classes de sortie (mise à disposition d'un plus grand nombre de données dans un environnement en pleine évolution) ; 3) proposer, grâce à la généralisation, des solutions fiables malgré des données incomplètes ou imprécises ; et 4) intégrer dans leur analyse des connaissances préalables et des contraintes physiques réalistes (Atkinson et Tatnall, 1997 ; Pal et Mather, 2003 ; Rogan *et al.*, 2008 ; Hansen, 2012 ; Jensen, 2016). Toutefois, les inconvénients que présentent les réseaux de neurones artificiels ont limité leur adoption à quelques applications concrètes décisives (Pal et Mather, 2003 ; Qiu et Jensen, 2004 ; Jensen, 2016). Leur principal défaut est sans doute que ces réseaux peuvent être assimilés à des « boîtes noires » en matière d'interprétation (Rodriguez-Galiano *et al.*, 2012 ; Gómez *et al.*, 2016). En effet, il est depuis toujours difficile d'expliquer de façon claire le processus ayant abouti au résultat, car les règles de classification et d'interprétation des images apprises par le réseau ne peuvent

pas être facilement consultées ou décrites (Qiu et Jensen, 2004 ; Jensen, 2016). En conséquence, on privilégie généralement d'autres méthodes de classification aux explications plus compréhensibles.

Machines à vecteurs de support

Les machines à vecteurs de support sont une technique d'apprentissage statistique supervisé non paramétrique utilisée pour résoudre les problèmes de classification (Smola et Schoelkopf, 1998 ; Vapnik, 2000) ; elles recèlent un grand potentiel pour la classification de données d'images obtenues par télédétection (Melgani et Bruzzone, 2004 ; Pal et Mather, 2005). Elles visent à résoudre un problème d'optimisation quadratique afin de déterminer les frontières de séparation optimales (hyperplans) entre deux classes dans un espace de caractéristiques multidimensionnel (Foody et Mathur, 2004). Pour ce faire, les machines à vecteurs de support travaillent uniquement sur les données d'apprentissage qui se trouvent à la limite entre plusieurs classes (qu'on appelle « vecteurs de support »). Lorsque les classes ne sont pas séparables, les données d'apprentissage sont projetées dans un espace de dimension supérieure grâce à une fonction noyau : la nouvelle distribution des données est alors mieux adaptée à la détermination d'un hyperplan linéaire (Van der Linden *et al.*, 2009). Cette procédure est répétée pour chaque paire de classes afin de répartir les données dans un nombre prédéfini de classes. Les règles de séparation optimale des classes sont alors utilisées afin d'affecter toutes les données d'images aux classes cibles prédéfinies. La classification par machine à vecteurs de support repose donc sur l'idée que seuls les échantillons d'apprentissage situés à la limite des classes sont nécessaires à la discrimination (Foody et Mathur, 2004).

L'avantage des machines à vecteurs de support réside dans leur capacité à surpasser les méthodes de classification traditionnelles lorsque l'on ne dispose que de jeux de données d'apprentissage limités (Foody et Mathur, 2004 ; Waske et Benediktsson, 2007). En effet, leur principe sous-jacent consiste à fonder le processus d'apprentissage sur une minimisation des risques structurels (Van der Linden *et al.*, 2009). Cela permet de minimiser les erreurs de classification concernant les données non vues sans formuler de suppositions préalables sur la distribution statistique de ces données (Mountrakis *et al.*, 2011). Le principal inconvénient de cette méthode concerne la sélection de la fonction noyau la plus adaptée et son paramétrage. Bien qu'il existe de nombreuses possibilités, certaines fonctions noyaux sont incapables de configurer les machines à vecteurs de support de façon optimale pour des applications liées à la télédétection (Mountrakis *et al.*, 2011). Il s'agit d'un problème non négligeable, car un choix inapproprié risque d'entraîner un surapprentissage ou un surlissage, susceptibles d'avoir un impact négatif important sur les performances des machines à vecteurs de support et la précision de leur classification (Ustuner, 2015 ; Martins *et al.*, 2016). En outre, cette méthode n'est pas optimisée pour gérer les données contenant du bruit, notamment les valeurs aberrantes souvent présentes dans les données de télédétection, dont la prise en compte risque de réduire considérablement les performances du classifieur (Mountrakis *et al.*, 2011). Malgré ces inconvénients, les machines à vecteurs de support restent une méthode populaire de classification de la couverture du sol.

Classification par arbres de décision

Les arbres de décision, méthode de classification supervisée fondée sur des partitions binaires récursives respectant un ensemble de règles optimisées, sont devenus une solution attrayante permettant d'extraire des informations sur les différentes classes à des fins de classification de la couverture du sol (Huang et Jensen, 1997; Friedl *et al.*, 2002). Un arbre de décision utilise une série de caractéristiques comme données d'entrée, et produit un résultat (ou décision) par le biais d'une série de tests. Les arbres déterminent des règles grâce à des régions de partitionnement binaire récursif (nœuds) de plus en plus homogènes par rapport à leur variable de classe (Breiman, 1984). La classification par arbres de décision crée des modèles multivariés fondés sur un ensemble de règles décisionnelles définies par des combinaisons de caractéristiques et un ensemble de fonctions linéaires discriminantes appliquées à chaque nœud (Champagne *et al.*, 2014). En règle générale, une fois que l'on a collecté des échantillons d'apprentissage en quantité suffisante (Lu et Weng, 2007), un algorithme d'apprentissage utilise ces données afin de générer des arbres de décision qui sont ensuite transformés en une autre forme de représentation des

connaissances, à savoir des règles de production. Celles-ci étant faciles à comprendre, elles peuvent être examinées par des experts, qui peuvent les modifier directement en faisant preuve de prudence (Jensen, 1986).

L'utilisation des arbres de décision pour la classification des images présente divers avantages : cette méthode permet notamment de gérer des données à différentes échelles de mesure (Brown de Colstoun *et al.*, 2003), des données d'entrée présentant une distribution statistique non normale (non paramétrique) (Friedl et Brodley, 1997 ; Hansen *et al.*, 1996), ou encore des relations non linéaires entre les données d'entrée et les classes de sortie (Friedl *et al.*, 2002). Ces avantages sont similaires à ceux des réseaux de neurones artificiels. Les arbres de décision présentent toutefois des atouts supplémentaires : ils sont faciles à appliquer, car ils nécessitent l'estimation d'un plus petit nombre de paramètres (Friedl *et al.*, 2010 ; Gómez *et al.*, 2016) ; ils possèdent une structure hiérarchique transparente et facile à interpréter (Hansen *et al.*, 1996 ; Rodriguez-Galiano et Chica-Olmo, 2012) ; et ils peuvent être entraînés avec une faible interaction humaine, en créant des règles et des conditions directement à partir des données d'apprentissage (Huang et Jensen, 1997). La caractéristique la plus importante des arbres de décision est leur capacité à s'adapter à de nouvelles données d'apprentissage. De même, le système peut être évalué de façon à examiner le processus suivi pour parvenir à une conclusion (Jensen, 2016). Les inconvénients de cette méthode sont ses faibles performances dans les espaces de caractéristiques de grande dimension (Pal et Mather, 2003), sa sensibilité au bruit (Ghimire *et al.*, 2012) et sa tendance au surapprentissage (Breiman, 1984). Des chercheurs travaillant dans le domaine de la télédétection tentent actuellement de mieux comprendre ce qui influence les performances de la classification par arbre de décision (Hansen, 2012), ce qui a entraîné la mise au point de méthodes fondées sur des ensembles d'arbres de décision (méthode des forêts aléatoires, notamment), qui permettent d'obtenir de meilleurs résultats en matière de classification en combinant un grand nombre d'arbres de décision différents.

Classification par forêts aléatoires

Les forêts aléatoires constituent une application améliorée des arbres de décision. Il s'agit d'un algorithme d'apprentissage fondé sur une méthode d'ensemble, qui combine plusieurs classifications des mêmes données de façon à produire une classification plus précise que celle obtenue avec d'autres formes d'arbres de décision (Cutler *et al.*, 2007 ; Ghimire *et al.*, 2012). Cette méthode a pour objectif de soumettre un même jeu de données à de nombreuses classifications par arbres de décision, puis d'utiliser une approche par règles afin de combiner les prédictions de l'ensemble des arbres. Les différents arbres sont construits à partir de divers sous-ensembles de données d'apprentissage à l'aide d'un processus de « bagging ». Le bagging consiste à sous-échantillonner de façon aléatoire les données originales (avec remplacement, ce qui signifie qu'un même élément peut être réutilisé plusieurs fois) afin de construire chaque arbre. Chaque arbre de la forêt est généralement construit en utilisant deux tiers des données d'apprentissage ; le tiers non utilisé (« out-of-bag », ou OOB) servira ensuite à estimer l'erreur de prédiction (Breiman, 2001). Chaque échantillon *bootstrap* est alors soumis à une classification ; cependant, le partitionnement binaire n'utilise à chaque nœud (division) qu'un petit nombre de variables prédictives sélectionnées aléatoirement (Rodriguez-Galiano et Chica-Olmo, 2012). Le processus de division se poursuit jusqu'à ce que l'indice de Gini ne puisse plus être réduit par une nouvelle subdivision (Cutler *et al.*, 2007). Chaque arbre contribue par un vote à affecter la classe la plus fréquente aux données d'entrée (Breiman, 1984 ; Rodriguez-Galiano et Chica-Olmo, 2012). Pour prédire la classe d'une observation, on calcule le vote majoritaire qui s'y rapporte, le partage des voix étant réparti de manière aléatoire (Cutler *et al.*, 2007).

Le principal avantage des forêts aléatoires est que cette technique est potentiellement plus précise et plus fiable que les méthodes paramétriques conventionnelles ou que les méthodes d'apprentissage automatique par arbres de décision (Rodriguez-Galiano et Chica-Olmo, 2012). En effet, un groupe de classificateurs permet d'obtenir des résultats plus précis qu'un classifieur pris isolément, tout en contournant les faiblesses propres à ces méthodes (Breiman, 1984 ; Ghimire *et al.*, 2010 ; Kotsiantis et Pintelas, 2004). Par ailleurs, les forêts aléatoires sont considérées comme étant relativement simples à paramétrer, dans la mesure où il suffit de définir deux paramètres (à savoir, le nombre d'arbres

de classification souhaités et le nombre de variables de prédiction utilisées dans chaque nœud afin de construire l'arbre) pour établir un modèle de prédiction (Rodriguez-Galiano et Chica-Olmo, 2012). D'autres avantages sont liés au fait que les forêts aléatoires ont recours au bagging pour construire différents arbres à partir de sous-ensembles de données d'apprentissage. Les arbres ainsi construits sont utilisés pour calculer la précision et le taux d'erreur de chaque observation à l'aide des prédictions OOB, dont on calcule ensuite la moyenne pour l'ensemble des observations. Les observations OOB n'étant pas utilisées pour construire un arbre, elles servent essentiellement à estimer la précision par validation croisée (Rodriguez-Galiano et Chica-Olmo, 2012). Les forêts aléatoires permettent également d'évaluer l'importance d'une seule variable : il convient alors de modifier l'une des variables d'entrée en maintenant le reste constant, et de mesurer la diminution de la précision qui en résulte au moyen de l'erreur OOB (Breiman, 1984 ; Rodriguez-Galiano et Chica-Olmo, 2012). Cette technique est utile lorsqu'il est important de connaître l'influence exercée par chaque variable prédictive sur le modèle de classification (Ghimire *et al.*, 2010 ; Gislason *et al.*, 2006 ; Pal et Mather, 2005). L'inconvénient des forêts aléatoires est que lorsqu'on dispose d'un grand nombre d'arbres, il est plus difficile d'examiner chaque arbre et de comprendre sa structure (Deschamps *et al.*, 2012), ce qui entraîne un effet « boîte noire » qui brouille les règles de décision (Gómez *et al.*, 2016).

Pour conclure, une synthèse exhaustive de cet ensemble de travaux a récemment été réalisée par Khatami *et al.* (2016), qui ont procédé à la méta-analyse statistique d'études sur la classification supervisée d'images de couverture du sol par pixels. Ce compte-rendu, fondé sur des travaux de recherche publiés entre 1998 et 2012 dans cinq revues de télédétection influentes, avait pour objectif de fournir des conseils cohérents concernant les performances relatives des différents processus de classification utilisés pour réaliser des cartes de couverture du sol. Malheureusement, il n'est pas possible de désigner une seule solution comme étant la meilleure dans toutes les situations ; toutefois, s'il ne faut choisir qu'une seule méthode, on aura tendance à privilégier divers modèles de forêts aléatoires ou de machines à vecteurs de support, en raison de la maturité de ces classifieurs par apprentissage automatique et de leur capacité à gérer des caractéristiques aux dimensions très importantes. Il est également possible de concevoir ponctuellement des stratégies de traitement plus élaborées, en combinant par exemple plusieurs classifieurs de façon à obtenir plusieurs résultats de classification et à procéder de manière hiérarchique (en d'autres termes, opérer une première discrimination entre les surfaces en eau, les sols nus, les zones urbaines, les forêts, les terres cultivées et autres, puis une deuxième classification visant à séparer les différentes classes agricoles).

Post-traitement

Les opérations de post-traitement peuvent améliorer la classification, puisqu'elles permettent d'appliquer différentes techniques de filtrage ou de fusionner plusieurs résultats de classification. Pour commencer, les erreurs macroscopiques peuvent être corrigées de manière interactive, car elles sont clairement identifiées par une inspection visuelle systématique. Des opérateurs de filtrage élémentaires sont appliqués sur une fenêtre glissante de 3 pixels × 3 pixels ou de 5 pixels × 5 pixels, dans laquelle un filtre majoritaire supprime l'effet « poivre et sel » induit par la classification par pixels. Plus intéressant encore, ce type de filtre majoritaire peut également être appliqué au résultat d'une classification par pixels en utilisant des objets obtenus à l'aide d'une segmentation d'image par réflectance multispectrale, ce qui permet d'obtenir une carte de couverture du sol nettement plus lisse.

Il convient d'utiliser des techniques de fusion pour combiner les résultats d'une classification d'ensemble. Il est possible d'obtenir une seule carte grâce à un vote majoritaire : la classe choisie est alors celle approuvée par l'ensemble des classifieurs (vote unanime), par au moins la moitié des classifieurs plus un (majorité absolue) ou par plusieurs classifieurs (majorité simple). On peut également avoir recours à un vote majoritaire pondéré lorsque certains classifieurs sont censés être plus fiables ou suivant la probabilité d'appartenance du résultat de classification.

Il convient de signaler que les différentes étapes décrites ci-dessus sont en grande partie interdépendantes, et que chaque décision doit tenir compte de toute la chaîne de production des cartes de couverture du sol afin de trouver une solution appropriée.

2.4. PRATIQUES ACTUELLES ET JEUX DE DONNÉES DE COUVERTURE DU SOL EXISTANTS

Plutôt que de créer une nouvelle carte de couverture du sol, on peut dans certains cas fusionner plusieurs cartes existantes pour parvenir à un meilleur résultat (en réduisant les incertitudes et en indiquant la légende de classification souhaitée) (Jung *et al.*, 2006). Une carte mondiale des terres cultivées a ainsi été réalisée en compilant toutes les informations disponibles issues de diverses sources concernant une même classe de couverture du sol ou d'utilisation des terres (Thenkabail *et al.*, 2009). Deux autres cartes mondiales ont également été produites afin de cartographier les terres cultivées en mettant l'accent sur la gestion de l'eau : la carte mondiale des terres aptes aux cultures pluviales et la carte mondiale des superficies d'irrigation. Toutefois, la résolution spatiale grossière de ces produits (10 km) ne répond pas aux besoins des applications opérationnelles et entraîne d'importantes incertitudes, en particulier concernant les systèmes agricoles complexes tels que ceux qui existent en Afrique.

Pour obtenir des informations sur la couverture du sol, les programmes nationaux ou régionaux utilisent souvent des images d'une résolution allant de 5 à 30 m, ou ont recours à l'intégration d'images multicateurs. Certains pays ont mis en place des projets visant à cartographier spécifiquement les types de cultures annuelles au niveau national grâce à la télédétection par satellite ; c'est le cas notamment du projet Cropland Data Layers (CDL) aux États-Unis ou de l'Inventaire annuel des cultures au Canada, tous deux réalisés à une résolution de 30 m. La plupart des États membres de l'Union européenne, quant à eux, tiennent à jour un système d'identification des parcelles agricoles (SIPA) fondé sur les déclarations des agriculteurs afin de gérer la répartition des subventions de la Politique agricole commune. Ces informations annuelles, qui couvrent la majeure partie du secteur agricole, se sont avérées très utiles dans le cadre de l'élaboration de cartes nationales de la couverture du sol et des terres cultivées, effectuée par exemple à partir des données fournies par le système Sentinel-2 pour le suivi de l'agriculture (voir la section 2.4.2 ci-dessous).

D'autres travaux, tels que le projet Africover et le programme du Réseau mondial sur le couvert végétal (GLCN, de l'anglais *Global Land Cover Network*), ont permis de réaliser des cartes nationales de couverture du sol détaillées, fondées sur l'interprétation visuelle d'images d'une résolution spatiale de 30 m plutôt que sur une classification automatique. Ces cartes sont par conséquent mises à jour moins fréquemment. En outre, diverses initiatives mondiales de cartographie du sol ont récemment publié des cartes mondiales de couverture du sol, notamment les projets GlobeLand30¹³, Globcover 2005 et 2009¹⁴, MODIS Land Cover¹⁵, et, plus récemment, la base de données CCI Land Cover de l'ESA¹⁶.

Bien que l'on dispose parfois de plusieurs cartes de couverture du sol pour un même pays, il n'est pas toujours évident de déterminer laquelle sera la plus utile pour des applications spécifiques, ni comment les combiner de façon à obtenir un meilleur jeu de données. Les normes en matière de documentation relative aux cartes de couverture du sol sont présentées ci-dessous, de même qu'un recensement systématique de tous les jeux de données actuellement disponibles.

2.4.1. Métadonnées, politiques sur les données et production participative

Les métadonnées sont des données qui fournissent des informations sur d'autres données. Dans le domaine géospatial, ces métadonnées sont essentielles pour utiliser de façon appropriée une image matricielle ou une couche vectorielle. La norme ISO 19115-1:2014 définit le schéma requis pour décrire des informations géographiques et des services au moyen de métadonnées. Les métadonnées doivent fournir des informations concernant l'identification, l'étendue, la qualité, les aspects spatiaux et temporels, le contenu, la référence spatiale cartographique, la représentation des données, les politiques de disponibilité et de distribution, ainsi que d'autres propriétés des données et services

13 <http://glc30.tianditu.com/>.

14 http://due.esrin.esa.int/page_globcover.php.

15 <http://modis-land.gsfc.nasa.gov/landcover.html>.

16 <http://maps.elie.ucl.ac.be/CCI/viewer/>.

géographiques numériques. Lorsque les archives de métadonnées sont structurées suivant une norme commune, cela facilite la localisation et la lisibilité des métadonnées tant pour les humains que pour les machines, ce qui leur permet d'être utilisées automatiquement pour des catalogues de données logicielles. La norme ISO 19144-2:2012, présentée dans la section 2.2.1 ci-dessus, porte plus spécifiquement sur les informations relatives à la couverture du sol : cette norme définit le cadre LCML, qui vise à décrire en détail la typologie de la couverture du sol.

Des politiques d'ouverture des données sont de plus en plus souvent appliquées aux données gouvernementales, ce qui permet à un plus grand nombre d'utilisateurs de consulter les informations disponibles et favorise éventuellement des efforts concertés dans le but de les tenir à jour. Les politiques d'ouverture adoptées aux États-Unis concernant les SIG et l'observation de la Terre ont montré la voie, et la Directive européenne INSPIRE¹⁷ recommande désormais à tous les États membres de l'Union européenne de décrire l'ensemble des informations géospatiales produites par l'administration et les organismes gouvernementaux à l'aide de métadonnées précises et de mettre ces données à la disposition du public. Les licences Creative Commons¹⁸ constituent également une très bonne solution pour partager des données géospatiales tout en veillant à ce que leurs auteurs soient cités, et en fixant éventuellement des restrictions concernant l'utilisation commerciale des données en question.

Plus récemment, des initiatives collaboratives de collecte de données de couverture du sol ont également connu un développement rapide. La plateforme Geo-Wiki¹⁹ propose différentes applications fondées sur la production participative permettant de comparer des cartes mondiales de couverture du sol ou de collecter des données de référence. Une telle plateforme peut collecter un volume impressionnant de données *in situ* ou d'échantillons photo-interprétés lorsqu'elle est bien conçue et qu'elle bénéficie d'une bonne promotion. Toutefois, il convient encore d'élaborer une méthodologie concernant l'exploitation scientifique des données obtenues par production participative. L'initiative Open Street Map constitue un bon exemple de collecte collaborative de données géospatiales librement exploitables²⁰ : celle-ci met à la disposition du grand public une carte mondiale de référence pouvant inspirer des initiatives de cartographie détaillée de la couverture du sol (conception de collecte de données *in situ*, par exemple).

2.4.2. Inventaire détaillé des jeux de données relatifs à la couverture du sol et aux terres cultivées existants

Waldner *et al.* (2015) proposent un cadre analytique visant à apprécier si les cartes de couverture du sol existantes sont adaptées à la cartographie des terres cultivées, en évaluant le jeu de données effectivement disponible pour chaque pays. Ce cadre procède à une évaluation quantitative des informations relatives aux terres cultivées présentes dans les cartes de couverture du sol en se fondant sur quatre critères : 1) utilité des informations thématiques pour la définition des terres cultivées ; 2) actualité ; 3) résolution spatiale ; et 4) niveau de confiance. Cette première analyse a permis d'identifier les zones prioritaires dans les initiatives de cartographie des terres cultivées à l'échelle mondiale.

L'identification et la collecte de cartes de couverture du sol nationales, régionales et mondiales constituent un travail de longue haleine à cause de la diversité des sources et des producteurs impliqués, mais également des nombreuses politiques différentes en matière de distribution des données. L'élaboration d'une base de données spatiale et d'un inventaire exhaustifs est un effort permanent, compte tenu des nouveaux produits qui inondent continuellement le marché et des mises-à-jours et modifications constamment apportées aux politiques d'accès aux données. Les jeux de données mondiaux, régionaux et nationaux ont été identifiés grâce à un examen systématique : séances de travail avec les principaux experts, analyses documentaires et recherches Internet. Lors de cette collecte, il a fallu distinguer les données mises gratuitement à la disposition du grand public des jeux de données existants ; en effet,

17 Directive 2007/2/CE du Parlement européen et du Conseil du 14 mars 2007 établissant une infrastructure d'information géographique dans la Communauté européenne (INSPIRE), disponible à l'adresse <http://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX%3A32007L0002> (consultée le 10 juin 2017).

18 <https://creativecommons.org/share-your-work/>.

19 <http://www.geo-wiki.org/>.

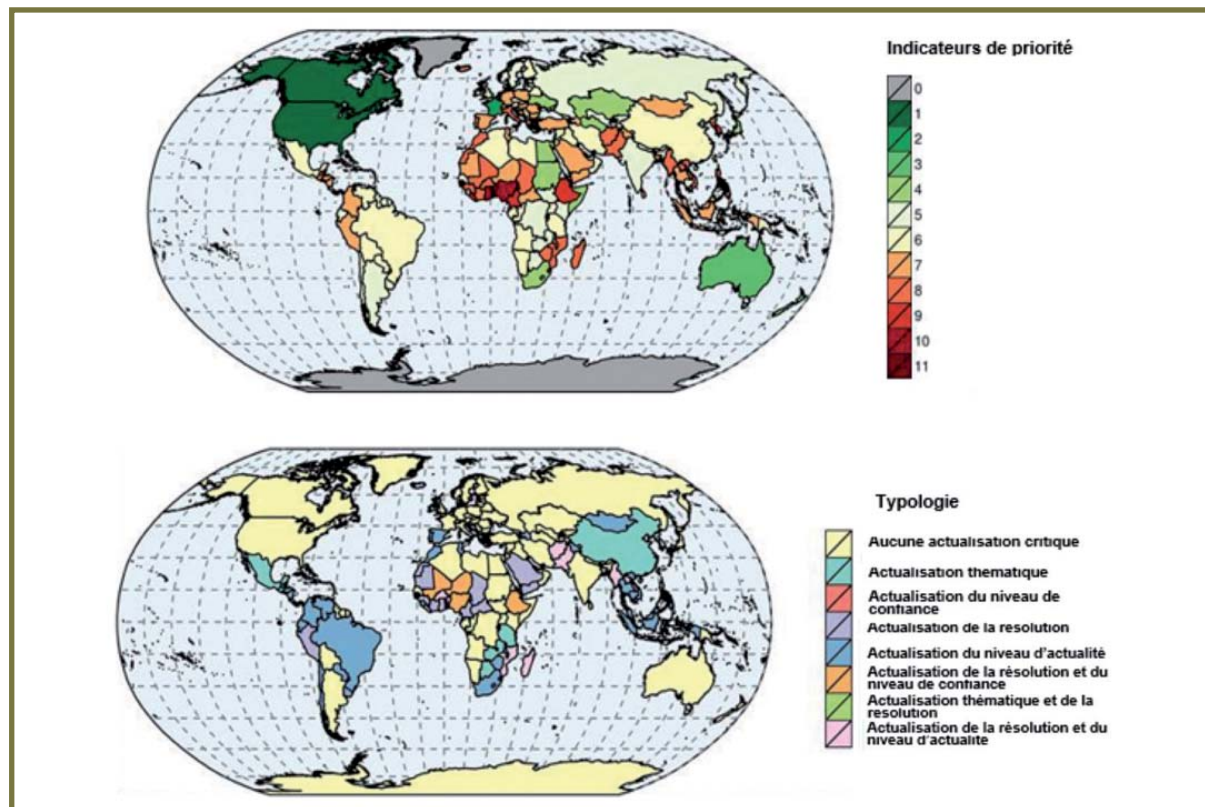
20 <https://www.openstreetmap.org/#map=5/51.500/-0.100>.

ces derniers sont souvent soumis à une politique de diffusion empêchant leur utilisation, ou il peut être difficile d'accéder au jeu de données géoréférencées à proprement parler (plutôt qu'à une simple image des cartes intégrée dans un site Internet). On a donc appliqué une règle générale selon laquelle un jeu de données non disponible est considéré comme inexistant et n'est pas pris en compte dans cet inventaire.

Les cartes mondiales de couverture du sol telles que celles réalisées dans le cadre des projets GLC 2000, GlobCover 2005 et 2009, MODIS Land Cover, GlobeLand30, et CCI Land Cover (ESA) ne ciblent pas spécifiquement la composante agricole du paysage. Une mauvaise définition sémantique des classes risque de compromettre leur utilisation pour des applications agricoles ; toutefois, le manque de précision observée dans la plupart de ces produits concernant la classe des terres cultivées peut s'expliquer par plusieurs autres raisons : 1) la nature hétérogène et dynamique des agrosystèmes mondiaux ; 2) la structure spatiale du paysage (taille des parcelles) et la diversité des cultures ; 3) les différences au niveau du cycle des cultures ; 4) les différences au niveau des pratiques et des calendriers de culture au sein d'une même classe ; 5) la similarité spectrale avec d'autres classes de couverture du sol ; et 6) pour les cartes des satellites optiques, la persistance de la couverture nuageuse.

Suite à une identification et une collecte exhaustives des cartes de couverture du sol, chaque critère a été quantifié pour chaque carte et une analyse multicritères a été réalisée à l'échelle nationale afin de repérer les régions prioritaires en ce qui concerne la cartographie des terres cultivées. Trois grandes régions prioritaires ont été identifiées : les pays d'Afrique (principalement l'Afrique de l'Ouest), l'Asie du Sud-Est (en particulier l'Indonésie) et l'Amérique du Sud (Brésil). D'autres pays, notamment l'Éthiopie, Madagascar, le Mozambique et le Pakistan, devront également être examinés avec attention.

FIGURE 3. (A) CARTE DES INDICATEURS DE PRIORITÉ ET (B) TYPOLOGIE DES ACTUALISATIONS POUR CHAQUE PAYS.



Les zones auxquelles on a attribué un indice de priorité élevé (tons rouges) représentent les régions prioritaires en ce qui concerne la cartographie des terres cultivées ; à l'inverse, les zones auxquelles on a attribué un indice peu élevé (tons verts) correspondent aux régions pour lesquelles on dispose de cartes à jour, exactes et précises. L'Afrique de l'Ouest, l'Éthiopie et l'Asie du Sud-Est (Indonésie) apparaissent clairement comme des régions prioritaires en matière de cartographie des terres cultivées (Waldner *et al.*, 2015).

Compte tenu de la richesse et de la qualité variable des cartes de couverture du sol existantes, l'étude s'est appuyée sur les travaux précédents afin de les harmoniser. Suivant l'analyse des priorités, une couche unifiée des terres cultivées²¹ a été créée en combinant les produits les plus adaptés. Cette couche a été évaluée en fonction des jeux de données de validation disponibles au niveau mondial, et sa précision globale se situe entre 84 % et 95 %, un score supérieur à la plupart des cartes mondiales de couverture du sol.

2.4.3. Détection des changements de la couverture du sol

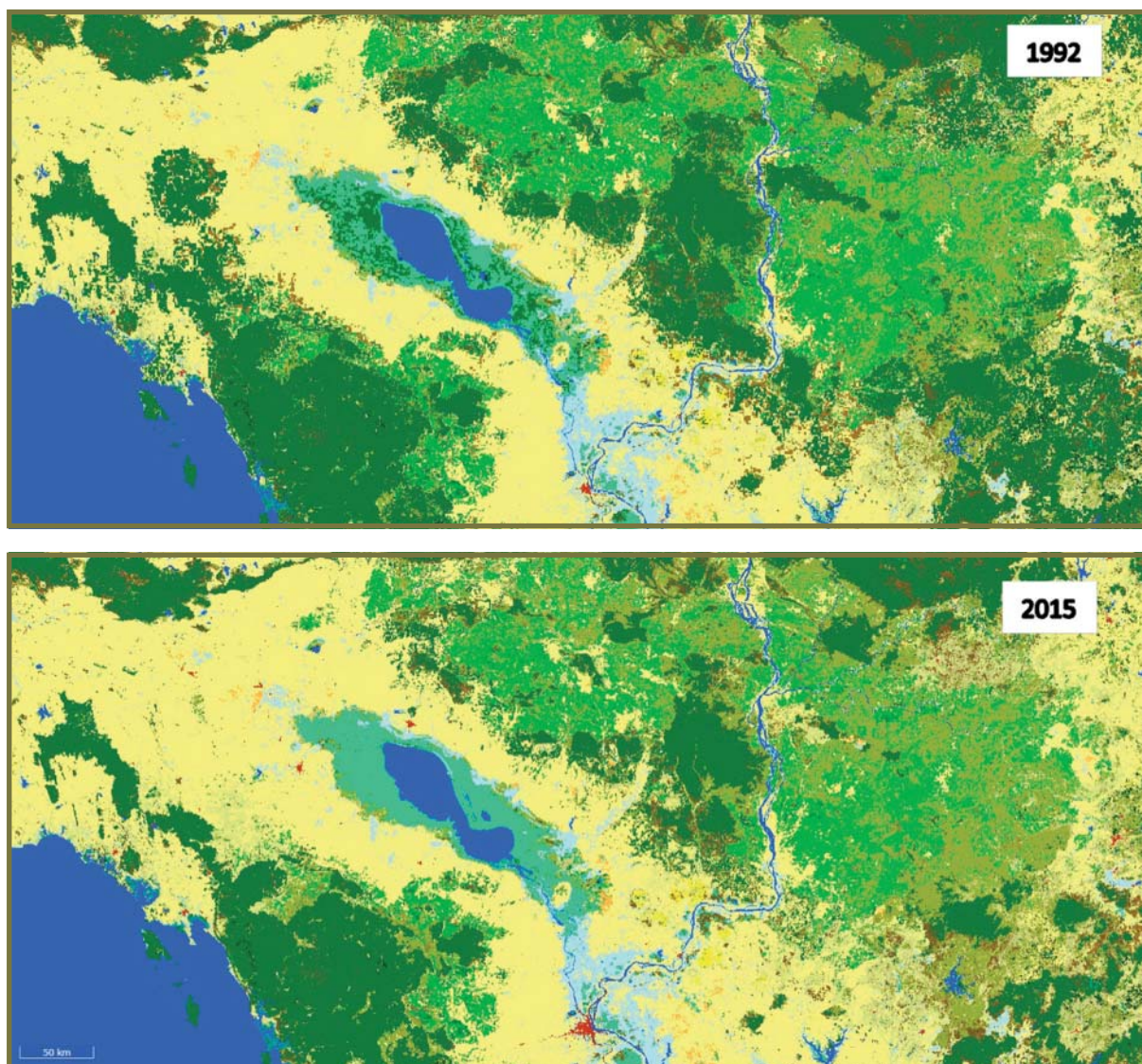
Les études existantes (voir par exemple Jung *et al.*, 2006 et McCallum *et al.*, 2006) expliquent souvent les divergences constatées entre plusieurs cartes de couverture du sol (parfois successives) par une incompatibilité entre les différentes typologies de couverture du sol et par le niveau de précision des cartes actuelles, qui varie de 70 % à 85 % pour les cartes mondiales de couverture du sol. Ce niveau de précision globale ne permet pas d'obtenir des informations relatives aux changements de couverture du sol en comparant directement ces cartes, car le taux de changement annuel de la couverture du sol est nettement inférieur au taux d'erreur des cartes en question.

Sur des cartes ayant une résolution grossière, on peut observer d'une année sur l'autre d'importantes variations des étiquettes de couverture du sol sans aucun lien avec de véritables changements de la couverture du sol. Ce problème semble s'expliquer en partie par le fait qu'à une résolution spatiale comprise entre 300 et 500 m, les paysages sont souvent constitués d'un mélange de différentes classes. En outre, étant donné la variabilité annuelle de la phénologie de la végétation et les éventuelles perturbations (incendies, sécheresses et invasions d'insectes, notamment), il est relativement difficile d'obtenir une caractérisation annuelle cohérente.

C'est pourquoi l'Initiative de l'ESA sur le changement climatique a mis au point, dans le cadre de son volet consacré à la couverture du sol, une stratégie visant à dissocier la cartographie de la couverture du sol de la détection des changements de couverture du sol. Il a ainsi été possible de produire une série temporelle unique et cohérente de 24 cartes mondiales annuelles de couverture du sol d'une résolution de 300 m entre 1992 et 2015, et de mettre en évidence les changements enregistrés d'une année sur l'autre entre les différents types de couvertures du sol (figure 4). Cette longue série de cartes n'a pu être réalisée que grâce à des observations de 1 km² issues des satellites AVHRR, SPOT-Vegetation et PROBA-V. Une telle résolution ne devrait permettre de détecter que les changements de couverture du sol les plus importants d'une année sur l'autre ; cependant, grâce aux capacités d'observation quotidienne de ces instruments, il est possible de traiter la variabilité interannuelle de la saisonnalité observée dans différents biomes. Une fois le changement de couverture du sol détecté, on utilise des images d'une résolution de 300 m afin de délimiter précisément la zone concernée.

21 <http://maps.elie.ucl.ac.be/geoportail/>.

FIGURE 4. CHANGEMENTS DE COUVERTURE DU SOL SURVENUS ENTRE 1992 ET 2015 AU CAMBODGE, DÉTECTÉS À 1 KM ET CARTOGRAPHIÉS À UNE RÉOLUTION DE 300 M.



On observe, entre autres phénomènes, un développement urbain (en rouge) ainsi qu'un important empiètement des terres cultivées (en jaune) sur les différents types de forêts (tons verts). Ces cartes issues du projet CCI Land Cover et leur légende complète sont disponibles à l'adresse <http://maps.elie.ucl.ac.be/CCI/viewer/>.

Même lorsque l'on utilise des cartes de couverture du sol à haute résolution, telles que celles réalisées dans le cadre du programme CORINE Land Cover, il est difficile de garantir leur stabilité au fil du temps, et une procédure spécifique de détection des changements est nécessaire. Cette instabilité entre les différents produits exige clairement l'adoption d'autres méthodes ou d'autres concepts.

Deux grandes difficultés doivent être prises en compte pour que la détection des modifications de la couverture du sol soit fiable. Premièrement, des méthodes spécifiques de détection des changements doivent être mises au point et faire l'objet de tests intensifs afin de pouvoir rendre compte du très faible pourcentage de changement annuel de la couverture du sol (le développement des terres cultivées le plus rapide jamais observé au Brésil et en Argentine atteignait moins de 4 % par an ; en Afrique, il reste généralement inférieur à 1 %). À ce jour, on ne dispose d'aucune méthode validée pour divers agro-écosystèmes permettant de détecter les évolutions annuelles des terres cultivées. Toutefois, plusieurs méthodes de détection des changements, qu'elles soient fondées sur une

approche par objets (Desclée *et al.*, 2006 ou Ernst *et al.*, 2013, par exemple) ou par pixels (Hansen *et al.*, 2013, par exemple), sont aujourd'hui parvenues à une certaine maturité et peuvent fournir des informations précises sur les changements concernant les forêts, liés essentiellement mais pas exclusivement au défrichement agricole. La plupart de ces méthodes comparent directement les valeurs de réflectance ou les ensembles de valeurs d'indice de végétation afin d'identifier les changements de la surface terrestre, plutôt que de classer la couverture du sol. En outre, la détection précise des changements survenus au niveau des terres agricoles a été compromise par le manque de séries temporelles denses de données haute résolution. Compte tenu de la récente disponibilité des séries temporelles des satellites Sentinel-1, Sentinel-2 et Landsat-8 et de leur continuité à long terme, il serait nécessaire d'élaborer de nouvelles méthodes de détection des changements annuels concernant les terres cultivées.

2.5. INITIATIVES DE CARTOGRAPHIE EN COURS À L'ÉCHELLE MULTINATIONALE OU MONDIALE

Plusieurs initiatives internationales cherchent actuellement à réaliser des avancées dans le domaine de la cartographie de la couverture du sol. Cette section présente trois de ces initiatives, spécifiquement axées sur les terres agricoles.

2.5.1. Carte mondiale des terres cultivées

La possibilité offerte par Google Engine d'accéder à d'importantes capacités de calcul informatique et à la totalité des archives Landsat en ligne ouvre la voie au lancement d'initiatives majeures, telles que le projet *Global Food Security Analysis-Support Data at 30 meters* (GFSAD30) piloté par l'USGS (Thenkabail *et al.*, 2012). Ce projet utilise des méthodes de classification fondées sur une analyse temporelle du signal en vue de réaliser une carte mondiale des terres cultivées d'une résolution de 30 m à partir d'un jeu de données multisources, et notamment des archives Landsat et MODIS (figure 5). Il propose également une application mobile et un accès à des jeux de données prêts à l'emploi, ouvrant ainsi la voie à des travaux collaboratifs.

FIGURE 5. ÉTAT D'AVANCEMENT ACTUEL (MARS 2017) DU PROJET GFSAD30 DE L'USGS, QUI VISE À CARTOGRAPHIER LES TERRES CULTIVÉES À UNE RÉOLUTION DE 30 M ([HTTPS://CROPLANDS.ORG/](https://croplands.org/)).

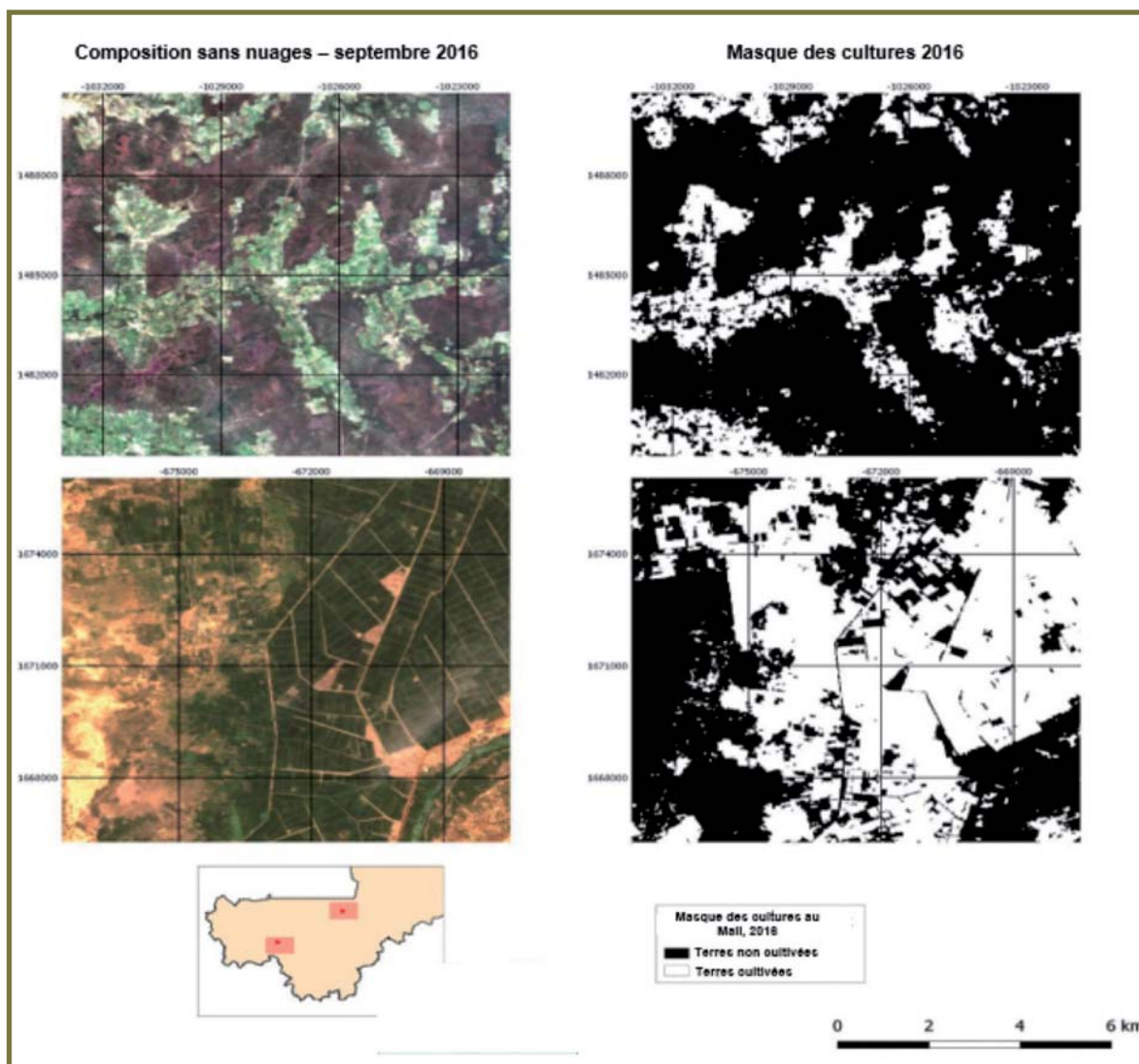


2.5.2. Système Sentinel-2 pour le suivi de l'agriculture

Le projet Sentinel-2 pour le suivi de l'agriculture (Sen2-Agri) de l'ESA, piloté par l'Université catholique de Louvain, entend répondre aux besoins définis par d'importantes parties prenantes internationales et soutenus par l'ESA en matière d'informations (Bontemps *et al.*, 2016). Il a mis au point et diffusé un logiciel libre de traitement permettant de télécharger automatiquement l'ensemble des images Sentinel-2 et Landsat-8 enregistrées durant la saison de végétation et d'obtenir quasiment en temps réel quatre produits cartographiques nationaux d'une résolution de 10 m : i) synthèses mensuelles sans nuages des valeurs de réflectance de surface dans les dix bandes de Sentinel-2, d'une résolution de 10 ou 20 m selon la bande ; ii) masques mensuels des cultures d'une résolution de 10 m, fournis au cours de la saison agricole afin de cartographier les cultures annuelles ; iii) cartes des types de cultures et indicateurs précoces pour cinq grands types de cultures, fournis à la mi-saison et en fin de saison ; iv) état de végétation indiquant la croissance des cultures, fourni tous les sept jours sous réserve d'observations sans nuages. L'état de végétation se compose d'un ensemble de cartes d'indicateurs où figurent l'indice différentiel normalisé de végétation, la surface foliaire et les mesures de phénologie issues des profils temporels du NDVI.

Le système Sen2-Agri a fait l'objet de démonstrations nationales couronnées de succès en Afrique du Sud, au Mali (voir encadré ci-dessus) et en Ukraine, ainsi que dans plus de neuf autres sites à travers le monde. Son fonctionnement nominal implique que les satellites Sentinel-2 A et B soient tous deux en orbite, de même que le satellite Landsat-8, afin d'améliorer la fréquence d'observation. Bien qu'il soit possible de produire un masque des cultures sans données *in situ*, la capacité à collecter rapidement des données *in situ* dont la qualité a été contrôlée s'est avérée essentielle à la réalisation de cartes des types de cultures au cours de la saison de végétation. Système de production opérationnel lancé en juin 2017 auprès du grand public, Sen2-Agri a vocation à devenir une source essentielle de produits de télédétection destinés à diverses applications en matière de surveillance en milieu agricole, mais également à une stratification précise du plan d'échantillonnage et à la compilation de statistiques sur les cultures précoces (Defourny *et al.*, 2016). La particularité de ce système de production réside dans son fonctionnement quasiment en temps réel et dans ses outils automatisés permettant le traitement évolutif de séries temporelles d'images télédétectées tout au long de la saison, d'une résolution de 10 m et couvrant intégralement certains territoires nationaux.

FIGURE 6. COMPOSITIONS EN COULEURS NATURELLES D'UNE RÉOLUTION DE 10 M DE SEN2-AGRI (À GAUCHE) ET CARTES DE TERRES CULTIVÉES CORRESPONDANTES (À DROITE) POUR LA SAISON DE VÉGÉTATION 2016 AU MALI.



Ces cartes ont été réalisées automatiquement en septembre à partir des séries temporelles du satellite Sentinel-2A : en haut, gros plan sur la région de Kita, et en bas, gros plan sur la zone de production de riz de l'Office du Niger.

2.5.3. Cartographie du riz irrigué

La cartographie des systèmes de riziculture irriguée constitue un domaine de recherche actif, en particulier lorsqu'elle a recours aux observations issues de radars à synthèse d'ouverture (SAR). Le riz possède une signature spectrotemporelle extrêmement spécifique, de sorte qu'il n'est généralement pas étudié dans une perspective de cartographie de la couverture du sol, mais plutôt dans un cadre de suivi de la riziculture. Compte tenu de l'apparition de plusieurs systèmes SAR opérationnels en orbite capables de fournir des séries temporelles denses et cohérentes, telles que celles issues des satellites PALSAR, Radarsat-2 ou Sentinel-1 A et B, la cartographie des terres irriguées est désormais au point. L'Institut international de recherche sur le riz (IRRI, de l'anglais *International Rice Research Institute*), par exemple, œuvre très activement à la détection des dates des semences et des zones plantées dans l'ensemble des Philippines (projet PRISM). On peut également citer les projets ambitieux

Asia-Rice et GEO-Rice, soutenus respectivement par l'Agence d'exploration aérospatiale japonaise et l'Agence spatiale européenne et actuellement menés dans le cadre de l'initiative GEOGLAM.

PERSPECTIVES

Depuis l'apparition de la flotte de satellites Sentinel (Sentinel-1 A et B, Sentinel-2 A et B), qui fournissent des images mondiales d'observation de la Terre (souvent) gratuites, d'une résolution comprise entre 10 et 20 m et à une fréquence régulière (temps de revisite de cinq jours pour Sentinel-2 et de 12 jours pour Sentinel-1), l'observation de la Terre est entrée dans une nouvelle ère d'applications pleinement opérationnelles mais également pérennes. Cette avancée est particulièrement utile pour les initiatives de cartographie de la couverture du sol régulièrement menées à l'échelle nationale à des fins de stratification des paysages agricoles. Gageons que d'autres disciplines (statistiques et agro-économie, notamment) ne tarderont pas à identifier clairement les avantages de cette évolution et à opérer les ajustements nécessaires en termes de flux d'informations.

Cette révolution progressive concerne non seulement la disponibilité des technologies spatiales et les politiques relatives aux données qui en résultent, mais également les infrastructures dématérialisées d'informatique et les progrès méthodologiques liés à l'apprentissage automatique et aux algorithmes d'intelligence artificielle. Même les stratégies de collecte de données sur le terrain peuvent tirer parti des évolutions informatiques, et notamment des initiatives de production participative, des applications mobiles et du développement de l'utilisation des UAV.

Les cartes statiques sont progressivement délaissées au profit des séries temporelles de cartes annuelles, dont la qualité est renforcée chaque année, de même que leur capacité à détecter les changements. Les cartes s'assimileront de plus en plus à des bases de données géoréférencées concernant la couverture du sol et l'utilisation des terres, constamment mises à jour grâce au flux d'observations satellite reçues. Dans le cadre des statistiques agricoles, l'important développement des capacités de calcul informatique et le renforcement des flux de données directement collectées sur le terrain contribuent fortement à produire des cartes de couverture du sol et des terres cultivées de façon rapide et rentable.

Cette évolution est également prise en considération dans le contexte de l'élaboration des politiques. La couverture du sol est considérée comme une source d'information essentielle que les pays peuvent utiliser pour rendre compte des indicateurs des objectifs de développement durable (ODD). Il convient également de mentionner que la classification de la couverture du sol et de l'utilisation des terres est une composante essentielle de la norme internationale désignée sous le nom de Cadre central du Système de comptabilité environnementale et économique (SCEE), et plus encore de son système de comptabilité destiné à l'agriculture, la sylviculture et la pêche (SCEE-ASP), mis en place conjointement par la FAO et la Division de statistique des Nations Unies. Le SCEE-ASP est un cadre statistique qui vise à fournir des statistiques agricoles et à élaborer des indicateurs agro-environnementaux, et la classification de la couverture du sol adoptée par le SCEE est elle-même fondée sur les classificateurs LCML.

Au-delà de l'utilisation de la couverture du sol à des fins de stratification par échantillonnage, la réalisation de cartes de couverture du sol conformes au SCEE à l'échelle nationale est censée garantir une communication des statistiques agro-environnementales respectant le Système de comptabilité nationale et favoriser l'élaboration de politiques agricoles susceptibles de mieux appréhender les aspects environnementaux des activités agricoles. Ces liens forcément étroits entre les processus d'élaboration de politiques, les statistiques agricoles, les indicateurs et les informations relatives à la couverture du sol constituent des arguments de poids en faveur d'un renforcement des capacités nationales qui permettrait de cartographier les sols de manière opérationnelle et efficace à partir des données de télédétection.

2.6. RÉFÉRENCES BIBLIOGRAPHIQUES

- Abdulaziz, A.M., Hurtado, J.M. et Al-Douri, R.** 2009. « Application of Multitemporal Landsat Data to Monitor Land Cover Changes in the Eastern Nile Delta Region, Egypt ». *International Journal of Remote Sensing*, vol. 30, n° 11, p. 2977–96.
- Abedi, M., Gholam-Hossain, N. et Fathianpour, N.** 2013. « Fuzzy Outranking Approach: A Knowledge-Driven Method for Mineral Prospectivity Mapping ». *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, p. 556–67.
- Al-Anazi, A.F. et Gates, I.D.** 2010. « Support Vector Regression for Porosity Prediction in a Heterogeneous Reservoir: A Comparative Study ». *Computers & Geosciences*, vol. 36, n° 12, p. 1494–1503.
- Atkinson, P.M. et Tatnall, A.R.L.** 1997. « Introduction Neural Networks in Remote Sensing ». *International Journal of Remote Sensing*, vol. 18, n° 4, p. 699–709.
- Baatz, M. et Schäpe, A.** 2000. « Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation ». In : Strobl, J., Blaschke, T. et Griesebner, G. (dir.), *Angewandte Geographische Informationsverarbeitung XII* (p. 12–23), Wichmann-Verlag : Heidelberg.
- Bagan, H., Kinoshita, T. et Yamagata, Y.** 2012. « Combination of AVNIR-2, PALSAR, and Polarimetric Parameters for Land Cover Classification ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, n° 4, p. 1318–28.
- Bagan, H. et Yamagata, Y.** 2010. « Improved Subspace Classification Method for Multispectral Remote Sensing Image Classification ». *Photogrammetric Engineering and Remote Sensing*, vol. 76, n° 11, p. 1239–51.
- Bhatta, B.** 2008. *Remote Sensing and GIS*. Oxford University Press : Oxford, Royaume-Uni.
- Bontemps, S., Arias, M. et al.** 2015. « Building a Data Set over 12 Globally Distributed Sites to Support the Development of Agriculture Monitoring Applications with Sentinel-2 ». *Remote Sensing*, vol. 7, n° 9, p. 16062–90.
- Bontemps, S., Herold, M., Kooistra, L., van Groenestijn, A., Hartley, A., Arino, O., Moreau, I. et Defourny, P.** 2012. « Revisiting land cover observation to address the needs of the climate modeling community ». *Biogeosciences*, vol. 9, n° 6, p. 2145–57.
- Breiman, L.** 1984. *Classification and Regression Trees*. CRC Press : Boca Raton, FL, États-Unis.
- Breiman, L.** 2001. « Random Forests ». *Machine Learning*, vol. 45, n° 1, p. 5–32.
- Brown de Colstoun, E.C., Story, M.H., Thompson, C., Commisso, K., Smith, T.G. et Irons, J.R.** 2003. « National Park Vegetation Mapping Using Multitemporal Landsat 7 Data and a Decision Tree Classifier ». *Remote Sensing of Environment*, vol. 85, n° 3, p. 316–27.
- Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J.H. et Rosen, D.B.** 1992. « Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps ». *IEEE Transactions on Neural Networks*, vol. 3, n° 5, p. 698–713. doi: 10.1109/72.159059.

Carreiras, J.M.B., Pereira, J.M.C., Campagnolo, M.L. et Shimabukuro, Y.E. 2006. « Assessing the Extent of Agriculture/Pasture and Secondary Succession Forest in the Brazilian Legal Amazon Using SPOT VEGETATION Data ». *Remote Sensing of Environment*, vol. 101, n° 3, p. 283–98.

Champagne, C., McNairn, H., Daneshfar, B. et Shang, J.L. 2014. « A Bootstrap Method for Assessing Classification Accuracy and Confidence for Agricultural Land Use Mapping in Canada ». *International Journal of Applied Earth Observation and Geoinformation*, vol. 29, p. 44–52.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M. et al. 2015. « Global land cover mapping at 30m resolution: A POK-based operational approach ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, p. 7–27.

Chen, C. et McNairn, H. 2006. « A Neural Network Integrated Approach for Rice Crop Monitoring ». *International Journal of Remote Sensing*, vol. 27, n° 7, p. 1367–93.

Congalton, R.G., Balogh, M., Bell, C., Green, K., Milliken, J.A. et Ottman, R. 1998. « Mapping and Monitoring Agricultural Crops and Other Land Cover in the Lower Colorado River Basin ». *Photogrammetric Engineering and Remote Sensing*, vol. 64, n° 11, p. 1107–14.

Cortes, C. et Vapnik, V. 1995. « Support-Vector Networks ». *Machine Learning*, vol. 20, n° 3, p. 273–97.

Cutler, D.R., Edwards, T.C., Beard, K.B., Cutler, A., Hess, K.T., Gibson, J. et Lawler, J.J. 2007. « Random Forests for Classification in Ecology ». *Ecology*, vol. 88, n° 11, p. 2783–92.

Davidson, A.M. 2016. *Review of satellite image classification methods*. Document interne. Agriculture et Agroalimentaire Canada : Ottawa.

DeFries, R.S., Hansen, M.C. et Townshend, J.R.G. 2000. « Global Continuous Fields of Vegetation Characteristics: A Linear Mixture Model Applied to Multi-Year 8 Km AVHRR Data ». *International Journal of Remote Sensing*, vol. 21, n° 6–7, 1^{er} janvier 2000, p. 1389–1414.

DeFries, R.S. et Chan, J.C.W. 2000. « Multiple Criteria for Evaluating Machine Learning Algorithms for Land Cover Classification from Satellite Data ». *Remote Sensing of Environment*, vol. 74, n° 3, p. 503–15.

Defourny, P. et Bontemps S. 2012a. « Revisiting Land-Cover Mapping Concepts ». In : Giri, C. (dir.), *Remote Sensing of Land Use and Land Cover* (p. 45–64), CRC Press : Boca Raton, FL, États-Unis.

Defourny, P., Bontemps, S., Mayaux, P., et Herold, M. 2012b. « Global Land-Cover Map Validation Experiences: Toward the Characterization of Uncertainty ». In : Giri, C. (dir.), *Remote Sensing of Land Use and Land Cover* (p. 207–224), CRC Press : Boca Raton, FL, États-Unis.

Defourny, P., Vancutsem, C., Bicheron, P., Brockmann, C., Nino, F., Schouten, L., Leroy, M., et al. 2006. « GLOBCOVER: a 300 m global land cover product for 2005 using Envisat MERIS time series ». In : Kerle, N. et Skidmore, A., *Proceedings of the ISPRS Commission VII mid-term symposium, Remote sensing: from pixels to processes* (p. 8–11).

Defourny, P., Bontemps, S., Bellemans, N., Cara, C. et al. 2016. « Nationwide demonstration cases of Sentinel-2 satellite exploitation towards early crop area indicator ». Article préparé pour la septième Conférence internationale sur les statistiques de l'agriculture (ICAS VII), 26–28 octobre 2016. Rome.

Denègre, J. 2013. *Thematic Mapping From Satellite Imagery: A Guidebook*. Elsevier.

Deschamps, B., McNairn, H., Shang, J. et Jiao, X. 2012. « Towards Operational Radar-Only Crop Type Classification: Comparison of a Traditional Decision Tree with a Random Forest Classifier ». *Journal canadien de télédétection*, vol. 38, n° 1, p. 60–68.

Desclée, B., Bogaert, P. et Defourny, P. 2006. « Forest change detection by statistical object-based method ». *Remote Sensing of Environment*, vol. 102, n° 1-2, p. 1–11.

Di Gregorio, A. 2013. *A cropland nomenclature conform to the FAO Land Cover Meta-Language*. Rapport technique SIGMA.

Di Gregorio, A. et Leonardi, U. 2016. *Land Cover Classification System: User manual. Software version 3*. Disponible à l'adresse : <http://www.fao.org/3/a-i5428e.pdf>. Consulté le 10 juin 2017.

Duro, D.C., Franklin, S.E. et Dubé, M.G. 2012. « A Comparison of Pixel-Based and Object-Based Image Analysis with Selected Machine Learning Algorithms for the Classification of Agricultural Landscapes Using SPOT-5 HRG Imagery ». *Remote Sensing of Environment* vol. 118, p. 259–72.

Duveiller G., Defourny P., Desclée B. et Mayaux P. 2008. « Deforestation in Central Africa: estimates at regional, national and landscape levels by advanced processing of systematically-distributed Landsat extracts ». *Remote Sensing of Environment*, vol. 112, p. 1969–81.

Duveiller, G. et Defourny, P. 2010. « A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing ». *Remote Sensing of Environment*, vol. 114, n° 11, p. 2637–50.

Enderle, W. et Weih, J.C. 2005. « Integrating Supervised and Unsupervised Classification Methods to Develop a More Accurate Land Cover Classification ». *Journal of the Arkansas Academy of Science*, vol. 59, p. 65–73.

Ernst, C., Mayaux, P., Verhegghen, A., Bodart, C., Musampa, C. et Defourny, P. 2013. « National forest cover change in Congo Basin : deforestation, reforestation, degradation and regeneration for the years 1990, 2000 and 2005 ». *Global Change Biology*, vol. 19, n° 4, p. 1173-87.

Foody, G.M. 2008. « RVM-based Multi-class Classification of Remotely Sensed Data ». *International Journal of Remote Sensing*, vol. 29, n° 6, p. 1817–23.

Foody, G.M. 2004. « Supervised Image Classification by MLP and RBF Neural Networks with and without an Exhaustively Defined Set of Classes ». *International Journal of Remote Sensing*, vol. 25, n° 15, p. 3091–3104.

Foody, G.M., Boyd D.S. et Sanchez-Hernandez, C. 2007. « Mapping a Specific Class with an Ensemble of Classifiers ». *International Journal of Remote Sensing*, vol. 28, n° 8, p. 1733–46.

Foody, G.M. et Mathur, A. 2004. « A Relative Evaluation of Multiclass Image Classification by Support Vector Machines ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, n° 6, p. 1335–43.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E. et al. 2002. « Global Land Cover Mapping from MODIS: Algorithms and Early Results ». *Remote Sensing of Environment*, vol. 83, n° 1–2, p. 287–302.

- Friedl, M.A. et Brodley, C.E.** 1997. « Decision Tree Classification of Land Cover from Remotely Sensed Data ». *Remote Sensing of Environment*, vol. 61, n° 3, p. 399–409.
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A. et Huang, X.** 2010. « MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets ». *Remote Sensing of Environment*, vol. 114, n° 1, p. 168–82.
- Fritz, F., See, L., McCallum, I. et al.** 2015. « Mapping global cropland and field size ». *Global Change Biology*, vol. 21, n° 5, p. 1980–92.
- Frizzelle, B.G. et Moody, A.** 2001. « Mapping Continuous Distributions of Land Cover ». *Photogrammetric Engineering and Remote Sensing*, vol. 67, n° 6, p. 693–705.
- Ghimire, B., Rogan, J. et Miller, J.** 2010. « Contextual Land-Cover Classification: Incorporating Spatial Dependence in Land-Cover Classification Models Using Random Forests and the Getis Statistic ». *Remote Sensing Letters* vol. 1, n° 1, p. 45–54.
- Ghimire, B., Rogan, J., Rodríguez Galiano, V., Panday, P. et Neeti, N.** 2012. « An Evaluation of Bagging, Boosting, and Random Forests for Land-Cover Classification in Cape Cod, Massachusetts, USA ». *GIScience & Remote Sensing*, vol. 49, n° 5, p. 623–43.
- Giri, C.P.** 2012. *Remote Sensing of Land Use and Land Cover: Principles and Applications*. CRC Press : Boca Raton, FL, États-Unis.
- Gislason, P.O., Benediktsson, J.A. et Sveinsson, J.R.** 2006. « Random Forests for Land Cover Classification ». *Pattern Recognition in Remote Sensing (PRRS 2004)*, vol. 27, n° 4, p. 294–300.
- Glanz, H., Carvalho, L., Sulla-Menashe, D. et Friedl, M.A.** 2014. « A Parametric Model for Classifying Land Cover and Evaluating Training Data Based on Multi-Temporal Remote Sensing Data ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 97, p. 219–28.
- Gómez, C., White, J.C. et Wulder, M.A.** 2016. « Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, p. 55–72.
- Hansen, M., Dubayah, R. et Defries, R.** 1996. « Classification Trees: An Alternative to Traditional Land Cover Classifiers ». *International Journal of Remote Sensing*, vol. 17, n° 5, p. 1075–81.
- Hansen, M.** 2012. « Classification Trees and Mixed Pixel Training Data ». In : Giri, C. (dir.), *Remote Sensing of Land Use and Land Cover* (p. 127–36), CRC Press : Boca Raton, FL, États-Unis.
- Huang, C., Davis, L.S. et Townshend, J.R.G.** 2002. « An Assessment of Support Vector Machines for Land Cover Classification ». *International Journal of Remote Sensing*, vol. 23, n° 4, p. 725–49.
- Huang, X. et Jensen, J.R.** 1997. « A Machine-Learning Approach to Automated Knowledge-Base Building for Remote Sensing Image Analysis with GIS Data ». *Photogrammetric Engineering and Remote Sensing*, vol. 63, n° 10, p. 1185–94.

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre Canto, G., Bontemps, S., Defourny, P. et Koetz, B. 2015. « Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery ». *Remote Sensing*, vol. 7, n° 9, p. 12356–79.

Joint Experiment of Crop Assessment and Monitoring (JECAM). 2014. *JECAM Guidelines: Definition of the Minimum Earth Observation Dataset Requirements*. Documents standard de la JECAM. Disponible à l'adresse http://www.jecam.org/JECAM_EO_Guidelines_v1_0.pdf. Consulté le 10 juin 2017.

Jensen, J.R. 1986. « Introductory Digital Image Processing: A Remote Sensing Perspective ». SciTech Connect, Université de Caroline du Sud, Columbus, États-Unis. 1^{er} janvier 1986. Disponible à l'adresse : <http://www.osi.gov/scitech/biblio/5166368>. Consulté le 10 juin 2017.

Jensen, J.R. 2016. *Introductory Digital Image Processing: A Remote Sensing Perspective*, 4^e éd. Pearson : Glenview, IL, États-Unis.

Ji, C.Y. 2000. « Land-Use Classification of Remotely Sensed Data Using Kohonen Self-Organizing Feature Map Neural Networks ». *Photogrammetric Engineering and Remote Sensing*, vol. 66, n° 12, p. 1451–60.

Jones, H.G. et Vaughan, R.A. 2010. *Remote Sensing of Vegetation: Principles, Techniques, and Applications*. Oxford University Press : Oxford, Royaume-Uni – New York, États-Unis.

Kamusoko, C. et Aniya, M. 2009. « Hybrid Classification of Landsat Data and GIS for Land Use/Cover Change Analysis of the Bindura District, Zimbabwe ». *International Journal of Remote Sensing*, vol. 30, n° 1, p. 97–115.

Khatami, R., Mountrakis, G. et Stehman, S.V. 2016. « A Meta-Analysis of Remote Sensing Research on Supervised Pixel-Based Land-Cover Image Classification Processes: General Guidelines for Practitioners and Future Research ». *Remote Sensing of Environment*, vol. 177, p. 89–100.

Kotsiantis, S.B. et Pintelas, P.E. 2004. « Combining Bagging and Boosting ». *International Journal of Computational Intelligence*, vol. 1, n° 4, p. 324–33.

Kumar, L., Priyakant, S., Brown, J.F., Ramsey, R.D., Rigge, M., Stam, C.A., Hernandez, A.J., Hunt Jr., E.R. et Reeves, M.C. 2016. « Characterization, Mapping and Monitoring of Rangelands: Methods and Approaches ». In : Thenkabail, P.S. (dir.), *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing*, vol. 2 (p. 309–50). Remote Sensing Handbook series. CRC Press : Boca Raton, FL, États-Unis.

Laba, M., Smith, S.D. et Degloria, S.D. 1997. « Landsat-Based Land Cover Mapping in the Lower Yuna River Watershed in the Dominican Republic ». *International Journal of Remote Sensing*, vol. 18, n° 14, p. 3011–25.

Lambert, M.-J., Waldner, F. et Defourny, P. 2016. « Cropland Mapping over Sahelian and Sudanian Agrosystems: A Knowledge-Based Approach Using PROBA-V Time Series at 100-m ». *Remote Sensing*, vol. 8, n° 3, p. 1–23.

Latham, J. et Rosati, I. 2016. *Land Information in the Context of Agricultural Statistics*. Rapport technique GO-15-2016 de la GSARS. Disponible à l'adresse : http://gsars.org/wp-content/uploads/2016/08/TR_Information-on-Land-in-the-Context-of-Ag-Statistics-180816.pdf. Consulté le 10 juin 2017.

Liu, C., Frazier, P. et Kumar, L. 2007. « Comparative assessment of the measures of thematic classification accuracy ». *Remote Sensing of Environment*, vol. 107, n° 4, p. 606–16.

- Lu, D. et Weng, Q.** 2007. « A Survey of Image Classification Methods and Techniques for Improving Classification Performance ». *International Journal of Remote Sensing*, vol. 28, n° 5, p. 823–70.
- Mannan, B., Roy, J. et Ray, A.K.** 1998. « Fuzzy ARTMAP Supervised Classification of Multi-Spectral Remotely-Sensed Images ». *International Journal of Remote Sensing*, vol. 19, n° 4, p. 767–74.
- Martínez Beltran, C. et Calera Belmonte, A.** 2001. « Irrigated Crop Area Estimation Using Landsat TM Imagery in La Mancha, Spain ». *Photogrammetric Engineering & Remote Sensing*, vol. 67, n° 10, p. 1177–84.
- Martins, S., Bernardo, N., Ogashawara, I. et Alcantara, E.** 2016. « Support Vector Machine Algorithm Optimal Parameterization for Change Detection Mapping in Funil Hydroelectric Reservoir (Rio de Janeiro State, Brazil) ». *Modeling Earth Systems and Environment*, vol. 2, n° 3.
- Matton, N., Sepulcre Canto, G., Waldner, F., Valero, S., Morin, D., Inglada, J., Arias, M., Bontemps, S., Koetz, B. et Defourny, P.** 2015. « An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series ». *Remote Sensing*, vol. 7, n° 10, p. 13208–32.
- Melgani, F. et Bruzzone, L.** 2004. « Classification of Hyperspectral Remote Sensing Images with Support Vector Machines ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, n° 8, p. 1778–90.
- Mountrakis, G., Watts, R., Luo, L. et Wang, J.** 2009. « Developing Collaborative Classifiers Using an Expert-Based Model ». *Photogrammetric Engineering and Remote Sensing*, vol. 75, n° 7, p. 831–43.
- Mountrakis, G., Im, J. et Ogole, C.** 2011. « Support Vector Machines in Remote Sensing: A Review ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, n° 3, p. 247–59.
- Pal, M. et Mather, P.M.** 2005. « Support Vector Machines for Classification in Remote Sensing ». *International Journal of Remote Sensing*, vol. 26, n° 5, p. 1007–11.
- Pal, M. et Mather, P.M.** 2006. « Some Issues in the Classification of DAIS Hyperspectral Data ». *International Journal of Remote Sensing*, vol. 27, n° 14, p. 2895–916.
- Pal, M. et Mather, P.M.** 2003. « An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Classification ». *Remote Sensing of Environment*, vol. 86, n° 4, p. 554–65.
- Pal, N.R., Laha, A. et Das, J.** 2005. « Designing Fuzzy Rule Based Classifier Using Self-organizing Feature Map for Analysis of Multispectral Satellite Images ». *International Journal of Remote Sensing*, vol. 26, n° 10, p. 2219–40.
- Qiu, F. et Jensen, J.R.** 2004. « Opening the Black Box of Neural Networks for Remote Sensing Image Classification ». *International Journal of Remote Sensing*, vol. 25, n° 9, p. 1749–68.
- Radoux, J. et Defourny, P.** 2008. « Quality assessment of segmentation results devoted to object-based classification ». In : Blaschke, T., Lang, S. et Hay, G.J. (dir.), *Object-Based Image Analysis : Spatial concepts for knowledge driven remote sensing applications* (p. 257–271), Springer-Verlag : Berlin–Heidelberg.
- Rigol-Sanchez, J.P., Chica-Olmo, M. et Abarca-Hernandez, F.** 2003. « Artificial Neural Networks as a Tool for Mineral Potential Mapping with GIS ». *International Journal of Remote Sensing*, vol. 24, n° 5, p. 1151–6.

Rodriguez-Galiano, V. et Chica-Olmo, M. 2012. « Land Cover Change Analysis of a Mediterranean Area in Spain Using Different Sources of Data: Multi-Seasonal Landsat Images, Land Surface Temperature, Digital Terrain Models and Texture ». *Applied Geography*, vol. 35, n° 1–2, p. 208–18.

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C. et Roberts, D. 2008. « Mapping Land-Cover Modifications over Large Areas: A Comparison of Machine Learning Algorithms ». *Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue*, vol. 112, n° 5, p. 2272–83.

Rumelhart, D.E., Hinton, G.E. et Williams, R.J. 1986. « Learning Representations by Back-Propagating Errors ». *Nature*, vol. 323, n° 6088, p. 533–36.

Sluiter, R. et Pebesma, E.J. 2010. « Comparing Techniques for Vegetation Classification Using Multi- and Hyperspectral Images and Ancillary Environmental Data ». *International Journal of Remote Sensing*, vol. 31, n° 23, p. 6143–61.

Smola, A. et Schoelkopf, B. 1998. *A Tutorial on Support Vector Regression*. Série de rapports techniques NeuroCOLT2, NC2-TR-1998-030. Disponible à l'adresse : <http://citeseer.ist.psu.edu/smola98tutorial.html>. Consulté le 10 juin 2017.

Thenkabail, P.S., Knox, J.W., Ozdogan, M., Gumma, M.K., Congalton, R.G., Wu, Z., Milesi, C., Finkral, A., Marshall, M., Mariotto, I., You, S. Giri, C. et Nagler, P. 2012. « Assessing future risks to agricultural productivity, water resources and food security: how can remote sensing help? ». *Photogrammetric Engineering and Remote Sensing, Special Issue on Global Croplands: Highlight Article, août 2012*, vol. 78, n° 8, p. 773–82.

Ustuner, M. 2015. « Application of Support Vector Machines for Landuse Classification Using High-Resolution RapidEye Images: A Sensitivity Analysis ». *European Journal of Remote Sensing*, vol. 48, n° 1, p. 403.

Valero, S., Morin, D., Inglada, J., Sepulcre Canto, G., Arias, M., Hagolle, O., Dedieu, G., Bontemps, S., Defourny, P. et Koetz, B. 2016. « Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions ». *Remote Sensing*, vol. 8, n° 55, p. 1–21.

Van der Linden, S., Rabe, A., Okujeni, A. et Hostert, P. 2009. *Image SVM Classification, Manual for Application: Image SVM Version 2*. Disponible à l'adresse : http://dev.geo.hu-berlin.de/trac/enmap-box/export/8/enmap-box/tags/Save20150115/SourceCode/applications/imageSVM/_help/svc_manual.pdf. Consulté le 10 juin 2017. Publication de l'Université Humboldt de Berlin, Allemagne.

Vapnik, V.N. 2000. *The Nature of Statistical Learning Theory*. 2^e édition. Statistics for Engineering and Information Science. Springer : New York, États-Unis.

Vincenzi, S., Zucchetta, M., Franzoi, P., Pellizzato, M. Pranovi, F., De Leo, G.A. et Torricelli, P. 2011. « Application of a Random Forest Algorithm to Predict Spatial Distribution of the Potential Yield of Ruditapes Philippinarum in the Venice Lagoon, Italy ». *Ecological Modelling*, vol. 222, n° 8, p. 1471–78.

Waldner, F., Fritz, S., Lamarche, C., Bontemps, S. et Defourny, P. 2016. « A Unified Cropland Layer at 250 m for Global Agriculture Monitoring ». *Data*, vol. 1(1), n° 3, p. 1–13.

Waldner, F., De Abelleira, D., Veron, S.R., Zhang, M., Wu, B., Plotnikov, D., Bartalev, S., Lavreniuk, M., Skakun, S., Kussul, N., Le Maire, G., Dupuy, S., Jarvis, I et Defourny, P. 2016. « Towards a set of agrosystem-

specific cropland mapping methods to address the global cropland diversity ». *International Journal of Remote Sensing*, vol. 37, n° 14, p. 3196–231.

Waldner, F., Fritz, S., Di Gregorio, A. et Defourny, P. 2015. « Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps ». *Remote Sensing*, vol. 7, n° 6, p. 7959–86.

Waske, B. et Benediktsson, J.A. 2007. « Fusion of Support Vector Machines for Classification of Multisensor Data ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, n° 12, p. 3858–66.

Waske, B. et Braun, M. 2009. « Classifier Ensembles for Land Cover Mapping Using Multitemporal SAR Imagery ». *ISPRS Journal of Photogrammetry and Remote Sensing, Theme Issue: Mapping with SAR: Techniques and Applications*, vol. 64, n° 5, p. 450–7.

Wilkinson, G.G. 1997. « Open Questions in Neurocomputing for Earth Observation ». In : Kanellopoulos, I., Wilkinson, G.G., Roli, F. et Austin, J. (dir.), *Neurocomputation in Remote Sensing Data Analysis* (p. 3–13). Springer : Berlin–Heidelberg.

Xiuwan, C. 2002. « Using Remote Sensing and GIS to Analyse Land Cover Change and Its Impacts on Regional Sustainable Development ». *International Journal of Remote Sensing*, vol. 23, n° 1, p. 107–24.

Zhong, Y., Zhang, L., Gong, J. et Li, P. 2007. « A Supervised Artificial Immune Classifier for Remote-Sensing Imagery ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, n° 12, p. 3957–66.

Zuo, R. et Carranza, E.J.M. 2011. « Support Vector Machine: A Tool for Mapping Mineral Prospectivity ». *Computers & Geosciences*, vol. 37, n° 12, p. 1967–75.

Annexe

Liste des meilleures cartes de couverture du sol disponibles à l'échelle nationale, régionale et mondiale pour les différents pays du monde. Les deux premières colonnes présentent respectivement la région couverte par le produit en question et ses références. La dernière colonne fournit l'année ou la période de référence du produit (modifié de Waldner *et al.*, 2015).



Chapitre 3

La télédétection au service de la conception de bases de sondage

Javier Gallego

Le sujet abordé dans ce chapitre reprend en grande partie le contenu des chapitres 4 et 6 du manuel publié récemment par la GSARS, intitulé « Bases de sondage principales pour les statistiques agricoles » (GSARS, 2015). Le chapitre 4 dudit manuel fournit des directives sur l'utilisation des systèmes d'information géographique (SIG), des systèmes mondiaux de navigation par satellite (GNSS, de l'anglais *Global Navigation Satellite System*, plus connus sous le nom de GPS) et de la télédétection, tandis que le chapitre 6 est plus spécifiquement consacré aux bases de sondage aréolaires (BSA). Le lecteur est invité à consulter ce manuel pour toute information complémentaire ; le présent chapitre a toutefois vocation à se suffire à lui-même.

La télédétection utilisée aux fins de statistiques agricoles est, dans la plupart des cas, effectuée à l'aide de capteurs optiques, qui sont généralement classés en fonction de leur résolution spatiale. Ceci dépend en grande partie de la répétitivité de l'acquisition des images : le coût des images satellite THR, dont les pixels mesurent entre 0,5 m et 2,5 m environ, est généralement trop élevé pour assurer une couverture exhaustive, à moins d'envisager l'utilisation de couches accessibles au grand public, telles que celles fournies par Google Earth ou Bing. L'année d'acquisition desdites images présente souvent des caractéristiques très variables. Elles présentent donc un intérêt limité pour l'estimation des surfaces cultivées au cours d'une année ou d'une saison données. Elles peuvent toutefois s'avérer très utiles à l'élaboration d'une base de sondage, sauf en cas d'évolution rapide du paysage agricole en question.

Plusieurs images d'une résolution de 10 à 30 m sont aujourd'hui disponibles gratuitement, permettant ainsi la réalisation d'une couverture complète. Elles ne peuvent toutefois contribuer à l'élaboration d'une base de sondage que si la superficie de la parcelle principale n'est pas excessivement réduite et s'il est possible d'en distinguer les contours dans le paysage.

D'autres types d'images, notamment l'imagerie par radar, semblent offrir une solution prometteuse en amélioration constante ; elles ne sont toutefois actuellement utilisées dans le secteur agricole que pour la surveillance du riz paddy.

3.1. BASES DE SONDAGE DE TYPE LISTE ET BASES ARÉOLAIRES

Les bases de sondage établies dans le but d'effectuer des statistiques agricoles sont souvent classées en deux catégories principales : les bases de type liste (BTL) et les bases aréolaires (BSA). Les premières peuvent prendre la forme de recensements agricoles ou de listes des ménages, des fermes ou des exploitations agricoles dressées à partir de registres administratifs. Dans les secondes, les unités d'échantillonnage sont des éléments de l'espace géographique : points, lignes (transects) ou polygones (segments).

Les différences entre BTL et BSA sont loin d'être évidentes :

- Souvent, les districts de recensement (DR) jouent le rôle d'unités primaires d'échantillonnage (UPE) dans le cadre des BTL. Si leurs frontières sont visibles (avec un peu de chance, elles font l'objet d'une couche SIG à part entière), la première étape de l'élaboration d'une BTL pourra être considérée comme une BSA;
- Lors de la toute première enquête menée dans le cadre d'une BSA, une liste d'exploitations agricoles est associée aux points ou aux segments échantillonnés. Si l'identification des exploitations agricoles associées aux points ou aux segments peut s'avérer difficile la première année, la liste ainsi établie peut tout de même être considérée comme un échantillon présentant une probabilité proportionnelle à la taille (PPT). Ainsi, à l'opposé d'une liste complète des exploitations résidant dans le DR, il sera plus facile de mettre à jour une liste constituée à partir des points échantillonnés d'une année sur l'autre.

La télédétection joue un rôle beaucoup plus important dans l'élaboration des BSA que dans celle des BTL, plus traditionnels. Elle n'en est pas moins utile pour ces derniers, notamment lorsqu'il existe une couche SIG représentant les frontières des DR et que la stratification peut intégrer les informations sur la couverture du sol obtenues par télédétection.

3.1.1. Stratification

La définition d'une stratification constitue probablement l'application la plus courante de la télédétection aux bases de sondage agricoles. Les stratifications effectuées à partir de données de télédétection peuvent prendre différentes formes dans le cadre des bases aréolaires. En ce qui concerne les BTL, en revanche, l'applicabilité des outputs de la télédétection se limite généralement à leur aspect « aréolaire ». Les images peuvent notamment servir à stratifier les DRs lorsque ceux-ci servent d'UPE et qu'il existe une couche SIG représentant leurs frontières.

Les images classées ou les cartes de couverture du sol élaborées à partir d'images peuvent servir à définir des strates répondant à certains critères, tels que le pourcentage total de terres cultivées ou la prédominance de certaines cultures. Par exemple, les strates peuvent être définies comme des ensembles d'unités composées de plus de 60 %, de 30 à 60 %, de 10 à 30 % ou de moins de 10 % de terres cultivées. Certaines strates peuvent également être associées à des cultures spécifiques. C'est le cas par exemple d'un ensemble d'unités comprenant plus de 50 % d'oliviers ou abritant un grand nombre de terres irriguées. Si les images classées ou les cartes de couverture du sol sont fortement biaisées, l'efficacité de la stratification en matière de variance s'en trouvera affectée ; cela n'introduira en revanche aucun biais dans les estimateurs produits dans le cadre d'une enquête utilisant ce type de stratification. Les biais relatifs aux estimateurs de surface peuvent être dus à la présence d'erreurs systématiques dans les observations des enquêteurs, mais non à une stratification erronée.

3.1.2. Erreurs de stratification

Il n'est pas rare que les observations effectuées sur une unité d'échantillonnage dans le cadre d'une enquête ne correspondent pas à la définition de la strate à laquelle appartient ladite unité. Par exemple, dans l'exemple de Sétif (section 0 ci-dessous), un segment i , composé de 40 à 70 % de terres cultivées, a été échantillonné dans la strate 3. Partons de l'hypothèse selon laquelle les observations menées sur le terrain indiquent qu'en réalité, la surface cultivée totale du segment ne s'élève qu'à 25 %. Le coordonnateur de l'enquête peut considérer que ce segment a été mal attribué au moment de la stratification, et que cette dernière sera de meilleure qualité s'il est réaffecté à la strate 2. Ce serait une erreur, qui pourrait gravement biaiser les estimateurs, la règle ne pouvant être appliquée à des

segments identiques hors de l'échantillon (sur lesquels on ne dispose donc d'aucune information). Il est important de respecter les principes selon lesquels 1) le processus de stratification est antérieur à l'échantillonnage et 2) il présente une qualité équivalente au sein de l'échantillon sélectionné et en dehors. La découverte de nombreuses erreurs de classification dans l'échantillon peut inciter les statisticiens à envisager une révision de la stratification ; dans ce cas, celle-ci doit être révisée en vue des enquêtes suivantes conformément à une règle générale et à un degré d'information applicables à l'ensemble du territoire. La stratification ne peut être révisée sur la seule base des informations collectées sur l'échantillon.

3.2. BASES DE SONDAGE ARÉOLAIRES

3.2.1. Bases contenant des segments délimités par des frontières physiques

Les segments contenus dans une base aréolaire peuvent être délimités par des frontières physiques (p. ex., rivières ou routes). C'est ainsi que procède l'USDA, par exemple, depuis les années 1930 (Davis, 2009). La délimitation des unités d'échantillonnage se fonde, dans une large mesure, sur la photo-interprétation. L'apparition de différentes couches de classification des cultures¹ a donné naissance à un nouvel outil de stratification (Boryan et Yang, 2012), qui n'exclut pas pour autant la nécessité de procéder à une identification visuelle des frontières, une activité qui mobilise des ressources conséquentes et exige un temps considérable. Les efforts de photo-interprétation peuvent être réduits par l'introduction d'une étape intermédiaire consistant à délimiter et à échantillonner les UPE. Les UPE sélectionnées sont ensuite divisées en segments, dont un seul est sélectionné lors de la deuxième étape (figure 1). Notons que cette approche n'est pas un échantillonnage à deux degrés à proprement parler dans la mesure où un seul segment par UPE étant sélectionné lors de la deuxième étape.

La photo-interprétation est essentielle à la détermination des unités d'échantillonnage dans les bases aréolaires contenant des segments délimités par des frontières physiques. Elle permet également de fournir des informations extrêmement utiles à la stratification. Ce type de base aréolaire constitue cependant rarement une solution adaptée dans les pays en développement caractérisés par des parcelles agricoles de petite taille.

FIGURE 1 UPE DÉLIMITÉE PAR DES FRONTIÈRES PHYSIQUES AU SEIN DE LAQUELLE UN SEGMENT A ÉTÉ SÉLECTIONNÉ.



¹ https://www.nass.usda.gov/Research_and_Science/Cropland/Release/index.php.

3.2.2. Bases fondées sur des grilles régulières

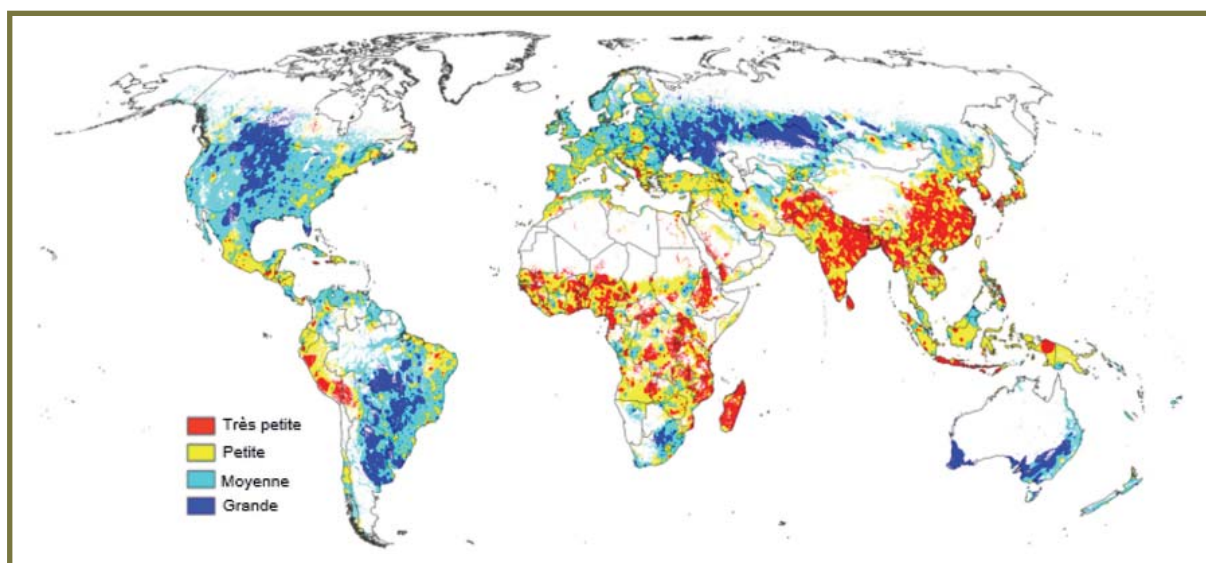
Une autre solution consiste à définir les segments en appliquant une grille géométrique habituellement composée de segments carrés à une projection cartographique donnée. Lors de la définition de la base aréolaire, les segments carrés présentent généralement les mêmes niveaux de précision que les segments délimités par des frontières physiques, pour un coût moins élevé. Une base aréolaire de ce type pourrait être définie dans un premier temps comme la simple superposition d'une grille composée de carrés à des frontières administratives. Les petits polygones situés aux frontières peuvent être supprimés s'ils sont de taille inférieure à la moitié de la superficie d'une cellule de la grille. La couverture sous-estimée ainsi produite peut être plus ou moins compensée en attribuant aux cellules incomplètes une pondération équivalente à celle d'une cellule entière. Dans la figure 2, par exemple, cette technique introduit un biais très faible, de l'ordre de 0,05 % seulement, pour l'ensemble de la zone. Ces opérations se fondent uniquement sur des SIG basiques, et non sur la télédétection.

FIGURE 2. SUPERPOSITION D'UNE GRILLE CARRÉE À UNE RÉGION ADMINISTRATIVE (GAUCHE) ET GRILLE CARRÉE APRÈS SUPPRESSION DES CELLULES DE PETITE TAILLE (DROITE).



Dans l'exemple ci-dessus, une grille de taille homogène a été utilisée. Un premier niveau d'affinage peut supposer d'adapter la taille de la grille à celle de la parcelle. Plusieurs méthodes permettent de parvenir à ce résultat. Il est possible, par exemple, de se fonder sur une carte obtenue par télédétection des parcelles dont les superficies sont dominantes. La figure 3 est une carte exploitable présentant les principales superficies des champs (Fritz *et al.*, 2015). Certains utilisateurs préféreront cependant utiliser leurs propres connaissances locales pour l'affiner ou l'améliorer.

FIGURE 3. CARTE MONDIALE DE LA SUPERFICIE LA PLUS COURANTE DES CHAMPS.



Une autre solution consiste à partir d'une grille à large maillage et à visualiser les cellules échantillonnées à l'aide d'une image THR. Il est possible de fixer une règle pour réduire la zone à observer si le nombre de champs dépasse un certain seuil (figure 4). Cette observation partielle n'introduit aucun biais dans l'estimation de la superficie tant que la règle reste fixe (p. ex., si l'on ne retient que la partie centrale d'un segment important) et appliquée indépendamment de la couverture du sol étudiée. À noter tout de même qu'elle complique le calcul de la variance des estimateurs.

La règle de base souvent appliquée en matière de taille des segments est la suivante : le nombre moyen de champs ou de parcelles de couverture du sol par segment doit être compris entre 10 et 20, de façon à ce qu'il ne soit pas nécessaire de consacrer plus d'une demi-journée à la visite sur le terrain (Taylor *et al.*, 1997).

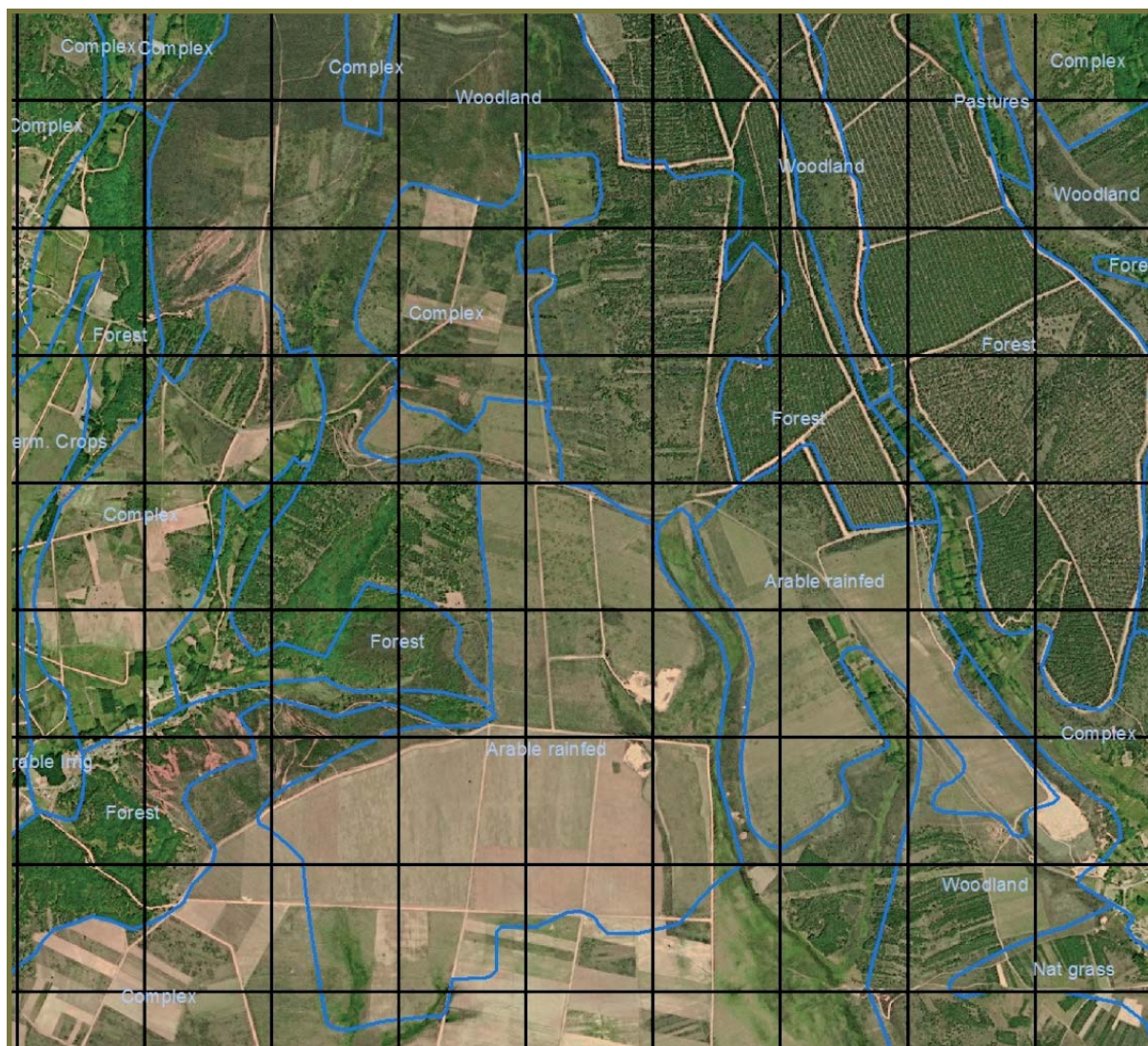
FIGURE 4. SEGMENT RÉDUIT COMPTE TENU DU NOMBRE TROP ÉLEVÉ DE CHAMPS QU'IL ENGLOBE.



3.2.2.1. Stratification à l'aide de cartes nationales ou régionales de couverture du sol

Les cartes de couverture du sol ou les images satellite classées permettent de calculer, à partir des informations disponibles, un indicateur de la surface cultivée pour chacune des cellules de la grille. Ce calcul produit généralement des résultats, qui ne peuvent être qualifiés de « statistiques » compte tenu de leur qualité médiocre, mais qui restent très utiles pour stratifier le territoire. Ce type de stratification s'avère particulièrement pratique pour l'élaboration de bases de sondage bon marché à partir de grilles régulières. Il ne présente en revanche généralement aucun intérêt pour les bases aréolaires élaborées à partir de segments délimités par des frontières physiques, puisque la proportion de terres cultivées ou de types particuliers de cultures au sein de chaque UPE peut être évaluée visuellement lorsque leur périmètre est clairement défini.

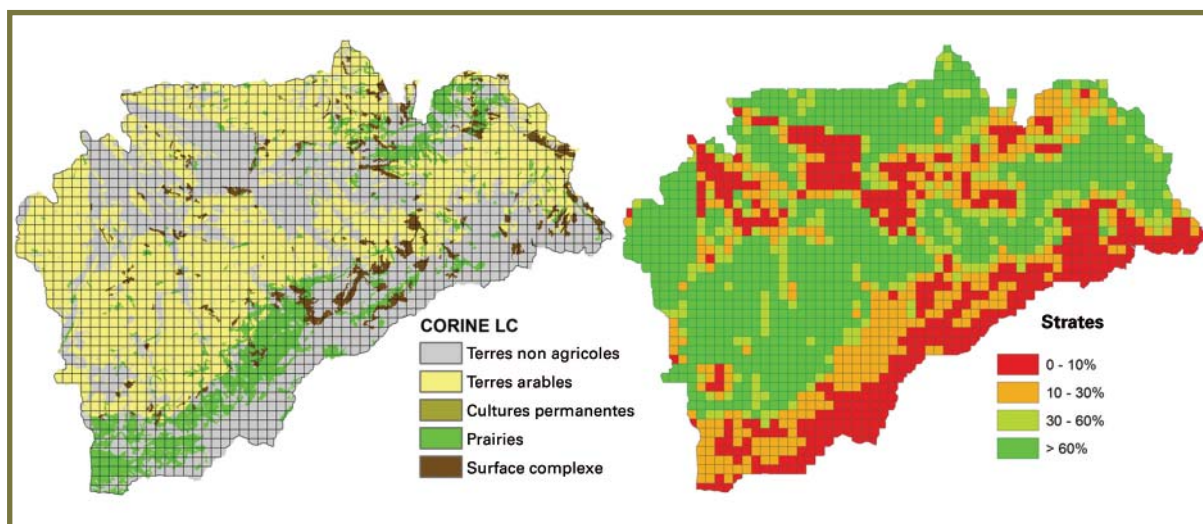
FIGURE 5. SUPERPOSITION D'UNE CARTE DE COUVERTURE DU SOL ET D'UNE GRILLE D'ÉCHANTILLONNAGE.



La figure 5 illustre la superposition d'une image et d'une carte de couverture du sol disponible, issue du projet CORINE Land Cover (CLC). Comme on peut le voir, la carte est imparfaite et propose des classes vagues, telles que «[surface] complexe». Dans tous les cas, ce type de carte peut servir à attribuer un indice relatif aux terres cultivées à chaque unité de la grille d'échantillonnage (dans le cas présent, des carrés), par exemple :

Indice relatif aux terres arables=terres arables+0.4 × complexe+0.2 × cultures permanentes+0.2 × prairies

FIGURE 6. SUPERPOSITION D'UNE GRILLE D'ÉCHANTILLONNAGE À UNE CARTE DE COUVERTURE DU SOL HAUTE RÉOLUTION (GAUCHE) ET STRATIFICATION QUI EN DÉCOULE (DROITE)



Dans la figure 6, ce type de stratification est appliqué à la base aréolaire de la figure 2. Dans cet exemple, l'indice est calculé à partir de la superficie occupée par les cultures permanentes, puisqu'il a été observé que les polygones étiquetés «cultures permanentes» par le projet CLC contiennent en réalité une certaine part de terres arables. L'indice peut être peaufiné davantage à l'aide des éventuelles connaissances de l'utilisateur sur la carte de couverture du sol et les seuils servant à définir les strates. Malgré les imperfections du CLC et de l'indice calculé, cette méthode de stratification, utilisée pour recenser les principales cultures en Espagne, présente une efficacité relative d'environ 2 (Gallego *et al.*, 1999).

D'autres cartes de couverture du sol similaires portant sur d'autres parties du monde sont disponibles. Par exemple, un grand nombre de cartes de pays africains a été produit dans le cadre du projet Africover (Latham, 2009). Bon nombre d'entre elles doivent être mises à jour, mais même les cartes obsolètes peuvent servir de base à la stratification², le principal risque étant que le plan d'échantillonnage exclue la strate étiquetée «terres non cultivées» alors que d'importantes surfaces ont été mises en culture entre temps. Ce facteur produira un biais important dans les estimations s'il est ignoré par le statisticien. Pour régler ce problème en l'absence de ressources permettant une mise à jour complète de la carte, il convient de recourir à un processus d'échantillonnage en deux phases pour les strates de terres non cultivées.

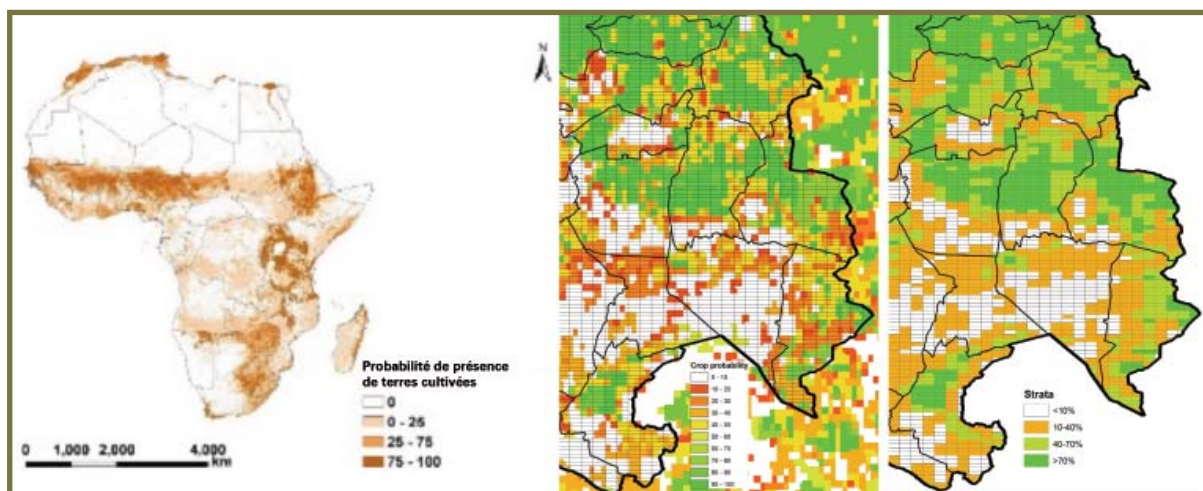
3.2.3. Stratification à partir de cartes mondiales des terres cultivées

La figure 7 illustre un processus de stratification effectué à Sétif (Algérie) à partir d'une carte de probabilité de présence de cultures obtenue en comparant différentes cartes de couverture du sol existantes (Fritz et See, 2008). Dans cet exemple, les cellules d'échantillonnage sont représentées par des rectangles de 1 500 m × 600 m. Deux rangées de six points ont dû faire l'objet d'une visite dans chacune des cellules échantillonnées (observation incomplète de segments dans un système d'échantillonnage à deux degrés). Un «indice de probabilité de présence de cultures», calculé à partir d'une moyenne pondérée sur la carte, a été attribué à chaque cellule de 1 500 m × 600 m aux fins de stratification. Le sens de cet indice n'est pas tout à fait évident et ne correspond pas

² <http://www.fao.org/geonetwork/srv/fr/main.present>.

à la proportion prévue de terres cultivées ; il a toutefois permis d'effectuer une stratification rationnelle définie par les intervalles 0-10 %, 10-40 %, 40-70 % et plus de 70 %.

FIGURE 7. STRATIFICATION À PARTIR D'UNE CARTE DE PROBABILITÉ DE PRÉSENCE DE CULTURES ET D'UNE GRILLE D'ÉCHANTILLONNAGE COMPOSÉE DE RECTANGLES, À SÉTIF (ALGÉRIE).



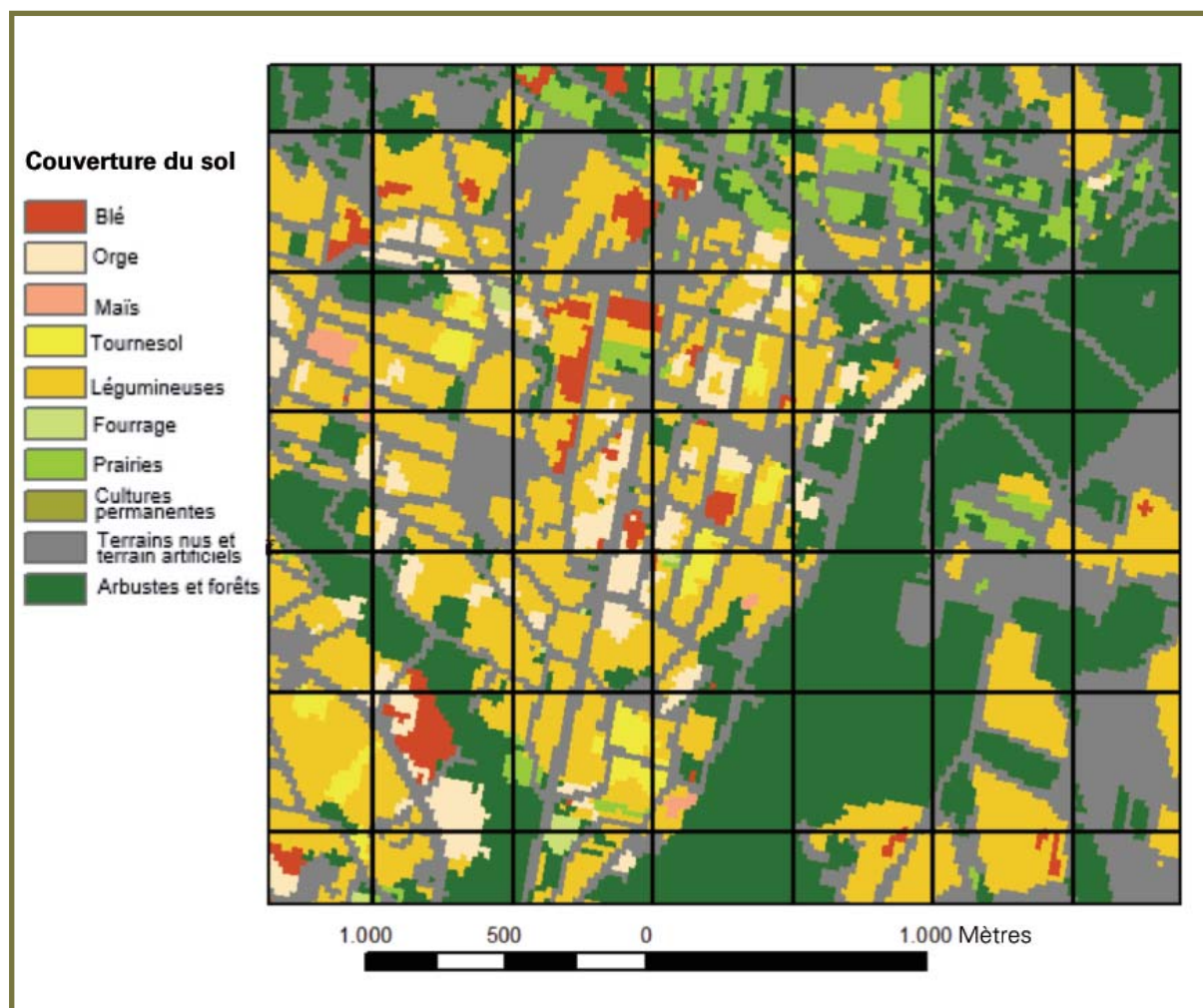
La base de données Global Land Cover-SHARE de la FAO³ constitue un autre exemple de système international de ce type. Elle peut paraître plus complète, puisqu'elle estime la part représentée par 11 types principaux de couvertures du sol pour chaque pixel.

3.2.4. Stratification fondée sur des données plus détaillées relatives à une année précédente

Il existe dans certains pays des couches SIG plus détaillées qui ne sont pas toujours à jour. L'USDA a mis au point une méthode de stratification fondée sur la couche de données relatives aux terres cultivées (Boryan et Yang, 2012), qui prend la forme d'une vaste mosaïque d'images classées. Dans certains cas, des cartes détaillées des parcelles ont été produites quelques années auparavant. Ce type de couche a, par exemple, été utilisé à des fins de stratification dans le cadre d'un projet pilote mené dans le gouvernorat de Siliana, en Tunisie (Sghaier, 2016). La procédure est identique à celle présentée ci-dessus pour les produits présentant une résolution plus grossière : pour chaque élément de la base de sondage (généralement une cellule), un indicateur de la proportion de terres arables ou cultivées est estimé à partir de la moyenne pondérée des proportions disponibles sur la carte ; une stratification est ensuite effectuée à partir des intervalles de l'indicateur ou sur la base d'autres indicateurs portant sur certaines cultures spécifiques. La figure 8 montre un exemple d'image classée superposée à une grille de segments carrés pouvant être utilisée à des fins de stratification.

³ http://www.glcn.org/databases/lc_glcshare_en.jsp.

FIGURE 8. GRILLE DE SEGMENTS CARRÉS SUPERPOSÉE À UNE IMAGE CLASSÉE.



3.2.5. Photo-interprétation d'images du domaine public et production participative

Les images THR proposées gratuitement par des outils tels que Google Earth ou Bing peuvent servir à recueillir des informations agricoles. La création d'une carte de couverture du sol à partir de ces images à l'aide de méthodes standard et d'une résolution spatiale relativement fine suppose toutefois un travail conséquent de la part de photo-interprètes qualifiés et, partant, un budget considérable.

La production participative peut constituer une solution alternative bon marché à envisager, si l'on évite de sous-estimer ses limites et si l'on définit des stratégies pour réduire leur impact. Le réseau Geo-wiki fait ainsi appel à des photo-interprètes bénévoles⁴ et réalise des progrès notables à cette fin :

- La communauté est capable de modifier les couches de polygones. Il peut toutefois être risqué de permettre à un grand nombre de personnes de modifier la géométrie des polygones. Les photo-interprètes peuvent être invités, sinon à délimiter des polygones, à évaluer des polygones prédéfinis (cellules de grilles ou polygones déterminés à l'aide d'un logiciel de reconnaissance des formes). La question posée aux bénévoles peut être très simple, comme le montre la figure 9 : « Voyez-vous des terres cultivées dans le rectangle ? » Il convient en revanche de poser des questions légèrement plus précises aux fins de stratification, telles que : « Le pourcentage de terres cultivées que vous voyez dans le rectangle est-il très élevé, élevé, moyen, faible ou nul ? ». Une

⁴ <https://laco-wiki.net/fr/Bienvenue>, <http://www.geo-wiki.org/>.

auto-évaluation du niveau de confiance peut en outre s'avérer utile : « Êtes-vous certain(e) de votre réponse, relativement confiant(e), ou n'êtes-vous pas sûr(e) du tout ? » Avec ce type de question, le temps consacré à la photo-interprétation varie de 30 secondes à une minute par cellule en présence d'une bonne connexion Internet (un frein qui s'avère déterminant). Un bénévole consacrant trois heures par jour à cet exercice peut ainsi évaluer entre 150 et 300 cellules.

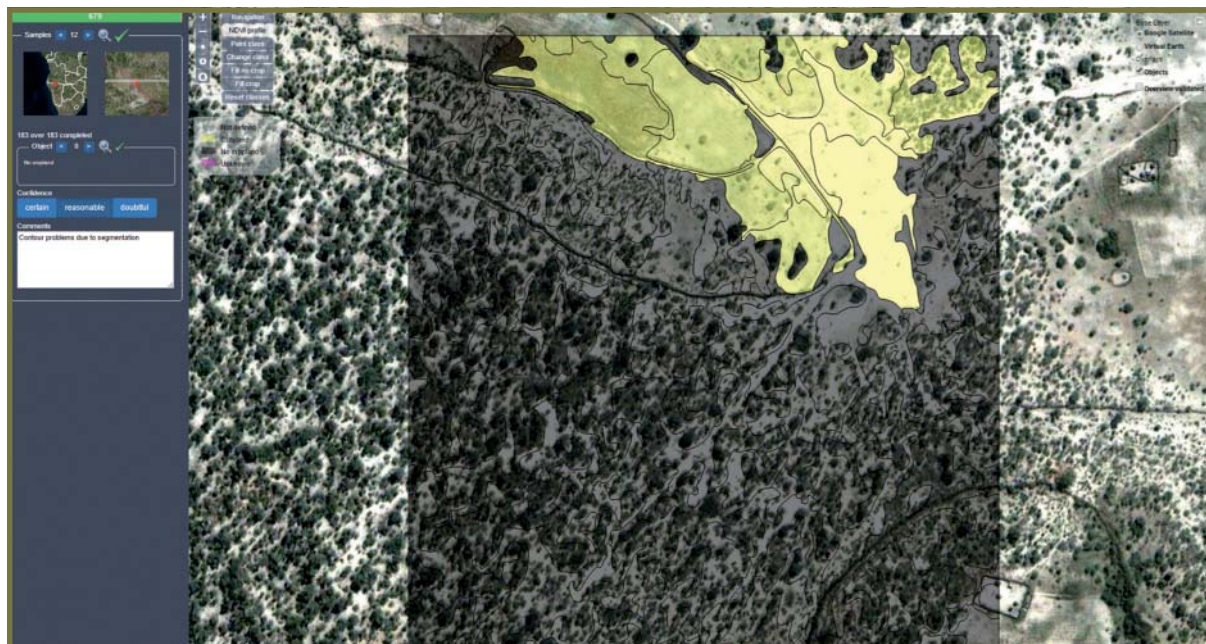
- La figure 10 présente un système plus élaboré. Le paysage a, au préalable, été divisé en polygones à l'aide d'un logiciel de segmentation automatique. Chaque recoupement entre ces polygones et une cellule de la grille actuelle est étiqueté « terre cultivée » ou « terre non cultivée » par le photo-interprète. Le degré global de fiabilité est également évalué. L'outil de photo-interprétation calcule ensuite un indice relatif à la proportion probable de terres cultivées qui servira à la stratification.
- Il est essentiel d'évaluer la fiabilité des bénévoles. Le réseau Geo-wiki examine les photo-interprétations effectuées par différents bénévoles ou par des bénévoles et des spécialistes afin de leur attribuer des notes de fiabilité. Lorsque les bénévoles obtiennent systématiquement d'excellents résultats, les notes qu'ils attribuent à leurs pairs bénéficient d'une pondération plus forte.
- Il est par ailleurs important de motiver les bénévoles. Des systèmes de récompenses (p. ex., offrir un smartphone au meilleur photo-interprète du mois) sont en cours de pilotage, et produisent des résultats encourageants.

Il n'est pas toujours possible de procéder à une stratification complète avec la seule contribution des bénévoles les plus performants disponibles. Supposons, par exemple, que la superficie la plus courante des champs mesure entre 0,5 et 1 ha. Les segments idéaux mesurent quant à eux 300 m × 300 m. Si les strates agricoles du pays occupent une surface de 300 000 km², il faudrait photo-interpréter plus de 3 millions de cellules pour obtenir une stratification complète. Un système d'échantillonnage en deux phases semblable à celui utilisé dans l'enquête LUCAS de l'UE pourrait s'avérer plus adapté dans ce cas. Il serait par exemple envisageable de procéder à un premier échantillonnage systématique au sein d'une grille de 5 × 5, ce qui ramènerait à environ 130 000 le nombre de cellules à photo-interpréter.

FIGURE 9. INTERFACE SIMPLE DE PHOTO-INTERPRÉTATION PARTICIPATIVE.



FIGURE 10. SYSTÈME DE PHOTO-INTERPRÉTATION PLUS ÉLABORÉ, COMPOSÉ DE POLYGONES PRÉALABLEMENT DÉLIMITÉS PAR UN LOGICIEL DE SEGMENTATION AUTOMATIQUE.



3.2.6. Bases aréolaires constituées de points en grappe

Pour les bases aréolaires constituées de points (en grappe ou non) ou de segments, les unités d'échantillonnage n'ont pas besoin d'être définies par imagerie. Celle-ci pourra en revanche s'avérer fondamentale pour :

- la stratification, en particulier en cas d'échantillonnage de points, en deux phases précédée d'une stratification basée sur la photo-interprétation du pré-échantillon, et ;
- les documents relatifs à l'enquête de terrain : lorsque la localisation d'un point fournie par GPS ne correspond pas à celle constatée sur l'image, cette dernière, plus stable et jouissant d'une meilleure traçabilité, est à privilégier, en particulier si l'identification des changements de couverture du sol constitue l'une des principales cibles.

Les bases aréolaires à points sont principalement constituées de points en grappe ou non. Les bases à points en grappes correspondent à l'échantillonnage traditionnel à deux degrés. Elles peuvent également être considérées comme des bases à segments ayant fait l'objet d'une observation incomplète : en effet, au lieu d'observer toute la surface occupée par le segment, seuls les types de couverture du sol (cultures ou autres classes de couverture du sol) correspondant aux points échantillonnés sont observés. L'échantillon des points contenus dans les segments réguliers prend souvent la forme d'une grille régulière (figure 11). Cette solution réduit considérablement la charge de travail tout en n'altérant que faiblement la précision de l'étude. L'enquête Teruti menée en France (FAO, 1998) depuis les années 1960 en constitue un bon exemple. Cette dernière utilise des grappes de points de 6×6 espacés de 300 m ; toutefois, si la superficie de la parcelle dominante est relativement petite, il peut s'avérer plus optimal d'espacer les points de 100 m seulement.

FIGURE 11. GRILLE RÉGULIÈRE DE POINTS APPLIQUÉE À UN SEGMENT CARRÉ (SYSTÈME D'ÉCHANTILLONNAGE À DEUX DEGRÉS).



La procédure présentée ci-dessus peut s'appliquer pour les bases aréolaires de segment fondées sur une grille régulière : l'ensemble du carré ou du rectangle est alloué à une strate en fonction des informations dont on dispose, même si l'enquête ne porte que sur une grille de points en son sein.

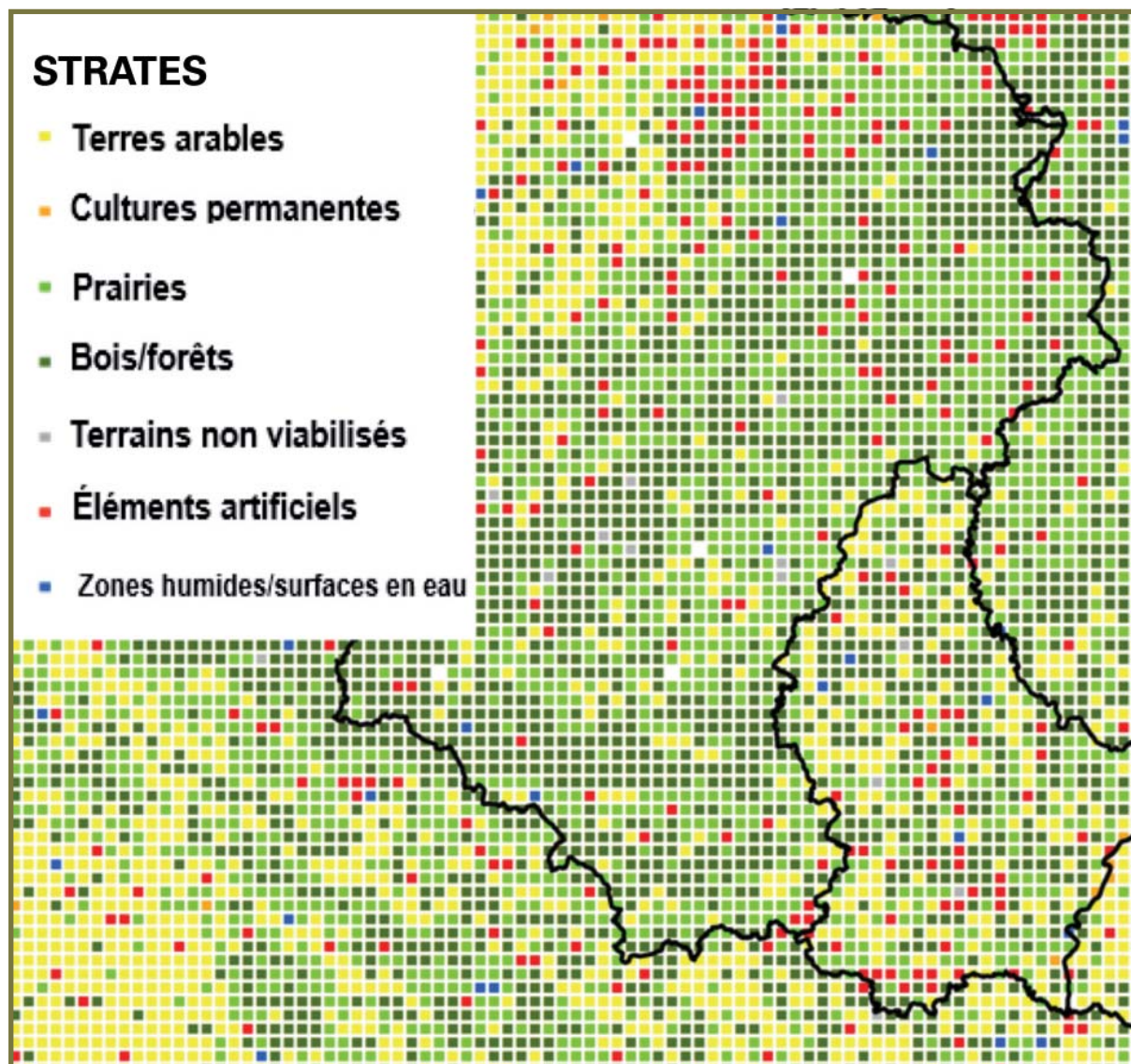
3.2.7. Bases aréolaires constituées de points non regroupés

Les points non regroupés sont généralement utilisés dans le cadre d'un processus d'échantillonnage en deux phases (à ne pas confondre avec l'échantillonnage à deux degrés). La première phase consiste à sélectionner un grand échantillon en vue d'une stratification incomplète, laquelle sert ensuite de base à la sélection de l'échantillon final. Citons comme exemples la version modifiée (2002) de l'enquête italienne Agrit (Martino, 2003) et l'enquête LUCAS d'Eurostat (Gallego et Delincé, 2010).

L'étude comparative des résultats et des coûts liés aux deux versions d'Agrit (Martino, 2003 ; Gallego et Delincé, 2010) montre que la technique des points non regroupés permet d'améliorer la rentabilité des enquêtes à base de sondage aréolaire dans le contexte européen (compte tenu de ses paysages et réseaux routiers) ; cette étude

produirait toutefois probablement des résultats très différents dans les pays en développement, étant donné l'impact plus important des frais de déplacement dans ces pays.

FIGURE 12. ÉCHANTILLON STRATIFIÉ DE LA PREMIÈRE PHASE (EUROPE CENTRALE).



Les principales étapes sont les suivantes :

- Sélection d'un vaste échantillon de points (pré-échantillon). Dans le cadre d'Agrit et de LUCAS, des grilles régulières constituées de plus d'un million de points chacune ont été sélectionnées.
- Stratification du pré-échantillon par photo-interprétation à l'aide d'une nomenclature simple (figure 12).
- Sélection d'un sous-échantillon à l'aide d'un taux d'échantillonnage adapté aux priorités de l'enquête.

3.2.8. Étude comparative bases constituées de segments/de points

Il est plus facile de collecter et de traiter les données relatives aux points que celles relatives aux segments. Le travail de terrain mené sur les segments suppose de délimiter les champs compris dans un segment donné et de les numériser avant d'en estimer la superficie. Ce processus prend du temps (de quelques semaines à plusieurs mois) pour les échantillons de grande taille, ce qui retarde d'autant la production d'estimations. Le perfectionnement

des appareils de navigation bon marché, tels que les smartphones, a davantage amélioré la rentabilité des bases constituées de points que celle des bases constituées de segments.

Les segments aréolaires fournissent des informations plus adaptées pour la coregistration géométrique, c'est-à-dire la superposition des informations obtenues sur le terrain aux images satellite ; ils sont également en mesure de fournir des informations de meilleure qualité sur la structure et la taille des parcelles. Ce point peut s'avérer utile pour la production d'indicateurs agro-environnementaux, tels que les indices paysagers. Les segments se prêtent également mieux au regroupement avec des images satellite avec estimateur par régression (Carfagna, 2007).

Une étude comparative a montré que le niveau de précision des segments délimités par des frontières physiques était très semblable à celui fourni par les segments carrés. Ces derniers sont toutefois à privilégier, le processus d'élaboration de la base de sondage y afférente étant beaucoup moins coûteux (González *et al.*, 1991).

3.2.9. Bases aréolaires constituées de transects

Un transect se définit comme une ligne droite d'une certaine longueur. Ce terme, qui correspond à la définition du mot «segment» en géométrie élémentaire, est utilisé dans le cadre de l'élaboration de bases de sondage aréolaires, le terme «segment» revêtant une autre signification dans le présent contexte.

Les transects peuvent être combinés avec l'échantillonnage par points. Cette solution, testée dans le cadre d'un projet pilote en Chine (Kerdiles *et al.*, 2013), peut permettre de surmonter la difficulté que représente la localisation d'un point dans les paysages principalement caractérisés par de fines bandes de cultures. La localisation d'un point sur une image détermine la sélection d'un ensemble de bandes. On attribue audit point la part de chaque culture le long d'un transect perpendiculaire aux bandes, estimée dans la pratique en comptant le nombre de pas effectués le long des limites d'une large parcelle (figure 13). Cette méthode présente certains inconvénients encore non évalués, mais peut offrir une solution intéressante pour ce type de paysage difficile.

FIGURE 13. TRANSECT GÉNÉRÉ À PARTIR D'UN POINT ÉCHANTILLONNÉ POUR ENQUÊTER SUR DES PARCELLES LONGUES ET FINES.



Les transects sont souvent utilisés dans les enquêtes environnementales pour estimer la longueur d'éléments linéaires (de Vries, 1986 ; Gallego et Delincé, 2010), mais sont rarement employés aux fins de statistiques agricoles.

En principe, un transect n'a qu'une seule dimension (sa largeur étant de 0) ; toutefois, ce terme est aussi utilisé pour décrire des polygones longs et fins, mesurant par exemple $100\text{ m} \times 50\text{ km}$. Ce type d'unités d'échantillonnage est bien adapté aux vols à basse altitude (Jolly et Watson, 1979). Elles ont notamment servi au recensement des cheptels nomades et pourraient être utilisées plus souvent si la loi autorisait les vols d'UAV (également connus sous le nom de drones) de relativement longue durée. Il conviendrait sinon d'envisager de recourir de nouveau à de petits avions pilotés, en tenant compte des améliorations apportées aux appareils photo (résolution de 3-5 cm) et à l'orthorectification des images (Gallego *et al.*, 2016). Les simulations effectuées à partir de données aux Pays-Bas montrent que les bandes longues et fines, plus adaptées aux vols de basse altitude, produisent bien davantage de résultats que les segments carrés (figure 14).

FIGURE 14. SYSTÈME D'ÉCHANTILLONNAGE PAR BANDES DANS LE CADRE D'UNE ÉTUDE DE SIMULATION (PAYS-BAS).



3.3. UTILISATION DE LA TÉLÉDÉTECTION DANS LES BASES DE SONDAGE DE TYPE LISTE

3.3.1. Localisation d'une exploitation agricole : un problème difficile à résoudre

Les informations tirées d'images prises par satellite ou par avion sont étroitement liées à la localisation précise d'un pixel. C'est pourquoi le recours potentiel à la télédétection dans le cadre des BTL dépend en grande partie des informations de localisation des éléments d'une BSA. Les unités élémentaires (individus statistiques) des BTL sont les exploitations agricoles ou les ménages. Bien que les sections suivantes font davantage référence aux premières, les analyses et résultats obtenus s'appliquent également, dans une large mesure, aux seconds. Les exploitations agricoles sont habituellement associées à un DR ou à d'autres types d'unités administratives de petite taille, qui constituent souvent la seule information disponible sur la localisation desdites exploitations. La représentation explicite des DR par SIG détermine la nécessité de recourir ou non à la télédétection.

3.3.2. Registres administratifs

Dans certains pays, les systèmes d'informations administratives fondés sur les SIG, intitulés «registres», contiennent une description géoréférencée détaillée de chaque exploitation (siège ou champs administrés). Les registres recensent les champs ou les parcelles agricoles en s'appuyant sur des documents orthophotographiques ou des images satellite. Ils peuvent également être utilisés pour la production directe et en temps utile de statistiques non biaisées, mais seulement dans certaines conditions particulières de temps, de précision et d'intégralité, difficiles à remplir dans la pratique. D'autre part, les parcelles définies dans un registre ne correspondent pas nécessairement à celles en monoculture. Cependant, même les registres obsolètes ou incomplets peuvent constituer une excellente source de données en vue de l'élaboration de bases de sondage — particulièrement aréolaires — fonctionnelles.

L'existence d'un registre peut offrir une solution satisfaisante au problème de représentation d'une exploitation agricole par SIG. Un procédé plus simple consisterait à attribuer à une exploitation les coordonnées de son siège. Le terme «siège» est toutefois difficile à définir et peut varier en fonction de la situation agricole de chaque zone géographique.

3.3.3. Utilisation des districts de recensement comme unités primaires d'échantillonnage

Les couches basées sur le SIG de petites zones administratives, telles que les districts de recensement (DR, utilisés dans le cadre de recensements), les communes, les villages, etc., se prêtent à des usages variés qui vont bien au-delà de la simple élaboration d'une base de sondage aux fins de statistiques agricoles. Bien que chère en termes absolus, la création d'une couche de petites unités administratives s'avère bien moins coûteuse que la tenue d'un registre administratif recensant les exploitations agricoles à partir du SIG. La création d'une couche de ce type dans le cadre d'un programme fiable de statistiques agricoles peut être une solution à la fois abordable et utile pour l'administration du territoire. En matière de statistiques agricoles, ce type de couche recensant les frontières des DR est utile à la fois pour les BTL et les BSA.

Dans ce chapitre, le terme «DR» est utilisé dans son sens générique. Ce terme peut correspondre aux DR tels qu'utilisés dans les recensements démographiques ou agricoles, ou à certaines unités administratives, telles que les communes ou les villages. La création, la mise à jour et l'amélioration de couches SIG représentant les frontières des DR peuvent constituer une étape importante dans le cadre d'une enquête agricole. Plusieurs cas de figure peuvent se présenter, qui dépendent de la définition légale ou opérationnelle du DR :

- Les DR sont définis sur des cartes topographiques imprimées. Ces dernières sont alors scannées afin de produire une image matricielle. Les limites des polygones sont ensuite numérisées et superposées aux cartes scannées en arrière-plan. Chaque DR est alors défini par une couche de polygones (généralement, un seul polygone). Toutes les couches de DR doivent être géoréférencées et fusionnées en une couche unique. La dernière étape consiste à corriger les chevauchements et à combler les espaces manquants afin d'obtenir une couche homogène. Il peut être utile pour cela de disposer d'une image présentant la résolution spatiale la plus fine possible en arrière-plan, puisque

les frontières des DR sont souvent délimitées par des routes, des cours d'eau et d'autres éléments paysagers visibles.

- Autre situation : il existe déjà une couche SIG constituée de polygones représentant les frontières. Elle doit toutefois être affinée, mise à jour ou corrigée. Dans ce cas, seules les dernières étapes décrites ci-dessus sont applicables. Si les limites des DR sont affinées à l'aide d'images satellite, il est important de vérifier au préalable la bonne qualité de leur géométrie, à la fois en matière de résolution spatiale et d'orthorectification. Notons en particulier qu'au moment de la rédaction de ce chapitre, la qualité de l'orthorectification des couches d'images fournies par Google Earth était insuffisante. Le DR présenté dans la figure 15 semble être délimité par des éléments paysagers physiques et pourra nécessiter quelques retouches. Il conviendra dans ce cas de procéder à des vérifications dans d'autres documents (plans cadastraux, par exemple) afin de garantir que les corrections appropriées sont effectuées.

FIGURE 15. EXEMPLE DE DR POUR LEQUEL IL EXISTE DES DOUTES QUANT À LA PRÉCISION GÉOMÉTRIQUE.



- La délimitation légale des DR découle de la description écrite de leurs frontières. Au mieux, leur définition formelle ou légale est dictée par des éléments paysagers visibles. Dans ce cas, l'imagerie est indispensable à la conversion des descriptions écrites en frontières géoréférencées. Ceci suppose à la fois un travail documentaire et des missions sur le terrain en collaboration avec des spécialistes locaux.
- Dans certains cas, la définition légale est plus vague et peut renvoyer aux ménages plutôt qu'au territoire. Par exemple, la définition du DR peut renvoyer spécifiquement à un petit groupe de population tout en mentionnant plus généralement les ménages voisins. Les administrateurs locaux ou agents de vulgarisation sont généralement capables d'identifier le DR auquel appartient un ménage donné, mais sa délimitation géographique précise peut en

revanche poser problème. Dans ce cas de figure, l'imagerie haute résolution s'avère très utile ; l'essentiel du travail effectué provient néanmoins des interactions entre les différents échelons administratifs locaux et nationaux. La création d'une couche SIG représentant les frontières des DR est essentielle pour une bonne administration du territoire ; ladite couche peut-être utilisée à de nombreuses autres fins que les seules enquêtes agricoles.

Les éléments paysagers définissant les frontières des DR doivent être clairement visibles sur les images servant de fondement à l'élaboration, à l'amélioration ou à la mise à jour des couches SIG afférentes. Ces images peuvent prendre différentes formes, par exemple :

- Couverture aérienne orthophotographique ;
- Couches publiques telles que celles fournies par Google Earth ou Bing — ces solutions s'avèrent fort intéressantes de par leur accessibilité et leur gratuité ; il convient toutefois de vérifier que leur précision géométrique soit suffisamment homogène et que la date à laquelle elles ont été prises soit bien connue (même si celle-ci est moins importante que leur résolution spatiale, qui doit être aussi fine que possible) ;
- Archives d'entreprises privées (p. ex., Digital Globe), accessibles sur inscription à un prix modéré.

3.3.4. Échantillonnage de Districts de Recensement avec une probabilité proportionnelle à la zone couverte

Nous partons de l'hypothèse que les DR à échantillonner se situent dans un pays ou dans une région pour lequel/laquelle il existe déjà une couche SIG représentant leurs frontières. Les statisticiens peuvent chercher à échantillonner les DR présentant une probabilité proportionnelle à la zone géographique D_i de chacun d'entre eux. L'approche type consiste à procéder à un échantillonnage aléatoire avec probabilité proportionnelle à la taille (PPT), tel qu'illustré dans la figure 17(a). Une méthode géographique plus simple consiste à échantillonner des points présentant une probabilité uniforme, puis de sélectionner les DR dans lesquels se situent ces points. À noter toutefois que la configuration spatiale de l'échantillon sera alors irrégulière, comme c'est souvent le cas dans le cadre d'échantillonnages aléatoires : plusieurs groupes de communes adjacentes seront sélectionnés, tandis que des zones relativement vastes demeureront vacantes. Cet inconvénient propre à l'échantillonnage aléatoire est bien connu et peut être facilement compensé grâce à un échantillonnage systématique. Dans certaines enquêtes à base de sondage aréolaire, la sélection des DR ou des UPE s'est appuyée sur un processus d'échantillonnage systématique unidimensionnel, dans le cadre duquel la population était répartie en fonction d'une règle portant sur leur localisation. Cette classification repose fréquemment sur une répartition en dents de scie. Cette stratégie a notamment été adoptée par l'USDA (Cotter et Tomczac, 1994). La figure 16 illustre une répartition en dents de scie de petites unités administratives rangées par bandes ouest-est. La disposition bidimensionnelle se transforme ainsi en un tableau unidimensionnel auquel peut être appliqué l'échantillonnage systématique traditionnel par PPT. Dans certains cas, la disposition spatiale demeurera irrégulière.

FIGURE 16. MODÈLE DE RÉPARTITION DE COMMUNES EN DENTS DE SCIE (ÉTAPE PRÉALABLE AU PROCESSUS D'ÉCHANTILLONNAGE SYSTÉMATIQUE).

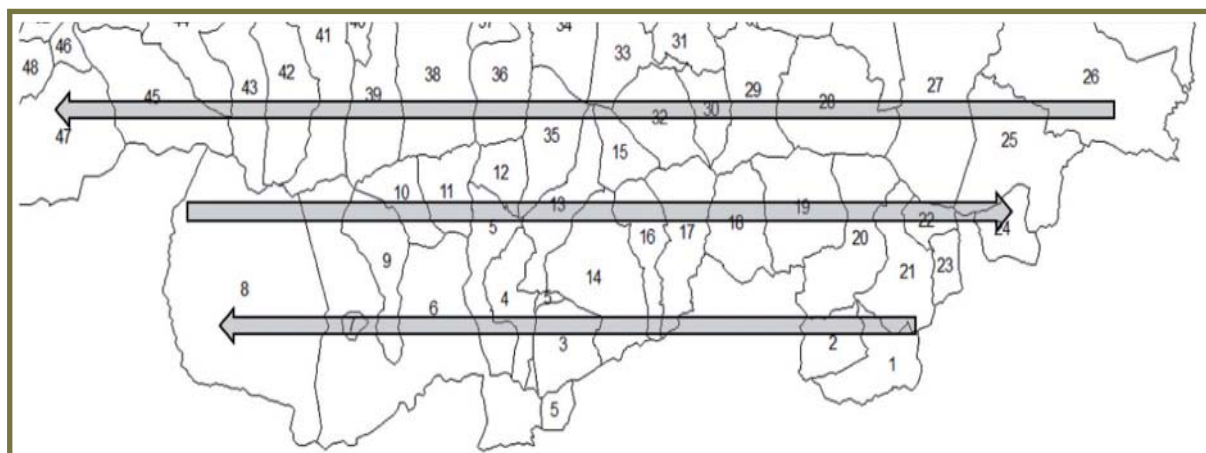
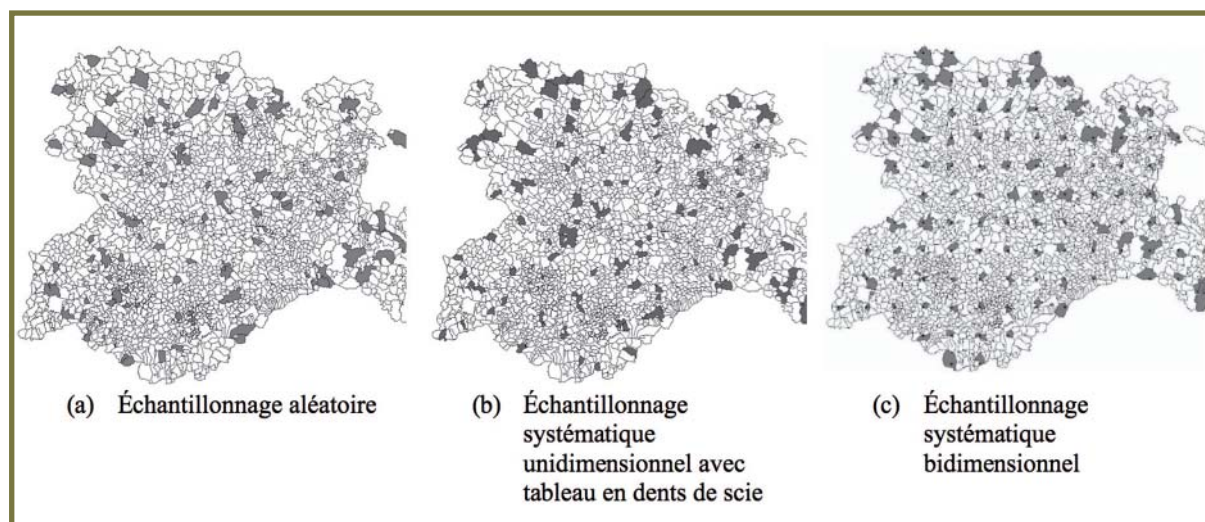


FIGURE 17. ÉCHANTILLONS AVEC PPT DE COMMUNES DE CASTILLE-ET-LEÓN (ESPAGNE) ET REPRÉSENTATION DE LA ZONE GÉOGRAPHIQUE CORRESPONDANTE (RÉGION) SERVANT D'OUTIL DE MESURE DE LA TAILLE.



L'échantillonnage bidimensionnel avec PPT peut être effectué à l'aide d'une grille systématique de points (figure 17(c)) offrant une répartition plus régulière. Ce type de disposition géographique régulière améliore l'efficacité du processus (Dunn et Harrison, 1993), à condition que la corrélation spatiale diminue avec l'éloignement. Le principal inconvénient tient au fait que l'estimation de la variance est nécessairement biaisée. Les formules standard de variance donnent généralement lieu à une surestimation, de telle façon que la mise au point peut être occultée par le biais d'estimation de la variance. La traçabilité du processus d'échantillonnage constitue sans doute le principal avantage de l'échantillonnage systématique. La figure 17(c) présente l'exemple de la communauté autonome de Castille-et-León (Espagne), où la sélection d'un échantillon de communes a été effectuée à partir d'une grille régulière composée de cellules espacées de 30 km.

L'échantillonnage avec probabilité proportionnelle à la zone géographique n'est pas optimal aux fins de statistiques agricoles. Il est recommandé de procéder à un échantillonnage avec probabilité proportionnelle à la surface cultivée au sein de chaque DR. D'autres critères, tels que les surfaces agricoles utilisées (SAU) ou les cultures annuelles du moment, peuvent également s'avérer intéressantes. Il suffit pour cela de photo-interpréter un échantillon de points, à condition que les terres cultivées soient identifiables sur les images. Cette solution peut être utile s'il n'existe aucune donnée récente sur le nombre total de terres cultivées par DR pouvant servir à l'échantillonnage avec PPT. La photo-interprétation d'un échantillon de points et la sélection du DR concerné si l'emplacement d'un point correspond à une terre cultivée (figure 18) donnent automatiquement lieu à un échantillonnage PPT de DR, même si la superficie cultivée dans chaque DR et dans la région tout entière est inconnue. Cette information est toutefois nécessaire pour calculer les estimateurs relatifs au total de toute variable supplémentaire Z , ce qui pose problème :

$$\hat{Z} = \sum_i z_i \frac{A}{n a_i}$$

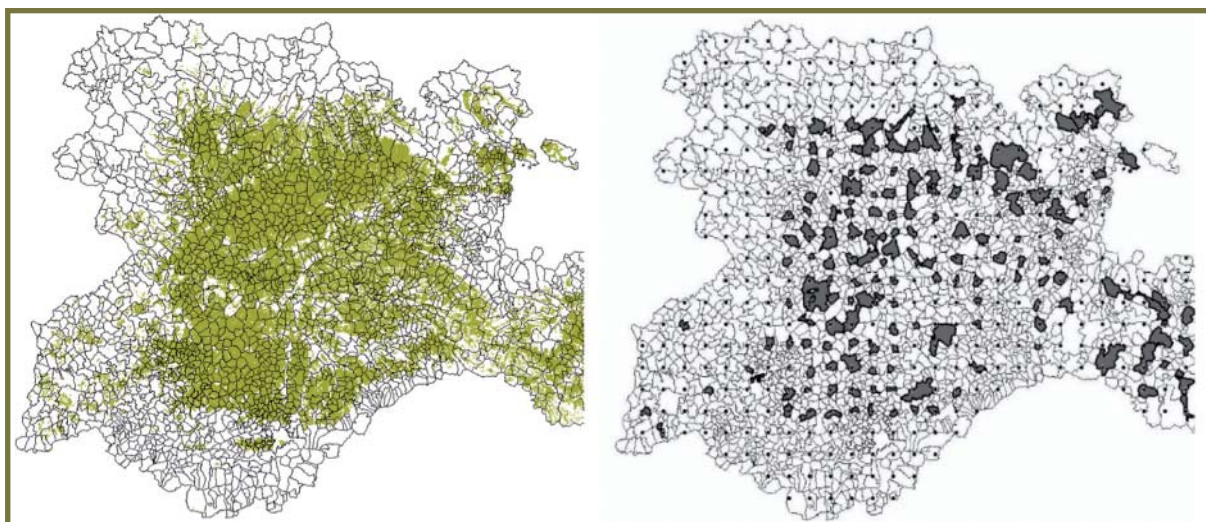
où A et a_i correspondent respectivement à la surface cultivée dans la région et dans le i^{e} DR. Pour pallier ce problème, il est possible de photo-interpréter un échantillon plus large — considéré alors comme un pré-échantillon — ou de passer par un intermédiaire, tel que des images classées ou des cartes de couverture du sol. En présence d'une carte de couverture du sol ou d'une image classée convenables, la superficie cultivée (ou tout autre concept similaire) d'après la carte pourrait servir de paramètre déterminant pour l'échantillonnage avec PPT.

FIGURE 18. POINT PHOTO-INTERPRÉTÉ COMME TERRE CULTIVÉE (GAUCHE) ET POINT PHOTO-INTERPRÉTÉ COMME TERRE NON CULTIVÉE (DROITE)



La figure 19 représente des terres arables telles que recensées par le projet CLC (Agence européenne pour l'environnement, 2007), qui porte sur l'UE et certains pays voisins. Dans cet exemple, les terres arables sont regroupées au cœur de la région de Castille-et-León. L'échantillonnage des DR avec probabilité proportionnelle aux terres arables se concentre donc principalement sur le centre de la région. Ici, l'échantillonnage s'est fait en deux temps : les statisticiens ont défini une grille systématique avant de sélectionner les DR abritant les points de la grille correspondant à des terres arables d'après le CLC.

FIGURE 19. RÉPARTITION DES TERRES ARABLES EN CASTILLE-ET-LEÓN (GAUCHE) ET ÉCHANTILLON SYSTÉMATIQUE DE COMMUNES PRÉSENTANT UNE PROBABILITÉ PROPORTIONNELLE À LA SURFACE DE TERRES ARABLES (DROITE).



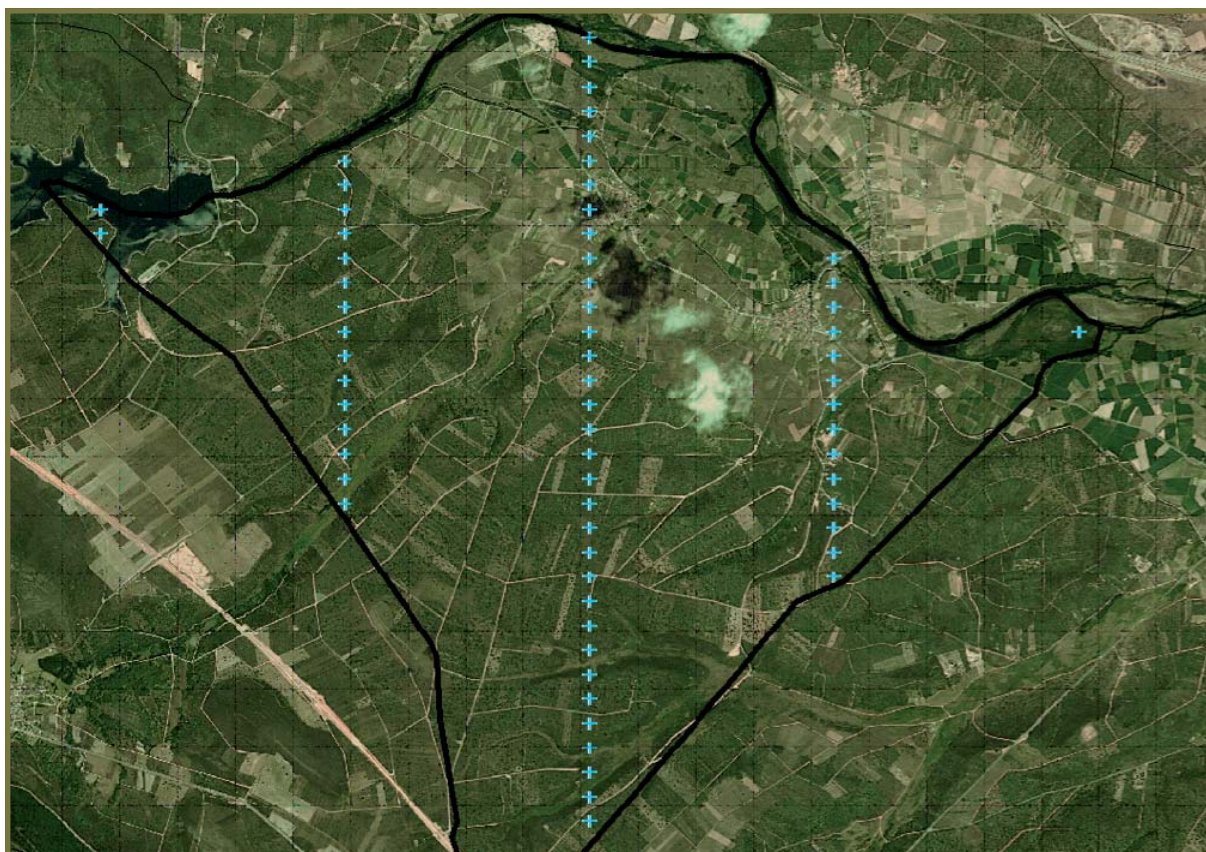
3.3.5. Stratification des districts de recensement

Traditionnellement, les DR sont stratifiés à partir de données tirées d'un recensement ou d'autres sources d'information, tels que les rapports administratifs fournis par les agents de vulgarisation locaux recensant le nombre total de terres cultivées ou de têtes de bétail. Si ces informations sont indisponibles, trop anciennes ou

considérées comme non fiables, il est alors possible de se tourner vers la télédétection. Les méthodes présentées dans la section 3.2 ci-dessus peuvent être appliquées pour stratifier les DR s'il existe une couche SIG représentant leurs frontières. Dans ce cas, la stratification peut être effectuée en procédant à une photo-interprétation directe ou en calculant des indicateurs à partir des couches SIG disponibles : cartes de couverture du sol, classifications détaillées d'images ou registres des exploitations agricoles.

Un système d'échantillonnage à deux degrés peut également être envisagé, les DR et les points servant respectivement d'unités primaires et secondaires d'échantillonnage (UPE et USE). Un système d'échantillonnage en deux phases — avec utilisation d'une grille régulière de points au cours de la première phase — peut être mis en œuvre lors de la deuxième étape (figure 20). Ces points sont alors photo-interprétés comme « terres cultivées », « terres non cultivées » ou « incertain », donnant lieu à une stratification simple menant à la sélection d'un échantillon final. L'identification des exploitations agricoles gérant les parcelles correspondant aux champs sélectionnés produit un échantillon avec probabilité proportionnelle à la surface cultivée (Gallego *et al.*, 1994). Ce système montre à quel point la distinction entre bases de type liste et bases aréolaires peut être floue : on part d'un DR, qui constitue un élément typique des BTL ; on introduit ensuite un échantillon de points, qui constitue quant à lui un élément typique des BSA ; enfin, celui-ci donne lieu à un échantillonnage d'exploitations agricoles, qui constitue l'unité de base des BTL. Une grille très déséquilibrée (représentée par un rectangle dont la longueur va d'ouest en est et la largeur du nord au sud) est représentée dans la figure 20 pour montrer que les grilles n'ont pas nécessairement besoin d'être carrées, ce qui peut présenter l'avantage de réduire la distance de marche entre les points lors d'une opération de terrain.

FIGURE 20. GRILLE DE POINTS SUPERPOSÉE À UN DR, POUVANT ÊTRE UTILISÉE COMME ÉCHANTILLON DE PREMIÈRE PHASE AU COURS DE LA DEUXIÈME ÉTAPE DU PROCESSUS D'ÉCHANTILLONNAGE À DEUX DEGRÉ.



3.4. MESURE DES PARCELLES

La mesure des parcelles agricoles constitue généralement une étape nécessaire des enquêtes agricoles, que celles-ci soient fondées sur des BTL ou des BSA. En effet, les informations fournies par les agriculteurs de la plupart des pays en développement sur la superficie d'une parcelle donnée sont rarement fiables, souvent en raison de la méconnaissance de l'information. La photo-interprétation peut constituer une alternative valable à la méthode traditionnelle reposant sur le mètre ruban et la boussole. Elle peut s'avérer plus ou moins précise que les mesures GPS en fonction des caractéristiques des images ou des frontières des parcelles mesurées. En principe, la précision des délimitations effectuées par photo-interprétation est à peu près du même ordre de grandeur que celle de la résolution de l'image, à condition que les contrastes entre les objets soient suffisamment marqués. Ainsi, la mesure des parcelles à l'aide d'une résolution inférieure à 3-4 m est préférable à celle effectuée à l'aide d'un appareil GPS ayant une précision de positionnement de l'ordre de 5 m environ. Le problème est cependant beaucoup plus complexe.

La figure 21 présente un exemple de surface occupée principalement par des parcelles agricoles d'une taille inférieure à 1 ha. L'image de gauche est une orthophotographie aérienne (résolution de 0,5 m) datant de 2007, tandis que celle de droite a été prise par le satellite SPOT (résolution de 2,5 m) en 2012. La parcelle « 1 » peut être à peu près délimitée sur les deux images, le contraste avec les parcelles voisines étant de bonne qualité. Elle présente toutefois des formes considérablement différentes, ce qui n'est que partiellement dû au fait que les images ont été prises à des dates différentes. La parcelle 2 est facile à distinguer dans l'orthophotographie, mais quasiment indiscernable sur l'image de droite. Une photo-interprétation sur le terrain de la parcelle concernée s'impose alors, afin d'identifier les éventuels changements ayant eu lieu entre la date où l'image a été prise et la date de l'enquête. Lors du choix de l'appareil portatif utilisé à cette fin, il conviendra de veiller à ce que son écran assure une bonne visibilité.

FIGURE 21. PHOTO-INTERPRÉTATION DE PARCELLES À L'AIDE D'UNE ORTHOPHOTOGRAPHIE (RÉSOLUTION DE 0,5 M) ET D'UNE IMAGE SATELLITE (RÉSOLUTION DE 2,5 M).



3.5. ÉCHANTILLONNAGE D'IMAGES SATELLITE

3.5.1. Expériences menées dans les années 1970

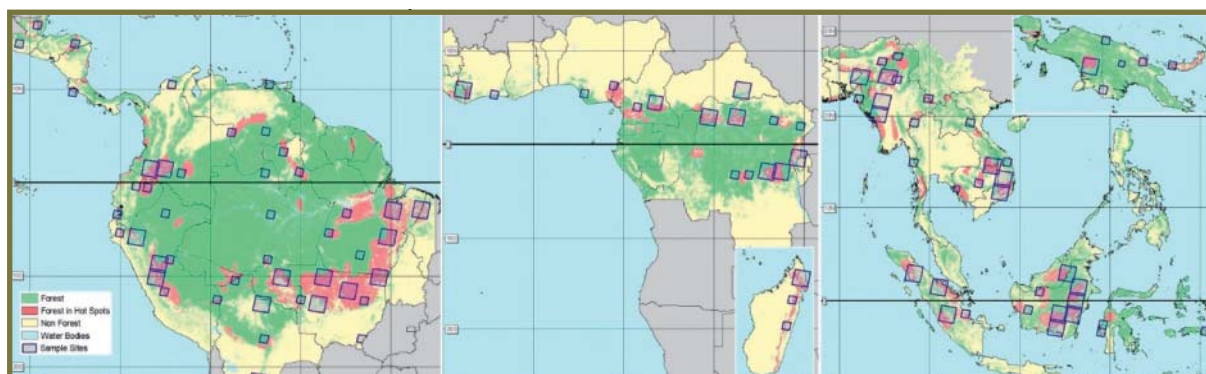
Dans le cadre du projet LACIE (de l'anglais *Large Area Crop Inventory Experiment*), le premier à étudier l'application de la télédétection aux statistiques agricoles, les images prises par le capteur MSS de Landsat ont été découpées en tuiles de 5×6 milles marins (McDonald et Hall, 1980), dont un échantillon a fait l'objet d'une classification. Les articles relatifs au projet n'ont pas clairement justifié cette sélection, qui pourrait toutefois s'expliquer par le fait que la classification d'une scène entière prise par le capteur MSS de Landsat exigeait des installations informatiques dont le Service national de statistiques agricoles (NASS, de l'anglais *National Agriculture Statistical Service*) de l'USDA ne disposait pas en interne (Hanuschak, communication personnelle, 2001). Celui-ci aurait pu demander de l'aide à l'Agence spatiale américaine (NASA, de l'anglais *National Aeronautics and Space Administration*), mais a au contraire décidé de concevoir une méthode lui permettant d'opérer sans aide extérieure. Les erreurs d'échantillonnage ont été calculées, mais l'impact de l'estimation par comptage des pixels des surfaces cultivées présentes dans chaque tuile a été ignoré. Cette lacune a été corrigée à la fin des années 1970, lorsque l'USDA a acquis la capacité de classer des images entières et a commencé à utiliser des estimateurs par régression associant les données de dénombrement de juin menée sur le terrain (environ 16 000 segments délimités par des frontières physiques) à des images classées en tant que données auxiliaires (Hanuschak *et al.*, 1979 ; Chhikara *et al.*, 1986 ; Allen, 1990).

3.5.2. Échantillonnage d'images à résolution moyenne

Le découpage d'une image en tuiles et l'échantillonnage de ces dernières en vue de leur classification automatique ont perdu tout leur sens dans les années 1980 ; l'échantillonnage des images satellite a toutefois continué à se justifier dans certains cas. Ces images, qui présentent une résolution spatiale de 20 à 60 m, étaient considérées à l'époque comme des « images à haute résolution ». L'échantillonnage des images satellite était une solution judicieuse lorsque la région étudiée était étendue et que la couverture totale des images nécessitait un budget trop important pour un projet spécifique (la politique de diffusion gratuite ne s'appliquait pas à ce type d'images). Par exemple, les « estimations de modification rapide des surfaces cultivées » du projet MARS de l'UE se sont appuyées sur un échantillon de 60 sites mesurant 40×40 km chacun. Le but était ici de couvrir chaque site à l'aide de quatre images prises par le capteur multibandes du satellite SPOT (SPOT-XS) tout au long de la période de végétation. À noter que le nombre maximum d'images n'a pas toujours été atteint. Pendant plusieurs années, le projet a produit des estimations conformes à celles fournies par d'autres sources d'information ; une étude plus poussée a cependant constaté que le recours à la technique du comptage des pixels comme base de calcul des estimations donnait aux analystes des images une marge de subjectivité d'environ 20 à 30 %, de telle façon que les estimations pouvaient être ajustées pour correspondre à celles calculées par des sources externes (Gallego, 2006).

Les images satellite ont souvent été échantillonnées dans le cadre d'évaluations mondiales des ressources forestières ou lorsqu'elles portaient sur des surfaces très vastes, comme la ceinture tropicale (figure 22). Dans le premier cas, il peut arriver qu'une image soit divisée en tuiles et que seule l'une d'entre elles — ou un échantillon — soit analysée lors du processus de photo-interprétation visuelle des images. Dans ce cas, le coût de chaque unité d'échantillonnage est proportionnel à sa taille, ce qui justifie le choix d'unités plus petites (Achard *et al.*, 2002).

FIGURE 22. ÉCHANTILLON STRATIFIÉ DE SCÈNES LANDSAT ET QUARTS DE SCÈNES UTILISÉS POUR ESTIMER L'ÉVOLUTION DE LA FORÊT TROPICALE

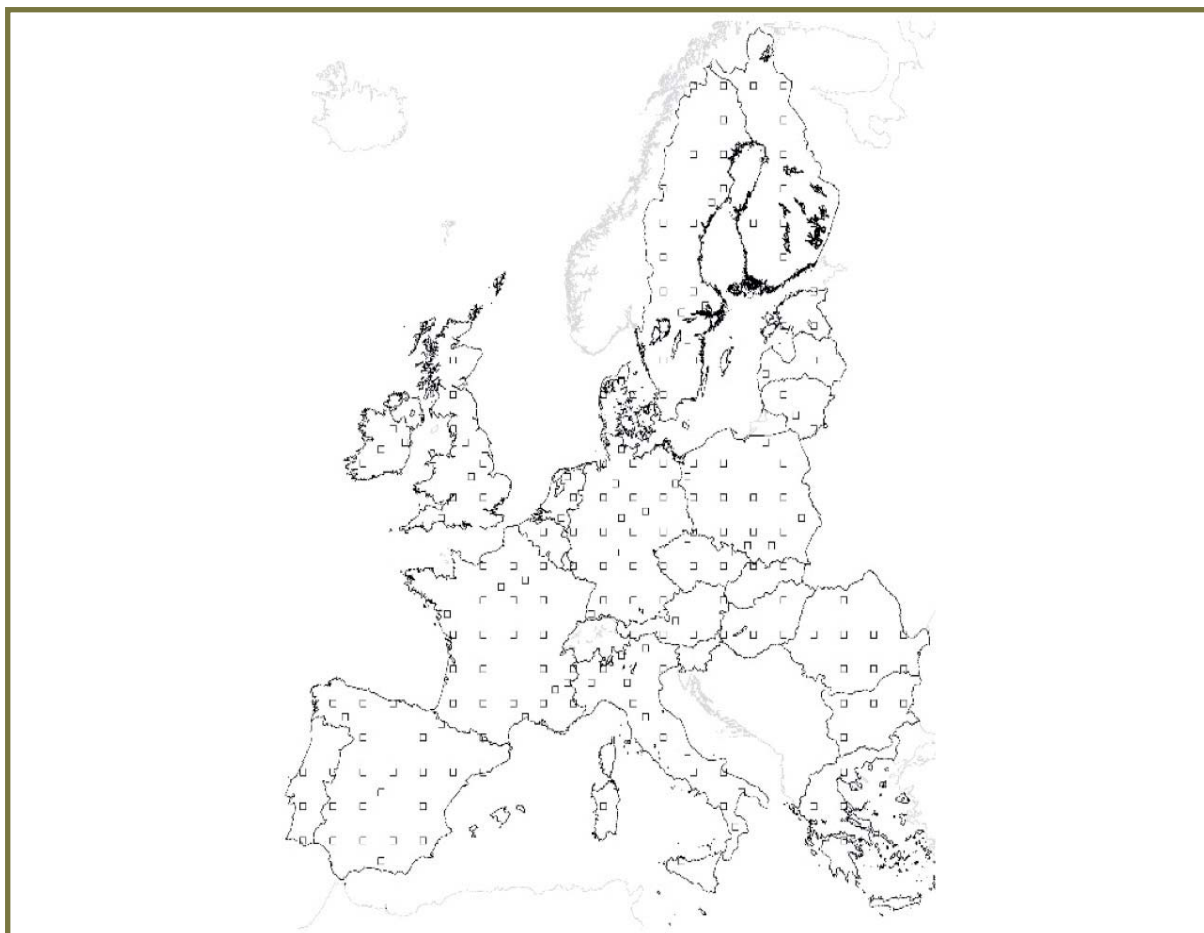


3.5.3. Échantillonnage d'images à très haute résolution

Les images THR (dont les pixels mesurent actuellement entre 0,3 et 3 m) sont de plus en plus fréquentes, et la question s'est posée de savoir si elles pouvaient jouer un rôle dans l'estimation des superficies. La couverture totale coûte extrêmement cher et est difficile à gérer. Il est donc possible de créer un échantillon d'unités dont la taille permette d'assurer une couverture par des images THR. La taille couverte par une image THR varie d'un satellite à un autre. Les carrés de 10×10 km ou de 15×15 km peuvent être considérés comme étant adaptés pour assurer la couverture par une image THR (figure 23). Le recours à ces échantillons aux fins d'estimations du changement de couverture du sol a été étudié dans le cadre du projet Geoland 2 de l'UE (Aleksandrowicz *et al.*, 2014).

L'efficacité de l'échantillonnage d'unités de 10×10 km a été analysée en partant de l'hypothèse optimiste selon laquelle les monocultures étaient correctement identifiables (Gallego et Stibig, 2013). Cette étude a conclu que l'identification des cultures par imagerie THR effectuée à partir de ce type d'échantillon pouvait être plus rentable que les enquêtes ponctuelles menées sur le terrain dans le cadre du projet LUCAS de l'UE, si le coût par unité de 10×10 km s'élevait à moins de 500 dollars US environ, achat et traitement des images compris. Ce seuil dépend de la structure de corrélation spatiale et varie par conséquent d'une culture à l'autre dans chaque paysage. Il est, à l'heure actuelle, impossible de parvenir à ce coût. Ce système pourra toutefois s'avérer intéressant à l'avenir pour les cultures facilement identifiables sur image, s'il est possible de réduire sensiblement le coût de l'achat et du traitement des images.

FIGURE 23. ÉCHANTILLON DE SITES DE 10 × 10 KM DEVANT ÊTRE ANALYSÉS À L'AIDE D'IMAGES THR DANS LE CADRE DU PROJET GEOLAND 2.



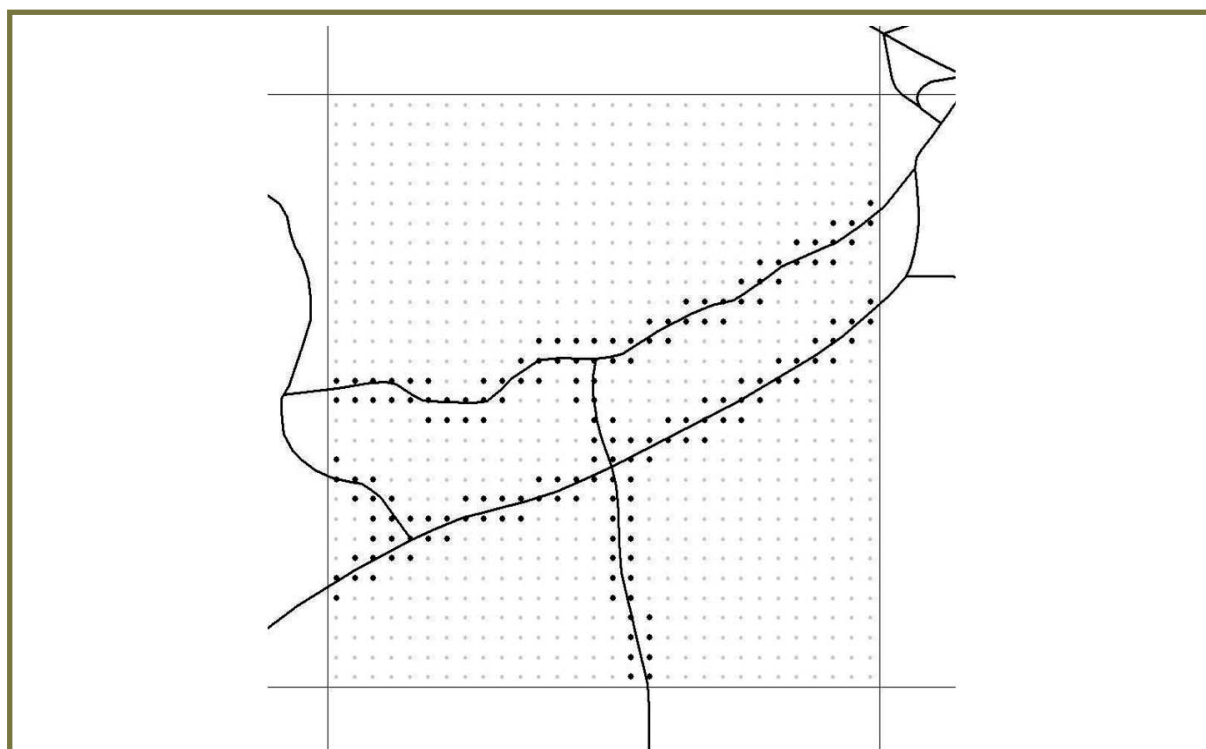
3.6. ENQUÊTES LE LONG DES ROUTES

De nombreuses enquêtes à base de sondage aréolaire se sont heurtées à des difficultés lorsque les unités d'échantillonnage se trouvaient loin des routes et ne pouvaient être atteintes qu'après une longue marche. Ce phénomène est particulièrement problématique dans les pays ne disposant pas d'un réseau routier important ou lorsque les contraintes de temps sont particulièrement strictes. Différents organismes internationaux mènent des enquêtes le long des routes pour évaluer la situation agricole de pays tiers. Dans la plupart des cas, cependant, il ne semble pas y avoir de note méthodologique décrivant les bonnes pratiques d'échantillonnage des routes le long desquelles les observations sont effectuées.

La concentration des observations le long des routes semble être pratique, et plusieurs solutions sont envisageables : Moisen (1996) propose par exemple d'effectuer une stratification à partir de zones tampon le long des routes. Cette option semble prometteuse dans le cadre de la méthode de suivi de la végétation basée sur la modélisation qu'il a mise en place ; elle est cependant peu susceptible de produire des résultats dans le cadre d'estimations des superficies cultivées, puisque la variance serait dans ce cas principalement calculée à partir des zones éloignées des routes, avec un taux d'échantillonnage plus faible.

Le dispositif CropWatch s'appuie sur un système qui part de l'hypothèse qu'il existe des données fiables sur les surfaces totales relevant d'une classe plus générale, telle que les terres arables ou le nombre total de terres cultivées (Wu et Li, 2012). Un chiffre plus général peut être obtenu, par exemple, à l'aide d'un modèle de régression simple à partir duquel les chiffres de FAOSTAT sont extrapolés, ou grâce à la télédétection, par exemple par la photo-interprétation d'un vaste échantillon de points comme «terres cultivées» ou «terres non cultivées». L'enquête effectuée le long de la route sert à estimer la part de chaque culture par rapport au nombre total de terres cultivées. L'hypothèse sous-jacente est la suivante : le ratio entre les surfaces occupées par différentes cultures situées près et loin des routes est comparable. Cette hypothèse a été testée au sein de l'UE par l'enquête sur le terrain menée en 2009 dans le cadre du projet LUCAS à l'aide d'une couche SIG représentant le réseau routier. L'échantillonnage de tronçons de route présentant une probabilité proportionnelle à leur longueur et de points dans une zone tampon de 100 m a entraîné un biais positif pour le blé (+ 2,5 %) et un biais négatif pour l'orge (-4,3 %), lequel peut être réduit de moitié grâce à la modification du système d'échantillonnage : de larges segments carrés de 3×3 km sont échantillonnés, une grille régulière est sélectionnée en leur sein, et seuls les points situés dans une zone tampon éloignée de 100 m des routes sont visités (figure 24). Grâce à ce nouveau système d'échantillonnage, le biais ne s'élève plus qu'à +1,4 % pour le blé et à -2,4 % pour l'orge. Ce biais est en effet partiellement dû au fait que l'orge est plus courante dans les zones moins fertiles, qui abritent généralement un réseau routier moins dense ; elles bénéficient par conséquent d'une pondération plus forte lors de l'échantillonnage de segments que lors de l'échantillonnage d'arcs routiers.

FIGURE 24. ÉCHANTILLON DE POINTS OBTENU DANS LE CADRE D'UNE ENQUÊTE LE LONG D'UNE ROUTE PAR L'ÉCHANTILLONNAGE DE LARGES SEGMENTS CARRÉS.



3.7. RÉFÉRENCES BIBLIOGRAPHIQUES

- Achard, F., Eva, H.D., Stibig, H.-J., Mayaux, P., Gallego, J., Richards, T. et Malingreau, J.P. 2002. «Determination of Deforestation Rates of the World's Humid Tropical Forests». *Science*, vol. 297, n° 5583, p. 999–1002.
- Agence européenne pour l'environnement (AEE). 2007. *CLC2006 technical guidelines*. Rapport technique n° 17/2007. Publication de l'AEE : Copenhague. Disponible à l'adresse : http://www.eea.europa.eu/publications/technical_report_2007_17. Consulté le 5 mai 2017.
- Aleksandrowicz, S., Turlej, K., Lewiński, S. et Bochenek, Z. 2014. «Change detection algorithm for the production of land cover change maps over the European Union countries». *Remote Sensing*, vol. 6, n° 7, p. 5976–94.
- Allen, J.D. 1990. «A Look at the Remote Sensing Applications Program of the National Agricultural Statistics Service». *Journal of Official Statistics*, vol. 6, n° 4, p. 393–409.
- Boryan, C.G. et Yang, Z. 2012. «A new land cover classification based stratification method for area sampling frame construction». In : *Proceedings of 2012 First International Conference on Agro-Geoinformatics*, 2–4 août 2012. Shanghai, Chine. Disponible à l'adresse : <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6311727&queryText%3DBoryan>. Consulté le 5 mai 2017.
- Carfagna E. 2007. « A comparison of area frame sample designs for agricultural statistics ». *Bulletin of the International Statistical Institute, 56^e session. Compte-rendu de la réunion STCPM11 sur les statistiques agricoles et rurales*, 22–29 août 2007. Lisbonne.
- Carfagna, E. et Gallego, F.J. 2005. «Using Remote Sensing for Agricultural Statistics». *International Statistical Review*, vol. 73, n° 3, p. 389–404.
- Chhikara, R., Lundgren, J. et Houston, A. 1986. «Crop Acreage Estimation Using a -Landsat-Based Estimator as an Auxiliary Variable». *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, n° 1, p. 157–68.
- Cotter, J. et Tomczac, C. 1994. «An Image Analysis System to Develop Area Sampling Frames for Agricultural Surveys». *Photogrammetric Engineering and Remote Sensing*, vol. 60, n° 3, p. 299–306.
- Davies, C. 2009. *Area frame design for agricultural surveys*. Rapport de recherche n° RDD-09-xx de la RDD. Publication de l'USDA-NASS : Washington. Disponible à l'adresse : http://www.nass.usda.gov/Publications/Methodology_and_Data_Quality/Advanced_Topics/AREA%20FRAME%20DESIGN.pdf. Consulté le 5 mai 2017.
- De Vries, P.G. 1986. *Sampling Theory for Forest Inventory. A Teach-yourself Course*. Springer-Verlag : Berlin–Heidelberg (Allemagne) – Wageningen (Pays-Bas).
- Dunn, R. et Harrison, A.R. 1993, «Two-dimensional systematic sampling of land use». *Journal of the Royal Statistical Society series: Applied Statistics*, vol. 42, n° 4, p. 585–601.
- FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture). 1996. *Enquêtes agricoles à base de sondages multiples, Volume 1 : Enquêtes courantes fondées simultanément sur des méthodes de sondages aréolaires et de sondages par listes d'exploitations*. Collection FAO : Développement statistique, n° 7. Disponible à l'adresse : http://www.fao.org/fileadmin/templates/ess/ess_test_folder/World_Census_Agriculture/Publications/FAO_SDS/SDS_7_Enqu%C3%AAtes_agricoles_%C3%A0_base_de_sondages_multiples.pdf. Consulté le 5 mai 2017.

- FAO.** 1998. *Multiple frame agricultural surveys, Volume 2: Agricultural survey programmes based on area frame or dual frame sample designs*. Collection FAO : Développement statistique, n° 10. Disponible à l'adresse : http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Publications/SDS/10_multiple_frame_agricultural_surveys.pdf. Consulté le 5 mai 2017.
- Fritz, S. et See, L.** 2008. « Identifying and quantifying uncertainty and spatial disagreement in the comparison of global land cover for different applications ». *Global Change Biology*, vol. 14, n° 5, p. 1057–75.
- Fritz, S., See, L., McCallum, L., You, L., Bun, A., Moltchanova, E. et Obersteiner, M.** 2015. « Mapping global cropland and field size ». *Global Change Biology*, vol. 21, n° 5, p. 1980–92.
- Gallego, F.J. et Stibig, H.J.** 2013. « Area estimation from a sample of satellite images: the impact of stratification on the clustering efficiency ». *Journal of Applied Earth Observation and Geoinformation*, vol. 22, p. 139–46.
- Gallego, F.J.** 2006. « Review of the Main Remote Sensing Methods for Crop Area Estimates ». In : *Workshop proceedings: Remote sensing support to crop yield forecast and area estimates*. Archives internationales de photogrammétrie, de télédétection et de sciences de l'information spatiale, XXXVI, 8/W48, p. 65–70. Disponible à l'adresse : http://www.isprs.org/publications/PDF/ISPRS_Archives_WorkshopStresa2006.pdf. Consulté le 5 mai 2017.
- Gallego, F.J., Peedell, S. et Carfagna, E.** 1999. « The use of CORINE Land Cover to improve area frame survey estimates ». *Research in Official Statistics*, vol. 2, n° 2, p. 99–122.
- Gallego, F.J., Delincé, J. et Carfagna, E.** 1994. « Two-Stage Area Frame Sampling on Square Segments for Farm Surveys ». *Survey Methodology*, vol. 20, n° 2, p. 107–15.
- Gallego, J., Giovacchini, A. et Downie, M.** 2016. « Sampling stripes (transects) for crop area estimation with drones ». Article préparé pour la septième Conférence internationale sur les statistiques de l'agriculture (ICAS VII), 26–28 octobre 2016. Rome.
- Gallego, F.J., Carfagna, E. et Baruth, B.** 2010. « Accuracy, Objectivity and Efficiency of Remote Sensing for Agricultural Statistics ». In : Benedetti, R., Bee, M., Espa, G. et Piersimoni, F. (dir.), *Agricultural Survey Methods* (ch. 12, p. 193–211). John Wiley & Sons : Chichester, Royaume-Uni.
- Gallego, F.J. et Delincé, J.** 2010. « The European Land Use and Cover Area-frame statistical Survey (LUCAS) ». In : Benedetti, R., Bee, M., Espa, G. et Piersimoni, F. (dir.), *Agricultural Survey Methods* (ch. 10, p. 151–68). John Wiley & Sons : Chichester, Royaume-Uni.
- González F., López, S. et Cuevas, J.M.** 1991. « Comparing Two Methodologies for Crop Area Estimation in Spain Using Landsat TM Images and Ground Gathered Data ». *Remote Sensing of Environment*, vol. 35, n° 1, p. 29–36.
- GSARS (Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales)** 2015. *Manuel sur les bases de sondage principales pour les statistiques agricoles — développement de la base de sondage, conception du plan d'échantillonnage et estimation*. Manuel de la GSARS : Rome. Disponible à l'adresse : <http://gsars.org/wp-content/uploads/2016/06/Handbook-on-MSF-FR-WEBFILE-280616.pdf>. Consulté le 5 mai 2017.
- Hanuschak, G., Sigman, R., Craig, M., Ozga, M., Luebbe, R., Cook, P., Kleweno, D. et Miller, Ch.** 1979. « Obtaining timely crop area estimates using ground gathered and Landsat data ». Bulletin technique n° 1609 de l'USDA. Disponible à l'adresse : <https://play.google.com/store/books/details?id=wYwoAAAAYAAJ&rdid=book-wYwoAAAAYAAJ&rdot=1>. Consulté le 5 mai 2017.

- Hendricks, W.A., Searls, D.T. et Horvitz, D.G.** 1965. «A comparison of three rules for associating farms and farmland with sample area segments in agricultural surveys». In : S.S. Zarkovich (dir.), *Estimation of areas in Agricultural Statistics* (p. 191–8). Publication de la FAO : Rome.
- Jolly, G.M. et Watson, R.M.** 1979. «Aerial sample survey methods in the quantitative assessment of ecological resources». In : Cormack, R.M., Patil, G.P. et Robson, D.S. (dir.), *Sampling Biological Populations* (p. 203–16). International Co-Operative Publishing House : Fairland, États-Unis.
- Kerdiles, H., Spyratos, S., Gallego, J. et Dong, Q.** 2013. «Assessing the crop acreage in Mengcheng county on the North China plain using adapted regression estimator method». In : *2nd International Conference on Agro-Geoinformatics: Information for Sustainable Agriculture, Agro-Geoinformatics 2013*, p. 577–582. Disponible à l'adresse : <http://iopscience.iop.org/1755-1315/17/1/012057>.
- Latham, J.** 2009. «FAO land cover mapping initiatives». In : Campbell, J.C. *et al.* (dir.), *North America Land Cover Summit* (p. 75–95). Publication de l'Association of American Geographers : Washington. Disponible à l'adresse : <http://www.aag.org/galleries/nalcs/CH6.pdf>. Consulté le 5 mai 2017.
- MacDonald, R.B. et Hall, F.G.** 1980. «Global crop forecasting». *Science*, vol. 208, n° 4445, p. 670–9.
- Martino, L.** 2003. «The Agrid system for short-term estimates in agriculture». Article présenté au colloque DRAGON, 9–11 juillet 2003. Cracovie/Balice, Pologne.
- Moisen, G.G.** 1996, «Generalized linear mixed models for analyzing error in a satellite-based vegetation map of Utah». In : Mowrer, H.T., Czaplewski, R.L. et Hamre, R.H. (dir.), *Spatial accuracy assessment in natural resources and environmental sciences* (p. 459–66). Rapport technique général RM-GTR-277 du Service des forêts de l'USDA.
- Taylor, J., Sannier, C., Delincé, J. et Gallego, F.J.** 1997. *Regional Crop Inventories in Europe Assisted by Remote Sensing: 1988-1993*. Rapport de synthèse, EUR 17319 EN. Publication de la Commission européenne : Luxembourg. Disponible à l'adresse : <http://mars.jrc.ec.europa.eu/mars/Bulletins-Publications/Regional-Crop-Inventories-in-Europe-Assisted-by-Remote-Sensing-1988-1993>.
- Wu, B. et Li, Q.** 2012. «Crop planting and type proportion method for crop acreage estimation of complex agricultural landscapes». *International Journal of Applied Earth Observation*, vol. 16, p. 101–12.



4

Chapitre 4

Réalisation d'une cartographie détaillée des cultures à partir de données de télédétection (couches de données sur les cultures)

Andrew M. Davidson, Thierry Fisette, Heather McNairn et Bahram Daneshfar

4.1. INTRODUCTION

La lutte contre les menaces qui pèsent à court et long terme sur la stabilité et la fiabilité de l'accès à la nourriture pour tous nécessite une bonne compréhension de la situation et des tendances de la production agricole à l'échelle d'un pays. La quantification de la disponibilité alimentaire peut toutefois être difficile, car la production végétale nationale, régionale et mondiale dépend des stratégies locales de gestion des terres et des changements météorologiques incessants (Fischer *et al.*, 2005). Les informations concernant les superficies ensemencées et leur rendement doivent être fréquemment actualisées pour prévoir la disponibilité (la production) alimentaire (Lobell et Field, 2007 ; Banque mondiale, ONU et FAO, 2010 ; Waldner *et al.*, 2015). Seul un travail de cartographie rapide, régulier, détaillé et suffisamment précis permet de collecter de telles informations sur les terres cultivées (Gallego *et al.*, 2008).

Le recours aux données des satellites d'observation de la Terre est la solution la plus économique pour collecter en temps voulu — grâce à leur fréquence de revisite élevée — un ensemble d'informations spatialement explicites, détaillées et fiables sur de vastes surfaces terrestres (Atzberger *et al.*, 2016). Une fois combinées aux statistiques nationales, aux observations de terrain (*in situ*) et aux données et informations auxiliaires (secondaires), ces informations peuvent faciliter le travail de cartographie des surfaces cultivées (Gallego, 2004 ; Kussul *et al.*, 2016). Pour réaliser un inventaire précis des cultures, il convient toutefois de sélectionner les données satellitaires utiles, de collecter des informations de qualité sur le terrain et d'adopter des méthodes adéquates pour le prétraitement

et le traitement ultérieur des données, ainsi que des procédures rigoureuses. Ce travail est loin d'être aisé, car les systèmes de culture sont souvent différents et complexes. De plus, les types de plantes cultivées, leur rythme de croissance et les pratiques de gestion varient d'une région à l'autre. Pour être efficaces, les méthodes de télédétection doivent donc être adaptées aux conditions environnementales et aux systèmes de culture locaux.

Au cours de la dernière décennie, plusieurs projets ont tenté d'exposer le besoin de données de télédétection spatialement explicites pour cartographier les systèmes de culture, en particulier le lieu (Fritz *et al.*, 2015), le moment (Whitcraft *et al.*, 2015a), la fréquence (Whitcraft *et al.*, 2015b), la raie spectrale et la résolution spatiale des relevés (Whitcraft *et al.*, 2015c). La sélection des meilleures données et méthodologies aux fins de cartographie des cultures, reste l'une des grandes priorités de la recherche internationale. Plusieurs initiatives internationales ont d'ailleurs tenté d'harmoniser les approches et de mettre au point des protocoles de suivi et de compte rendu, ainsi que des bonnes pratiques, pour divers systèmes agricoles mondiaux (p. ex., l'initiative GEOGLAM, qui comprend l'expérience JECAM, l'initiative Asia-RiCE, le projet SIGMA et des contributions du projet Sen2-Agri).

Ce chapitre vise globalement à fournir une vue d'ensemble des approches en matière de télédétection utilisées aux fins de cartographie nationale annuelle détaillée (au niveau des parcelles) des cultures. Dans la première partie, nous présenterons les méthodes existantes de cartographie par télédétection des terres cultivées, avec notamment un bref aperçu de la classification supervisée des images et de la classification fondée sur les pixels et les objets. Dans la deuxième partie, nous décrirons les différents types de données satellitaires, de données de terrain et de données secondaires servant à la cartographie détaillée des terres cultivées. Dans la troisième partie, nous décrirons la mise en œuvre opérationnelle d'un programme de cartographie nationale des cultures en nous appuyant sur des exemples spécifiques tirés de l'inventaire annuel des cultures du Canada. Dans la dernière partie, nous présenterons les principales difficultés et perspectives auxquelles s'attendre à l'avenir en matière de cartographie nationale des types de cultures.

4.2. CLASSIFICATION DE L'IMAGERIE SATELLITE AUX FINS DE RÉALISATION D'UNE CARTOGRAPHIE DÉTAILLÉE DES CULTURES

4.2.1. Classification supervisée de l'imagerie satellite

La classification de l'imagerie satellite occupe une place centrale dans de nombreuses applications de télédétection. Au sens large, elle désigne l'utilisation de techniques automatiques pour identifier différents types de surfaces comprises dans une ou plusieurs bandes spectrales, puis produire une carte thématique sans intervention ou presque de l'utilisateur (Jones et Vaughan, 2010). Selon leurs caractéristiques spectrales, les données des images (pixels ou objets, voir la section 4.2.2) sont ainsi classées automatiquement dans l'une des classes cibles prédéfinies.

La plupart du temps, la préparation des cartes de cultures se fait au moyen de la classification supervisée (Martínez Beltran et Calera Belmonte, 2001 ; Congalton *et al.*, 1998). Cette méthode, qui nécessite une intervention humaine préalable conséquente, comprend généralement quatre grandes étapes : i) l'analyste d'image utilise les observations de terrain pour identifier dans les données de l'image les sites qui correspondent à chaque type de surface (classes cibles) à cartographier (cet ensemble de données de référence est souvent appelé « échantillon d'apprentissage », car il sert ensuite à apprendre à l'ordinateur à reconnaître les zones présentant des similarités spectrales attribuables à chaque classe) ; ii) l'algorithme de classification établit un lien entre ces sites de référence et les données de l'image (voire les données secondaires, le cas échéant) pour définir statistiquement les caractéristiques spectrales (et non spectrales, si des données secondaires sont utilisées) de chaque classe cible ; iii) l'algorithme compare ensuite chaque pixel de l'image avec ces signatures, puis l'attribue à la classe cible dont il est le plus proche ; et iv) l'exactitude de la classification finale est évaluée (validée) à l'aide d'un ensemble de points de référence relevés sur le terrain et n'ayant pas servi lors de la phase d'apprentissage de l'ordinateur de classification (c'est ce que l'on

appelle l'«échantillon de validation»). Le processus de classification supervisée servant à la cartographie détaillée des cultures est entièrement résumé à la figure 1.

Jusqu'à récemment, la méthode de classification par maximum de vraisemblance était la plus couramment utilisée en matière de classification supervisée des données de télédétection (Lu et Weng, 2007 ; Bhatta, 2008 ; Kumar *et al.*, 2016). Les données d'apprentissage sont ici utilisées pour décrire statistiquement les classes cibles selon leur fonction de densité de probabilité multidimensionnelle. Chaque fonction de densité représente la probabilité selon laquelle la structure spectrale d'une classe est présente dans une région donnée de l'espace spectral multidimensionnel (Denègre, 2013). Les données de l'image sont ensuite attribuées à la classe d'apprentissage à laquelle elles sont les plus susceptibles d'appartenir (Jensen, 1986). La méthode du MV a été amplement utilisée dans différentes études sur la classification de l'imagerie satellite des régions agricoles (Laba *et al.*, 1997 ; Xiuwan, 2002 ; Abdulaziz *et al.*, 2009 ; Kamusoko et Aniya, 2009 ; Rogan *et al.*, 2008). Toutefois, compte tenu de ses limites actuelles — particulièrement sa dépendance vis-à-vis de l'hypothèse d'une distribution gaussienne des données d'entrée qui s'avère souvent erronée pour les données d'images multitemporelles (Gislason *et al.*, 2006 ; Glanz *et al.*, 2014) — les chercheurs spécialisés dans le domaine de la télédétection agricole continuent à réfléchir à de nouvelles techniques de classification. Parmi ces nouvelles méthodes, plusieurs s'annoncent très prometteuses, notamment les réseaux de neurones artificiels (Rumelhart *et al.*, 1986 ; Rigol-Sanchez *et al.*, 2003), les machines à vecteurs de support (Abedi *et al.*, 2013 ; Al-Anazi et Gates, 2010 ; Cortes et Vapnik, 1995 ; Ghimire *et al.*, 2012 ; Zuo et Carranza, 2011), les arbres de décision (Breiman, 1984) et les ensembles d'arbres de classification tels que les forêts aléatoires (Breiman, 2001 ; Vincenzi *et al.*, 2011 ; Waske et Braun, 2009 ; Ghimire *et al.*, 2012 ; Rodriguez-Galiano et Chica-Olmo, 2012).

et d'hétérogénéité afin de produire des objets spatialement et spectralement hétérogènes (Haralick et Shapiro, 1985 ; Comaniciu et Meer, 2002 ; Benz *et al.*, 2004). L'extraction des caractéristiques passe par l'identification des attributs spectraux, texturaux, morphologiques et contextuels associés à chaque objet créé lors de la segmentation (Blaschke, 2010). Les objets d'image peuvent ensuite être classifiés en complétant les données du classifieur avec les informations sur les objets qui composent les éléments. Cette méthode porte le nom de classification orientée objet (Jensen, 1986). Son efficacité pour l'identification et la cartographie des cultures a été démontrée par Evans *et al.*, (2002), Brown de Coulston *et al.*, (2003), Castillejo-González *et al.*, (2009), McNairn *et al.*, (2009), Peña-Barragán *et al.*, (2011) et Vieira *et al.*, (2012).

4.3. COUCHES DE DONNÉES D'ENTRÉE SERVANT À LA CLASSIFICATION DES CULTURES

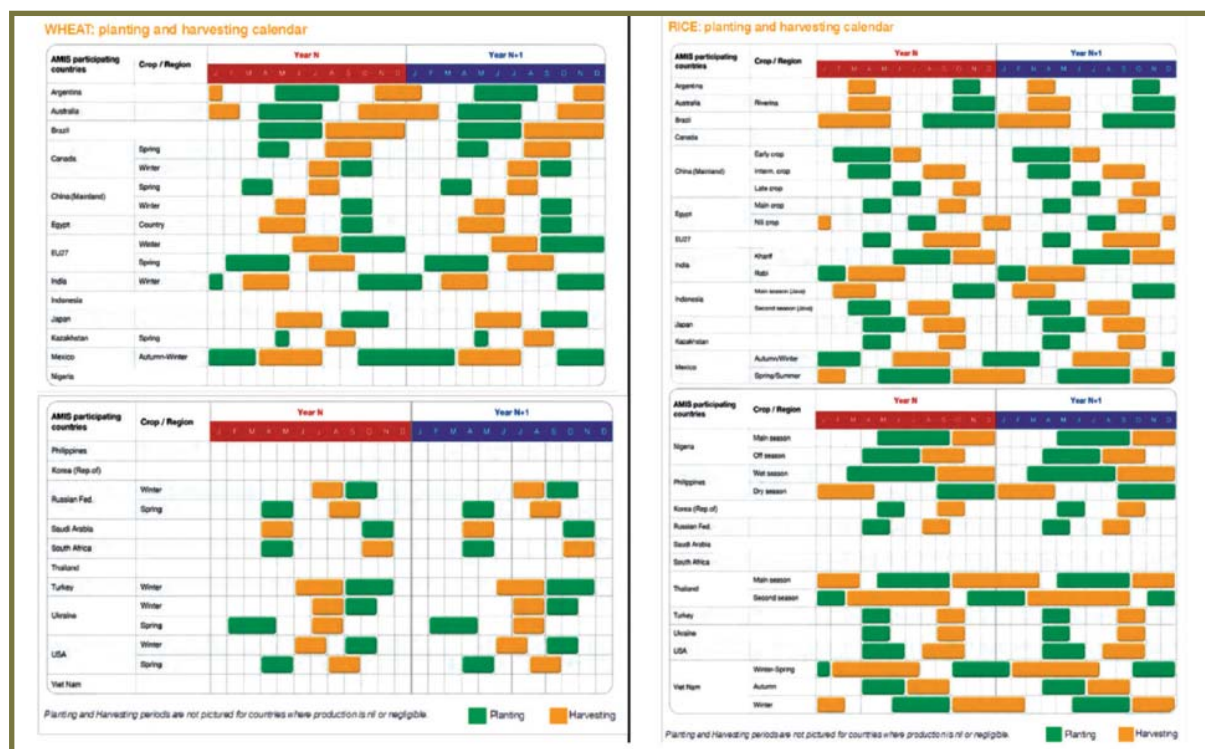
Quelle que soit précisément la méthode utilisée, la classification supervisée des types de cultures nécessite plusieurs types de données, à savoir : i) des données de terrain, ii) des données satellitaires et iii) des données et des informations auxiliaires (secondaires) (Teleguntla *et al.*, 2016). Lorsque ces données sont de très bonne qualité et peuvent être associées, les cartes produites offrent la plus haute précision possible (Thenkabail *et al.*, 2009a et 2009b). Dans cette section, nous présenterons les principaux points à prendre en considération lors de la collecte de données de terrain, satellitaires et auxiliaires de haute qualité aux fins de classification agricole.

4.3.1. Données de terrain (*in situ*)

La qualité des observations de terrain joue un rôle essentiel dans la cartographie des terres cultivées, car ces données sont utilisées lors de la phase d'apprentissage du processus de classification supervisée des images, à partir duquel sont produites les cartes de cultures, mais aussi lors de la phase de validation indépendante de ces cartes, pour laquelle elles servent de référence. La qualité des données de terrain étant déterminante pour la précision de la classification, les praticiens doivent donc prêter une attention particulière à la façon dont ces données sont collectées. Traditionnellement, les données de terrain servant à la classification des cultures sont recueillies lors des opérations de terrain. Les praticiens se rendent sur les parcelles où ils collectent des informations sur les cultures — p. ex., le type de culture, le cultivar, le stade de croissance, etc. — ainsi que l'emplacement géographique de l'observation au moyen d'un système GPS. La disponibilité de ces informations permettra de constituer un échantillonnage aussi représentatif que possible de la diversité des cultures dans la région étudiée. Un nombre suffisant d'échantillons doit donc être collecté pour chaque classe cible.

La qualité des cartes de culture basées sur des données satellitaires dépend des critères spatiotemporels du plan de sondage utilisé pour les observations de terrain. La chronologie des levers de terrain s'aligne généralement sur le calendrier cultural local, qui décrit les grandes phases de croissance des différentes cultures tout au long de l'année (voir la figure 2). Pour optimiser les résultats, les visites sur site doivent coïncider avec le stade de croissance pendant lequel la culture est la plus facile à identifier, c'est-à-dire généralement la floraison, la fructification et la reproduction. Dans les régions ayant une seule campagne agricole, une observation par parcelle et par saison suffit généralement. Si une même parcelle est utilisée pour plusieurs cultures au cours d'une année (simultanément ou séparément), une visite sur site est nécessaire pour chaque cycle de culture. Les méthodes les plus adaptées pour localiser des échantillons représentatifs de parcelles, non biaisés et spatialement indépendants sont connues, mais rarement exploitables. La base de sondage est constituée à partir « d'enquêtes superficielles » qui consistent à relever les données depuis des véhicules motorisés le long des routes. Bien que cette méthode permette au collecteur de données de recenser facilement et rapidement la diversité des cultures de toutes les parcelles visibles, le caractère non aléatoire des données collectées peut introduire un biais statistique qui doit être soigneusement évalué avant toute exploitation desdites données.

FIGURE 2. CALENDRIERS CULTURAUX DU BLÉ (GAUCHE) ET DU RIZ (DROITE).



Les dates de semis et de récolte déterminent la période d'échantillonnage au cours de laquelle les observations de télédétection et de terrain doivent être effectuées aux fins de cartographie des cultures fondée sur des données spatiales (AMIS, 2012).

La taille d'échantillon nécessaire pour la phase d'apprentissage du classifieur d'images est souvent difficile à déterminer *a priori*. Même si la précision globale de la classification augmente proportionnellement au nombre de parcelles utilisées lors de la phase d'apprentissage, la nature exacte de cette relation est complexe et varie selon les classes de cultures. Le nombre et la qualité des images utilisées en entrée pour la classification influent aussi sur le nombre de parcelles à échantillonner. Par exemple, de grands échantillons sont nécessaires si la variabilité de la réponse spectrale de chaque classe de culture n'est pas suffisamment caractérisée pour permettre de discriminer les cultures. Quoiqu'il en soit, puisque cette situation ne peut être prédite *a priori*, il en résulte souvent un sous-échantillonnage des données de terrain pour certains types de cultures et un suréchantillonnage pour d'autres.

Les principaux désavantages des opérations de terrain viennent du fait qu'elles sont laborieuses, complexes et prennent du temps ; surtout lorsqu'il s'agit de collecter des données représentatives à l'échelle nationale ou régionale.

D'autres sources de données fiables doivent donc généralement être utilisées pour compléter (ou remplacer entièrement) les opérations de terrain. Les données collectées par les organismes gouvernementaux locaux, régionaux ou nationaux peuvent remplir cette fonction, de même que les données collectées selon une enquête dite « grand public » plus informelle. Les données de terrain fournies par les organismes gouvernementaux sont généralement les plus précises, les plus détaillées et les plus complètes aux fins de cartographie des types de cultures, mais leur pertinence pour la classification doit souvent être soigneusement vérifiée. Bien que prometteuse, l'enquête grand public — qui consiste à obtenir des informations en faisant appel à un grand nombre de personnes, rémunérées ou non, via Internet — n'est pas encore répandue parmi les méthodes de collecte des données de terrain servant à classer les cultures.

4.3.2. Données d'observation de la Terre

4.3.2.1. Sources de données d'observation de la Terre servant à la classification des cultures

La coordination des données d'observation de la Terre destinées au Réseau Sentinelles en milieu agricole suppose de connaître le besoin de données de télédétection spatialement explicites, en particulier le lieu (Fritz *et al.*, 2015), le moment (Whitcraft *et al.*, 2015a), la fréquence (Whitcraft *et al.*, 2015b), la raie spectrale et la résolution spatiale des relevés (Whitcraft *et al.*, 2015c). En raison de la diversité et de la complexité des systèmes de culture ainsi que de la variabilité régionale des cultures et de leur rythme de croissance, le type de capteurs à utiliser, le nombre optimal d'images et le calendrier d'acquisition des images varient souvent selon le lieu.

La décision de recourir à des capteurs optiques ou des radars à synthèse d'ouverture (SAR, de l'anglais *Synthetic Aperture Radar*) se fonde souvent sur un compromis entre les facteurs suivants : a) la nature hétérogène et dynamique intrinsèque de l'écosystème agricole étudié ; b) l'étendue de la zone géographique à cartographier ; c) l'UCM nécessaire pour représenter les parcelles individuelles et les autres unités écologiques utiles (zones humides, zones boisées, etc.) ; d) les différences entre les cycles des cultures ; e) les différences entre les pratiques et les calendriers culturels au sein de la même classe ; f) la similarité spectrale avec d'autres classes de couverture du sol ; g) les contraintes techniques des systèmes de télédétection (taille de la fauchée ; résolutions spatiales, temporelles, spectrales et radiométriques ; couverture nuageuse pour les systèmes optiques) ; et h) la disponibilité des données (payantes ou en libre accès) (Waldner *et al.*, 2015). Ces points à prendre en compte dans le choix du système de télédétection sont résumés dans le tableau 1 — destiné à l'initiative GEOGLAM — qui présente brièvement les caractéristiques des capteurs nécessaires pour produire des informations cartographiques sur les cultures.

TABLEAU 1. CONDITIONS RELATIVES À L'UTILISATION DES DONNÉES SATELLITAIRES D'OBSERVATION DE LA TERRE, CONÇUES PAR UNE ÉQUIPE SPÉCIALE DU COMITÉ SUR LES SATELLITES D'OBSERVATION DE LA TERRE (CEOS, DE L'ANGLAIS COMMITTEE ON EARTH OBSERVATION SATELLITES) POUR GEOGLAM (CEOS, 2014; GEOGLAM, 2014).

Résolution spatiale	Raie spectrale	Fréquence d'observation efficace (sans nuage)	Taille des parcelles	Produits et applications d'information recherchés						
				Masque des cultures	Type de culture, surface et calendrier cultural	État des cultures	Rendement des cultures	Variables biophysiques des cultures	Variables environnementales	Pratiques agricoles / Systèmes de cultures
Échantillonnage à basse résolution (> 100 m)										
>500-2000m	Optique	Journalier	Tout			X		L		
100-500m	Optique	2 à 5 semaines	Tout	X	X	X	L	L	X	L
5-50Km	Micro-ondes	Journalier	Tout			X	X	X	X	
Échantillonnage à résolution moyenne (10 à 100 m)										
10-70m	Optique	mensuelle (min. 2 hors saison et 3 en saison) ; à reconduire tous les 1 à 3 ans	Tout	X	L/M					X
10-70m	Optique	~hebdomadaire (8 jours ; au moins 1 fois tous les 16 jours)	Tout	X	X	X	X	X	X	X
10-100m	SAR à double polarisation	~hebdomadaire (8 jours ; au moins 1 fois tous les 16 jours)	Tout	X	X	X	X	X	X	X
Échantillonnage à haute résolution (5 à 10 m)										
5-10m	Visible, NIR, SWIR	mensuelle (min. 3 en saison)	M/S	M/S	M/S					
5-10m	Visible, NIR, SWIR	~hebdomadaire (8 jours ; au moins 1 fois tous les 16 jours)	Tout		M/S	X		X	X	X
5-10m	SAR à double polarisation	Mensuelle								
Échantillonnage à très haute résolution (< 5 m)										
<5m	Visible, NIR	3 fois par an (2 en saison + 1 hors saison) ; à reconduire tous les 3 ans	S	S	S					
<5m	Visible, NIR	1 à 2 fois par mois	Tout		X		X			X

Les conditions d'utilisation incluent la résolution spatiale, la raie spectrale, la fréquence à laquelle on peut espérer faire des acquisitions sans nuage, l'étendue géographique et le type d'application ou de produit qui pourra être créé à partir des données. Elles sont affinées selon la taille des parcelles dont l'image doit être acquise ou pour lesquelles certains types de données seraient utiles. Les parcelles sont classées par taille : grandes (G – plus de 15 ha), moyennes (M – de 1,5 à 15 ha) et petites (P – moins de 1,5 ha). La lettre « X » et le mot « toutes » signifient que ces données permettent de créer des produits valables pour toutes les tailles de parcelle (Whitcraft et al., 2015c).

4.3.2.2. Données satellitaires optiques

Les systèmes de télédétection optique collectent des informations spectrales dans les longueurs d'onde visibles, le NIR et le SWIR. Afin de former des images de la surface terrestre, ces capteurs enregistrent la façon dont les éléments cibles au sol réfléchissent le rayonnement solaire. Ils sont recommandés pour cartographier la végétation, car les informations tirées de leurs principales bandes spectrales (généralement le bleu, le vert, le rouge, le NIR et le SWIR) permettent de distinguer facilement les signatures spectrales uniques des végétaux parmi les autres types de surfaces. Dans les parties visibles, le NIR et le SWIR du spectre électromagnétique, la quantité d'énergie

solaire ambiante absorbée, réfléchi et transmise par la végétation dépend principalement et respectivement de la pigmentation des plantes, de la structure interne de leur feuillage et de leur teneur en humidité (Jensen, 1986). Ces réactions physiques et chimiques (au niveau de l'atome) sont propres à chaque culture et permettent de connaître le stade de croissance et l'état de la plante. La forme et l'intensité de la courbe de réflectance sur ces longueurs d'onde sont donc souvent utilisées pour différencier les types de végétaux. L'utilisation des capteurs optiques pour la surveillance des cultures ne se fait toutefois pas sans difficulté. Dans la mesure où ces capteurs sont passifs, au sens où ils utilisent le soleil comme seule source de rayonnement, leurs images ne peuvent être collectées que pendant la journée et sous un ciel dégagé. Par ailleurs, la combinaison de facteurs autres que la présence et la quantité de végétation verte (à savoir, la végétation sénescente, la biomasse ligneuse, les sols et les ombres) produit souvent des spectres composites qui compliquent la distinction de la végétation verte (Colwell, 1974). Ces difficultés ont conduit à l'élaboration et à l'application d'indices spectraux de végétation qui combinent deux bandes spectrales ou plus pour renforcer le signal des végétaux tout en minimisant les effets perturbateurs.

De l'échelle nationale à l'échelle internationale, la cartographie agricole se fonde principalement sur des observations réalisées par ciel dégagé avec des capteurs à basse résolution (plus de 100 m) tels que le satellite SPOT-VGT (Bartholomé et Belward, 2005 – Global Land Cover 2000), le système AVHRR embarqué sur le satellite NOAA (NOAA-AVHRR) (Loveland *et al.*, 2000 – IGBP DISCover ; Hansen *et al.*, 2000 – UMD Global Land Cover ; Defries *et al.*, 2003 – Global AVHRR Land Cover ; Defries et Townshend, 2003 – Global AVHRR NDVI Land Cover), le système MODIS de la NASA (NASA-MODIS) (Friedl *et al.*, 2002 – MODIS Global Land Cover ; Tateishi *et al.*, 2011 – Global Land Cover by National Mapping Organizations) et le capteur MERIS monté sur le satellite ENVISAT (ENVISAT-MERIS) (Arino *et al.*, 2008 – GlobCover). Bien que la fréquence de revisite quotidienne de ces systèmes et leur large fauchée permettent de fournir les vues synoptiques panoramiques nécessaires à la classification phénologique de la végétation, leur utilité pour la classification détaillée des cultures des petites exploitations est limitée par les pixels à basse résolution qui ne permettent généralement pas de représenter les parcelles individuelles (Wardlow *et al.*, 2007 ; Wardlow et Egbert, 2008 ; Ozdogan, 2010). Ces systèmes ont donc permis de cartographier des zones agricoles selon une ou deux classes très générales ne suffisant toutefois pas à estimer les surfaces cultivées de façon précise et détaillée. Les données phénologiques produites peuvent néanmoins servir à compléter les études utilisant des observations à plus haute résolution. Des exceptions existent toutefois dans les régions agricoles où les parcelles sont étendues ou dans celles où les techniques de déconvolution peuvent être appliquées aux composants de surface présentant un spectre de pixels mixtes (Jakubauskas *et al.*, 2002 ; Lunetta *et al.*, 2006 ; Chang *et al.*, 2007 ; Fritz *et al.*, 2008 ; Galford *et al.*, 2008 ; Lunetta *et al.*, 2010 ; Sakamoto *et al.*, 2010 ; Pan *et al.*, 2012 ; Brown *et al.*, 2013 ; Waldner *et al.*, 2016).

De l'échelle régionale à l'échelle nationale, la cartographie agricole se fonde principalement sur des observations multispectraux réalisées par des capteurs à résolution moyenne (10 à 100 m) — particulièrement Landsat, dont les données sont accessibles gratuitement (Badhwar, 1984 ; Reese *et al.*, 2002 ; Guerschman *et al.*, 2003 ; Turker et Arikan, 2005 ; McNairn *et al.*, 2009 ; Fiset *et al.*, 2013 ; Johnson, 2013 ; Fiset *et al.*, 2014 ; Yan et Roy, 2014 ; Fiset *et al.*, 2015). Toutefois, bien que la résolution spatiale de ces capteurs permette de représenter les parcelles individuelles, leur fréquence de revisite est nettement inférieure aux survols quotidiens des orbiteurs NOAA-AVHRR, SPOT-VGT et NASA-MODIS (les missions Landsat-7 et Landsat-8 ont un temps de revisite de huit jours, chacune réalisant un survol tous les seize jours). La longueur de ces temps de revisite est problématique dans les régions où le temps est souvent nuageux, car elle permet rarement d'y obtenir le nombre d'images sans nuages nécessaires à la réalisation d'une cartographie détaillée des cultures (Jewell, 1989 ; McNairn *et al.*, 2002 ; Blaes *et al.*, 2005 ; McNairn *et al.*, 2009). Dans de tels cas, les données tirées d'autres capteurs multispectraux à résolution moyenne viennent combler les lacunes, mais certaines présentent un coût (p. ex., celles tirées des satellites SPOT, du capteur AWIFS et des satellites DMC), une résolution spatiale basse (p. ex., 56 m pour le capteur AWIFS), des bandes spectrales moins nombreuses (p. ex., SPOT, DMC) ou d'autres contraintes techniques ou d'accessibilité. À l'avenir, les synergies entre les programmes opérationnels tels que Landsat et Sentinel-2 (Wulder *et al.*, 2015) seront très utiles pour compléter les données manquantes grâce à l'augmentation de la fréquence des acquisitions.

géométriquement et radiométriquement compatibles (Gómez *et al.*, 2016) ainsi qu'à la disponibilité d'observations mondiales tous les deux à cinq jours (Drusch *et al.*, 2012 ; Irons *et al.*, 2012 ; Wulder *et al.*, 2015).

À l'échelle locale, la cartographie agricole se fonde principalement sur des observations réalisées par des capteurs à haute résolution (moins de 10 m), tels que RapidEye (Tapsall *et al.*, 2010 ; Conrad *et al.*, 2011 ; Kim et Yeom, 2014 ; Ustuner *et al.*, 2014 ; Beyer *et al.*, 2015 ; Schuster *et al.*, 2015 ; Lussem *et al.*, 2016 ; Xu *et al.*, 2016) et, dans une moindre mesure, IKONOS (Xie *et al.*, 2007 ; Turker et Ozdarici, 2011), Quickbird (Yang *et al.*, 2007) et WorldView (Alabi *et al.*, 2016). S'il a été démontré que la télédétection à haute résolution augmentait la précision de la cartographie des cultures au niveau des parcelles dans les régions comptant des champs de petites tailles (Salehi, *et al.*, 2013), ces capteurs ne sont toutefois généralement pas exploités pour le travail de cartographie agricole couvrant la totalité d'une région ou d'un pays (principalement en raison de leur coût et de la charge supplémentaire que représente le traitement informatique de ces importants volumes de données). En revanche, l'utilité de ces données pour décrire la variabilité des conditions de croissance au sein même des parcelles justifie leur exploitation dans l'agriculture de précision et la gestion des intrants agricoles tels que les engrais, les herbicides, les semences et les combustibles (utilisés pour la préparation du sol, les semailles, la pulvérisation, etc.).

4.3.2.3. Données SAR

Les systèmes de télédétection SAR fonctionnent en émettant de l'énergie dans le domaine des micro-ondes, puis en mesurant l'intensité et la phase de l'énergie diffusée suite à l'interaction avec la cible. Ces capteurs, autonomes sur le plan énergétique, fonctionnent jour et nuit. Les micro-ondes à faible fréquence qu'ils utilisent ne sont pas sensibles aux conditions atmosphériques et permettent donc de collecter des données en présence de nuages. En captant les micro-ondes du spectre électromagnétique, les SAR permettent de distinguer les types de végétation, car la diffusion de ces longueurs d'onde — plus longues que celles utilisées par les capteurs optiques — varie selon la structure de chaque culture (taille, forme et orientation des feuilles, des tiges et des fruits) et la teneur en eau du couvert végétal (au niveau moléculaire). Toutefois, comme dans le cas des données optiques, l'utilisation des données SAR pour la surveillance des cultures comporte certaines embûches. La principale difficulté tient à la perturbation du signal radar par les caractéristiques du sol (humidité et rugosité). L'ampleur de cet effet varie selon le type de culture, son stade de croissance, ainsi que la configuration du capteur de micro-ondes. Les études (Blaes *et al.*, 2005 ; McNairn *et al.*, 2009 ; Larrañaga *et al.*, 2011) et la pratique (Fisette *et al.*, 2015) ont toutefois montré que la combinaison des données optiques et SAR offrait une meilleure précision que les données optiques seules.

Les radars SAR se caractérisent par leur fréquence de fonctionnement, leur angle d'incidence et leur polarisation. La fréquence des ondes émises (en GHz ou en longueur d'ondes exprimées en cm) détermine la profondeur de pénétration des micro-ondes dans le couvert cultural ainsi que les éléments de ce couvert (feuillage, fruits, tiges) qui interagissent avec le signal, et comment. Aucune fréquence unique ne permet de distinguer toutes les cultures, car la pénétration dans le couvert cultural doit être assez profonde pour que les micro-ondes atteignent la structure générale de la plante, sans pour autant que les caractéristiques du sol perturbent les informations reçues par le SAR. La profondeur optimale de pénétration ainsi que la longueur d'onde adaptée à la dimension de chaque structure culturale varient selon chaque type de plante et son stade de croissance. La plupart des chercheurs suggèrent donc d'utiliser plusieurs fréquences lors du travail de classification des cultures (Skriver *et al.*, 1999 ; Hill *et al.*, 2005 ; Baghdadi *et al.*, 2009 ; McNairn *et al.*, 2009 ; Shang *et al.*, 2009 ; Hoekman *et al.*, 2011 ; Haldar *et al.*, 2012 ; Jia *et al.*, 2012 ; Skriver, 2012). Cette méthode reste néanmoins actuellement difficile à mettre en œuvre en raison du manque de données exploitables parmi les relevés des capteurs couvrant plusieurs fréquences. Jusqu'à présent, les méthodes ont donc généralement consisté à compléter les données optiques avec des données SAR sur une seule fréquence (voir par exemple McNairn *et al.*, 2009 ; Hütt *et al.*, 2016).

Les polarisations optimales pour la classification des cultures au moyen des radars SAR sont faciles à définir. La polarisation linéaire croisée (horizontale-verticale — HV — ou verticale-horizontale — VH) est la plus adaptée pour distinguer les cultures (McNairn *et al.*, 2009 ; Koppe *et al.*, 2013 ; Sonobe *et al.*, 2014). La repolarisation

qui se produit lorsque le signal est diffusé plusieurs fois dans le couvert cultural est nécessaire pour produire une réponse HV ou VH valable. La polarisation verticale-verticale (VV) permet aussi d'identifier les cultures, car les ondes de transmission verticales sont atténuées ou diffusées par la structure verticale de nombreuses cultures. La polarisation horizontale-horizontale (HH) est la moins utile. L'angle d'incidence influe aussi sur la pénétration dans le couvert cultural, mais sa sélection est de moindre importance.

Comme pour les capteurs optiques, la sélection de la fauchée et de la résolution du SAR dépend de la région, et plus précisément du compromis à atteindre entre la taille de la parcelle et la zone à cartographier. L'orbite du SAR (ascendante ou descendante) doit aussi être déterminée. Dans les régions tempérées, l'humidité associée aux basses températures nocturnes peut entraîner la formation de rosée matinale sur les couverts, qui peut affecter le phénomène de rétrodiffusion (Gillespie *et al.*, 1990). Certaines études ont toutefois indiqué que l'effet d'accroissement de la rétrodiffusion causé par la rosée était toutefois commun à tous les types de cultures et donc peu susceptible de compromettre leur distinction (Wood *et al.*, 2002). Toutefois, les recherches sur les effets de la rosée étant limitées, les acquisitions réalisées tôt le matin doivent être évitées autant que possible.

4.3.2.4. Prétraitement des données optiques et SAR

Les données optiques et SAR nécessitent généralement diverses procédures de prétraitement afin de pouvoir être utilisées en entrée dans les algorithmes de classification servant à la cartographie des paysages agricoles. La nature exacte de ces procédures — qui corrigent les distorsions radiométriques et géométriques des données — dépend des capteurs et plateformes utilisés pour acquérir les données, des conditions atmosphériques au moment de l'acquisition des données et des méthodes de classification. Les distorsions radiométriques peuvent provenir des variations dans l'illumination et dans la géométrie de visée de la scène, des conditions atmosphériques, et du bruit et de la réponse du capteur. Les distorsions géométriques peuvent provenir du mouvement de la plateforme et des systèmes de balayage, des variations d'altitude, d'attitude et de vitesse de la plateforme, du relief du terrain et de la courbure et de la rotation de la Terre. Les corrections ont pour objectif de compenser ces distorsions pour que les représentations géométriques et radiométriques de l'imagerie soient aussi proches que possible de la réalité. Dans la pratique, l'ampleur du travail de prétraitement de l'image à la charge des utilisateurs dépend de l'usage qui sera fait de ces données. Ainsi, le prétraitement effectué à la source est plus important pour les produits de données de haut niveau, souvent fournis par les agences spatiales sous forme de données immédiatement analysables et directement exploitables dans les publications scientifiques, tandis que dans le cas des produits de données plus modestes, l'utilisateur final assure une partie du travail de prétraitement.

4.3.3. Données et informations auxiliaires (secondaires)

Les données et informations auxiliaires provenant d'autres sources que les systèmes de télédétection permettent parfois d'améliorer la précision de la classification. Les données auxiliaires sont des variables contextuelles supplémentaires qu'il est possible d'ajouter aux données de télédétection à différents stades du processus de classification des images. Elles portent, entre autres éléments, sur l'élévation, la pente, l'aspect, l'hydrologie, la géologie, les sols, les réseaux de transport, les frontières politiques et les cartes de végétation (Jensen, 2016). Elles peuvent être intégrées avant la classification des images (p. ex., comme dans le cas de la stratification *a priori* des images ; voir Jensen, 2016), pendant (p. ex., en données d'entrée pour la classification) (Maselli *et al.*, 1995 ; Huang et Jensen, 1997 ; Stow *et al.*, 2003 ; Qiu et Jensen, 2004) ou après (p. ex., pour le tri des données après classification ; voir Rocha et Queluz, 2002). Les informations auxiliaires désignent les connaissances spécialisées qui peuvent être intégrées dans le processus de classification. Il peut s'agir d'éléments probants sur les configurations de paysage les plus plausibles à l'issue de la classification. Par exemple, il a été démontré que l'intégration d'informations auxiliaires par les réseaux bayésiens (Jensen et Nielsen, 2007) permettait de guider des tâches de classification complexes grâce à des connaissances spécialisées et de caractériser des phénomènes grâce à des inférences plausibles basées sur des éléments probants (Atzberger *et al.*, 2016). Bien qu'utile, l'intégration des données et informations auxiliaires ne doit toutefois pas introduire de nouvelles erreurs dans le système de classification.

4.4. RÉALISATION D'UNE CARTOGRAPHIE DES CULTURES À L'ÉCHELLE NATIONALE

4.4.1. De la recherche à la pratique

Les systèmes opérationnels de surveillance des cultures des pays doivent fournir en temps voulu des informations standard et interchangeables sur les cultures. Ces informations doivent offrir une précision valide au sens statistique du terme (Atzberger *et al.*, 2016) et s'appuyer sur des données et des méthodes solides, cohérentes et reproductibles (Franklin et Wulder, 2002). La mise en œuvre de ces systèmes doit se fonder sur des travaux de recherche et développement et nécessite de s'assurer que les méthodes élaborées (et les données utilisées) à l'échelle régionale sont suffisamment solides pour s'appliquer dans des zones géographiques nettement plus étendues où l'accès aux données (d'observation de la Terre ou de terrain) est plus difficile (Bontemps *et al.*, 2012 ; Gong *et al.*, 2013). Dans les faits, le passage de la recherche à la pratique est souvent délicat, car les scénarios de transition de la recherche vers les opérations et les applications présentent plusieurs difficultés et obstacles potentiels, dont notamment : a) le manque de connaissances scientifiques ; b) les difficultés rencontrées pour intégrer les connaissances scientifiques et les capacités techniques aux besoins opérationnels ; et c) les limites liées aux technologies d'observation, à l'efficacité de l'utilisation de ces observations ou à la puissance de calcul nécessaire pour utiliser les observations dans des modèles opérationnels (Conseil national de la recherche des Académies nationales des États-Unis, 2003). Il convient aussi de noter que la relation entre la recherche et la pratique n'est pas unidirectionnelle. Les systèmes opérationnels doivent être dynamiques et continuellement ajustés selon l'évolution des besoins, les dernières études, ainsi que l'amélioration continue et l'évolution constante des flux de données satellitaires et des réseaux de données sur le terrain. La recherche reste donc au cœur du système opérationnel pendant et après la mise en œuvre (GEOGLAM, 2015).

4.4.2. Les systèmes de surveillance en milieu agricole

Au cours des quarante dernières années, de nombreuses initiatives se sont intéressées à l'utilisation de l'imagerie satellite dans l'étude des terres. De nombreuses stratégies de cartographie ont ainsi été mises en place à l'échelle mondiale, régionale et nationale (tableau 2 ; Waldner *et al.*, 2015). La plupart de ces initiatives ont été de courte durée, à l'exception de certaines qui se poursuivent aujourd'hui dans le contexte opérationnel.

Les principaux systèmes de suivi mondiaux sont les suivants : a) le Réseau Sentinelles d'alerte précoce contre la famine (FEWS-NET) de l'Agence des États-Unis pour le développement international (USAID) ; b) le Système mondial d'information et d'alerte précoce sur l'alimentation et l'agriculture (SMIAR) de la FAO ; c) le projet MARS du Centre commun de recherche (JRC, de l'anglais *Joint Research Centre*) de la Commission européenne ; d) le Programme de surveillance des cultures de l'Institut des applications de la télédétection (IRSA, de l'anglais *Institute of Remote Sensing Application*) de l'Académie des sciences de Chine (ASC) et e) le Service agricole à l'étranger (FAS, de l'anglais *Foreign Agricultural Service*) de l'Initiative de suivi satellitaire de l'agriculture mondiale (GLAM) de l'USDA (Becker-Reshef *et al.*, 2010). Bien qu'utiles pour les travaux de suivi de l'agriculture mondiale, ces initiatives sont généralement limitées par leur résolution d'échantillonnage relativement basse (souvent plus de 10 km), le faible niveau de détail dans les classes de cultures, le manque de validation et des incertitudes non négligeables (Vancutsem *et al.*, 2012 ; Waldner *et al.*, 2015) qui les rendent inadéquates pour réaliser une cartographie nationale détaillée des cultures (au niveau des parcelles).

À l'échelle régionale ou nationale, le recours à une imagerie nettement plus précise (30 m au maximum) a permis de cartographier les cultures de façon plus précise et plus détaillée au niveau des parcelles. Voici quelques exemples d'initiatives de ce type :

- a. aux États-Unis, le Cropland Data Layer est préparé chaque année par le service NASS de l'USDA (à partir des images optiques captées par le satellite britannique DMC 2 et Landsat-8) ;

- b. au Royaume-Uni, l'ensemble de données « Land Cover Plus: Crops » (LC+ Crops) est préparé chaque année par le Centre britannique pour l'écologie et l'hydrologie (CEH, de l'anglais *Centre for Ecology & Hydrology*) en collaboration avec la société Remote Sensing Applications Consultants Ltd (RSAC) (à partir des images optiques captées par Sentinel-2 et de l'imagerie SAR en bande C captée par Sentinel-1);
- c. en Chine, le Système de suivi de l'agriculture par télédétection (CHARMS, de l'anglais *Chinese Agriculture Remote Sensing Monitoring System*) et le Système de surveillance des cultures (CCWS, de l'anglais *China Crop Watch System*) sont utilisés par le ministère de l'Agriculture et l'ASC pour diffuser plusieurs fois par mois des produits de données (à partir d'images à résolution moyenne [moins de 30 m, p. ex., HJ-1, CCD, GF-1, Landsat-5, -7 et -8, SPOT, IRS, Envisat] et d'images à basse résolution [de 250 à 1 000 m, p. ex., MODIS, AVHRR]);
- d. au Pakistan, le Système d'information agricole (AIS, de l'anglais *Agricultural Information System*) est utilisé par l'Agence spatiale nationale (SUPARCO) et le ministère de l'Alimentation et de l'Agriculture pour diffuser régulièrement des produits de données en collaboration avec la FAO (à partir des images optiques de SPOT-5);
- e. en Inde, le programme d'estimation des superficies consacrées aux cultures et de la production végétale (CAPE, de l'anglais *Crop Acreage and Production Estimation Program*) est utilisé par le ministère de l'Agriculture pour réaliser chaque année une cartographie de l'utilisation des terres (à partir de l'imagerie optique captée par RESOURCESAT-1 et -2 et des données SAR de RISAT-1 et -2);
- f. en Russie, le service satellitaire de surveillance de la végétation VEGA-PRO est utilisé par l'Académie des sciences de Russie pour communiquer des données plusieurs fois par an, avec le soutien de la Fondation Skolkovo;
- g. en Europe, le projet Sentinel-2 pour le suivi de l'agriculture (Sen2-Agri) est utilisé par l'ESA pour communiquer régulièrement des données pendant la période de végétation (à partir des images optiques captées par Sentinel-2).

TABLEAU 2. RECENSEMENT DES CARTES NATIONALES, RÉGIONALES ET MONDIALES DE LA COUVERTURE DU SOL (SOURCE : WALDNER ET AL., 2015). CE TABLEAU NE CONTIENT QUE DES DONNÉES ACCESSIBLES À TOUS. LES ENSEMBLES DE DONNÉES DONT L'UTILISATION OU L'ACCÈS EST LIMITÉ PAR UNE POLITIQUE DE DIFFUSION N'ONT PAS ÉTÉ RECENSÉS.

Échelle	Nom du produit et référence	Période
Mondiale		
	FROM-GLC (Gong <i>et al.</i> , 2013)	2013
	Global Cropland Extent (Pittman <i>et al.</i> , 2010)	2000-2008
	GlobCover 2009 (Arino <i>et al.</i> , 2008)	2009
	Climate Change Initiative Land Cover (CCI) (Defourny <i>et al.</i> , 201X)	2000, 2005, 2010
	MODIS Land Cover Type MOD12Q1, 2005 (NASA)	2005
	GLC-SHARE, FAO (Latham <i>et al.</i>)	1990-2014
	IIASA-IFPRI Global Cropland Map (Fritz <i>et al.</i> , 2015)	1990-2012
	GLC2000 (Bartholomé et Belward, 2005)	1999-2000
	Programme international géosphère-biosphère (PIGB) (Eidenshink et Faundeen, 1994)	1992-1993
	Global Land Cover by National Mapping Organizations (GLCNMO) (Tateishi <i>et al.</i> , 2011)	2007-2009
Régionale		
	Corine Land Cover, Agence européenne pour l'environnement (AEE)	2006
	Base de données sur la couverture du sol de la Communauté de développement de l'Afrique australe, Conseil sud-africain de la recherche scientifique et industrielle (CSIR, de l'anglais <i>Council for Scientific and Industrial Research</i>)	2002
	Masque de cultures d'Afrique, JRC (Vancutsem <i>et al.</i> , 2012)	2012
	Atlas environnemental de l'Afrique du Nord, Commission de coopération environnementale (CCE)	2005

Échelle	Nom du produit et référence	Période
	Carte de la couverture du sol pour l'Amérique latine et les Caraïbes (Blanco <i>et al.</i> , 2013)	2008
Burundi, Congo, Égypte, Érythrée, Kenya, Ouganda, République-Unie de Tanzanie, Rwanda, Somalie, Soudan	Carte du bassin du Congo (Verhegghen <i>et al.</i> , 2012)	2000-2007
	Africover, FAO	1999-2001
Bhoutan, Népal, Sénégal	Réseau mondial sur le couvert végétal (GLCN)	2005-2007
Belgique, France, Pays-Bas	Système d'identification des parcelles agricoles	2012-2014
Barbade, Dominique, Grenade, Îles Vierges, Porto Rico, Saint-Kitts-et-Nevis, République dominicaine	USGS	2000-2001
Fidji, Îles Cook, Îles Marshall, Îles Salomon, Kiribati, Micronésie, Nioué, Palaos, Timor-Leste, Tonga, Tuvalu, Vanuatu	Division des géosciences et technologies appliquées (Communauté du Pacifique)	1999-2010
Botswana, Malawi, Namibie, République-Unie de Tanzanie, Rwanda, Zambie	Land Cover Scheme II, Système régional de visualisation et de surveillance (ICIMOD-SERVIR)	2010
Nationale		
Chine	GlobeLand30 (Chen <i>et al.</i> , 2015)	2009-2010
Japon	High Resolution Land Use-Land Cover Map, Agence japonaise d'exploration aérospatiale (JAXA) (Takahashi <i>et al.</i> , 2013)	2006-2011
Tadjikistan	(Thenkabail <i>et al.</i> , 2012)	2010
Burkina Faso	Base de données de l'occupation des terres du Burkina Faso	2000
Canada	Inventaire annuel des cultures, AAC (Fisette <i>et al.</i> , 2013, 2014 et 2015)	2011-présent
Canada	Ressources naturelles Canada (Latifovic <i>et al.</i> , 2004)	2005
États-Unis d'Amérique	Cropland Data Layer, USDA (NASS, 2015)	2008-présent
Chine	Carte de la couverture du sol de Chine (Liu <i>et al.</i> , 2005)	1995-1996
Australie	Base de données numérique sur la couverture du sol (Lymburner <i>et al.</i> , 2011)	2011
Cambodge	Couverture du sol du Cambodge, Agence japonaise de coopération internationale (AJCI)	2002
Nouvelle-Zélande	Base de données sur la couverture du sol, version 4, ministère de l'Environnement	2004
Afrique du Sud	Couverture du sol de l'Afrique du Sud, CSIR	2000-2001
Afrique du Sud	Couverture du sol de l'Afrique du Sud, Institut national d'Afrique du Sud sur la biodiversité (SANBI)	2009
Uruguay	Couverture du sol de l'Uruguay, FAO	2010
Mexique	Couverture du sol du Mexique, Commission nationale du Mexique pour la connaissance et l'usage de la biodiversité (CONABIO)	1999
Argentine	Couverture et usage du sol, Institut national de technologie agricole (INTA)	2006
Équateur	Usage du sol, département de l'Information environnementale	2001
Thaïlande	Département des Forêts royales de Thaïlande	2000
Chili	Corporation nationale forestière du Chili	1999
Inde	Utilisation des terres et couverture du sol de l'Inde, Centre national de télédétection de l'Inde (Sreenivas <i>et al.</i> , 2015)	2012
Gambie	(Holecz <i>et al.</i> , 2013)	2013
Royaume-Uni	Ensemble de données « Land Cover Plus: Crops » du Royaume-Uni (Conseil de recherche pour l'environnement naturel, 2015)	2015-présent
Ukraine	Couverture du sol de l'Ukraine (Lavreniuk <i>et al.</i> , 2015)	2010
Fédération de Russie	TerraNorte Arable Lands of Russia (Bartalev <i>et al.</i> , 2011)	2014

4.4.3. Étude de cas : l'inventaire annuel des cultures du Canada fondé sur des données spatiales

4.4.3.1. Exploitation des données d'observation de la Terre par le gouvernement canadien

Agriculture et Agroalimentaire Canada (AAC) — le ministère en charge du secteur canadien de l'agriculture et de l'agroalimentaire — a exploité les récents progrès technologiques accomplis dans le domaine des satellites et des capteurs ainsi que la multiplication des missions spatiales d'observation de la Terre pour cartographier et suivre l'utilisation des terres agricoles et son évolution. Les images utilisées par AAC proviennent de plusieurs satellites et capteurs (notamment Landsat-5, -7 et -8, RADARSAT-2, SPOT, DMC, RapidEye et Ressourcesat-1) qui balayent les régions optiques et micro-ondes du spectre électromagnétique selon diverses résolutions spatiales (de pixels, comprises entre 5 et 56 m). De nombreux travaux publiés par les chercheurs d'AAC ont déjà démontré que les satellites d'observation de la Terre permettaient d'obtenir en temps voulu et de façon rentable des informations exactes et évolutives sur l'utilisation des terres à l'échelle nationale (McNairn *et al.*, 2009; Shang *et al.*, 2009; Deschamps *et al.*, 2012; Fisette *et al.*, 2013 et 2014; Champagne *et al.*, 2014; Jiao *et al.*, 2014; Liu *et al.*, 2016) et son évolution (Pei *et al.*, 2011; Li *et al.*, 2013; El-Khoury *et al.*, 2014 et 2015; Huffman *et al.*, 2015). À l'avenir, AAC devrait encore mieux répondre aux besoins informationnels du secteur grâce au lancement et à la mise en service puis à disposition de nouvelles technologies de capteurs satellitaires (telles que les systèmes Sentinel-1, -2 et -3 et la constellation RADARSAT). L'utilisation des données d'observation de la Terre jouera un rôle essentiel dans la préparation de la nouvelle génération de grands produits informationnels officiels d'AAC.

4.4.3.2. L'inventaire annuel des cultures du Canada basé sur des données spatiales

L'inventaire annuel des cultures du Canada (IAC, voir la figure 3) est l'un des principaux produits informationnels préparés par AAC à partir de données d'observation de la Terre. Cet inventaire est actualisé chaque année et accessible gratuitement au public sur le portail de données ouvertes du gouvernement canadien. Il prend la forme d'une carte quadrillée (appelée « image matricielle ») qui rend compte de l'utilisation des terres agricoles et de la couverture du sol non agricole sur le territoire agricole du Canada (Fisette *et al.*, 2013 et 2014). Les premières versions de l'IAC (2009 et 2010) couvraient les provinces des Prairies (Alberta, Saskatchewan et Manitoba), puis la carte a été étendue à la totalité du pays à partir de 2012. L'IAC contient des informations très précises au niveau des parcelles sur les différents types de cultures, pâturages et prairies. Il constitue dès lors un ensemble de données de base crucial pour l'élaboration des programmes et des politiques d'AAC. L'IAC offre également des informations importantes sur l'utilisation des terres aux autres ministères du gouvernement canadien, à différents gouvernements provinciaux, aux producteurs, aux organisations non gouvernementales, aux universités et collèges, ainsi qu'aux entités du secteur public et privé.

[illegible]

Les méthodes de classification utilisées pour la création de l'IAC sont élaborées et personnalisées par AAC (McNairn *et al.*, 2009 ; Shang *et al.*, 2009 ; Deschamps *et al.*, 2012 ; Fisette *et al.*, 2013 et 2014 ; Champagne *et al.*, 2014 ; Jiao *et al.*, 2014 ; Liu *et al.*, 2016). Pour les versions originales de l'IAC, une méthode de classification fondée sur les arbres de décision avait été utilisée. Ce choix se fondait principalement sur son efficacité par rapport à d'autres approches de classification supervisée, démontrée dans le cadre de plusieurs travaux de classification de l'utilisation des terres (voir, par exemple, Friedl et Brodley, 1997 ; Brown de Colstoun *et al.*, 2003 ; Pal et Mather, 2003 ; Xu *et al.*, 2005 ; Peña-Barragán *et al.*, 2011 ; Deschamps *et al.*, 2012 ; Vieira *et al.*, 2012 ; Liu *et al.*, 2016), et particulièrement dans l'élaboration du CDL annuel aux États-Unis (Boryan *et al.*, 2011). Plus récemment, AAC s'est tourné vers les approches de classification par forêts aléatoires. Cette méthode a permis d'obtenir des résultats particulièrement bons dans le domaine du surveillance en milieu agricole ainsi qu'un niveau élevé de précision cartographique au Canada (McNairn *et al.*, 2009 ; Deschamps *et al.*, 2012 ; Fisette *et al.*, 2013, 2014 et 2015) et ailleurs dans le monde (p. ex., aux États-Unis [Xie *et al.*, 2007 ; Watts *et al.*, 2009 ; Zhong *et al.*, 2014], en Argentine [Valero *et al.*, 2016 ; Waldner *et al.*, 2016], au Brésil [Müller *et al.*, 2015 ; Waldner *et al.*, 2016], au Pérou [Tatsumi *et al.*, 2015], en Chine [Valero *et al.*, 2016 ; Waldner *et al.*, 2016 ; Hütt *et al.*, 2016], en Russie [Valero *et al.*, 2016 ; Waldner *et al.*, 2016], en Ukraine [Valero *et al.*, 2016 ; Waldner *et al.*, 2016], au Japon [Sonobe *et al.*, 2014], en Iran [Eisavi *et al.*, 2015], en Australie [Pringle *et al.*, 2012], en France [Valero *et al.*, 2016], au Burkina Faso [Valero *et al.*, 2016] et en Afrique du Sud [Valero *et al.*, 2016]).

Depuis 2009, AAC réalise son IAC en appliquant la méthode des forêts aléatoires à diverses combinaisons d'images optiques (Landsat-5, -7 et -8; Resourcesat-1; DMC; SPOT) et SAR (RADARSAT-2) au moyen de fonctions discriminantes estimées de façon empirique sur des centaines de milliers d'échantillons de données de terrain servant pour la phase d'apprentissage. La méthode des forêts aléatoires a été privilégiée en raison de ses avantages comparatifs, dont notamment, dans le contexte particulier d'AAC, son efficacité pour traiter des données discrètes, sa vitesse de traitement (18 fois supérieure aux arbres de décision, d'après Deschamps *et al.*, 2012), son indépendance vis-à-vis de la répartition des signatures de classe, l'interprétabilité de ses règles de classification, sa rentabilité et sa précision accrue (Friedl et Brodley, 1997; Pal et Mather, 2003). Pour améliorer davantage la précision de l'algorithme, AAC a enrichi le processus de classification par arbres de décision avec des options avancées, telles que l'élagage et le boosting.

Au sein d'AAC, la classification annuelle par forêts aléatoires est réalisée région par région, car l'évolution des rotations de culture, de la croissance des cultures et des modèles de récoltes génère d'importantes différences de réflectance entre les scènes adjacentes captées par le satellite au cours de la période de disponibilité de la scène. Chaque région de classification doit combiner des images optiques et SAR captées à différentes dates, mais la combinaison réelle des images par région peut toutefois varier selon les limites de disponibilité des données. La classification région par région doit donc être envisagée par chaque organisme souhaitant réaliser une cartographie nationale à grande échelle.

4.4.3.4. Données de terrain (*in situ*)

Les informations sur la réalité de terrain servant à l'apprentissage et à la classification des modèles sont déterminantes pour la réussite de la classification des cultures. Actuellement, les données annuelles de l'assurance-récolte fournies par quatre provinces canadiennes (l'Alberta, la Saskatchewan, le Manitoba et le Québec, soit collectivement 87 % du territoire agricole total) sont les sources d'informations géospatiales les plus exactes, les plus détaillées et les plus complètes sur les types de cultures du Canada. Lorsque les provinces ne peuvent fournir ces données d'assurance, les informations sur la réalité de terrain proviennent d'observations réalisées sur le terrain par le personnel d'AAC. Avant de servir pour la classification par forêts aléatoires, les données *in situ* et leur biais éventuel sont analysés. Les observations de terrain présentent deux principales sources de biais. Le premier type de biais peut apparaître dans les données tirées de l'assurance-récolte, certaines cultures étant moins fréquemment assurées que d'autres. La non-représentativité des cultures non assurées dans les bases de données de l'assurance-récolte entraîne leur sous-représentativité dans les données d'apprentissage utilisées pour la classification par forêts aléatoires. Le second type de biais peut aussi survenir si le calendrier des opérations de terrain réalisées par AAC ne coïncide pas avec les étapes de la croissance des cultures qui permettent le mieux d'identifier précisément les types de cultures. Les erreurs d'identification des cultures sont alors plus fréquentes et les données d'apprentissage utilisées pour la classification par forêts aléatoires peuvent contenir des erreurs systématiques. Ces deux exemples de biais peuvent réduire la précision des cartes sur l'utilisation des terres.

4.4.3.5. Besoin et prétraitement des données satellitaires optiques et SAR

Les études indiquent que les données satellitaires optiques et SAR sont autant nécessaires les unes que les autres pour caractériser correctement les grandes étapes (phénologiques) de croissance des cultures et réaliser une cartographie de haute précision à l'échelle du Canada (McNairn *et al.*, 2009; Shang *et al.*, 2009; Deschamps *et al.*, 2012; Fisette *et al.*, 2013 et 2014; Champagne *et al.*, 2014; Jiao *et al.*, 2014; Liu *et al.*, 2016). En effet, ces deux types de flux de données fournissent des informations uniques et précieuses sur le type de plante et leur croissance. Dans le cas de l'imagerie optique, il a été démontré que les observations acquises dans les régions NIR et SWIR du spectre électromagnétique s'avéraient extrêmement utiles pour différencier les types de cultures (Clark *et al.*, 1995). Les données SAR à double polarisation sont nettement plus sensibles à la structure des plantes et permettent de combler les lacunes présentes dans les images optiques lorsque les conditions météorologiques n'étaient pas idéales pendant les principales étapes de la croissance. Les résultats des études (McNairn *et al.*, 2009) et des applications pratiques (Fisette *et al.*, 2015) ont montré que la combinaison des données optiques et SAR offrait une meilleure précision que les données optiques seules. Par exemple, la combinaison entre les données optiques

et les données à double polarisation de RADARSAT-2 a permis une augmentation de plus de 16 % de la précision globale (Fisette *et al.*, 2015).

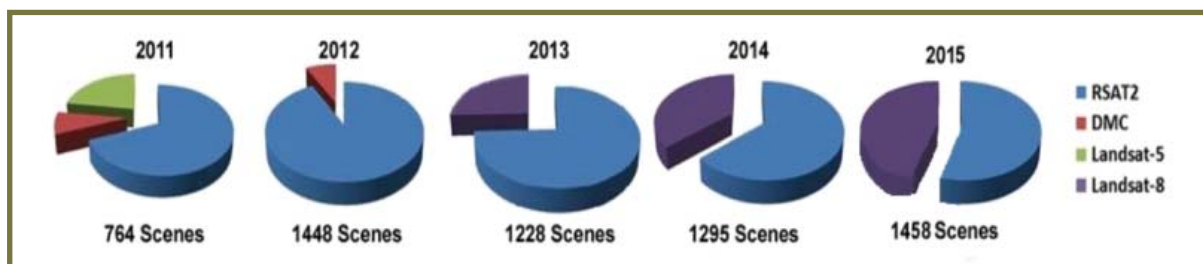
Les données optiques et SAR doivent être prétraitées avant de pouvoir servir de données d'entrée pour la réalisation de l'IAC au moyen de la classification par forêts aléatoires. Les images n'ayant pas encore été corrigées géométriquement sont orthorectifiées au moyen d'un modèle physique tridimensionnel multi-capteurs (Toutin, 2005). Étant donné qu'AAC réalise ses classifications annuelles par forêts aléatoires région par région, le mosaïquage et la correction atmosphérique des données optiques multidates ne sont pas nécessaires. Néanmoins, la présence possible de nuages sur ces images justifie l'application à chaque image optique d'une technique automatique de masquage des nuages et des ombres (Zhu et Woodcock, 2012) afin de corriger les sections de mauvaise qualité. Le bruit (chatolement) présent sur les données radar est éliminé au moyen d'un filtre gamma selon la méthode du maximum *a posteriori*. Les données sont ensuite rééchantillonnées sous un format 16-bits afin de réduire le temps de traitement. Les images monodates captées le long de la trajectoire sont mosaïquées ensemble afin de créer des régions de classification étendues contenant un nombre plus important de sites d'apprentissage.

4.4.3.6. Évolution des images optiques et SAR utilisées en entrée

Depuis la première version de l'IAC publiée en 2009, les données satellitaires utilisées en entrée du processus de classification (figure 4) ont évolué. Tandis que les données SAR exploitées dans ce cadre proviennent depuis toujours de RADARSAT-2, les satellites et capteurs fournissant les données optiques ont changé au fil du temps sous l'effet d'une combinaison de facteurs incitatifs et attractifs. Les facteurs incitatifs correspondent aux inconvénients des flux de données dont l'insuffisance nécessite de chercher d'autres sources. Les facteurs attractifs correspondent quant à eux aux avantages des flux de données nouveaux (ou nouvellement disponibles) qui les rendent préférables aux flux existants. Ces deux types de facteurs concernent généralement la disponibilité et le coût des données, les résolutions spatiale et spectrale, et la fréquence de revisite.

En 2009 et 2010, l'IAC couvrait uniquement les provinces canadiennes des Prairies (l'Alberta, la Saskatchewan et le Manitoba) en combinant des données optiques et SAR multidates captées par Ressourcesat-1 et RADARSAT-2. Ses cartes de cultures, quadrillées selon la résolution spatiale de 56 m de Ressourcesat-1, offraient une précision acceptable (supérieure à 75 %) pour les parcelles relativement étendues de l'Ouest canadien. Ce niveau de précision n'a toutefois pas pu être conservé lors des essais d'extrapolation de cette approche aux régions de l'est du pays — où les parcelles sont globalement plus petites, plus longues et/ou plus étroites. Il a donc été conclu que le flux de données de Ressourcesat-1 n'était pas adapté pour cartographier les cultures à l'échelle du Canada.

FIGURE 4. SOURCES ET VOLUMES DES DONNÉES UTILISÉES POUR LES INVENTAIRES ANNUELS DES CULTURES D'AAC, 2011-2015.



En 2011 et 2012, l'étendue géographique de l'IAC a été élargie à toutes les provinces canadiennes, à l'exception de la Colombie-Britannique et de Terre-Neuve-et-Labrador. Les inconvénients liés à la basse résolution spatiale des données de Ressourcesat-1 ont été directement évités en utilisant des données de plus haute résolution tirées

d'autres capteurs satellitaires optiques. En 2011, les flux de données englobaient des observations multidates réalisées par Landsat-5 et des observations monodates captées par la Disaster Monitoring Constellation (DMC), qui ont ensuite été intégrées aux images de RADARSAT-2. Les cartes de cultures de l'IAC 2011, quadrillées selon une résolution spatiale de 30 m, offraient par conséquent une précision élevée pour l'Ouest canadien ainsi que pour la majeure partie de l'est du pays (à l'exception des provinces de l'Île-du-Prince-Édouard et de la Nouvelle-Écosse, pour lesquelles la précision des cartes était inférieure à 70 %). En 2012, suite à la panne qui a touché le satellite Landsat-5 en milieu d'année, les données d'observation de la Terre utilisées pour créer l'IAC provenaient uniquement des images multidates de RADARSAT-2 et des observations monodates de DMC. La baisse non négligeable des données optiques disponibles a été partiellement compensée par le doublement des acquisitions réalisées par RADARSAT-2 par rapport à l'année précédente. Les cartes de cultures de l'IAC 2012, quadrillées selon une résolution spatiale de 30 m (comme celles de 2011), offraient une précision élevée (plus de 75 %) pour toutes les provinces cartographiées, y compris celles qui présentaient une qualité médiocre l'année précédente. L'augmentation de la précision des cartes de l'Île-du-Prince-Édouard et de la Nouvelle-Écosse en 2012 s'explique par la compensation des données optiques manquantes par les données SAR et par la meilleure qualité de l'ensemble de données de terrain servant pour l'apprentissage.

En 2013, la couverture géographique de l'IAC a encore été élargie jusqu'à couvrir toutes les provinces agricoles du Canada. Cette même année, le nouveau satellite Landsat-8 est devenu l'unique flux de données optiques alimentant l'IAC. L'intégration de ces données aux images SAR captées par RADARSAT-2 a permis d'obtenir une précision élevée (plus de 75 %) pour toutes les provinces cartographiées. Au moment de la rédaction de cette publication (2016), la combinaison des images captées par Landsat-8 et RADARSAT-2 reste une source fiable de données et de haute qualité pour la réalisation de l'IAC. D'après les estimations, la précision cartographique devrait dépasser 85 % pour toutes les provinces en 2016. Cette source de données d'observation de la Terre désormais stable montre que l'amélioration de la précision cartographique de l'IAC depuis 2013 s'explique en grande partie par des facteurs non liés à l'imagerie, tels que le renforcement de la qualité et de la fiabilité de l'échantillonnage de terrain et l'amélioration des résultats de la classification permise par l'ajustement des approches fondées sur les forêts aléatoires.

4.4.3.7. Traitement post-classification

La méthode des forêts aléatoires utilisée par AAC ne produit pas directement une carte de l'utilisation des terres et/ou de la couverture du sol qui soit utilisable par tous. Trois étapes de traitement ultérieur — le filtrage, le mosaïquage et l'ajout de classes permanentes — sont nécessaires avant d'évaluer la précision du produit final, de le valider et de publier les données (figure 1). La première étape consiste à appliquer des filtres post-classification aux résultats de la classification par forêts aléatoires. En effet, les cartes résultant de la classification fondée sur les pixels contiennent généralement des pixels isolés et mal classifiés. Les erreurs de classification des pixels sont souvent dues aux variations de croissance des cultures d'une même parcelle. Or, leur non-correction peut donner l'impression (souvent fausse) que l'agriculteur a planté plusieurs cultures sur cette parcelle. Les filtres post-classification corrigent directement ce problème en attribuant la valeur de la classe majoritaire de chaque parcelle à chaque pixel de cette parcelle. Il a été démontré que ce procédé améliorait de 5 à 10 % la précision de classification de certaines classes de cultures dans l'IAC (McNairn *et al.*, 2009; Fisette *et al.*, 2014 et 2015). La deuxième étape consiste à mosaïquer les différentes régions de classification afin de produire des cartes de cultures pour chaque province agricole du Canada. Le mosaïquage se fait au moyen d'un processus automatique qui sélectionne les zones les plus précises contenant des zones de superposition entre les régions de classification. La troisième et dernière étape de préparation de la carte consiste à ajouter des classes permanentes à l'IAC. Chaque année, AAC actualise une base de données contenant des éléments « permanents » (p. ex., parcours de golf, terrains de sport, pistes de ski et aéroports) peu susceptibles de changer d'une année sur l'autre. L'identification de ces classes évite de les confondre avec des classes agricoles présentant des similitudes spectrales. Les éléments permanents sont ajoutés *a posteriori* aux classifications filtrées et mosaïquées de l'utilisation des terres et/ou de la couverture du sol.

4.4.3.8. Évaluation de la précision et validation du produit

L'évaluation de la précision de l'IAC a lieu après sa création. Pour cela, AAC utilise la matrice de confusion (d'erreur), adoptée par les professionnels de la télédétection comme convention standard d'évaluation de la précision cartographique (Congalton, 1991). En plus d'évaluer la précision globale de la technique de classification, la matrice de confusion calcule les erreurs d'exclusion (erreurs par omission), les erreurs d'inclusion (erreurs de commission) et les F-scores associés à chaque classe d'une classification. La précision globale de l'IAC est calculée en divisant le nombre total de parcelles classifiées correctement par le nombre total de parcelles dans la matrice de confusion. La précision du producteur (erreur par omission) est la probabilité pour une parcelle de référence d'une classe de culture X d'être correctement classifiée comme classe X. La précision de l'utilisateur (erreur de commission) est la probabilité pour une parcelle classifiée X d'être bien une classe X dans la réalité de terrain. Si aucune parcelle n'est classifiée comme l'une des classes de référence prédéfinies, la précision de l'utilisateur n'est pas évaluée. Enfin, le F-score — un indicateur spécifique à chaque classe qui n'est pas affecté par les informations des autres classes — correspond à la moyenne harmonique de la précision de l'utilisateur et du producteur (Waldner *et al.*, 2015).

Dans le cadre de la classification des images, la précision globale des cultures dans l'IAC est validée pour chaque combinaison image-date. Une fois cartographiées, ces données fournissent des informations détaillées sur les variations de précision des données spatiales sur les cultures pour l'ensemble du territoire agricole du Canada (figure 5). Bien que l'IAC présente systématiquement une précision globale moyenne de 85 % à l'échelle nationale, la précision des cartes varie selon les cultures, les régions et les années, en fonction de la disponibilité des données satellitaires et de la représentativité géographique des données de terrain servant pour l'apprentissage (tableau 3). En général, la meilleure précision cartographique (soit plus de 90 %) est obtenue dans les régions où les cultures présentent des caractéristiques spectrales assez différentes lors de l'acquisition des données d'observation de la Terre, comme dans les provinces canadiennes des Prairies de la Saskatchewan et de l'Alberta. Ailleurs, toutefois, la précision est moindre et peut varier de 70 à 80 % environ. Cette variabilité s'explique principalement par deux facteurs tenant aux limites des données satellitaires et de terrain utilisées par la classification.

FIGURE 5. COUVERTURE GÉOGRAPHIQUE ET PRÉCISION GLOBALE DE LA CARTOGRAPHIE DANS L'INVENTAIRE ANNUEL DES CULTURES 2016 DU CANADA.

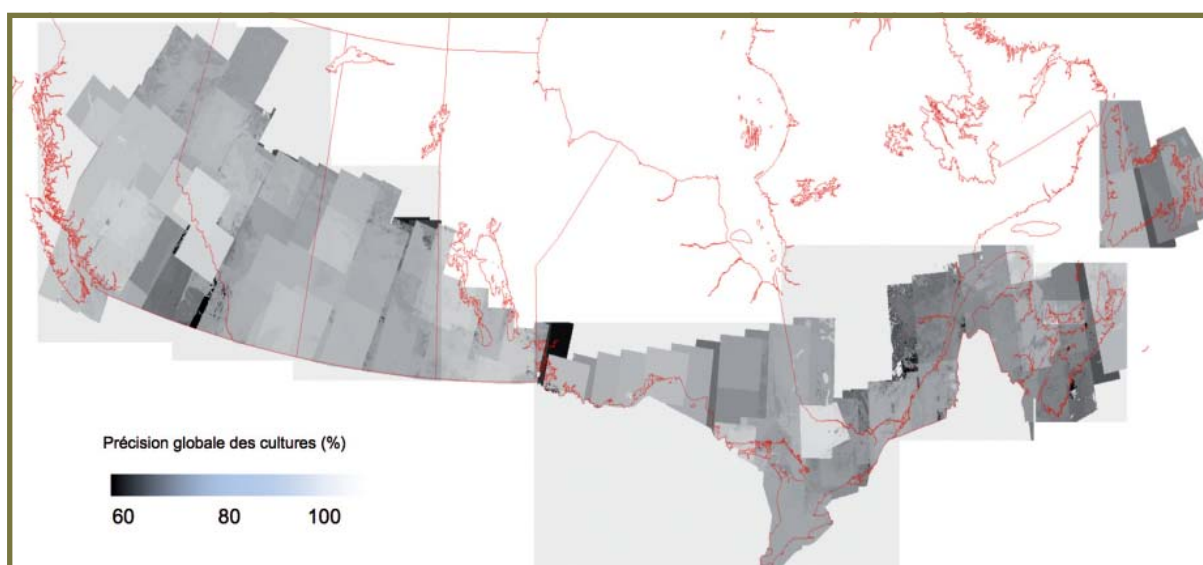


TABLEAU 3. VARIABILITÉ PAR PROVINCE DE LA PRÉCISION DE LA CLASSIFICATION DE L'INVENTAIRE ANNUEL DES CULTURES D'AAC, 2015

Province	2009	2010	2011	2012	2013	2014	2015
T.-N.-L.	--	--	--	--	--	--	--
Î.-P.-É.	--	--	67,5	78,7	87,9	81,0	83,7
N.-B.	--	--	88,1	88,0	87,3	89,1	86,1
N.-É.	--	--	64,2	89,9	74,2	64,4	85,2
Qc	--	--	81,4	81,8	87,5	83,9	87,1
Ont.	--	--	80,8	76,2	88,2	87,9	89,6
Man.	80,0	85,1	79,0	85,9	85,4	90,3	90,0
Sask.			87,1	82,4	86,5	85,9	89,7
Alb.			87,7	88,4	89,9	89,4	88,9
C.-B.	--	--	--	--	79,2	88,4	--
CANADA	80,0	85,1	85,3	83,9	86,0	87,4	85,0

(T.-N.-L. = Terre-Neuve-et-Labrador, Î.-P.-É. = Île-du-Prince-Édouard, N.-B. = Nouveau-Brunswick, N.-É. = Nouvelle-Écosse, Qc = Québec, Ont. = Ontario, Man. = Manitoba, Sask. = Saskatchewan, Alb. = Alberta, C.-B. = Colombie-Britannique). La précision n'est pas calculée pour les provinces et les années pour lesquelles il n'existe pas suffisamment de données de terrain destinées à la validation (ou lorsque le produit ne doit pas nécessairement être de haute qualité, comme dans le cas de la province de T.-N.-L. qui compte 95 % de pâturages).

En premier lieu, le nombre, la densité et la qualité (détail et précision) des données de terrain servant pour l'apprentissage et la validation de la classification par forêts aléatoires peuvent varier selon les provinces en raison de la pluralité des sources de données. Généralement, les données de meilleure qualité : a) offrent la meilleure représentativité spatiale des paysages agricoles ; b) sont les plus précises dans l'espace et dans le temps ; c) utilisent les classes (thématiques) de cultures les plus détaillées ; et d) contiennent les grands échantillons nécessaires pour l'apprentissage et la validation des classifications par forêts aléatoires. Les données fournies par les gouvernements provinciaux de l'Alberta, de la Saskatchewan, du Manitoba et du Québec répondent globalement à tous ces critères (bien que détaillées, les données apportées par le gouvernement provincial de la Colombie-Britannique portent uniquement sur certaines classes et n'offrent donc pas de représentativité spatiale complète). En revanche, les données relatives aux autres provinces (Ontario, Nouveau-Brunswick, Nouvelle-Écosse, Île-du-Prince-Édouard et Terre-Neuve-et-Labrador) sont collectées par le personnel d'AAC au moyen d'enquêtes superficielles qui offrent une moins bonne représentativité spatiale du paysage agricole dans son ensemble. Les cartes de cultures préparées à partir de ces données satisfont généralement aux exigences de précision cartographique définies par AAC, mais seulement au prix d'une généralisation qui consiste à regrouper les classes détaillées (p. ex., blé, orge et avoine) en une seule classe générale (p. ex., les céréales). Un renforcement de l'accès aux données de terrain collectées par toutes les provinces éviterait de telles incohérences dans le niveau de détail des classes à travers le pays.

En second lieu, la variation souvent significative de la couverture nuageuse entre les régions peut limiter la quantité de données satellitaires optiques (actuellement captées par Landsat-8) disponibles pour la classification par forêts aléatoires. Par exemple, si, par ciel dégagé, Landsat-8 est capable de capter trois à quatre images par période de végétation dans les provinces canadiennes des Prairies, les conditions météorologiques des régions côtières (à savoir principalement les régions agricoles des Provinces atlantiques – Terre-Neuve-et-Labrador, le Nouveau-Brunswick, la Nouvelle-Écosse et l'Île-du-Prince-Édouard) l'empêchent souvent de capter ne serait-ce qu'une seule image. Bien qu'AAC s'appuie davantage sur les images micro-ondes captées chaque mois (par RADARSAT-2) pour conduire la classification par forêts aléatoires dans ces régions, le manque ou l'absence de données optiques entraîne néanmoins une chute de 10 à 15 % de la précision cartographique par rapport aux autres régions pour lesquelles des données optiques sont disponibles. Ce problème devrait toutefois s'atténuer à l'avenir avec le renforcement de l'accès aux nouveaux flux d'images optiques (p. ex., Sentinel-2) et SAR (p. ex., Sentinel-1 et la mission de la Constellation RADARSAT qui sera lancée en 2018).

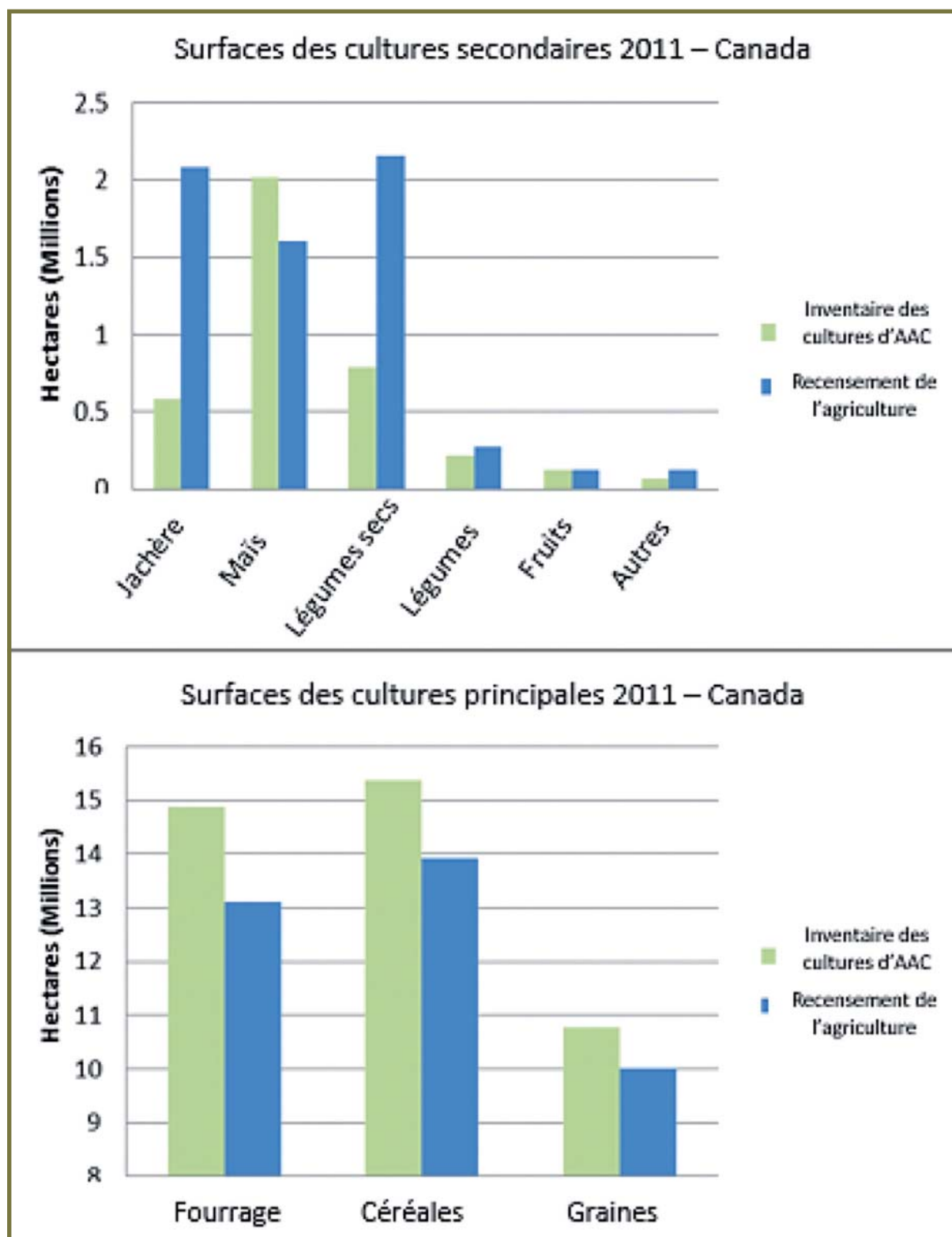
Il faut enfin noter que les matrices de confusion, bien que couramment utilisées, sont limitées par leur incapacité à expliquer plus précisément la source de l'erreur, qui peut provenir de la performance de l'algorithme de classification ou de la résolution spatiale des données de télédétection (Boschetti *et al.*, 2004). AAC envisage donc actuellement de se tourner vers une autre méthode — le triangle d'entropie de De Finetti (Valverde-Albacete et Peláez-Moreno, 2010) — pour évaluer ses erreurs de classification. Bien qu'elle n'ait pas encore véritablement fait sa place dans les études sur la classification des données de télédétection, cette méthode solidement basée sur la théorie de l'information pourrait fournir des indications plus fidèles sur les résultats de la classification que les précisions calculées à partir des matrices de confusion (Valverde-Albacete et Peláez-Moreno, 2014).

4.4.3.9. Comparaison avec les estimations des surfaces cultivées tirées des enquêtes statistiques (Recensement de l'agriculture)

Tous les cinq ans, l'IAC d'AAC fait l'objet d'une validation supplémentaire à partir des informations statistiques provenant du Recensement de l'agriculture réalisé à l'échelle nationale par Statistique Canada, l'organisme public chargé de recueillir, compiler, analyser, dépouiller et publier des renseignements statistiques portant sur les activités commerciales, industrielles, financières, sociales, économiques et générales de la population et sur l'état de celle-ci (Gouvernement canadien, 2008). La comparaison entre l'IAC et le Recensement de l'agriculture (figure 6) pour l'année de recensement la plus récente (2011) (Fisette *et al.*, 2014) montre que : a) à l'échelle nationale, l'IAC surestime de 15 % le territoire agricole du Canada par rapport au Recensement de l'agriculture 2011 ; et b) à une échelle plus régionale, l'IAC tend à surestimer les surfaces des principales cultures (p. ex., le fourrage, les céréales et les graines oléagineuses) et, à l'exception du maïs et des fruits, à sous-estimer les surfaces des cultures secondaires, parfois de façon drastique (p. ex., les zones en jachère et les légumes secs). Les différences entre les estimations fournies par ces deux sources d'information doivent être rapprochées.

Les écarts entre l'IAC et le Recensement de l'agriculture pour l'année 2011 sont dus à la variabilité de la taille des échantillons et de la répartition géographique des données de terrain utilisées pour la phase d'apprentissage de la classification par forêts aléatoires. La précision des classes de cultures cartographiées dépend fortement de la justesse des données d'apprentissage représentant chaque classe (Pal et Mather, 2003). La question de la sur- ou sous-représentativité d'une classe cartographiée dans les échantillons est particulièrement importante. Les classes de l'IAC surreprésentées dans les données d'apprentissage le sont généralement aussi dans la carte produite, tandis que les surfaces des cultures marginales (secondaires) y sont sous-représentées. Ces biais tendent à confirmer l'observation selon laquelle l'estimation des surfaces cultivées par un simple comptage de pixels dans les données de télédétection ne suffit pas à fournir des estimations non biaisées sur les superficies ensemencées (Gallego, 2004 et 2006).

FIGURE 6. COMPARAISON ENTRE LES ESTIMATIONS DES SURFACES CULTIVÉES DE L'INVENTAIRE ANNUEL DES CULTURES 2011 D'AAC ET LES DONNÉES DU RECENSEMENT DE L'AGRICULTURE RÉALISÉ LA MÊME ANNÉE PAR STATISTIQUE CANADA.



Plusieurs techniques statistiques ont été proposées (et appliquées) pour réduire le biais du simple comptage de pixels pour l'estimation des surfaces cultivées. Parmi ces méthodes, l'estimateur par régression combine des informations non biaisées mesurées dans un échantillon (au moyen d'observation de terrain) et des informations exhaustives, mais imprécises et biaisées (provenant des images satellitaires classifiées) (Gallego, 2004 et 2006; Gallego *et al.*, 2010 et 2014). Les images satellitaires classifiées servent de variable auxiliaire pour réduire le biais des estimations découlant des observations de terrain. Bien que cette méthode ait été largement utilisée dans d'importants projets opérationnels pour estimer les surfaces cultivées à l'aide de la télédétection (voir par exemple Wall *et al.*, 1984; Chhikara *et al.*, 1986; González-Alonso et Cuevas, 1993), la valeur ajoutée de la télédétection dépend fortement de la précision de la classification et se trouve directement proportionnelle à la qualité du levé de terrain (Gallego *et al.*, 2010).

AAC a conçu sa propre version modifiée de l'estimateur par régression afin de corriger directement le biais dans ses estimations des surfaces cultivées fondées sur l'IAC. Cette méthode — qui n'est pas encore opérationnelle — complète les informations des levés de terrain avec des cartes de cultures locales de très haute précision (24 km x 24 km) tirées des observations de télédétection à haute résolution et des points de levés de terrain de haute qualité. Grâce à leur grande qualité, ces cartes peuvent compenser statistiquement la couverture éparse et discontinue des points des levés de terrain menés dans ces régions et, de fait, être considérées comme l'équivalent des opérations de terrain non biaisés. L'appariement entre les estimations de surfaces tirées de ces cartes locales de haute précision et celles tirées des emplacements correspondant dans l'IAC permet d'identifier des relations mathématiques spécifiques aux cultures pour les estimations de surfaces fondées sur l'échantillonnage (les estimations de surfaces tirées des cartes de télédétection de haute qualité) et le comptage de pixels (les estimations de surface tirées de l'IAC). Ces relations permettent ensuite d'ajuster les estimations de surface basées sur l'IAC de façon à augmenter la précision des estimations pour chaque zone cultivée ainsi que pour l'ensemble de la surface cultivée.

4.4.3.10. Standardisation des légendes de carte

Les classes thématiques utilisées pour la classification des cultures peuvent varier selon les applications et les pays. Cette variabilité crée des incompatibilités qui ralentissent la constitution d'ensembles de données plus larges à l'échelle régionale et mondiale (Giri, 2012). Les classes utilisées pour identifier et cartographier les cultures suivent donc généralement des modèles de classification nationaux ou internationaux qui garantissent une interopérabilité sémantique optimale avec les autres classifications de la couverture du sol et facilitent la comparaison directe entre les classifications (Jensen, 1986; Teleguntla *et al.*, 2016). Ainsi, l'IAC d'AAC utilise une légende harmonisée fondée sur la classification de la couverture du sol effectuée dans le cadre du projet Observation de la Terre pour le développement durable des forêts (OTDD) du Service canadien des forêts (SCF) (Wulder et Nelson, 2001). Ce modèle est compatible avec plusieurs autres légendes reconnues à l'échelle internationale, notamment le système LCCS de la FAO (Di Gregorio et Jansen, 2000).

4.4.3.11. Publication et utilisation des données

Conformément à son engagement en faveur de la science ouverte, le gouvernement canadien assure un accès gratuit et universel aux données de l'IAC, qui sont diffusées sous format d'image matricielle (GeoTIFF) selon la projection conique équivalente d'Albers. Les métadonnées d'AAC — c'est-à-dire, les informations qui accompagnent et décrivent l'ensemble de données — respectent les normes ISO 19131 et sont publiées dans les deux langues officielles du Canada (l'anglais et le français). Jusqu'à présent, les cartes des cultures annuelles ont généralement été publiées huit mois environ après la fin de la période de végétation (c'est-à-dire au mois de mars suivant la fin de la période végétative cartographiée). Depuis 2009, le nombre de téléchargements annuels de l'IAC n'a cessé d'augmenter et l'inventaire est désormais l'un des ensembles de données géospatiales les plus téléchargés du gouvernement canadien.

Une fois publié, l'IAC remplit de nombreux usages. Il est très fréquemment utilisé par l'ensemble du secteur public canadien. Au niveau fédéral, les données de l'IAC permettent d'étayer les applications et les évaluations concernant diverses questions, notamment : a) l'estimation des surfaces cultivées; b) la modélisation des rendements régionaux;

c) les risques de production liés aux conditions climatiques ; d) la modélisation de l'érosion des sols ; e) les émissions et les absorptions de gaz à effet de serre liés aux activités agricoles ; f) les changements dans les populations d'oiseaux ; g) l'établissement de rapports sur les indicateurs de durabilité environnementale ; h) les évolutions urbaines dans les zones métropolitaines du Canada ; et i) l'évaluation de la qualité de l'air. Au niveau des provinces, les données de l'IAC servent notamment pour : a) l'évaluation des tendances agricoles générales ; b) l'évaluation de la biodiversité dans les zones humides ; c) l'élaboration des programmes de gestion agricole ; d) la cartographie de l'utilisation des terres riveraines ; e) la modélisation du transport des nutriments dans les bassins versants ; et f) la planification de l'expansion urbaine et de l'utilisation des terres. D'autres segments de l'économie font aussi amplement usage de cet ensemble de données. Par exemple, les groupes de producteurs utilisent l'IAC pour identifier et mettre en œuvre des stratégies qui répondent aux besoins régionaux spécifiques des cultivateurs, ainsi que pour analyser la conformité aux critères internationaux de certification de durabilité (p. ex., la directive de l'Union européenne sur les sources d'énergie renouvelables en ce qui concerne les changements d'affectation des sols). Les universités et collèges se servent aussi de l'IAC pour appuyer leurs activités de recherche sur la présence de ravageurs et la façon de les combattre, les tendances en matière de rotation des cultures ainsi que la répartition et la protection des pollinisateurs. Enfin, l'agro-industrie utilise l'IAC pour mener des validations indépendantes sur l'utilisation des terres agricoles afin d'enregistrer des crédits de compensation carbone dans ce secteur, pour cartographier les zones inondées par type de cultures afin d'étayer les demandes d'indemnisation adressées aux compagnies d'assurance, pour calibrer les modèles d'estimation des superficies ensemencées immédiatement après la récolte et pour analyser les zones de collecte des silos de grains.

4.4.3.12. Applicabilité des méthodes de l'IAC dans les autres pays

Les méthodologies employées pour créer l'IAC sont déjà utilisées dans l'élaboration de normes internationales sur les produits de données et l'établissement de rapports. Elles guideront aussi la création d'un système mondial de systèmes pour l'évaluation et la surveillance des cultures agricoles dans le cadre de l'expérience JECAM de l'initiative GEOGLAM. Ainsi, et pour faciliter le transfert de méthodes aux pays ayant des systèmes de culture similaires, AAC prépare actuellement une documentation détaillée sur ses méthodes de traitement, qui sera rendue accessible à tous. Bien que les packages et modules des logiciels de traitement d'image et de SIG utilisés dans la chaîne de traitement des données d'AAC soient principalement propriétaires, cette procédure pourrait néanmoins s'exécuter avec d'autres logiciels, y compris des logiciels libres. Ces derniers — dont le code source est rendu accessible en vertu d'une licence qui autorise son détenteur à étudier, modifier et distribuer le logiciel à toute personne et à toute fin — sont particulièrement intéressants, car leur utilisation est généralement gratuite. L'attractivité de l'approche canadienne en matière de cartographie des cultures tient à ses besoins relativement modestes en termes de ressources humaines (cinq employés à temps plein) et de coûts (150 000 dollars canadiens annuels) par rapport à d'autres grands pays (tels que les États-Unis ou la Chine).

4.4.3.13. Perspectives futures

Malgré la performance indéniable de l'IAC, certains de ses aspects doivent encore être améliorés pour répondre aux besoins futurs d'AAC et de ses clients d'une façon économe en coûts et en ressources informatiques. Il serait notamment particulièrement intéressant de parvenir à estimer les surfaces cultivées pendant la période de végétation afin de diffuser ces informations avant la publication de la version finale de l'IAC. La concrétisation d'un tel objectif nécessiterait : a) de négocier un accès anticipé aux observations de terrain des organismes provinciaux utilisées pour l'apprentissage et la validation de l'IAC ; b) d'optimiser les procédures de classification — p. ex., en passant de la méthode de cartographie par pixel à la méthode par objet — afin de réduire les opérations de calcul qui empêchent la réalisation d'estimations pendant la période de végétation ; et c) d'intégrer les flux de données nouveaux et futurs (tels que Sentinel-1, Sentinel-2 et Sentinel-3 et la mission de la constellation RADARSAT — MCR — dont le lancement est prévu en 2018). Par ailleurs, l'extrapolation des méthodes de l'IAC à des systèmes de culture plus complexes — tels que les systèmes de cultures multiples (qui consistent à planter plusieurs types de végétaux sur la même parcelle au cours d'une même année) et les systèmes de cultures associées (qui consistent à planter simultanément plusieurs types de végétaux dans la même parcelle) nécessite des études et des essais supplémentaires.

4.5. CONCLUSION

Ce chapitre a résumé les progrès réalisés récemment dans les applications avancées des technologies de télédétection aux fins de cartographie nationale détaillée (au niveau des parcelles) des cultures.

La création de cartes des cultures précises et détaillées nécessite des données de terrain *et* des données satellitaires multitudes de haute qualité. Les premières servent à l'apprentissage et à la validation de l'approche de classification supervisée utilisée pour classer l'imagerie satellite. Celles qui sont fournies par les organismes publics locaux, régionaux et nationaux sont généralement les plus précises, les plus détaillées et les plus complètes. À défaut, d'autres sources de données sont nécessaires. Un tel besoin peut accroître la difficulté de cartographier les cultures dans les parties du monde où il est impossible d'obtenir des informations officielles spatialement explicites au niveau des parcelles (c'est-à-dire, où les données ne sont pas collectées ou pas utilisables) et, particulièrement, où la collecte de données de terrain au moyen de méthodes alternatives est difficile, voire impossible (p. ex., en raison de ressources limitées [argent, temps, personnel], de problèmes d'accessibilité [mauvaises infrastructures routières et régions agricoles reculées] ou de conflits armés). Néanmoins, quelle que soit leur provenance, les données de terrain doivent être soigneusement évaluées avant de pouvoir servir à la classification. Des données satellitaires de haute qualité peuvent être captées par divers capteurs optiques ou SAR, dont la résolution spatiale doit toutefois être suffisante pour représenter les parcelles individuelles. De plus, les observations doivent être réalisées pendant la période de végétation afin de coïncider avec les grandes étapes de croissance des cultures ciblées. Les cartes nationales détaillées offrant la meilleure précision sur les cultures sont généralement produites au moyen d'une série d'images multitudes à résolution moyenne, spectralement riches et sans trous.

L'étude approfondie des travaux antérieurs ne recommande aucune méthode particulière de classification des images aux fins de cartographie détaillée des cultures, mais montre toutefois que les méthodes fondées sur l'apprentissage automatique sont généralement plus performantes. Parmi ces méthodes, la classification par forêts aléatoires a déjà été intégrée dans des stratégies de cartographie nationale utilisant des données optiques uniquement (États-Unis), des données optiques et SAR (Canada), et des données SAR uniquement (Royaume-Uni). D'autres approches, telles que les réseaux de neurones artificiels et les machines à vecteurs de support, sont aussi prometteuses, mais n'ont encore globalement été utilisées que dans des applications appartenant au domaine de la recherche (non opérationnelle).

L'estimation des surfaces cultivées au moyen de cartes des cultures fondées sur des données satellitaires doit être réalisée avec soin et inclure des méthodes d'évaluation et de correction des biais. En effet, dans le cas du comptage de pixels, les superficiesensemencées sont souvent sous-estimées pour les cultures secondaires et surestimées pour les cultures principales. Ces écarts sont dus à la variabilité de la taille des échantillons et de la répartition géographique des données de terrain utilisées pour la phase d'apprentissage des méthodes de classification. Ils tendent à confirmer l'observation selon laquelle l'estimation des surfaces cultivées d'après les données de télédétection (telle que pratiquée actuellement) ne suffit pas à fournir des estimations non biaisées sur les superficiesensemencées. Le biais peut être réduit au moyen d'outils statistiques tels que l'estimateur par régression, une méthode qui combine des informations non biaisées mesurées dans un échantillon (au moyen d'observation de terrain) et des informations exhaustives, mais imprécises et biaisées (provenant des images satellitaires classifiées).

La cartographie détaillée des cultures présente d'excellentes perspectives d'avenir en raison de l'adoption prochaine de flux améliorés de données satellitaires optiques et SAR qui, conjointement ou isolément, aideront à mieux caractériser les cycles de croissance spécifiques aux cultures au niveau des parcelles. Des difficultés devront cependant être surmontées. Pour parvenir à intégrer ces informations de façon rapide et efficace, les agences de cartographie nationales devront investir massivement dans les infrastructures informatiques nécessaires au traitement de données considérablement plus volumineuses.

4.6. RÉFÉRENCES BIBLIOGRAPHIQUES

Abdulaziz, A.M., Hurtado, J.M. et Al-Douri, R. 2009. «Application of Multitemporal Landsat Data to Monitor Land Cover Changes in the Eastern Nile Delta Region, Egypt». *International Journal of Remote Sensing*, vol. 30, n° 11, p. 2977–96.

Abedi, M., Gholam-Hossain, N. et Fathianpour, N. 2013. «Fuzzy Outranking Approach: A Knowledge-Driven Method for Mineral Prospectivity Mapping». *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, p. 556–67.

Alabi, T., Haertel, M. et Chiejile, S. 2016. «Investigating the Use of High Resolution Multi-spectral Satellite Imagery for Crop Mapping in Nigeria - Crop and Landuse Classification using WorldView-3 High Resolution Multispectral Imagery and LANDSAT8 Data». Actes de la 2^e Conférence internationale sur la théorie, la gestion et les applications des systèmes d'information géographique (GISTAM 2016), (p. 109–120), SCITEPRESS — Science and Technology Publications.

Al-Anazi, A.F. et Gates, I.D. 2010. «Support Vector Regression for Porosity Prediction in a Heterogeneous Reservoir: A Comparative Study». *Computers & Geosciences*, vol. 36, n° 12, p. 1494–1503.

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R. et Weber, J.-L. 2008. «GlobCover: The most detailed portrait of Earth». *European Space Agency Bulletin*, n° 136, p. 24–31.

Atzberger, C., Vuolo, F., Klisch, A., Rembold, F., Meroni, M., Mello, M.P. et Formaggio, A.R. 2016. «Agriculture». In : Thenkabail, P.S. (dir.), *Land Resources Monitoring, Modeling and Mapping with Remote Sensing*, (p. 71–112). CRC Press Inc. : Boca Raton, FL, États-Unis.

Badhwar, G.D. 1984. « Automatic corn-soybean classification using landsat MSS data. I. Near-harvest crop proportion estimation ». *Remote Sensing of Environment*, vol. 14, n° 1-3, n° 15–29.

Baghdadi, N., Boyer, N., Todoroff, P., El Hajj, M. et Bégué, A. 2009. «Potential of SAR sensors TerraSAR-X, ASAR/ENVISAT and PALSAR/ALOS for monitoring sugarcane crops on Reunion Island». *Remote Sensing of Environment*, vol. 113, n° 8, p. 1724–38.

Banque mondiale, Nations Unies (ONU) et Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO). 2010. Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales. Rapport numéro 56719— GLB. Publication de la Banque mondiale : Washington, D.C. Disponible à l'adresse : <http://www.fao.org/docrep/017/am082f/am082f.pdf>. Consulté le 3 juillet 2017.

Bartalev, S., Egorov, C., Loupian, E., Plotnikov, D. et Uvarov, I. 2011. «Recognition of arable lands using multi-annual satellite data from spectroradiometer MODIS and locally adaptive supervised classification». *Computer Optics*, vol. 35, n° 1, p. 103–16.

Bartholomé, E. et Belward, A.S. 2005. «GLC2000: a new approach to global land cover mapping from Earth observation data». *International Journal of Remote Sensing*, vol. 26, n° 9, p. 1959–77.

Becker-Reshef, I., Justice, C., Sullivan, M., Vermote, E., Tucker, C., Anyamba, A., Small, J., Pak, E., Masuoka, E., Schmaltz, J. et al. 2010. « Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project ». *Remote Sensing*, vol. 2, n° 6, p. 1589–1609.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. et Heynen, M. 2004. «Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, n° 3-4, p. 239–58.

Beyer, F., Jarmer, T., Siegmann, B. et Fischer, P. 2015. «Improved crop classification using multitemporal RapidEye data». Article présenté lors du 8^e atelier international sur l'analyse des images issues de la télédétection multitudes (Multi-Temp), 22–24 juillet 2015. Annecy, France, IEEE.

Bhatta, B. 2008. *Remote Sensing and GIS*. Oxford University Press : Oxford, Royaume-Uni.

Blaes, X., Vanhalle, L. et Defourny, P. 2005. «Efficiency of crop identification based on optical and SAR image time series». *Remote Sensing of Environment*, vol. 96, n° 3-4, p. 352–65.

Blanco, P.D., Colditz, R.R., López Saldaña, G., Hardtke, L.A., Llamas, R.M., Mari, N.A., Fischer, A., Caride, C., Aceñolaza, P.G. et del Valle, H.F. 2013. «A land cover map of Latin America and the Caribbean in the framework of the SERENA project». *Remote Sensing of Environment*, vol. 132, p. 13–31.

Blaschke, T. 2010. «Object based image analysis for remote sensing». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, n° 1, p. 2–16.

Bontemps, S., Arino, O., Bicheron, P., Brockman, C.C., Leroy, M., Vancutsem, C. et Defourny, P. 2012. «Operational Service Demonstration for Global Land-Cover Mapping: The GlobCover and GlobCorine Experiences for 2005 and 2009». In : Giri, C.P. (dir.), *Remote Sensing of Land Use and Land Cover: Principles and Applications* (p. 243–265), CRC Press : Boca Raton, FL, États-Unis.

Boryan, C., Yang, Z., Mueller, R. et Craig, M. 2011. «Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program». *Geocarto International*, vol. 26, n° 5, p. 341–58.

Boschetti, L., Flasse, S.P. et Brivio, P.A. 2004. «Analysis of the Conflict between Omission and Commission in Low Spatial Resolution Dichotomic Thematic Products: The Pareto Boundary». *Remote Sensing of Environment*, vol. 91, n° 3–4, p. 280–92.

Breiman, L. 1984. *Classification and Regression Trees*. CRC Press : Boca Raton, FL, États-Unis.

Breiman, L. 2001. «Random Forests». *Machine Learning*, vol. 45, n° 1, p. 5–32.

Brown, J.C., Kastens, J.H., Coutinho, A.C., de Castro Victoria, D. et Bishop, C.R. 2013. «Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data». *Remote Sensing of Environment*, vol. 130, p. 39–50.

Brown de Colstoun, E.C., Story, M.H., Thompson, C., Commisso, K., Smith, T.G. et Irons, J.R. 2003. «National Park Vegetation Mapping Using Multitemporal Landsat 7 Data and a Decision Tree Classifier». *Remote Sensing of Environment*, vol. 85, n° 3, p. 316–27.

Castillejo González, I.L., López Granados, F., García Ferrer, A., Peña Barragán, J.M., Jurado Expósito, M., Sánchez de la Orden, M. et González Audicana, M. 2009. «Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery». *Computers and Electronics in Agriculture*, vol. 68, n° 2, p. 207–15.

Chang, J., Hansen, M.C., Pittman, K., Carroll, M. et DiMiceli, C. 2007. «Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets». *Agronomy Journal*, vol. 99, n° 6, p. 1654-64.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M. et al. 2015. « Global land cover mapping at 30m resolution: A POK-based operational approach ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, p. 7–27.

Chhikara, R., Lundgren, J. et Houston, A. 1986. «Crop Acreage Estimation Using a Landsat-Based Estimator as an Auxiliary Variable». *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, n° 1, p. 157–68.

Clark, R.N., King, T.V., Ager, C. et Swayze, G.A. 1995. *Initial vegetation species and senescence/stress indicator mapping in the San Luis Valley, Colorado using imaging spectrometer data*. USGS Publication : Denver, CO, États-Unis.

Colwell, J.E. 1974. « Vegetation canopy reflectance ». *Remote Sensing of Environment*, vol. 3, n° 3, p. 175–83.

Comaniciu, D. et Meer, P. 2002. « Mean shift: a robust approach toward feature space analysis ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, n° 5, p. 603–19.

Congalton, R.G. 1991. « A review of assessing the accuracy of classifications of remotely sensed data ». *Remote Sensing of Environment*, vol. 37, n° 1, p. 35–46.

Congalton, R.G., Balogh, M., Bell, C., Green, K., Milliken, J.A. et Ottman, R. 1998. «Mapping and Monitoring Agricultural Crops and Other Land Cover in the Lower Colorado River Basin». *Photogrammetric Engineering and Remote Sensing*, vol. 64, n° 11, p. 1107-14.

Conrad, C., Machwitz, M., Schorcht, G., Löw, F., Fritsch, S. et Dech, S. 2011. «Potentials of RapidEye time series for improved classification of crop rotations in heterogeneous agricultural landscapes: experiences from irrigation systems in Central Asia». In : Neale, C.M.U. et Maltese, A. (dir.), *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIII* (p. 817412), SPIE Proceedings Vol. 8174, Publication de la Society of Photo-Optical Instrumentation Engineers : Washington, D.C.

Cortes, C. et Vapnik, V. 1995. «Support-Vector Networks». *Machine Learning*, vol. 20, n° 3, p. 273–97.

Defries, R.S. et Townshend, J.R.G. 2003. «LBA Regional Land Cover from AVHRR, 1-Degree, 1987 (Defries and Townshend)». ORNL DAAC, Oak Ridge, TN, États-Unis.

Defries, R.S., Hansen, M., Sohlberg, R. et Townshend, J.R.G. 2003. «LBA Regional Land Cover from AVHRR, 8-km, 1984» (Defries *et al.*). ORNL DAAC, Oak Ridge, TN, États-Unis.

Denègre, J. 2013. *Cartographie thématique dérivée des images satellitaires : un guide*. Elsevier : Kidlington, Oxford, Royaume-Uni.

Deschamps, B., McNairn, H., Shang, J. et Jiao, X. 2012. «Towards Operational Radar-Only Crop Type Classification: Comparison of a Traditional Decision Tree with a Random Forest Classifier». *Journal canadien de télédétection*, vol. 38, n° 1, p. 60–68.

Di Gregorio, A. et Jansen, L.J.M. 2000. *Land cover classification system: Classification concepts and user manual*. Publication de la FAO : Rome.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P. et al. 2012. « Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services ». *The Sentinel Missions - New Opportunities for Science*, vol. 120, p. 25–36.

Eidenshink, J.C. et Faundeen, J.L. 1994. « The 1 km AVHRR global land data set: first stages in implementation ». *International Journal of Remote Sensing*, vol. 15, n° 17, p. 3443–62.

Eisavi, V., Homayouni, S., Yazdi, A.M. et Alimohammadi, A. 2015. Land cover mapping based on random forest classification of multitemporal spectral and thermal images. *Environmental Monitoring and Assessment*, vol. 187, n° 5, p. 291.

El-Khoury, A., Seidou, O., Lapen, D.R., Sunohara, M., Zhenyang, Q., Mohammadian, M. et Daneshfar, B. 2014. « Prediction of land-use conversions for use in watershed-scale hydrological modeling: a Canadian case study ». *The Canadian Geographer/Le Géographe Canadien*, vol. 58, n° 4, p. 499–516.

El-Khoury, A., Seidou, O., Lapen, D.R., Que, Z., Mohammadian, M., Sunohara, M. et Bahram, D. 2015. « Combined impacts of future climate and land use changes on discharge, nitrogen and phosphorus loads for a Canadian river basin ». *Journal of Environmental Management*, vol. 151, p. 76–86.

Evans, C., Jones, R., Svalbe, I. et Berman, M. 2002. « Segmenting multispectral Landsat TM images into field units ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, n° 5, p. 1054–64.

Fischer, G., Shah, M., N. Tubiello, F. et van Velhuizen, H. 2005. « Socio-economic and climate change impacts on agriculture: an integrated assessment, 1990–2080 ». *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, n° 1463, p. 2067–83.

Fisette, T., Rollin, P., Aly, Z., Campbell, L., Daneshfar, B., Filyer, P., Smith, A., Davidson, A., Shang, J. et Jarvis, I. 2013. « AAFC annual crop inventory ». In : IEEE, *2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* (p. 270–274), Publication de l'IEEE.

Fisette, T., Davidson, A., Daneshfar, B., Rollin, P., Aly, Z. et Campbell, L. 2014. « Annual space-based crop inventory for Canada: 2009–2014 ». In : IEEE, *2014 IEEE International Geoscience and Remote Sensing Symposium* (p. 5095–98), Publication de l'IEEE.

Fisette, T., McNairn, H. et Davidson, A. 2015. *An Operational Annual Space-Based Crop Inventory Based on the Integration of Optical and Microwave Remote Sensing Data: Protocol Document*. Publication d'Agriculture et Agroalimentaire Canada : Ottawa.

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

Friedl, M.A. et Brodley, C.E. 1997. « Decision Tree Classification of Land Cover from Remotely Sensed Data ». *Remote Sensing of Environment*, vol. 61, n° 3, p. 399–409.

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A. et al. 2002. « Global Land Cover Mapping from MODIS: Algorithms and Early Results ». *Remote Sensing of Environment*, vol. 83, n° 1–2, p. 287–302.

Fritz, S., Massart, M., Savin, I., Gallego, J. et Rembold, F. 2008. «The use of MODIS data to derive acreage estimations for larger fields: A case study in the south-western Rostov region of Russia». *International Journal of Applied Earth Observation and Geoinformation*, vol. 10, n° 4, p. 453–66.

Fritz, F., See, L. et McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C. et Perger, C. 2015. «Mapping global cropland and field size». *Global Change Biology*, vol. 21, n° 5, p. 1980–92.

Galford, G.L., Mustard, J.F., Melillo, J., Gendrin, A., Cerri, C.C. et Cerri, C.E.P. 2008. «Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil». *Remote Sensing of Environment Special Issue*, vol. 112, n° 2, p. 576–87.

Gallego, F.J. 2004. « Remote sensing and land cover area estimation ». *International Journal of Remote Sensing*, vol. 25, n° 15, p. 3019–47.

Gallego, F.J. 2006. « Review of the Main Remote Sensing Methods for Crop Area Estimates ». In : *Workshop proceedings: Remote sensing support to crop yield forecast and area estimates*. Archives internationales de photogrammétrie, de télédétection et de sciences de l'information spatiale, XXXVI, 8/W48, 65–70. Disponible à l'adresse : http://www.isprs.org/publications/PDF/ISPRS_Archives_WorkshopStresa2006.pdf. Consulté le 5 mai 2017.

Gallego, F.J., Craig, M., Michaelsen, J. Bossyns, B. et Fritz S. (dir.). 2008. *Best practices for crop area estimation with Remote Sensing*. GEOSS Community of Practice Ag 0703a. Office des publications officielles des Communautés européennes : Luxembourg.

Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A. et Kussul, O. 2014. «Efficiency assessment of using satellite data for crop area estimation in Ukraine». *International Journal of Applied Earth Observation and Geoinformation*, vol. 29, p. 22–30.

Ghimire, B., Rogan, J., Rodríguez Galiano, V., Panday, P. et Neeti, N. 2012. «An Evaluation of Bagging, Boosting, and Random Forests for Land-Cover Classification in Cape Cod, Massachusetts, USA». *GIScience & Remote Sensing*, vol. 49, n° 5, p. 623–43.

Gillespie, T.J. 1990. « Radar detection of a dew event in wheat ». *Remote Sensing of Environment*, vol. 33, n° 3, p. 151–6.

Giri, C.P. 2012. *Remote Sensing of Land Use and Land Cover: Principles and Applications*. CRC Press : Boca Raton, FL, États-Unis.

Gislason, P.O., Benediktsson, J.A. et Sveinsson, J.R. 2006. «Random Forests for Land Cover Classification». *Pattern Recognition Letters*, vol. 27, n° 4, p. 294–300.

Glanz, H., Carvalho, L., Sulla-Menashe, D. et Friedl, M.A. 2014. «A Parametric Model for Classifying Land Cover and Evaluating Training Data Based on Multi-Temporal Remote Sensing Data». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 97, p. 219–28.

Gómez, C., White, J.C. et Wulder, M.A. 2016. «Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, p. 55–72.

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S. et al. 2013a. « Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data ». *International Journal of Remote Sensing*, vol. 34, n° 7, p. 2607–54.

González-Alonso, F. et Cuevas, J.M. 1993. « Remote sensing and agricultural statistics: crop area estimation through regression estimators and confusion matrices ». *International Journal of Remote Sensing*, vol. 14, n° 6, p. 1215–19.

Gouvernement canadien. Statistique Canada. 2008. *Mandat et objectifs*. Disponible à l'adresse : <http://www.statcan.gc.ca/fra/aperçu/mandat>. Consulté le 10 juin 2017.

Guerschman, J.P., Paruelo, J.M., Bella, C.D., Giallorenzi, M.C. et Pacin, F. 2003. « Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data ». *International Journal of Remote Sensing*, vol. 24, n° 17, p. 3381–402.

Haldar, D., Das, A., Mohan, S., Pal, O., Hooda, R.S. et Chakraborty, M. 2012. « Assessment of L-Band SAR Data at Different Polarization Combinations for Crop and Other Landuse Classification », *Progress in Electromagnetics Research B*, vol. 36, p. 303–21.

Hansen, M.C., DeFries, R.S., Townshend, J.R. et Sohlberg, R. 2000. « Global land cover classification at 1 km spatial resolution using a classification tree approach ». *International Journal of Remote Sensing*, vol. 21, n° 6-7, p. 1331–64.

Haralick, R.M. et Shapiro, L.G. 1985. « Image Segmentation Techniques ». In : Gilmore, J.F. (dir.), *SPIE Proceedings Vol. 0548: Applications of Artificial Intelligence II* (p. 2–9). Publication de la SPIE : Bellingham, WA, États-Unis.

Hill, M.J., Ticehurst, C.J., Lee, J.-S., Grunes, M.R., Donald, G.E. et Henry, D. 2005. « Integration of optical and radar classifications for mapping pasture type in Western Australia ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, n° 7, p. 1665–81.

Hoekman, D.H., Vissers, M.A.M. et Tran, T.N. 2011. « Unsupervised Full-Polarimetric SAR Data Segmentation as a Tool for Classification of Agricultural Areas ». *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, n° 2, p. 402–11.

Holecz, F., Collivignarelli, F., Barbieri, M., Gatti, L., Boschetti, M., Manfron, G., Brivio, P.A., Abukari, M. et Bondo, T. 2013. « Establishing National Baseline Land Cover Map Including Annual and Seasonal Variations for the Understanding of Current Agricultural Practices in the Gambia ». Document du FIDA. Non publié. Disponible à l'adresse : <https://operations.ifad.org/documents/654016/fd0979c5-e1f2-483c-8247-922d4ac27c2d>. Consulté le 3 juillet 2017.

Hütt, C., Koppe, W., Miao, Y. et Bareth, G. 2016. « Best Accuracy Land Use/Land Cover (LULC) Classification to Derive Crop Types Using Multitemporal, Multisensor, and Multi-Polarization SAR Satellite Images ». *Remote Sensing*, vol. 8, n° 8, p. 684.

Irons, J.R., Dwyer, J.L. et Barsi, J.A. 2012. « The next Landsat satellite: The Landsat Data Continuity Mission ». *Remote Sensing of Environment*, vol. 122, p. 11–21.

Jakubauskas, M.E., Legates, D.R. et Kastens, J.H. 2002. «Crop identification using harmonic analysis of time-series AVHRR NDVI data». *Computers and Electronics in Agriculture*, vol. 37, n° 1-3, p. 127–39.

Jensen, J.R. 1986. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Pearson Education, Inc. : Glenview, IL, États-Unis.

Jensen, F.V. et Nielsen, T.D. 2007. *Bayesian Networks and Decision Graphs*. Springer : New York, NY, États-Unis.

Jewell, N. 1989. « An evaluation of multi-date SPOT data for agriculture and land use mapping in the United Kingdom ». *International Journal of Remote Sensing*, vol. 10, n° 6, p. 939–51.

Jia, K., Li, Q., Tian, Y., Wu, B., Zhang, F. et Meng, J. 2012. « Crop classification using multi-configuration SAR data in the North China Plain ». *International Journal of Remote Sensing*, vol. 33, n° 1, p. 170–83.

Jiao, X., Kovacs, J.M., Shang, J., McNairn, H., Walters, D., Ma, B. et Geng, X. 2014. « Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 96, p. 38–46.

Johnson, D.M. 2013. «A 2010 map estimate of annually tilled cropland within the coterminous United States». *Agricultural Systems*, vol. 114, p. 95–105.

Jones, H.G. et Vaughan, R.A. 2010. *Remote Sensing of Vegetation: Principles, Techniques, and Applications*. Oxford University Press : Oxford, Royaume-Uni — New York, États-Unis.

Kamusoko, C. et Aniya, M. 2009. « Hybrid Classification of Landsat Data and GIS for Land Use/Cover Change Analysis of the Bindura District, Zimbabwe ». *International Journal of Remote Sensing*, vol. 30, n° 1, p. 97–115.

Kim, H.-O. et Yeom, J.-M. 2014. «Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data». *International Journal of Remote Sensing*, vol. 35, n° 19, p. 1–23.

Kumar, L., Priyakant, S., Brown, J.F., Ramsey, R.D., Rigge, M., Stam, C.A., Hernandez, A.J., Hunt Jr., E.R. et Reeves, M.C. 2016. «Characterization, Mapping and Monitoring of Rangelands: Methods and Approaches». In : Thenkabail, P.S. (dir.), *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing* (p. 309–350). CRC Press Inc. : Boca Raton, FL, États-Unis.

Kussul, N., Lemoine, G., Gallego, F.J., Skakun, S.V., Lavreniuk, M. et Shelestov, A.Y. 2016. «Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data». *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, n° 6, p. 2500–08.

Laba, M., Smith, S.D. et Degloria, S.D. 1997. «Landsat-Based Land Cover Mapping in the Lower Yuna River Watershed in the Dominican Republic». *International Journal of Remote Sensing*, vol. 18, n° 14, p. 3011–25.

Larrañaga, A., Álvarez-Mozos, J. et Albizua, L. 2011. «Crop classification in rain-fed and irrigated agricultural areas using Landsat TM and ALOS/PALSAR data». *Canadian Journal of Remote Sensing*, vol. 37, n° 1, p. 157–170.

Latham, J., Cumani, R., Rosati, I. et Bloise, M. 2014. *Global Land Cover SHARE (GLC-SHARE) Database Beta-Release Version 1.0-2014*. Disponible à l'adresse : <http://www.fao.org/uploads/media/glc-share-doc.pdf>. Consulté le 3 juillet 2017.

Latifovic, R., Zhu, Z.-L., Cihlar, J. et Olthof, I. 2004. «Land cover mapping of North and Central America - Global Land Cover 2000». *Remote Sensing of Environment*, vol. 89, n° 1, p. 116–27.

Lavreniuk, M., Kussul, N., Skakun, S., Shelestov, A. et Yailymov, B. 2015. «Regional retrospective high resolution land cover for Ukraine: Methodology and results». In : *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (p. 3965–8). Publication de l'IEEE.

Li, Z., Huffman, T., McConkey, B. et Townley-Smith, L. 2013. «Monitoring and modeling spatial and temporal patterns of grassland dynamics using time-series MODIS NDVI with climate and stocking data». *Remote Sensing of Environment*, vol. 138, p. 232–44.

Liu, J., Liu, M., Tian, H., Zhuang, D., Zhang, Z., Zhang, W., Tang, X. et Deng, X. 2005. «Spatial and temporal patterns of China's cropland during 1990–2000: an analysis based on Landsat TM data». *Remote Sensing of Environment*, vol. 98, n° 4, p. 442–456.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. et Merchant, J.W. 2000. «Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data». *International Journal of Remote Sensing*, vol. 21, n° 6-7, p. 1303–30.

Lu, D. et Weng, Q. 2007. «A Survey of Image Classification Methods and Techniques for Improving Classification Performance». *International Journal of Remote Sensing*, vol. 28, n° 5, p. 823–70.

Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G. et Worthy, L.D. 2006. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sensing of Environment*, 105, n° 2, p. 142–54.

Lunetta, R.S., Shao, Y., Ediriwickrema, J. et Lyon, J.G. 2010. «Monitoring agricultural cropping patterns across the Laurentian Great Lakes Basin using MODIS-NDVI data». *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, n° 2, p. 81–88.

Lussem, U., Hütt, C. et Waldhoff, G. 2016. «Combined Analysis of Sentinel-1 and Rapideye Data for Improved Crop Type Classification: An Early Season Approach for Rapeseed and Cereals». In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B8 (p. 959–63). Publication de l'ISPRS.

Lymburner, L., Geoscience Australia et Australian Bureau of Agricultural and Resource Economics. (2011). *The National Dynamic Land Cover Dataset*. Publication de Geoscience Australia : Canberra.

Martínez Beltran, C. et Calera Belmonte, A. 2001. «Irrigated Crop Area Estimation Using Landsat TM Imagery in La Mancha, Spain». *Photogrammetric Engineering & Remote Sensing*, vol. 67, n° 10, p. 1177–84.

Maselli, F., Conese, C., De Filippis, T. et Romani, M. 1995. «Integration of ancillary data into a maximum-likelihood classifier with nonparametric priors». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 50, n° 2, p. 2–11.

McNairn, H., Ellis, J., Van Der Sanden, J.J., Hirose, T. et Brown, R.J. 2002. «Providing crop information using RADARSAT-1 and satellite optical imagery». *International Journal of Remote Sensing*, vol. 23, n° 5, p. 851–70.

McNairn, H., Champagne, C., Shang, J., Holmstrom, D. et Reichert, G. 2009. «Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, n° 5, p. 434–49.

National Research Council of the National Academies. 2003. *Satellite Observations of the Earth's Environment: Accelerating the Transition of Research to Operations*. National Academies Press : Washington, D.C.

Pal, M. et Mather, P.M. 2003. «An Assessment of the Effectiveness of Decision Tree Methods for Land Cover Classification». *Remote Sensing of Environment*, vol. 86, n° 4, p. 554–65.

Pan, Y., Li, L., Zhang, J., Liang, S., Zhu, X. et Sulla-Menashe, D. 2012. «Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index». *Remote Sensing of Environment*, vol. 119, p. 232–42.

Pei, Z., Zhang, S., Guo, L., McNairn, H., Shang, J. et Jiao, X. 2011. «Rice identification and change detection using TerraSAR-X data». *Journal canadien de télédétection*, vol. 37, n° 1, p. 151–6.

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E. et Six, J. 2011. «Object-based crop identification using multiple vegetation indices, textural features and crop phenology». *Remote Sensing of Environment*, vol. 115, n° 6, p. 1301–16.

Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V. et Justice, C.O. 2010. «Estimating Global Cropland Extent with Multi-year MODIS Data». *Remote Sensing*, vol. 2, p. 1844–63.

Pringle, M.J., Denham, R.J. et Devadas, R. 2012. «Identification of cropping activity in central and southern Queensland, Australia, with the aid of MODIS MOD13Q1 imagery». *International Journal of Applied Earth Observations and Geoinformation*, vol. 19, p. 276–85.

Reese, H.M., Lillesand, T.M., Nagel, D.E., Stewart, J.S., Goldmann, R.A., Simmons, T.E., Chipman, J.W. et Tessar, P.A. 2002. «Statewide land cover derived from multiseasonal Landsat TM data». *Remote Sensing of Environment*, vol. 82, n° 2-3, p. 224–37.

Rigol-Sanchez, J.P., Chica-Olmo, M. et Abarca-Hernandez, F. 2003. «Artificial Neural Networks as a Tool for Mineral Potential Mapping with GIS». *International Journal of Remote Sensing*, vol. 24, n° 5, p. 1151–56.

Rocha, J. et Queluz, M.P. 2002. «Integration of census data, remote sensing, and GIS techniques for land-use and cover classification». In : Ehlers, M. (dir.), *SPIE Proceedings: Remote Sensing for Environmental Monitoring, GIS Applications, and Geology* (p. 73–83), Conference Volume 4545. Publication de la SPIE.

Rodriguez-Galiano, V. et Chica-Olmo, M. 2012. «Land Cover Change Analysis of a Mediterranean Area in Spain Using Different Sources of Data: Multi-Seasonal Landsat Images, Land Surface Temperature, Digital Terrain Models and Texture». *Applied Geography*, vol. 35, n° 1–2, p. 208–18.

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C. et Roberts, D. 2008. «Mapping Land-Cover Modifications over Large Areas: A Comparison of Machine Learning Algorithms». *Remote Sensing of Environment*, vol. 112, n° 5, p. 2272–83.

Rumelhart, D.E., Hinton, G.E. et Williams, R.J. 1986. «Learning Representations by Back-Propagating Errors». *Nature*, vol. 323, n° 6088, p. 533–36.

Sakamoto, T., Wardlow, B.D., Gitelson, A.A., Verma, S.B., Suyker, A.E. et Arkebauer, T.J. 2010. «A Two-Step Filtering approach for detecting maize and soybean phenology with time-series MODIS data». *Remote Sensing of Environment*, vol. 114, n° 20, p. 2146–59.

Schuster, C., Schmidt, T., Conrad, C., Kleinschmit, B. et Förster, M. 2015. «Grassland habitat mapping by intra-annual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data». *International Journal of Applied Earth Observations and Geoinformation*, vol. 34, n° 25–34.

Shang, J., McNairn, H., Champagne, C. et Jiao, X. 2009. «Application of Multi-Frequency Synthetic Aperture Radar (SAR) in Crop Classification». In : Jedlovec, G. (dir.), *Advances in Geoscience and Remote Sensing* (p. 557–568), InTech : Rijeka, Croatie.

Skriver, H. 2012. «Crop Classification by Multitemporal C- and L-Band Single- and Dual-Polarization and Fully Polarimetric SAR». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, n° 6, p. 2138–49.

Skriver, H., Svendsen, M.T. et Thomsen, A.G. 1999. «Multitemporal C-and L-band polarimetric signatures of crops». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, n° 5, p. 2413–29.

Sonobe, R., Tani, H., Wang, X., Kobayashi, N. et Shimamura, H. 2014. «Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data». *Remote Sensing Letters*, vol. 5, n° 2, p. 157–64.

Sreenivas, K., Sekhar, N.S., Saxena, M., Paliwal, R., Pathak, S., Porwal, M.C., Fyze, M.A., Rao, S.K., Wadodkar, M. et Anasuya, T. 2015. «Estimating inter-annual diversity of seasonal agricultural area using multi-temporal resourcesat data». *Journal of Environmental Management*, vol. 161, p. 433–42.

Takahashi, M., Nasahara, K.N., Tadono, T., Watanabe, T., Dotsu, M., Sugimura, T. et Tomiyama, N. 2013. *High-Resolution Land Use and Land Cover Map of Japan*. Agence d’exploration aérospatiale japonaise, p. 2384–7.

Tapsall, B., Milenov, P. et Tasdemir, K. 2010. «Analysis of RapidEye imagery for annual landcover mapping as an aid to European Union (EU) common agricultural policy». In : Wagner, W. et Székely, B. (dir.), *Proceedings of the ISPRS TC VII Symposium – 100 Years ISPRS*, (p. 568–73), Publication de l’ISPRS.

Tateishi, R., Uriyangqai, B., Al-Bilbisi, H., Ghar, M.A., Tsend-Ayush, J., Kobayashi, T., Kasimu, A., Hoan, N.T., Shalaby, A., Alsaadeh, B. et al. 2011b. «Production of global land cover data – GLCNMO». *International Journal of Digital Earth*, vol. 4, n° 1, p. 22–49.

Tatsumi, K., Yamashiki, Y., Canales Torres, M.A. et Taïpe, C.L.R. 2015. «Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data». *Computers and Electronics in Agriculture*, vol. 115, p. 171–9.

Teleguntla, P., Thenkabail, P.S., Xiong, J., Gumma, M.K., Giri, C., Milesi, C., Ozdogan, M., Congalton, R.G. et Tilton, J. 2016. «Global food security support analysis data at nominal 1km (GFSAD1km) derived from remote sensing in support of food security in the twenty-first century: current achievements and future possibilities». In : Thenkabail, P.S. (dir.), *Land Resources Monitoring, Modeling and Mapping with Remote Sensing*, (p. 131–159), CRC Press : Boca Raton, FL, États-Unis.

Thenkabail, P.S. et Wu, Z. 2012. «An automated cropland classification algorithm (ACCA) for Tajikistan by combining Landsat, MODIS, and secondary data». *Remote Sensing*, vol. 4, n° 10, p. 2890–18.

Thenkabail, P.S., Biradar, C.M., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O.R.P., Turral, H., Cai, X. et al. 2009. « Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium ». *International Journal of Remote Sensing*, vol. 30, n° 14, p. 3679–733.

Toutin, T. 2005. Multisource Data Fusion with an Integrated and Unified Geometric Modelling. *EARSeL Advances in Remote Sensing*, vol. 4, p. 118–29.

Turker, M. et Arikan, M. 2005. «Sequential masking classification of multi-temporal Landsat7 ETM+ images for field-based crop mapping in Karacabey, Turkey». *International Journal of Remote Sensing*, vol. 26, n° 17, p. 3813–30.

Turker, M. et Ozdarici, A. 2011. «Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: a comparison study». *International Journal of Remote Sensing*, vol. 32, n° 24, p. 9735–68.

Ustuner, M., Sanli, F.B., Abdikan, S., Esetlili, M.T. et Kurucu, Y. 2014. «Crop Type Classification Using Vegetation Indices of RapidEye Imagery». In : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-7 (p. 195–198). Publication de l'ISPRS.

Valero, S., Morin, D., Inglada, J., Sepulcre Canto, G., Arias, M., Hagolle, O., Dedieu, G., Bontemps, S., Defourny, P. et Koetz, B. 2016. «Production of a Dynamic Cropland Mask by Processing Remote Sensing Image Series at High Temporal and Spatial Resolutions». *Remote Sensing*, vol. 8, n° 55, p. 1–21.

Valverde-Albacete, F.J. et Peláez-Moreno, C. 2010. «Two Information-Theoretic Tools to Assess the Performance of Multi-Class Classifiers». *Pattern Recognition Letters*, vol. 31, n° 12, p. 1665–71.

Valverde-Albacete, F.J. et Peláez-Moreno, C. 2014. «100% Classification Accuracy Considered Harmful: The Normalized Information Transfer Factor Explains the Accuracy Paradox». *PLoS ONE*, vol. 9, n° 1 : e84217. Disponible à l'adresse : <https://doi:10.1371/journal.pone.0084217>. Consulté le 10 juin 2017.

Vancutsem, C., Marinho, E., Kayitakire, F., See, L. et Fritz, S. 2012. «Harmonizing and Combining Existing Land Cover/Land Use Datasets for Cropland Area Monitoring at the African Continental Scale». *Remote Sensing*, vol. 5, n° 1, p. 19–41.

Verhegghen, A., Mayaux, P., de Wasseige, C. et Defourny, P. 2012. «Mapping Congo Basin vegetation types from 300 m and 1 km multi-sensor time series for carbon stocks and forest areas estimation». *Biogeosciences*, vol. 9, p. 5061–79.

Vieira, M.A., Formaggio, A.R., Rennó, C.D., Atzberger, C., Aguiar, D.A. et Mello, M.P. 2012. «Object Based Image Analysis and Data Mining applied to a remotely sensed Landsat time-series to map sugarcane over large areas». *Remote Sensing of Environment*, vol. 123, p. 553–62.

Vincenzi, S., Zucchetta, M., Franzoi, P., Pellizzato, M. Pranovi, F., De Leo, G.A. et Torricelli, P. 2011. « Application of a Random Forest Algorithm to Predict Spatial Distribution of the Potential Yield of Ruditapes Philippinarum in the Venice Lagoon, Italy ». *Ecological Modelling*, vol. 222, n° 8, p. 1471–78.

Waldner, F., Fritz, S., Di Gregorio, A. et Defourny, P. 2015. « Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps ». *Remote Sensing*, vol. 7, n° 6, p. 7959–86.

Waldner, F., De Abelleira, D., Veron, S.R., Zhang, M., Wu, B., Plotnikov, D., Bartalev, S., Lavreniuk, M., Skakun, S., Kussul, N. et al. 2016. « Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity ». *International Journal of Remote Sensing*, vol. 37, n° 14, p. 3196–231.

Wall, S.L., Thomas, R.W., Brown, C.E. et Bauer, E.H. 1984. « Landsat-based inventory system for agriculture in California ». *Remote Sensing of Environment*, vol. 14, n° 1-3, p. 267–78.

Wardlow, B.D. et 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*, vol. 112, n° 3, p. 1096–116.

Wardlow, B., Egbert, S. et Kastens, J. 2007. « Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains ». *Remote Sensing of Environment*, vol. 108, n° 3, p. 290–310.

Waske, B. et Braun, M. 2009. « Classifier ensembles for land cover mapping using multitemporal SAR imagery ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, n° 5, p. 450–7.

Watts, J.D., Lawrence, R.L., Miller, P.R. et Montagne, C. 2009. « Monitoring of cropland practices for carbon sequestration purposes in north central Montana by Landsat remote sensing ». *Remote Sensing of Environment*, vol. 113, n° 9, p. 1843–52.

Whitcraft, A.K., Becker-Reshef, I. et Justice, C.O. 2015a. « Agricultural growing season calendars derived from MODIS surface reflectance ». *International Journal of Digital Earth*, vol. 8, n° 3, p. 173–97.

Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I. et Justice, C.O. 2015b. « Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations ». *Remote Sensing of Environment*, vol. 156, p. 438–47.

Whitcraft, A.K., Becker-Reshef, I. et Justice, C. 2015c. « A Framework for Defining Spatially Explicit Earth Observation Requirements for a Global Agricultural Monitoring Initiative (GEOGLAM) ». *Remote Sensing*, vol. 7, p. 1461–81.

Wood, D., McNairn, H., Brown, R. et Dixon, R. 2002. « The effect of dew on the use of RADARSAT-1 for crop monitoring ». *Remote Sensing of Environment*, vol. 80, n° 2, p. 241–7.

Wulder, M.A. et Nelson, T. 2001. *EOSD Land Cover Classification Legend Report*. Publication du Service canadien des forêts : Victoria, BC, Canada.

Wulder, M.A., Hilker, T., White, J.C., Coops, N.C., Masek, J.G., Pflugmacher, D. et Crevier, Y. 2015. « Virtual constellations for global terrestrial monitoring ». *Remote Sensing of Environment*, vol. 170, p. 62–76.

Xiuwan, C. 2002. « Using Remote Sensing and GIS to Analyse Land Cover Change and Its Impacts on Regional Sustainable Development ». *International Journal of Remote Sensing*, vol. 23, n° 1, p. 107–24.

Xu, X., Doktor, D. et Conrad, C. 2016. « Phenological Metrics Extraction for Agricultural Land-use Types Using RapidEye and MODIS ». *Geophysical Research Abstracts*, vol. 18 : 10375.

Yan, L. et Roy, D.P. 2014. «Automated crop field extraction from multi-temporal Web Enabled Landsat Data». *Remote Sensing of Environment*, vol. 144, p. 42–64.

Yang, C., Everitt, J.H., Fletcher, R.S. et Murden, D. 2007. «Using high resolution QuickBird imagery for crop identification and area estimation». *Geocarto International*, vol. 22, n° 3, p. 219–33.

Zhu, Z. et Woodcock, C.E. 2012. «Object-based cloud and cloud shadow detection in Landsat imagery». *Remote Sensing of Environment*, vol. 118, p. 83–94.

Zuo, R. et Carranza, E.J.M. 2011. «Support Vector Machine: A Tool for Mapping Mineral Prospectivity». *Computers & Geosciences*, vol. 37, n° 12, p. 1967–75.



5

Chapitre 5

Estimation des surfaces cultivées à partir des données de télédétection

Shibendu S. Ray et Neetu

5.1. INTRODUCTION SUR L'ESTIMATION DES SURFACES CULTIVÉES

Les informations sur la production végétale – qui jouent un rôle essentiel pour diverses activités de planification économique et de gestion des marchés agricoles (Gallego *et al.*, 2014) – peuvent concerner la surface cultivée et le rendement agricole. Si l'on considère toujours que l'estimation des surfaces cultivées est plus facile et plus rapide que celle du rendement agricole, cet exercice présente toutefois des difficultés et une complexité qui ne sont pas nécessairement visibles au premier abord (Craig et Atkinson, 2013).

Parmi les facteurs qui compliquent l'estimation des surfaces cultivées, on peut notamment citer les parcelles de taille réduite, les systèmes de culture dispersés et diversifiés, les différences phénologiques dans les systèmes de culture mixte, les périodes d'ensemencement prolongées (p. ex., en Inde, les semences du riz durent de juin à septembre), l'évolution des systèmes de cultures, les cultures à cycle court, les cultures sur les sites habités, la complexité physiographique (p. ex., cultures en terrasses ou en courbes de niveau dans les zones d'altitude) et la complexité saisonnière (p. ex., repousse de la canne à sucre après la récolte de la canne vierge, ou cultures présentant différentes saisons de végétation en fonction des conditions climatiques). L'évolution des surfaces ensemencées sous l'effet des dommages causés par des facteurs biotiques (parasites et maladies) et abiotiques (climat) doit aussi être prise en considération. De surcroît, l'estimation des surfaces cultivées doit se faire à différentes étapes : avant l'ensemencement (particulièrement le jour des premiers semis, qui varie selon les profits générés par la culture l'année précédente et selon les prévisions météorologiques), au début de l'ensemencement, à la mi-saison et avant la récolte (Vogel et Bange, 1999).

Craig et Atkinson (2013) ont recensé les méthodes d'estimation des surfaces cultivées. Le plus souvent, l'estimation se fait sur la base d'échantillons ou d'un dénombrement exhaustif de toutes les exploitations agricoles. L'échantillonnage peut s'appuyer sur une base aréolaire, sur une base de type liste – où seront recensées les exploitations agricoles – ou sur une prise en compte des deux, auquel cas on parlera d'échantillonnage à bases

multiples. L'échantillonnage est aussi parfois fondé sur l'opinion de spécialistes des cultures agissant à titre volontaire. Les enquêtes administratives, les unités de transformation des cultures (p. ex., les usines de coton ou de jute) et les marchés peuvent aussi fournir des informations sur les surfaces cultivées. L'estimation finale peut s'appuyer directement sur les données de l'enquête ou être produite par un panel d'experts qui vérifie les données des différentes sources et valide les estimations.

La méthode conventionnelle d'estimation des surfaces cultivées est toutefois longue, coûteuse, fastidieuse et sujette aux erreurs humaines. Elle est aussi extrêmement difficile à appliquer sur certains types de terrain, notamment les zones d'altitude. Face à ces problèmes, la télédétection par satellite offre une solution pour estimer les surfaces cultivées, soit directement soit indirectement en complément aux systèmes d'échantillonnage aréolaire. En plus de fournir des images temporelles, synoptiques, multibandes et multirésolution sur l'utilisation des terres et la couverture du sol, elle s'avère aussi utile pour classifier les différentes cultures.

C'est au début des années 1970 que les données de télédétection satellitaires ont été utilisées pour la première fois aux fins d'estimation des surfaces cultivées quand, en 1971, l'USDA, la NASA et plusieurs universités ont mis en œuvre un programme conjoint de surveillance de l'helminthosporiose du maïs (*Corn Blight Watch Experiment*) (Sharples, 1973). En 1972, suite au lancement réussi du satellite ERTS-A, la NASA et l'USDA ont mené des expériences communes pour déterminer si les technologies spatiales de télédétection multibandes pouvaient permettre d'étudier les principaux types de cultures (Bryan, 1974). Des programmes tels que CITARS (*Crop Identification Technology Assessment for Remote Sensing*) – mis en œuvre pour évaluer les procédures de classification et les techniques d'analyse alternatives des cultures du maïs et du soja – et LACIE (*Large Area Crop Inventory Experiment*) ont été mis en œuvre pour démontrer le potentiel de la télédétection aux fins de recensement des cultures (MacDonald, 1984).

Le programme LACIE est le premier à avoir été financé par le gouvernement américain pour déterminer s'il était possible d'exploiter les données satellitaires de télédétection – particulièrement les données Landsat – pour estimer la production de blé sur de vastes étendues géographiques (Nellis *et al.*, 2009). Tout d'abord mis en œuvre en 1974 dans les Grandes Plaines des États-Unis d'Amérique, il a ensuite été élargi au Canada et à l'ancienne Union soviétique (MacDonald, 1984). Ses résultats fructueux ont justifié la création en 1980 d'un second projet intitulé AgRISTARS (*Agricultural and Resources Inventory Surveys Through Aerospace Remote Sensing*). Ce nouveau programme s'inscrivant dans la continuité de LACIE devait permettre de surveiller d'autres cultures telles que l'orge, le maïs, le coton, le riz, le soja et le blé (Holmes *et al.*, 1979). Il a permis de démontrer l'intérêt de la disponibilité rapide des données et de la diminution de la quantité d'informations de terrain nécessaires pour identifier les cultures et prévoir les rendements.

Le programme LACIE a servi de modèle à d'autres programmes mondiaux mis en place pour mesurer des communautés végétales terrestres différentes à l'aide de la télédétection satellitaire (Erickson, 1984 ; MacDonald et Hall, 1980). Établi conjointement par la NASA, la NOAA et l'USDA, ce programme d'évaluation agricole en trois phases a été le premier à démontrer les usages qui pouvaient être faits des données Landsat. La phase I (1974-1975) a consisté à tester une méthode dans la région des Grandes Plaines (l'estimation des surfaces cultivées a été réalisée en mode quasi opérationnel tandis que le rendement et la production ont été estimés dans des conditions d'étude de faisabilité) ; la phase II (1975-1976) a évalué la méthode dans la région des Grandes Plaines, au Canada et dans des régions « indicatives » de l'ancienne URSS (le mode quasi opérationnel d'estimation des surfaces de culture du blé a été étendu à l'estimation du rendement et de la production) ; la phase III (1976-1977) s'est appuyée sur une technologie de deuxième génération élaborée au cours des phases I et II pour réaliser des prévisions sur la culture du blé à l'échelle nationale de l'URSS pour l'année 1977. Le projet a aussi mené des études exploratoires en Inde, en Chine, en Australie, en Argentine et au Brésil (MacDonald et Hall, 1980). Dans ces pays, les surfaces ont été estimées à l'aide d'une sélection de segments échantillonnés à partir de données Landsat, et le rendement au moyen de modèles météorologiques utilisant des données de l'Organisation météorologique mondiale (OMM).

L'objectif de performance de 90/90 fixé dans le cadre du programme LACIE signifiait que dans 90 % des cas, la marge d'erreur ne devait pas dépasser 10 %. C'est dans l'ex-URSS que le programme LACIE a obtenu les meilleurs résultats. Les critères d'exactitude de 90/90 ont aussi été satisfaits dans les Grandes Plaines. En revanche, bien qu'encourageants, les résultats pour le Canada, l'Inde, la Chine, l'Australie, le Brésil et l'Argentine n'ont pas atteint l'objectif d'exactitude de 90/90 (MacDonald et Hall, 1980).

Hanuschak *et al.* (1982) ont décrit l'efficacité avec laquelle le service statistique de l'USDA (SRS, de l'anglais *Statistical Reporting Service*) a utilisé les données Landsat entre 1972 et 1982 pour améliorer i) la base de sondage aréolaire (BSA) et ii) l'estimation par régression des surfaces cultivées.

Suite à cela, de nombreux programmes d'élaboration et de démonstration de méthodes ont été mis en œuvre dans plusieurs pays pour étudier comment exploiter les données de télédétection satellitaires aux fins d'estimation des surfaces cultivées (Dadhwal *et al.*, 2002).

De nombreux pays ont actuellement recours aux données de télédétection satellitaires pour différents aspects de l'estimation des surfaces cultivées. Le tableau 1 contient un résumé des données satellitaires utilisées par plusieurs programmes nationaux d'estimation des surfaces cultivées. Les différents types de données satellitaires utilisées par les pays ainsi que les méthodes adoptées par ceux-ci sont présentés par la suite. Afin de permettre au lecteur de mieux comprendre les différents aspects de l'exploitation des données satellitaires, la sixième section de ce chapitre présente deux programmes nationaux (le programme CDL du NASS de l'USDA et le programme FASAL de l'Inde), un programme régional (le programme MARS du JRC de l'Union européenne) et deux programmes mondiaux (le FAS de l'USDA aux États-Unis et le programme CropWatch en Chine).

5.2. MÉTHODES D'ESTIMATION DES SURFACES CULTIVÉES PAR TÉLÉDÉTECTION

Le recours à la télédétection aux fins d'identification et de classification des cultures se justifie principalement par la différence d'apparence des cultures dans les données multibandes (c'est-à-dire par le principe de signature spectrale). Cette variation s'explique par les caractéristiques structurelles, physiologiques et phénologiques des plantes ainsi que par la diversité des pratiques culturales. L'identification des cultures est rendue possible par l'association de ces données de télédétection et de certaines informations de terrain – la « réalité de terrain ». Ce même concept est utilisé dans les quatre grandes méthodes d'estimation des surfaces cultivées basées sur les données de télédétection : i) la conception de BSA ; ii) l'estimation directe ou le comptage de pixels ; iii) l'estimateur par régression et iv) l'estimateur par calage.

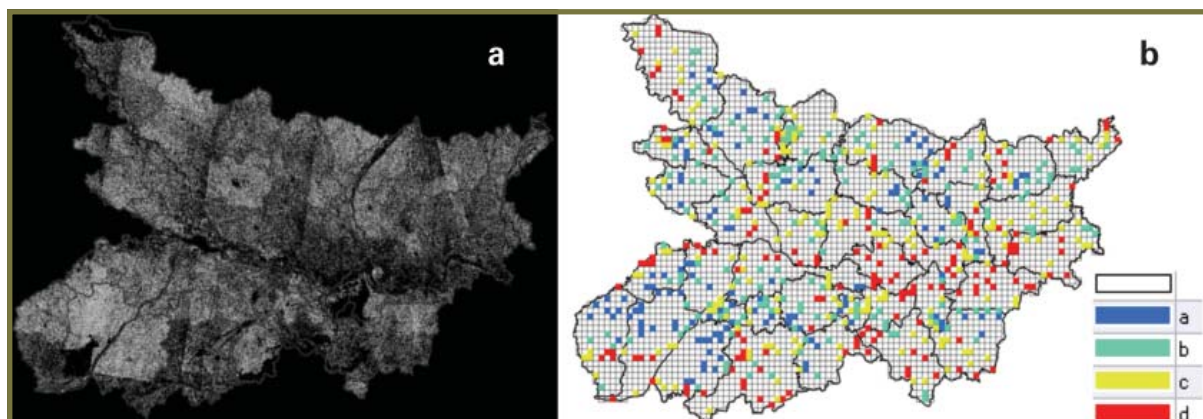
5.2.1. Conception de Bases de Sondage Aréolaire

Le chapitre 3 du présent manuel décrit le rôle que la télédétection peut jouer dans la conception des bases de sondage. L'imagerie satellite est ici d'une importance capitale, puisqu'elle fournit un ensemble de références pour l'élaboration de la base de population. Elle est particulièrement utile pour subdiviser les zones d'intérêt en districts de recensement (DR), dont les exploitations agricoles sont listées, et pour définir les bases aréolaires comprenant des unités primaires et secondaires d'échantillonnage (UPE et USE) aisément identifiables dans l'imagerie numérique (Delincé, 2015). La proportion de cultures identifiée par interprétation visuelle ou classification numérique des données de télédétection permet de décrire la variabilité spatiale et constitue également un paramètre de stratification pour la conception de l'échantillonnage aréolaire. Les bases aréolaires peuvent être délimitées i) soit par des éléments physiques ou naturels ; ii) soit par une forme régulière (échantillonnage par grille).

La figure 1 présente un exemple de base aréolaire de forme régulière (grille composée de segments de 5×5 km) utilisée pour l'évaluation de l'État du Bihar, en Inde, dans le cadre du projet FASAL. La carte classifiée des cultures (figure 1a) est utilisée pour stratifier les segments de 5×5 km en quatre classes (type A : plus de 50 % de surfaces

cultivées ; type B : de 30 à 50 % ; type C : de 15 à 30 % ; type D : de 5 à 15 %). Un identifiant unique est attribué à chaque segment de 5×5 km. Une fois la stratification terminée, environ 15 % des segments échantillonnés (tableau 2) sont sélectionnés dans chaque type aux fins de l'analyse finale. Environ 50 % de ces segments feront l'objet d'une visite pour collecter les données sur la réalité de terrain.

FIGURE 1. CONCEPTION DE L'ÉCHANTILLONNAGE ARÉOLAIRE POUR L'ÉTAT DU BIHAR (INDE) DANS LE CADRE DU PROGRAMME FASAL.



1a) Carte classifiée des cultures de l'État du Bihar. 1b) Superposition à l'image classifiée d'une grille composée de segments de 5×5 km stratifiés en quatre classes (A, B, C et D) d'après la proportion des surfaces cultivées. La figure montre les segments échantillonnés sélectionnés (environ 15 %) pour chaque type.

L'utilisation de segments carrés n'est pas la seule approche possible pour l'élaboration des bases aréolaires. La FAO a par exemple utilisé, dans le cadre de son étude sur l'estimation des zones rizicoles en Afghanistan (FAO, 2017), des segments irréguliers délimités par des frontières physiques. En raison de la complexité du paysage local, la base de sondage aréolaire contenait plusieurs degrés, à savoir des UPE (de 500 à 700 ha), des USE (de 200 à 300 ha) et des unités d'échantillonnage finales (de 25 à 35 ha). La stratification était fondée sur l'intensité de culture (plus de 75 %, de 50 à 75 %, de 25 à 50 % et moins de 25 %). La définition de la base d'échantillonnage et l'optimisation du sondage se sont appuyées sur l'imagerie captée par les instruments Pleiades et MODIS.

L'exploitation des données de télédétection lors de l'élaboration de la base de sondage aréolaire permet d'obtenir une efficacité de stratification élevée. Carfagna (2013) a observé que l'efficacité de la stratification était généralement comprise entre 1,1 et 1,6 % dans les zones pilotes du projet MARS. Gallego *et al.* (1999) ont constaté une efficacité comprise entre 1,7 et 2,2 % pour les grandes cultures d'Espagne. En Inde, dans le cadre de l'estimation des surfaces rizicoles à partir de données micro-ondes (captées par SAR), l'efficacité de stratification était comprise entre 1,0 et 2,68 % (tableau 2).

Le paramètre du coefficient de variation (CV) décrit aussi la pertinence de la stratification. En Inde, dans les zones rizicoles d'une superficie comprise entre un et trois millions d'hectares (au niveau de l'État), les CV variaient de 1,15 à 3 % respectivement pour les échantillons de 1 229 et 450 ha. Dans le cas des cultures rizicoles, une forte corrélation négative a été constatée entre le CV et la surface cultivée à un taux d'échantillonnage fixe (environ 15 %). Dans le cadre de son étude sur la conception de bases de sondage aréolaires dans différents pays, Delincé (2015) a lui aussi constaté que plus les surfaces cultivées étaient réduites, plus le coefficient de variation augmentait.

TABLEAU 1. EXPLOITATION DE LA TÉLÉDETECTION AUX FINS D'ESTIMATION DES SURFACES CULTIVÉES DANS DIFFÉRENTS PAYS.

Pays	Organisation	Nom du programme	Données satellitaires	Échelle	Cultures	Approche	Réf.
Afghanistan	FAO		ProbaV, Aqua/Terra, Landsat-8, Sentinel-1, Sentinel-2, SPOT-5, -6 et -7 et Pleiades-1A et -1B	District, province	Riz	Conception de BSA, classification des images et estimateur par régression	FAO, 2017
Afrique du Sud	Consortium national des statistiques sur les cultures (NCSC, de l'anglais <i>National Crop Statistics Consortium</i>)	Système d'estimation des cultures non dépendant des producteurs (PICES, de l'anglais <i>Producer Independent Crop Estimate System</i>)	Landsat	Province	Tournesol, maïs	Classification des images	Ferreira <i>et al.</i> , 2006
Argentine	Secrétariat à l'Agriculture, à l'Élevage, à la Pêche et à l'Alimentation (SAGPyA)		Landsat		Blé, maïs, soja	Échantillonnage aréolaire et classification	Justice et Becker-Reshef, 2007
Asie (6 pays)	Institut international de recherche sur le riz (IRRI)	Information et assurance des cultures par la télédétection dans les pays émergents (RIICE, de l'anglais <i>Remote Sensing based Information and Insurance for Crops in Emerging economies</i>)	Images SAR en bande X captées par COSMO-SkyMed ; TerraSAR-X	Certains sites sélectionnés	Riz	Classification des images	Nelson <i>et al.</i> , 2014
Australie	Université du Queensland		MODIS EVI		Blé, orge, pois chiche	Analyse harmonique, analyse en composantes principales	Potgieter <i>et al.</i> , 2007
Brésil	Compagnie nationale d'approvisionnement alimentaire (CONAB, du portugais <i>Companhia Nacional de Abastecimento</i>)	GeoSafras	Landsat et MODIS		Maïs, soja et blé	Analyse par régression	Fontana <i>et al.</i> , 2006
Canada	Statistique Canada	Encore au stade de la recherche	Landsat-8			Classification	Brisbane et Mohl, 2014

Pays	Organisation	Nom du programme	Données satellitaires	Échelle	Cultures	Approche	Réf.
Chine	Bureau national des statistiques, Ministère de l'Agriculture	Estimation par télédétection et sondage des superficies cultivées (CAERSS, de l'anglais <i>Crop Acreage Estimation by using Remote Sensing and Sample Survey</i>) Système de suivi de l'agriculture par télédétection (CHARMS, de l'anglais <i>China Agricultural Remote Sensing Monitoring System</i>)	Landsat TM	Province	Mais, riz et soja	Système d'échantillonnage aréolaire et régression/calage	Pan <i>et al.</i> , 2012
Espagne	Ministère de l'Agriculture de la communauté autonome	Carte des surfaces naturelles et des cultures de Castille-et-León	Deimos-1 et Landsat-8		Adaptation du CDL des États-Unis	Classification au moyen d'un algorithme d'apprentissage automatique	Del Blanco Medina et Nafria García, 2015
États-Unis d'Amérique	USDA/NASS	CDL	Resourcesat AWiFS, Landsat ETM+	État	Coton, blé, sorgho, riz, soja, etc.	Estimateur par régression	Bailey et Boryan, 2010
Éthiopie	Université de Californie	Projet de recherche	Ikonos, Landsat	District	Surfaces cultivées	Échantillonnage aréolaire et classification	Husak <i>et al.</i> , 2008
Fédération de Russie	Ministère de l'Agriculture	Agrocosmos	NDVI de MODIS produit tous les 16 jours	Oblast (unité administrative russe), district	En complément des recensements agricoles		Temnikov et Sergey, 2007
Hongrie	Centre de télédétection de l'Institut de géodésie, cartographie et télédétection (FÖMI)	Programme de prévision de la production agricole et de suivi des cultures (CROPMON, de l'anglais <i>Crop Monitoring and Production Forecast Program</i>)	Landsat et IRS-1C/1D	Comté	Blé, maïs	Classification des images	Csornai <i>et al.</i> , 2006

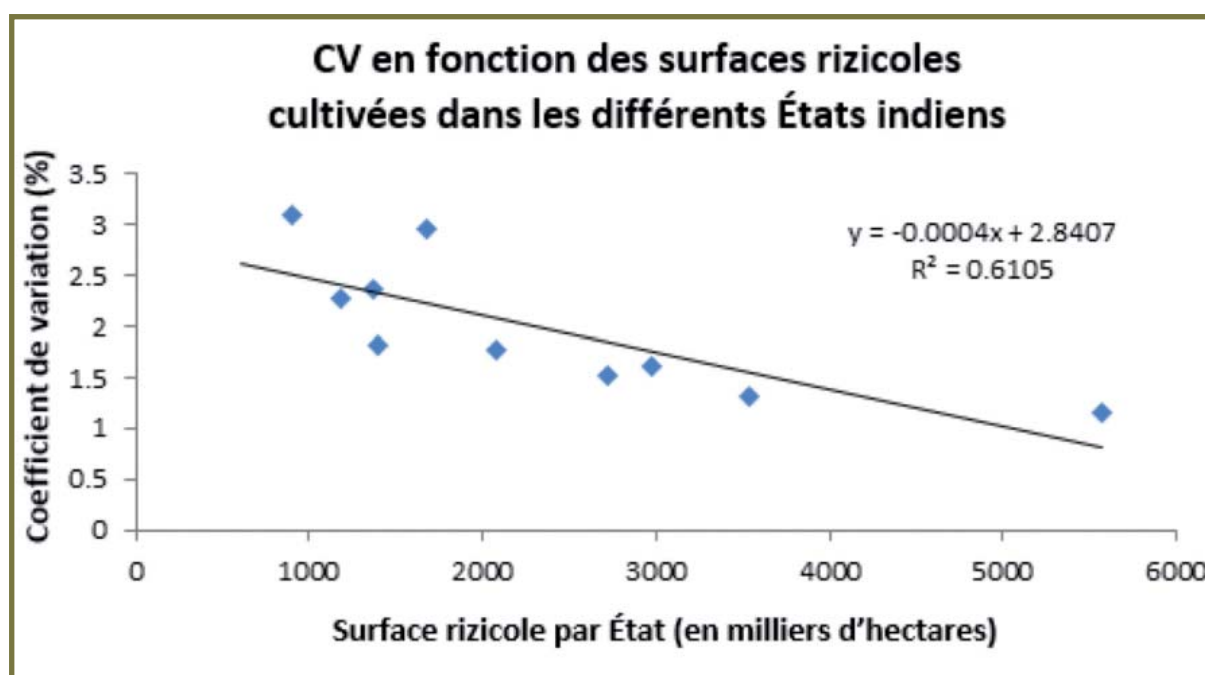
Pays	Organisation	Nom du programme	Données satellitaires	Échelle	Cultures	Approche	Réf.
Inde	Centre national Mahalanobis de prévision des récoltes (NCFC, de l'anglais <i>Mahalanobis National Crop Forecast Centre</i>) (ministère indien de la Coopération agricole et de la Protection sociale des agriculteurs, DAC&FW, de l'anglais <i>Department of Agricultural Cooperation & Farmers Welfare</i>) Commission de recherche sur l'espace et la haute atmosphère (SUPARCO, de l'anglais <i>Space and Upper Atmosphere Research Commission</i>)	FASAL	Resourcesat-2 : AWiFS et LISS III ; Landsat et Sentinel-2 ; SAR RISAT-1	District, État, pays	Riz, blé, coton, canne à sucre, sorgho, jute, colza/moutarde et pomme de terre	Conception de BSA et classification des images	Ray <i>et al.</i> , 2015
Pakistan			SPOT-5	Province	Blé, riz, coton, canne à sucre, maïs et pomme de terre	Classification des images et système d'échantillonnage aréolaire	Ahmad <i>et al.</i> , 2015
Union européenne (28 pays)	JRC	MARS	Landsat TM et SPOT XS	UE, État membre	Blé, orge, maïs, riz, légumes secs, colza, tournesol, betterave à sucre	Stratification et estimateur par régression	Gallego, 2000 et 2006

TABLEAU 2. PLAN D'ÉCHANTILLONNAGE DES ZONES RIZICOLES, COEFFICIENTS DE VARIATION (CV) ET EFFICACITÉ DE LA STRATIFICATION DANS PLUSIEURS ÉTATS INDIENS COUVERTS PAR LE PROJET FASAL.

État	Population de segments de 5 x 5 km				Échantillons des segments de 5 x 5 km				Population totale		Nombre total d'échantillons	Taux de sondage	CV (%)	Efficacité de la stratification
	A	B	C	D	a	b	c	d	N	n				
Andhra Pradesh	316	581	736	1 201	49	91	111	174	2 834	425	15	2,37	2,68	
Assam	524	840	937	681	96	120	125	84	3 000	425	14	1,77	1,46	
Bihar	646	989	1 091	784	109	145	166	109	3 510	529	15	1,60	1,230	
Chhattisgarh	711	979	1 042	1 347	111	154	159	203	4 079	642	15	1,31	1,56	
Haryana	273	265	339	499	46	53	61	93	1 376	253	18,4	2,28	2,66	
Jharkhand	233	552	747	872	47	89	127	141	2 404	404	16,8	1,81	1,234	
Karnataka	195	521	741	1 350	41	91	123	203	2 807	458	16,3	3,09	1,22	
Madhya Pradesh	273	460	826	848	59	79	130	128	2 407	396	16,4	2,95	1,00	
Punjab	462	554	503	387	72	105	93	78	1 906	348	18,2	1,52	1,53	
Uttar Pradesh	1 332	2 218	2 154	1 826	227	364	367	318	7 530	1 276	16,9	1,15	1,40	

Type A : > 50 % de surfaces cultivées ; type B : 30 à 50 % ; type C : 15 à 30 % ; type D : 5 à 15 %
L'efficacité de la stratification correspond au ratio entre les variances de l'échantillonnage aléatoire simple et de l'échantillonnage stratifié.

FIGURE 2. COEFFICIENTS DE VARIATION DES ESTIMATIONS EN FONCTION DES SURFACES RIZICOLES CULTIVÉES DANS LES DIFFÉRENTS ÉTATS INDIENS.



5.2.1.1 Estimation directe ou comptage de pixels

Cette approche consiste à classer l'imagerie satellite d'après la réalité de terrain collectée sur les sites échantillonnés. Le nombre de pixels associés à chaque culture dans une zone administrative est multiplié par la taille de pixel afin d'obtenir la superficie de la surface cultivée.

L'analyse des images est réalisée dans les segments échantillonnés ou d'après des scènes complètes (on parle alors de dénombrement exhaustif). Dans le premier cas, la superficie de chaque segment est estimée et statistiquement agrégée pour obtenir la superficie totale. Dans le cas du dénombrement exhaustif, les frontières administratives (district/État/comté/province) sont apposées sur l'image et tous les pixels associés à une culture sont dénombrés et multipliés par la taille de pixel pour obtenir la superficie de la surface cultivée. Une carte des cultures est aussi produite qui pourra servir à d'autres fins, telles que l'échantillonnage à des fins d'évaluation du rendement.

La classification peut être supervisée (les classes sont alors définies selon la réalité de terrain) ou non supervisée (la classification exploite alors la tendance inhérente des différentes classes à former des grappes dans l'espace de l'élément). Les algorithmes les plus fréquemment utilisés pour la classification supervisée sont l'algorithme « distance minimale », l'algorithme du parallélépipède et l'algorithme « maximum de vraisemblance ». L'algorithme des k-moyennes et la méthode ISODATA constituent quant à eux des exemples de classificateurs utilisés pour la classification non supervisée. D'autres approches de classification des cultures ont été utilisées récemment, telles que les classificateurs hiérarchiques (arbres de décision), les machines à vecteurs de support, les réseaux de neurones artificiels et les ensembles flous. Prasad *et al.* (2015) ont recensé les techniques existantes de classification d'image.

La classification est réalisée à partir de données satellitaires multites de résolution moyenne (MODIS, Resourcesat AWiFS ou SPOT-VGT) ou de données monodates à haute résolution (Landsat OLI, Resourcesat LISS III ou Sentinel-2 MSI). Les coûts des différentes données satellitaires optiques généralement utilisées pour la classification des cultures sont présentés dans le tableau 3.

TABLEAU 3. EXEMPLES DE COÛTS DES DIFFÉRENTES DONNÉES SATELLITAIRES OPTIQUES ET SAR UTILISÉES POUR ESTIMER LES SURFACES CULTIVÉES.

Satellite	Capteur	Caractéristiques du produit	Prix (en euros)*
EO1	MODIS (Terra et Aqua)	Résolution de 250 m/500 m/1 km	Gratuit
SPOT-5	HRS	Résolution multibande de 20 m, 60 × 60 km	1 900
		Résolution multibande de 10 m, 60 × 60 km	2 700
		Résolution multibande de 5 m, 60 × 60 km	5 400
	VEGETATION-2	Résolution de 1 km	Gratuit
Landsat-8	OLI		Gratuit
Resourcesat-2@	AWiFS	Résolution multibande de 56 m, 740 × 740 km	222
	LISS III	Résolution multibande de 24 m, 140 × 140 km	96
	LISS IV	Résolution multibande de 5,8 m, 70 × 70 km	147
Sentinel-2	MSI		Gratuit
RapidEye	Multibande	Produits Basic/Ortho, 3 500 km ² contigus	3 325
RISAT 1@	SAR en bande C	Mode ScanSAR à résolution moyenne (MRS), résolution de 18 m, 115 × 115 km ²	69
Sentinel-1	SAR en bande C		Gratuit
Radarsat-2	SAR en bande C	Mode large, résolution de 30 m ; 150 × 150 km ²	2 590
COSMO-SkyMed	SAR en bande X	Mode large ScanSAR, résolution de 30 m ; 100 × 100 km ²	1 650

Source : <http://www.e-geos.it/products/pdf/prices.pdf>. Ces prix s'appliquent aux nouvelles acquisitions.
Remarque : Voir le chapitre 1 pour plus d'informations sur les caractéristiques des capteurs.

5.2.1.2. Analyse de données multidates

L'analyse de données multidates s'appuie sur les différences phénologiques (types de croissance) entre les plantes cultivées dans une même zone. On utilise généralement une résolution spatiale modérée produisant des données à haute fréquence temporelle, telles que celles proposées par MODIS (250 m, chaque jour ou tous les huit jours), SPOT VGT (1 km, chaque jour ou tous les dix jours) ou Resourcesat-2 AWiFS (56 m, tous les cinq jours). La classification des cultures s'appuie sur sept à dix dates de co-enregistrement couvrant la majeure partie de la période de végétation. Une méthode de classification basée sur des règles de décision (hiérarchique) est appliquée aux produits multidates de l'indice NDVI pour classer les cultures selon leur cycle de croissance.

La figure 3 présente un exemple d'images composites de l'indice NDVI bimensuel provenant des données de Resourcesat-2 AWiFS pour l'État de l'Uttar Pradesh, en Inde, pendant la saison des cultures rabi d'hiver (de novembre/décembre à mars/avril). Cet ensemble de données est utilisé pour classer les principales cultures de l'État, à savoir le blé, le colza/la moutarde, la pomme de terre et les légumes secs. Les profils temporels des indices NDVI de toutes ces cultures sont représentés à la figure 4. L'indice NDVI de la pomme de terre augmente fortement pour atteindre son maximum à la fin du mois de décembre et pendant les quinze premiers jours de janvier, celui de la moutarde croît à partir de la fin du mois de décembre et diminue vers la fin du mois de février et celui du blé augmente à partir de janvier puis tout au long de la période de végétation (décembre à avril). Les différentes signatures temporelles des légumes secs sont comparables à celles des autres cultures principales.

5.2.2. Analyse de données monodates

L'analyse de données monodates part du principe que la plante peut être distinguée des autres classes de cultures et d'utilisation des terres/couverture du sol lorsque sa croissance végétative est à son maximum, à condition de disposer d'informations suffisantes sur la réalité de terrain. Des données satellitaires à haute résolution sont

utilisées à cette fin (p. ex., Resourcesat-2 LISS III, Landsat-8 OLI ou Sentinel-2 MSI). Des informations sur la réalité de terrain sont utilisées pour générer les signatures des cultures. La classification supervisée (p. ex., au moyen du classifieur par MV) permet de classifier les pixels pour une culture particulière. La surface des pixels est multipliée par le nombre de pixels pour obtenir la superficie cultivée au sein d'une zone administrative. La figure 5 présente un exemple de classification de la moutarde et du blé fondée sur des données Landsat.

FIGURE 3. IMAGES COMPOSITES DE L'INDICE NDVI HEBDOMADAIRE/BIMENSUEL POUR L'ÉTAT DE L'UTTAR PRADESH, EN INDE.

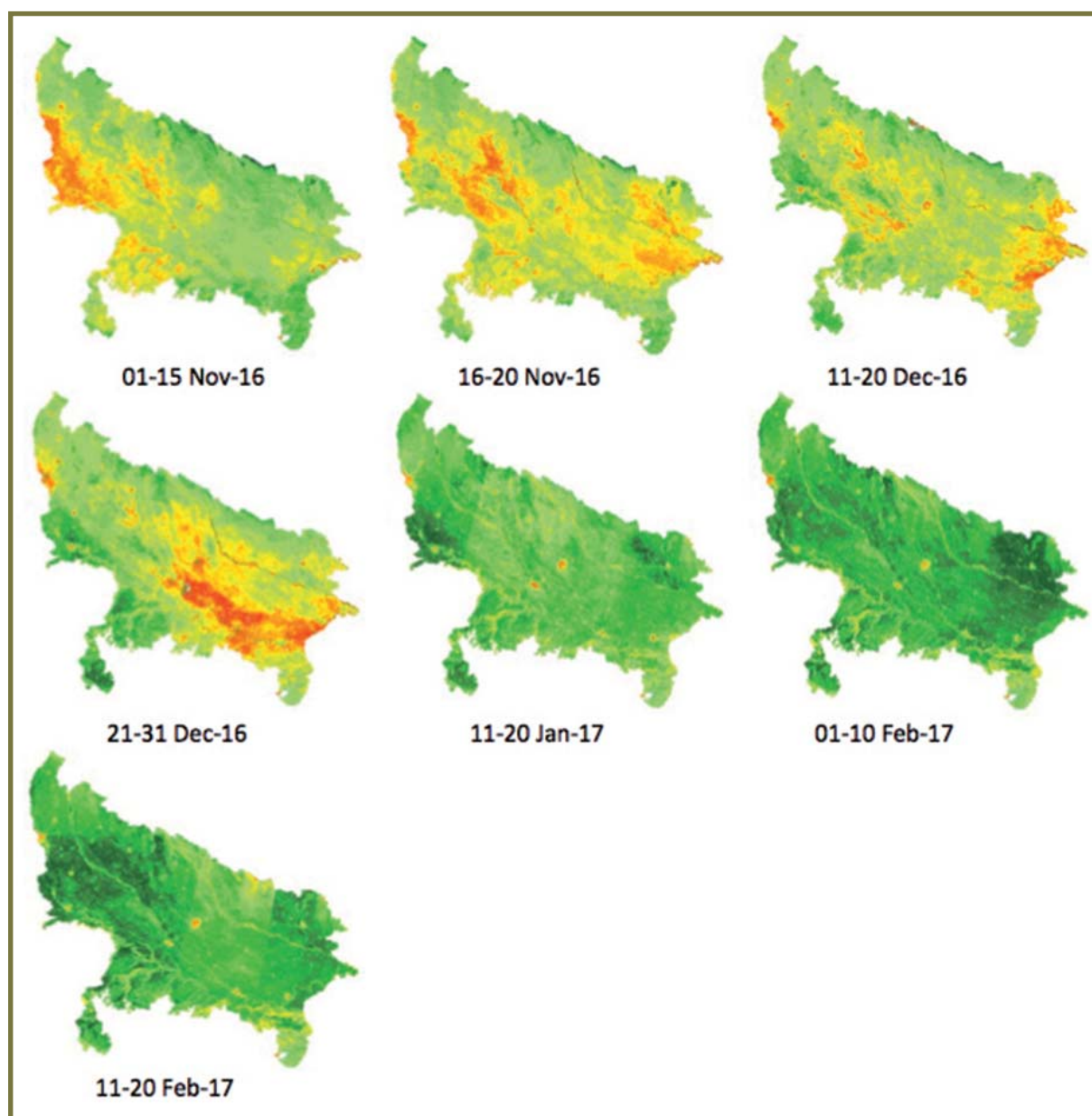


FIGURE 4. PROFILS TEMPORELS (GRADUÉS) DES INDICES NDVI DE PLUSIEURS CLASSES DE CULTURES POUR L'ÉTAT DE L'UTTAR PRADESH, EN INDE.

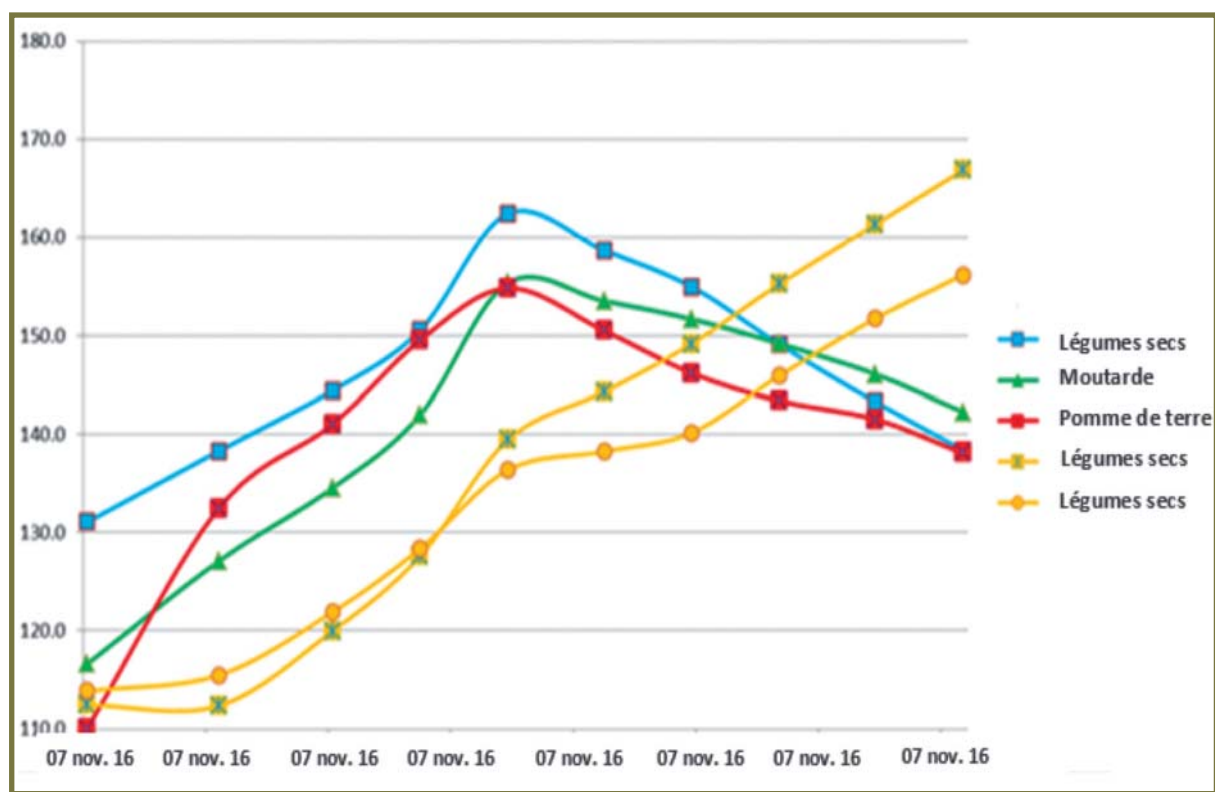
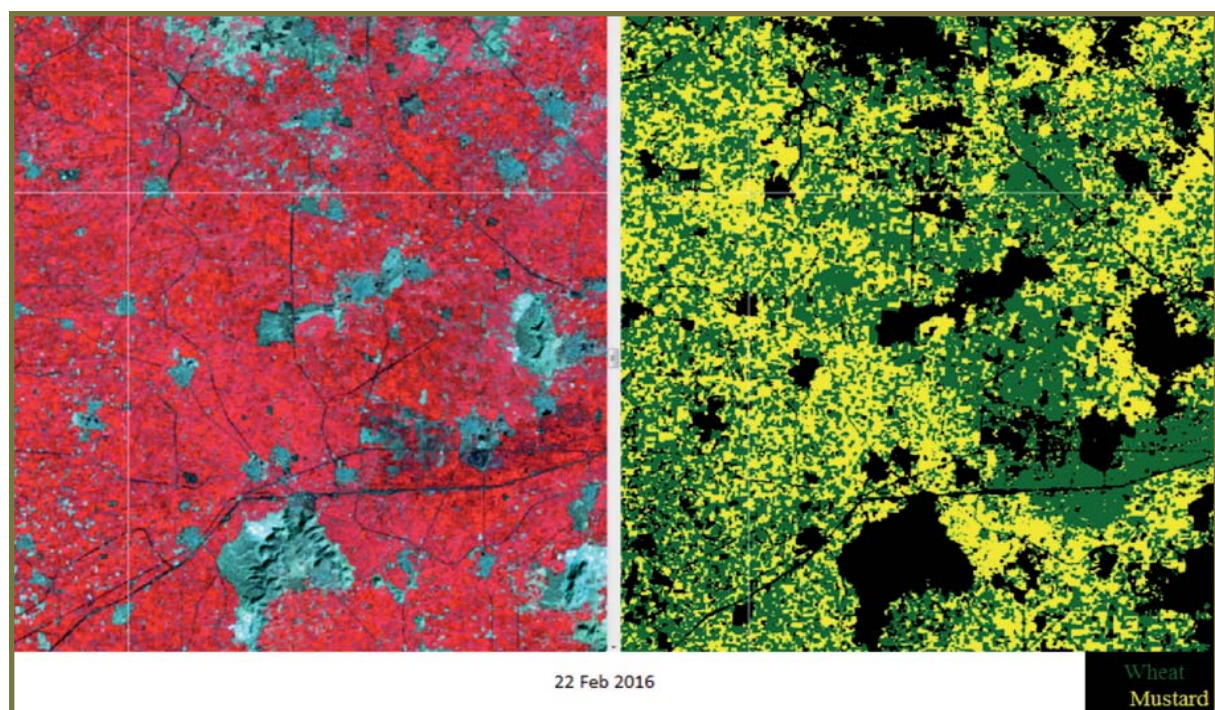


FIGURE 5. IMAGES DE LANDSAT FCC (GAUCHE) ET IMAGES CLASSIFIÉES (BLÉ EN VERT ET MOUTARDE EN JAUNE) (DROITE) POUR LE DISTRICT DE BHIWANI, DANS L'ÉTAT D'HARYANA, EN INDE.



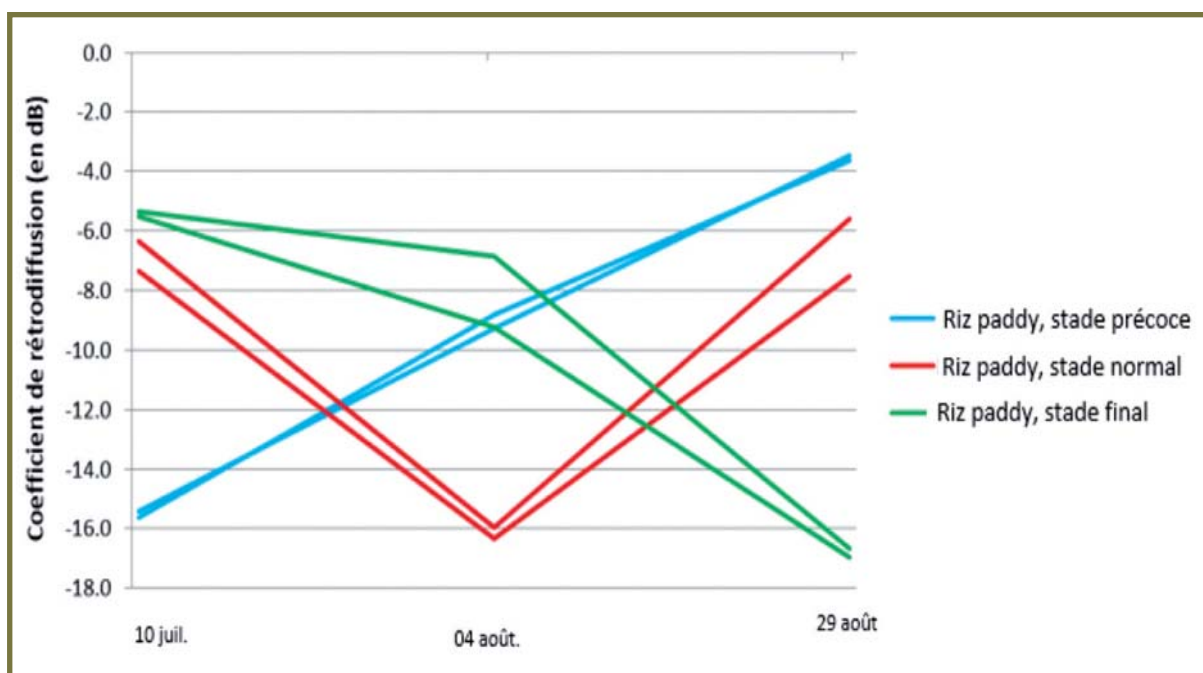
5.2.3. Exploitation des données SAR aux fins d'estimation des surfaces cultivées

Dans la plupart des pays d'Asie du Sud et du Sud-Est où l'on pratique la culture irriguée du riz, la couverture nuageuse persistante empêche la plupart du temps d'obtenir des données optiques pendant la saison des pluies (la principale saison de végétation du riz). Les données micro-ondes captées par les SAR offrent une solution à ce problème, ce type de radar étant sensible à la rugosité du sol. Or, habituellement, les jeunes plants sont repiqués dans des parcelles inondées. Peu après le repiquage, les plants émettent une valeur de rétrodiffusion très faible en raison de la réflexion spéculaire des micro-ondes sur l'eau stagnante des parcelles (Choudhury et Chakraborty, 2006 ; Suga et Konishi, 2008). Ensuite, la rétrodiffusion des ondes radar augmente à mesure de la croissance du plant et de la pousse des talles, jusqu'à ce que la plante atteigne son stade de reproduction. Ce phénomène est dû à la diffusion de volume par la végétation et aux multiples reflets entre les plants et la surface de l'eau (Chakraborty *et al.*, 2006 ; Nelson *et al.*, 2014). Une fois atteint le stade de reproduction, la rétrodiffusion du radar reste pratiquement constante (Chakraborty *et al.*, 1997). Par conséquent, l'estimation des surfaces rizicoles à partir des données SAR suppose habituellement d'acquérir des données à trois moments distincts – avant, pendant et après la mise en terre (semis) – en respectant un intervalle de 20 à 25 jours entre chaque date d'acquisition.

Les étapes de traitement des données SAR concernant les cultures rizicoles sont généralement les suivantes :

- géoréférencement des images ;
- étalonnage des images et suppression du bruit au moyen d'un filtre passe-bas adaptatif prédéfini ;
- co-enregistrement des images multitudes (trois dates) et préparation de l'ensemble de données ;
- conversion des données numériques des pixels en valeurs de rétrodiffusion ;
- superposition des sites de la réalité de terrain et identification des sites rizicoles ;
- élaboration d'une règle de décision fondée sur le profil temporel des valeurs de rétrodiffusion des cultures rizicoles (figure 6) ; et
- classification des cultures rizicoles et estimation des surfaces cultivées.

FIGURE 6. PROFIL TEMPOREL DES DIFFÉRENTES CLASSES DE RIZ POUR LE DISTRICT DE MIRZAPUR, DANS L'ÉTAT D'UTTAR PRADESH, EN INDE.



L'encadré ci-après présente un exemple typique de règle de décision pour la classification du riz.

ENCADRÉ 2

- Zone urbaine : $L1 > -6,0$ ET $L2 > -6,0$ ET $L3 > -6,0$
- Surface aquatique : $L1 < -17,0$ ET $L2 < -17,0$ ET $L3 < -17,0$
- Stade précoce du riz repiqué : $(-18,0 \leq L1 \leq -14,0)$ ET $(-12,0 < L2 \leq -8,0)$ ET $(-6,0 < L3 \leq -2,0)$ ET $(L2 > L1 + 1,0)$ ET $(L3 > L2 + 1,0)$
- Stade normal du riz repiqué : $(-10,0 \leq L1 \leq -4,0)$ ET $(-18,0 < L2 < -12,0)$ ET $(-10,0 < L3 \leq -3,0)$ ET $(L1 > L2 + 1,0)$ ET $(L3 > L2 + 1,0)$
- Stade final du riz repiqué : $(-8,0 \leq L1 \leq -3,0)$ ET $(-11,0 < L2 < -5,0)$ ET $(-18,0 < L3 \leq -14,0)$ ET $(L2 > L3 + 1,0)$

Où

L1 = coefficient de rétrodiffusion (db) à la première date

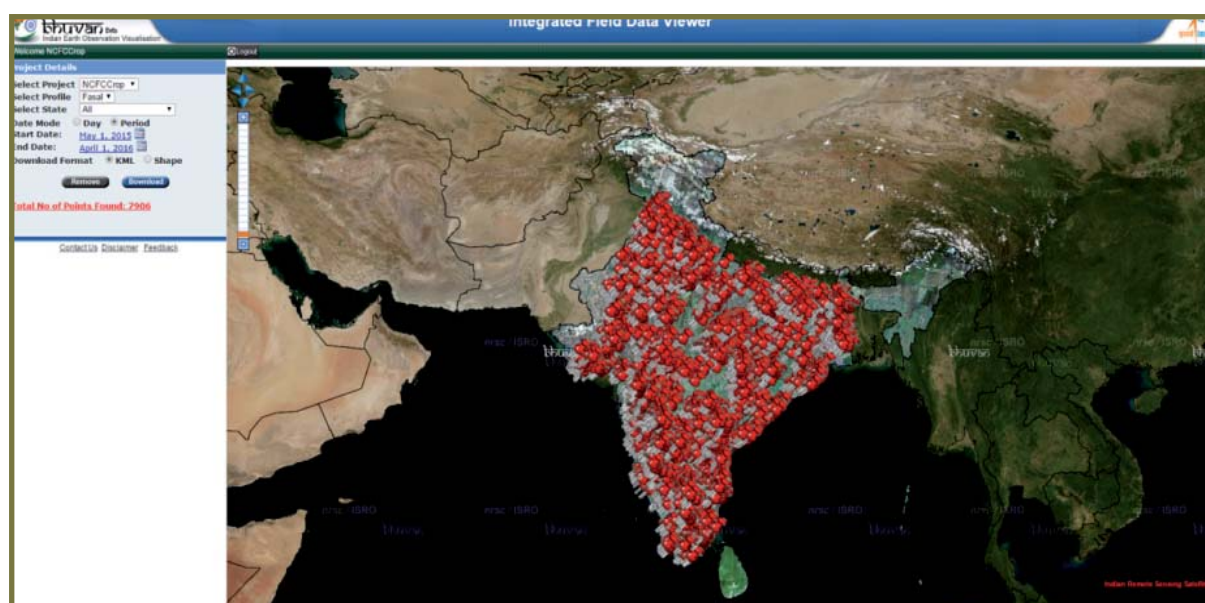
L2 = coefficient de rétrodiffusion (db) à la deuxième date

L3 = coefficient de rétrodiffusion (db) à la troisième date

5.2.4. Données sur la réalité de terrain

La réalité de terrain joue un rôle central dans la classification des cultures, soit pour élaborer l'algorithme soit pour valider le résultat. Les observations de réalité de terrain décrivent l'utilisation des terres et la couverture du sol. Les données de terrain recueillies aux fins de classification des cultures incluent généralement l'emplacement géographique, le village/district/État, le nom de la culture, la couverture, l'état, le stade de croissance, le modèle de culture – irriguée ou pluviale –, le rendement attendu, les dates d'ensemencement et de récolte, etc., ainsi que deux photographies (plan rapproché et plan large) de la parcelle. La réalité de terrain est recueillie à partir d'une sélection de sites échantillonnés répartis dans l'ensemble de la région d'étude afin d'en capter la diversité. Plusieurs applications Android pour smartphone ont été développées pour faciliter la collecte des données de terrain. Les données ainsi recueillies sont transférées en temps réel vers un serveur central où les analystes d'image peuvent les récupérer. En Inde, les données de terrain sont recueillies par des agents du département agricole de chaque gouvernement d'État, puis transférées vers le serveur Bhuvan (un géoportail de l'Organisation indienne pour la recherche spatiale [ISRO, de l'anglais *Indian Space Research Organisation*], voir la figure 7).

FIGURE 7. DONNÉES DE TERRAIN RECUEILLIES PAR SMARTPHONE ET DISPONIBLES SUR LE GÉOPORTAIL BHUVAN



5.2.5. Estimation de la précision : matrice de confusion et écart relatif

La matrice de confusion, aussi appelée matrice d'erreur, permet de représenter la performance de la classification. Dans le tableau 4, les classes réelles/de référence sont indiquées dans les colonnes et les pixels prévus/classifiés dans les lignes. La matrice de confusion fournit une synthèse des résultats et permet d'approfondir l'analyse de la classification.

Les tableaux 4 et 5 ci-après présentent des matrices de confusion utilisées dans deux scénarios de classification (le premier concernant la superficie totale des surfaces cultivées au Kazakhstan et le second les cultures individuelles en Inde). Dans son étude pilote réalisée au Kazakhstan, Gallego (2008) a montré que l'estimation par comptage de pixels de la superficie cultivée totale (céréales et jachère) présentait une marge de subjectivité d'environ 5 %. En Inde, la précision de la classification des cultures individuelles variait entre 70 et 90 %, avec une précision globale de 81,09 % et un indice Kappa de 0,7606. Dans le cadre d'un travail de cartographie de la culture irriguée du riz en Chine à partir des séries temporelles d'images MODIS, Clauss *et al.* (2016) ont conclu à une précision globale de 0,90 et une précision de l'utilisateur de 0,90 pour la classe « no_rice » et de 0,89 pour la classe « rice ».

TABLEAU 4. EXEMPLE DE MATRICE DE CONFUSION UTILISÉE POUR UNE ÉTUDE PILOTE AU KAZAKHSTAN.

Référence					
Classification		Céréales + jachère	Herbage + abandon	Total	Précision du producteur
	Cultures	1 470 + 152	57	1 679	96,6 %
	Herbage + abandon	39 + 68	353	460	76,7 %
	Total	1 729	410	2 139	
	Précision de l'utilisateur	93,8 %	86,1 %		

Source : Gallego, J., 2008., « Crop Area Estimation with Remote Sensing: Some considerations and experiences for the application to general agricultural statistics », présentation préparée en vue de l'atelier sur la mesure de la culture et de la production de feuilles de coca, 25-27 novembre 2008, Bogotá, Colombie. Accessible à l'adresse suivante : https://www.unodc.org/documents/crop-monitoring/Workshop_coca_leaves/Javier_Gallego1.pdf
Données satellitaires utilisées : MODIS (résolution de 250 m)

TABEAU 5. EXEMPLE DE MATRICE DE CONFUSION UTILISÉE POUR L'ÉTAT DU MADHYA PRADESH, EN INDE.

Données de référence										
Données classifiées	Pois chiche	Blé	Moutarde	Pomme de terre	Pois	Jachère	Urbain	Lentille	Total de la ligne	Précision de l'utilisateur
Pois chiche	153	28	0	3	6	5	2	0	197	78 %
Blé	1	474	62	4	43	2	1	0	587	81 %
Moutarde	0	56	525	6	27	1	4	2	621	85 %
Pomme de terre	0	25	28	334	48	0	0	0	435	77 %
Pois	0	30	23	11	323	2	2	0	391	83 %
Jachère	0	0	3	12	9	102	3	2	131	78 %
Urbain	3	2	0	1	6	4	112	5	133	84 %
Lentille	0	12	9	0	1	0	1	99	122	81 %
Total de la colonne	157	627	650	371	463	116	125	108	2 617	
Précision du producteur	97 %	76 %	81 %	90 %	70 %	88 %	90 %	92 %	81,09	

Précision globale de la classification = 81,09 %

Indice Kappa = 0,7606

Données satellitaires utilisées : Resourcesat-2 LISS III (résolution de 23,5 m)

Les matrices de confusion permettent de connaître la précision interne de la classification. En revanche, l'évaluation de la précision vis-à-vis d'une estimation standard (p. ex., les estimations d'un ministère de l'Agriculture) se fait au moyen de divers paramètres tels que l'écart relatif, l'erreur quadratique moyenne (EQM) et le coefficient de corrélation.

En Inde, le coefficient de corrélation entre les estimations basées sur les données de télédétection et les estimations ministérielles concernant les surfaces occupées par quatre cultures à l'échelle de l'État variait de 0,986 à 0,999. Dans le cadre du travail de cartographie de la culture irriguée du riz en Chine à partir des séries temporelles d'images MODIS, Clauss *et al.* (2016) ont trouvé un coefficient de détermination compris entre 0,91 et 0,93 par rapport aux estimations du gouvernement.

5.3. ESTIMATEUR PAR RÉGRESSION

Les méthodes telles que les estimateurs par régression, par calage et sur petits domaines associent des informations exhaustives mais imprécises provenant de l'imagerie satellite et des informations précises extraites d'un échantillon issues le plus souvent de mesures de terrain (Gallego, 2006).

Les estimateurs par régression sont décrits dans les textes sur les statistiques standards (voir par exemple Cochran, 1963).

Au niveau régional, l'estimateur est le suivant (Sud *et al.*, 2015) :

$$\hat{Y}_R = \sum_{i=1}^L N_h \bar{y}_{h(Reg)}$$

$$\text{où } \bar{y}_{h(Reg)} = \bar{y}_h + \hat{b}_h (\bar{X}_h - \bar{x}_h)$$

\bar{y}_h = pour chaque segment échantillonné de la strate h , valeur de la moyenne des surfaces cultivées relevées sur le terrain, à savoir

$$\bar{y}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} y_{hj}$$

\hat{b}_h = pour chaque segment de la strate h , coefficient de régression des surfaces relevées sur le terrain par rapport aux surfaces mesurées à partir des données de télédétection

\bar{X}_h = moyenne de la surface de l'ensemble des unités de tuile de la strate h , calculée à partir des données de télédétection (toute la surface doit donc être classifiée pour obtenir cette moyenne de population, à savoir

$$\bar{X}_h = \frac{1}{N_h} \sum_{j=1}^{N_h} X_{hj}$$

\bar{x}_h = pour chaque segment échantillonné de la strate h , valeur de la moyenne des surfaces cultivées mesurées à partir des données de télédétection.

$$\bar{x}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} x_{hj}$$

De nombreux pays, tels que les États-Unis, le Brésil et la Chine, utilisent l'estimateur par régression pour estimer les surfaces cultivées à partir de données de télédétection. La FAO (2017) a adopté une méthode hybride consistant à intégrer la base de sondage aréolaire à la classification des images pour renforcer la précision des statistiques sur les cultures. Elle a notamment utilisé cette approche pour estimer la superficie occupée par les cultures rizicoles en Afghanistan. La base de sondage aréolaire a été élaborée à partir d'images satellite enregistrées par Sentinel-2 et SPOT-5 (dont la résolution spatiale est de 10 m). Les terres agricoles contenues dans la zone du projet pilote ont été stratifiées et des segments choisis par tirage aléatoire systématique ont été interprétés visuellement, en même temps que les informations de terrain, pour estimer les statistiques relatives aux cultures à partir de la base de sondage aréolaire. L'interprétation visuelle des images satellite a servi d'échantillon d'apprentissage à l'algorithme de classification supervisée des images afin de générer des estimations de cultures basées sur les pixels. Le coefficient de détermination (R^2) de la régression linéaire entre les variables des pixels spécifiques au riz et les surfaces rizicoles dans les segments s'élevait à 0,96, ce qui témoigne de la haute précision obtenue par la combinaison de ces deux systèmes.

En Chine, dans le cadre du projet CAERSS, l'estimation des surfaces cultivées suit une méthode similaire. Dans un premier temps, une base de sondage aréolaire est élaborée à partir de données de télédétection multisources et multitemps puis adaptée pour définir le plan d'échantillonnage des cultures. Ensuite, une stratégie de sélection des échantillons est mise au point pour réaliser une stratification raisonnable et sélectionner les échantillons à sonder sur le terrain. Enfin, un modèle linéaire fondé sur l'association des données de terrain aux superficies cultivées classifiées tirées des images satellitaires (comme données auxiliaires) est adopté pour estimer les superficies cultivées avec une précision satisfaisante. Les superficies plantées des principales cultures peuvent ensuite être estimées à l'échelle des provinces et des comtés. Pour plus de détails sur la méthodologie et la procédure, voir Zhou (2013). Les coefficients de variation des estimations par régression linéaire pour les trois principales cultures (maïs, riz et soja) variaient entre 4,2 et 7,0 %.

5.4. ESTIMATEUR PAR CALAGE

Les estimateurs par calage intègrent des informations auxiliaires – les données de télédétection – dans le processus d'estimation (Benedetti *et al.*, 2015). Les erreurs de commission et d'omission de la matrice de confusion peuvent servir à corriger le biais. Stehman (2009) a indiqué que les informations de référence tirées d'un échantillon utilisé pour élaborer la matrice de confusion pouvaient aussi servir à déduire la superficie d'intérêt, soit directement soit à l'aide d'un modèle. L'estimateur direct et l'estimateur inverse sont deux méthodes qui s'appuient sur la matrice de confusion pour ajuster la superficie estimée par comptage de pixels. Leur différence tient au fait que l'estimateur direct utilise la précision de l'utilisateur tandis que l'estimateur inverse a recours à la précision du producteur. Les principales précautions consistent à calculer la matrice de confusion au moyen d'informations de terrain tirées d'un échantillon statistique composé de points ou segments (éléments) et à exécuter correctement l'extrapolation en tenant compte du plan d'échantillonnage (Gallego *et al.*, 2008).

Afin d'accroître la précision de l'estimation des surfaces par classification, Zhu *et al.* (2014) ont étudié la performance et la stabilité de plusieurs estimateurs assistés par un modèle. Les estimateurs direct et inverse basés sur la matrice de confusion, l'estimateur inverse par calage de la matrice de confusion, l'estimateur par le ratio et l'estimateur par régression simple ont ainsi été utilisés pour déduire les superficies occupées par plusieurs classes de couverture du sol au moyen d'un échantillonnage aléatoire simple sans remise. Une étude comparative a montré que les estimateurs par calage de la matrice de confusion, les estimateurs par le ratio et les estimateurs par régression simple fournissaient des estimations plus précises et plus stables que l'estimateur par échantillonnage aléatoire simple.

5.5. ESTIMATEUR SUR PETITS DOMAINES

L'estimation sur petits domaines est essentielle pour analyser les données d'enquête par sondage lorsque les tailles des échantillons de domaines (sous-population) sont trop réduites pour fournir une précision adaptée aux estimateurs directs sur domaine. Dans les estimations basées sur des données de télédétection, la précision peut être bonne pour les zones étendues (pays et États ou provinces) grâce aux échantillons de grande taille. En revanche, cela n'est pas toujours le cas pour les plus petites zones (districts ou comtés). Pour accroître la précision de l'estimation dans de plus petites zones, la technique de l'estimation sur petits domaines (SAE, de l'anglais *Small Area Estimation*) est appliquée. Les diverses méthodes élaborées pour appliquer la méthode SAE peuvent être regroupées comme des méthodes statistiques et des méthodes de microsimulation spatiale aux approches divergentes (Whitworth, 2013). L'approche statistique se base sur le modèle de régression qui permet d'évaluer formellement la relation entre une caractéristique présentant un intérêt et une ou des variables explicatives. Zhou (2016) a associé diverses méthodes de classification des cultures – imagerie satellite, données de terrain et techniques SAE – pour produire des estimations des surfaces occupées par des cultures de riz et de maïs à l'échelle des comtés et des villes.

5.6. EXEMPLES DE PROGRAMMES NATIONAUX, RÉGIONAUX ET MONDIAUX D'ESTIMATION DES SURFACES CULTIVÉES

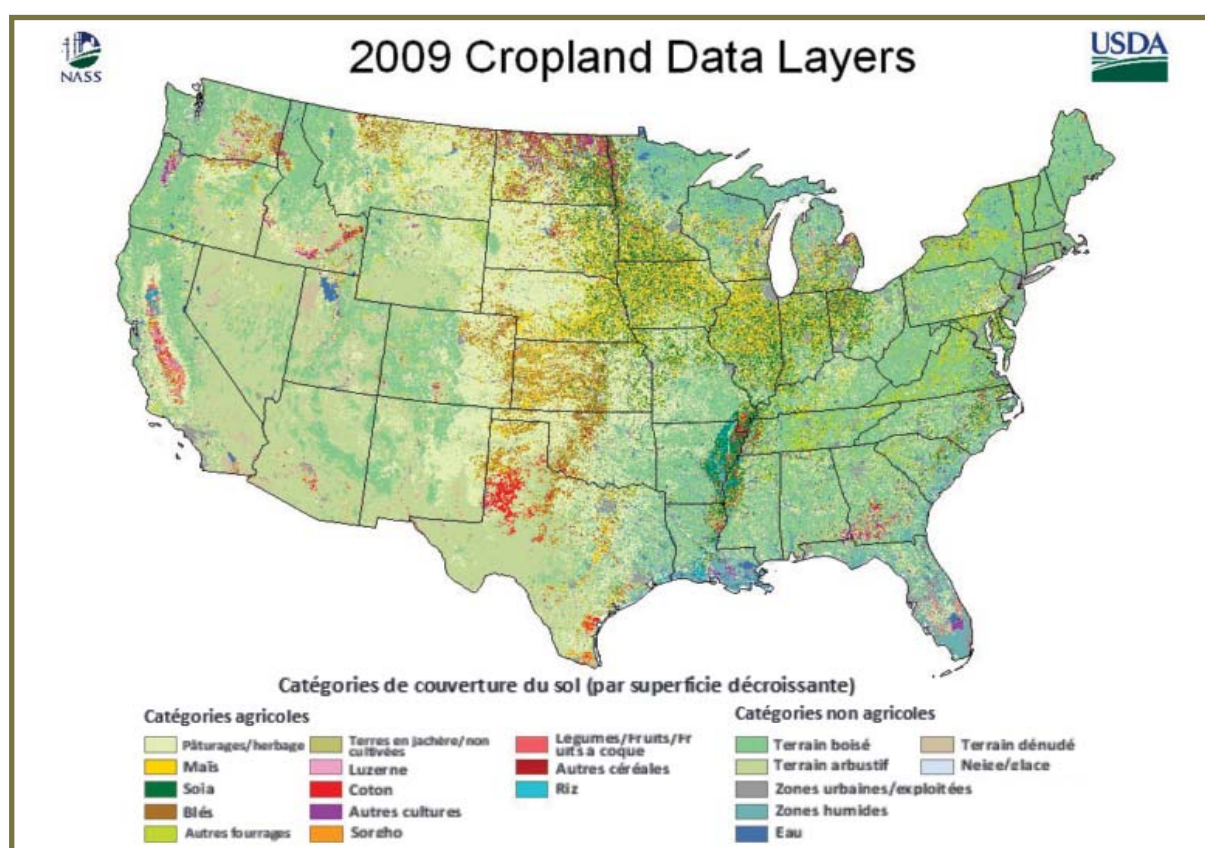
5.6.1. Programmes nationaux

5.6.1.1. CDL du NASS de l'USDA

Afin de publier régulièrement des statistiques précises et pratiques sur le secteur agricole aux États-Unis, le service NASS de l'USDA réalise de nombreuses études visant à collecter des informations sur plusieurs aspects des activités agricoles. En 2010, son programme CDL a grandement facilité sa mission en exploitant les techniques de télédétection pour fournir au Bureau des statistiques agricoles et aux bureaux de terrain du NASS des estimations opérationnelles sur les superficies. Ces estimations ont été communiquées pendant la saison agricole elle-même pour 16 types de cultures

et 27 États (Baily et Boryan, 2010). Le NASS est à l'origine de nombreux programmes novateurs visant à démontrer l'intérêt des données de télédétection pour estimer les superficies cultivées, comme LACIE et AgRISTARS. Son programme CDL a été lancé en 1997 avec un logiciel interne et des données Landsat. En 2006, le CDL a été fortement modifié suite à l'adoption d'un logiciel commercial et des données captées par Resourcesat-1 AWiFS. Le produit CDL est une carte matricielle de la couverture du sol, géoréférencée et spécifique aux cultures (Boryan *et al.*, 2011). En 2009, le programme CDL s'est avéré très utile pour fournir, pendant la saison agricole, des estimations opérationnelles sur les superficies occupées par 15 cultures dans 27 États. La présentation générale du programme CDL par Boryan *et al.* (2011) décrit les différents types de données d'entrée, les procédures de traitement, de classification et de validation, l'évaluation de la précision, les spécifications des produits CDL, les lieux de diffusion et la méthode d'estimation des superficies cultivées – qui consiste à utiliser un estimateur par régression en se basant sur le CDL.

FIGURE 8. CROPLAND DATA LAYER 2009.



Source : NASS de l'USDA (<https://nassgeodata.gmu.edu/CropScape/>)

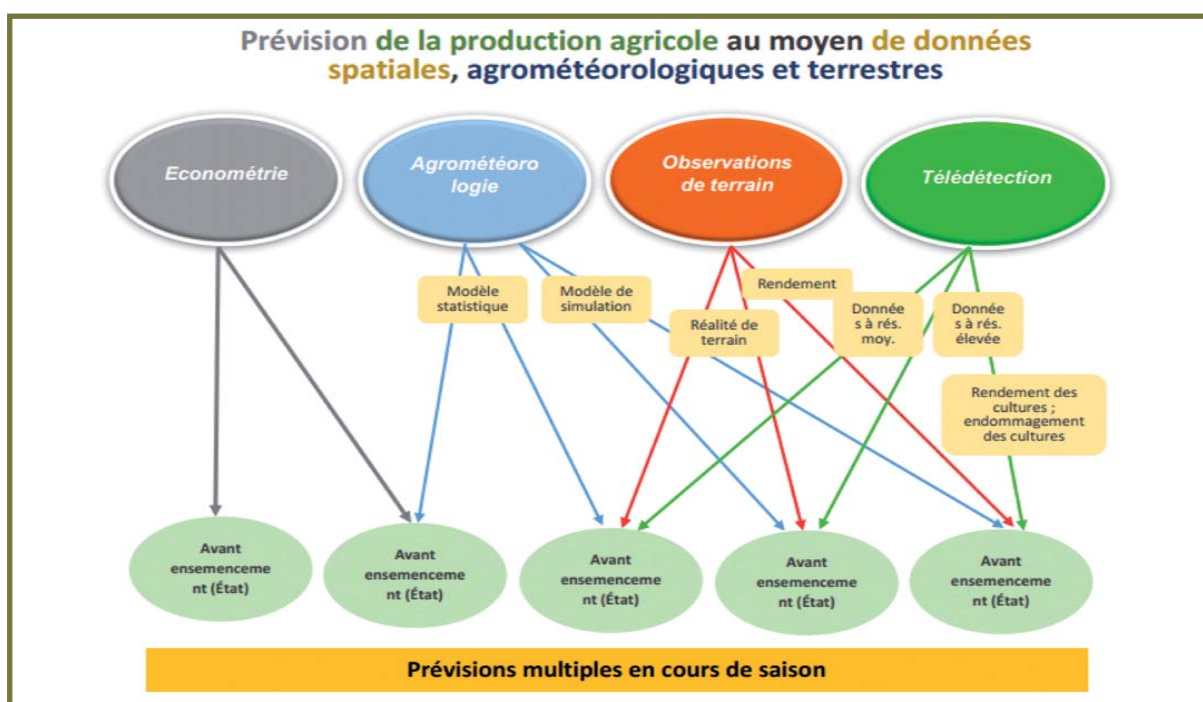
5.6.1.2. Programme indien FASAL

En Inde, l'estimation des cultures à partir des données de télédétection a débuté à la fin des années 1970 avec la réalisation d'une étude systématique visant à dresser l'inventaire des cultures à l'aide de données aériennes infrarouges couleur. Cette étude conjointe de l'ISRO et du Conseil indien de la recherche agricole (ICAR, de l'anglais *Indian Council of Agricultural Research*) s'inscrivait dans le cadre du projet *Agricultural Resource Inventory and Survey Experiment* (ARISE) (Sahai *et al.*, 1977). Par la suite, de nombreuses études expérimentales ont été menées à partir de données aériennes puis des données Landsat-1. Ces premières études sont présentées dans les publications de Bhavsar (1980), Navalgund et Sahai (1985), et Sahai et Dadhwal (1990). Le lancement de la mission IRS 1A s'est accompagné de la mise en place d'un programme national majeur, le CAPE, qui consistait à estimer les superficies et la production des principales cultures à l'échelle des districts et des États. Les études menées sous l'égide du programme CAPE ont permis de mettre au point un plan d'échantillonnage optimal, des

spécifications sur les capteurs, des chiffres sur la précision, une procédure d'analyse optimale et un logiciel interne d'évaluation des cultures (Dadhwal *et al.*, 2002). Parallèlement, des méthodes ont été élaborées pour parvenir à exploiter les données SAR (initialement émises par ENVISAT, puis par Radarsat) aux fins d'estimation des superficies des cultures rizicoles, dans le but d'éliminer le problème de la couverture nuageuse pendant la saison kharif (saison des pluies) (Patel *et al.*, 1995 ; Panigrahy *et al.*, 2000). L'expérience acquise dans le cadre du projet CAPE et de diverses études pilotes, de même que les exigences du ministère de l'Agriculture indien, ont été prises en compte lors de l'élaboration d'un projet d'inventaire exhaustif des cultures : le programme FASAL.

L'objectif du programme FASAL, officiellement lancé en 2007, est de fournir, avant la récolte, plusieurs prévisions de production concernant 11 grandes cultures à l'échelle du district, de l'État et du pays. Les prévisions s'appuient sur plusieurs approches (relevant de l'économétrie, de la télédétection et de l'agrométéorologie) et diverses données satellitaires (optiques et SAR) (Parihar et Oza, 2006). Les méthodes ont été élaborées puis optimisées pour le travail de prévision de la production végétale par le centre d'applications spatiales de l'ISRO. Leur mise en œuvre a ensuite été confiée au NCFC relevant du ministère de l'Agriculture (Ray *et al.*, 2015). Les stratégies suivies par le programme FASAL pour prévoir la production végétale sont illustrées à la figure 9.

FIGURE 9. STRATÉGIES DE PRÉVISION DE LA PRODUCTION VÉGÉTALE DU PROJET FASAL.



La méthode d'estimation des superficies adoptée par le projet FASAL exploite les données de télédétection pour : i) élaborer le plan d'échantillonnage et ii) produire des estimations au moyen de la classification des images basée sur des données monodates à haute résolution ou des données multirates à résolution moyenne. Les données SAR de RISAT-1 sont utilisées pour le riz et le jute, tandis que les estimations des autres cultures (blé, colza/moutarde, sorgho, coton et canne à sucre) se basent sur des données captées par Resourcesat-2 AWiFS, LISS III, Landsat OLI ou Sentinel-2 MSI. L'analyse des cultures de riz et de jute se fonde sur une approche de segmentation utilisant des échantillons de 5×5 km, un taux de sondage de 15 % et quatre strates (la stratification dépend de la couverture culturale dans chaque segment). En revanche, c'est la méthode du dénombrement exhaustif qui est suivie pour les autres cultures. Les données sur la réalité de terrain sont recueillies au moyen d'applications Android sur smartphone. L'estimation du rendement se base sur des modèles agrométéorologiques (modèles empiriques et modèles de simulation des cultures), des modèles fondés sur la télédétection et des données tirées de la méthode

de coupe échantillon, celle-ci s'appuyant sur un plan d'échantillonnage basé sur la télédétection. Les estimations produites dans le cadre du programme FASAL sont intégrées aux données utilisées pour produire les estimations finales du DAC&FW. Ray *et al.* (2016) ont réalisé un examen approfondi de la méthode, des résultats et de la précision du programme FASAL. Devant les bons résultats du projet d'estimation de la production végétale à partir des données de télédétection, le DAC&FW a lancé un nouveau programme d'évaluation et de gestion coordonnées de l'horticulture au moyen de la géoinformatique (CHAMAN, de l'anglais *Coordinated Horticulture Assessment and Management using geoinformatics*) pour réaliser l'inventaire des cultures horticoles (mangue, banane, agrumes, pomme de terre, tomate, piment fort et oignon) et planifier le développement horticole (Ray *et al.*, 2015).

5.6.2. Programmes régionaux: Programme MARS du JRC européen

Le projet MARS a été institué pour mettre au point des outils opérationnels de grande envergure dans le domaine des informations agricoles afin d'appuyer l'analyse des images satellitaires et des domaines connexes, telles que l'échantillonnage aréolaire et la conception de modèles agrométéorologiques (Gallego, 2000). La première période (1988-1997) du projet MARS a principalement porté sur l'estimation des surfaces cultivées sous deux grands volets : 1) les inventaires régionaux des surfaces cultivées et 2) les estimations rapides des changements des surfaces cultivées au sein de l'Union européenne. Les progrès accomplis pendant cette période ont été résumés par Gallego (2000). Les inventaires régionaux des surfaces cultivées utilisaient à la fois des images satellitaires à haute résolution et des levés de terrain selon un modèle statistique classique basé sur l'échantillonnage aréolaire et des visites de terrain fournissant la principale variable d'estimation au moyen d'un estimateur par régression. Les estimations rapides des changements des surfaces cultivées visaient à fournir une estimation précoce des changements au sein de l'UE d'après un échantillon de 60 sites de 40 × 40 km chacun, sans informations de terrain. Les deux programmes ont été abandonnés, le premier car son coût approchait le seuil de rentabilité, et le second en raison de sa marge de subjectivité trop importante.

Le programme d'échantillonnage AGRI4CAST du projet MARS tourne désormais essentiellement autour de l'enquête aréolaire sur l'utilisation et l'occupation des sols (LUCAS) (Gallego et Delincé, 2010). Les estimations des surfaces correspondant à différents types d'utilisation des terres ou de couverture du sol ne sont pas calculées au moyen d'une cartographie exhaustive de la zone étudiée, mais à partir d'observations réalisées sur environ 270 000 points issus d'un échantillonnage par points répartis dans l'ensemble de l'UE. La répétition de l'enquête à quelques années d'intervalle permet d'identifier les changements dans l'utilisation des terres¹. L'enquête LUCAS est un processus d'échantillonnage stratifié systématique composé de deux phases : la photo-interprétation de l'échantillon dans un premier temps, puis la collecte des données de terrain dans les échantillons dans un second temps. La dernière enquête LUCAS a été réalisée en 2012 dans 27 pays de l'UE, dont 24,7 % du territoire total est recouvert par des terres cultivées.

5.6.3. Programmes mondiaux

5.6.3.1. CropWatch (Chine)

Le programme CropWatch a été fondé par l'Institut de télédétection et Terre numérique (RADI, de l'anglais *Institute of Remote Sensing and Digital Earth*) de l'Académie des sciences de Chine. Ce programme réalise des évaluations à l'échelle mondiale (avec 65 unités homogènes de surveillance et d'information sur les cultures), à l'échelle régionale (dans sept zones majeures de production), à l'échelle nationale (pour 31 pays clés) et à l'échelle infranationale (dans des subdivisions de neuf grands pays) (Wu *et al.*, 2015). À chaque échelle, divers indicateurs sont sélectionnés pour mieux identifier les informations environnementales et agricoles pertinentes. Ces indicateurs englobent la pluviométrie, la température de l'air, le rayonnement photosynthétiquement actif, le potentiel agroclimatique de production de biomasse, l'intensité de culture, la fraction de terres arables cultivées

¹ <http://esdac.jrc.ec.europa.eu/projects/lucas>

(CALF, de l'anglais *Cropped Arable Land Fraction*), l'indice de la santé végétale, l'indice de l'état de la végétation maximale et la proportion des types de cultures. Les méthodes d'estimation des surfaces cultivées utilisées par CropWatch diffèrent selon les pays. En Chine, les estimations par télédétection de la superficie des terres arables et de la CALF sont associées à des estimations de la proportion des types de cultures basées sur des enquêtes de terrain (Wu et Li, 2012).

Pour les 30 autres pays et provinces ou États de neuf grands pays, la CALF est utilisée pour estimer chaque surface cultivée selon l'équation suivante :

$$\text{Surface cultivée} = a \times b \times \text{CALF}$$

où a et b sont les coefficients de régression linéaire entre les surfaces cultivées tirées de FAOSTAT ou, idéalement, les données infranationales éventuellement disponibles sur le site Internet du ministère de l'Agriculture ou du Bureau national des statistiques de Chine.

5.6.3.2. FAS de l'USDA (États-Unis)

Le FAS de l'USDA fournit des évaluations mensuelles sur l'état des cultures ainsi que des services de surveillance et d'estimation des cultures pour 17 produits de base mondiaux, 159 pays, 1 020 paires pays-culture (par exemple, Australie-Blé) et trois attributs : la superficie, le rendement et la production (Hoffman, 2016). Pour remplir sa mission de surveillance et d'évaluation des cultures mondiales, le FAS s'appuie sur diverses informations telles que des images satellitaires, les rapports de ses attachés, les visites de terrain, les données officielles et les bulletins d'information. Il exploite les données de 8 des 18 satellites d'observation de la Terre de la NASA, ainsi que des satellites de l'ESA, de l'ISRO et d'organisations privées. Le portail Internet « Crop Explorer » du FAS publie tous les dix jours de nombreux ensembles de données sur l'état de la végétation et les conditions météorologiques dans les grandes régions de cultures.

5.6.3.3. GEOGLAM (G20)

Dans leur Déclaration relative au Plan d'action sur la volatilité des prix alimentaires et sur l'agriculture adopté suite à la hausse des prix des denrées alimentaires à travers le monde en 2007-2008 et 2010, les chefs d'État des pays du G20 ont soutenu en 2011 les initiatives GEOGLAM et le Système d'information sur les marchés agricoles (AMIS, de l'anglais *Agricultural Market Information System*). L'initiative GEOGLAM renforce la capacité de la communauté internationale à produire et à diffuser en temps opportun des prévisions pertinentes et précises sur la production agricole à l'échelle nationale, régionale et mondiale à l'aide de données d'observation de la Terre, y compris des observations satellitaires et de terrain². Elle soutient aussi l'élaboration de rapports de suivi des cultures (« Crop Monitor »), dont les évaluations sur l'état des cultures mondiales sont utiles aux activités de suivi des marchés de l'initiative AMIS³. L'initiative Asia-RiCE est le fruit du travail d'une équipe *ad hoc* de parties prenantes convaincues de la nécessité d'intégrer à l'initiative GEOGLAM un volet sur l'estimation et le suivi des cultures rizicoles en Asie. Lors de sa phase 1 (2013-2015), le programme Asia-RiCE a installé des sites de démonstration technique à Taïwan, en Indonésie, au Japon, en Malaisie, en Thaïlande et au Viet Nam. La phase 2 de l'initiative doit désormais préparer des prévisions sur la croissance rizicole en Indonésie, aux Philippines, en Thaïlande et au Viet Nam, pour ensuite les communiquer à l'AMIS (Asia-RiCE, 2016).

² <https://www.earthobservations.org/geoglam.php>

³ <https://cropmonitor.org/>

5.7. RENTABILITÉ DE L'ÉVALUATION DES SURFACES AU MOYEN DE LA TÉLÉDÉTECTION

L'analyse de la rentabilité d'un système suppose de comparer son coût avec celui d'autres systèmes (tels que ceux employés pour la collecte traditionnelle de données agricoles) produisant le même résultat final (Radhakrishnan *et al.*, 1991). Dans le cas de l'estimation des surfaces à partir de la télédétection, les coûts concernent l'acquisition des données satellitaires, la collecte des données de terrain et le travail d'analyse. Outre les possibles économies directes (qui permettent d'obtenir la même estimation à moindre coût), les avantages incluent aussi l'amélioration de la rapidité de mise à disposition et de la précision des estimations. La précision peut s'évaluer en termes de réduction de la variance et d'augmentation de l'efficacité (dans le cas des plans d'échantillonnage ayant recours à la télédétection). D'après Carfagna (2001), la question de la rentabilité de la télédétection aux fins de statistiques agricoles a déjà fait l'objet de nombreux débats et dépend de plusieurs paramètres, tels que le niveau de fragmentation du paysage, les conditions météorologiques, le niveau d'optimisation et d'automatisation du projet et la structure de coûts. Ainsi, les résultats varient selon les situations.

Delincé (2015) a réalisé une étude approfondie sur la rentabilité des statistiques agricoles utilisant la télédétection dans quatre systèmes nationaux : en Haïti (échantillonnage aréolaire par points), au Maroc (échantillonnage aréolaire), en Chine (échantillonnage aréolaire et analyse par régression) et en Inde (échantillonnage aréolaire et comptage de pixels). Les conclusions de cette étude sont les suivantes :

- En Haïti, l'analyse a porté sur le système d'échantillonnage par points du Centre national de l'information géo-spatiale (CNIGS). Le coût de la stratification était visible dans l'augmentation de 3 % du coût des levés de terrain, mais les baisses de variance allaient jusqu'à 50 % au niveau régional.
- Au Maroc, la stratification réalisée au moyen de cartes de couverture du sol couvrant 66 000 km² a nécessité d'acquérir des images SPOT onéreuses, ce qui a entraîné une augmentation de 30 % des coûts de sondage annuels. L'efficacité gagnée justifie toutefois ces investissements.
- En Chine, une stratification fondée sur la télédétection a été réalisée sur une surface de 1,65 million de km². Les coûts ont augmenté de 3 % seulement, mais, dans la province d'Anhui, l'efficacité relative de la stratification a été de 2,8 pour le riz et de 1,4 pour le maïs.
- En Inde, des images radar et optiques sont utilisées pour surveiller 90 % de la production des huit principales cultures. L'efficacité de la stratification obtenue pour le riz était comprise entre 1,2 et 3,3 ; le biais introduit par le comptage de pixels a pu être évalué.

Pour le NASS de l'USDA, dont la plupart des applications opérationnelles de premier plan sont basées sur des enquêtes aréolaires et des données de télédétection pour les statistiques agricoles, l'analyse de la rentabilité a donné des résultats positifs (Carfagna, 2013). D'après cette étude, « les applications de la télédétection dans le domaine des statistiques agricoles peuvent être viables si leur coût total cadre avec le budget sans compromettre la faisabilité des enquêtes qui ne peuvent pas être menées au moyen des technologies satellitaires ».

Toutefois, compte tenu de la disponibilité de nombreuses données satellitaires gratuites (p. ex., Landsat, Sentinel, etc.) et de la diminution notable des prix des données satellitaires indiennes, la rentabilité de la télédétection pour les statistiques agricoles a fortement augmenté.

5.8. DIFFICULTÉS ET LIMITES

Dans leur présentation sommaire du programme LACIE, MacDonald et Hall (1980) ont indiqué que « les informations agricoles doivent être à la fois objectives, fiables, rapidement disponibles, assez complètes, efficaces et rentables. Dans la plupart des grands pays agricoles, les statistiques sur la production ne répondent à aucun de ces critères ». Ce constat reste valable aujourd'hui dans de nombreux pays. Les exemples ci-dessus démontrent que les données de télédétection provenant de divers satellites sont utiles dans le cadre de plusieurs méthodes de

constitution de statistiques agricoles (conception de plans d'échantillonnage aréolaire, estimation par classification ou par régression). Cependant, de nombreuses difficultés continuent de limiter le recours aux données de télédétection dans les opérations d'évaluation des surfaces cultivées, dont les suivantes :

- les parcelles de petites tailles, particulièrement en Asie et en Afrique (tableau 6). L'identification des cultures nécessite dans ces cas des données de télédétection à haute résolution ;
- la persistance de la couverture nuageuse pendant la saison des pluies. Les nuages limitent fortement l'intérêt de l'imagerie optique pour les applications agricoles (Eberhardt *et al.*, 2016). Des données SAR doivent être collectées pour contourner ce problème, mais leur utilité pour estimer d'autres cultures que le riz reste à prouver ;
- la diversité des pratiques culturales et agronomiques ;
- les systèmes de cultures mixtes et associées ;
- les petites surfaces incluant une grande variété de cultures, particulièrement dans le cas de l'horticulture.

D'autres problèmes techniques surviennent dans l'analyse et l'exploitation des données de télédétection, tels que la nécessité d'une quantité suffisante de données de terrain de qualité décrivant la variabilité réelle ; la puissance de calcul, les logiciels avancés et les équipements de stockage de données nécessaires pour analyser les données multitudes à haute résolution ; et la disponibilité de données satellitaires dans un court délai.

Malgré ces difficultés, plusieurs études ont démontré que l'exploitation des données de télédétection pour l'estimation des surfaces était rentable (Delincé, 2015 ; Gallego *et al.*, 2014).

La disponibilité actuelle et future de nombreuses constellations de satellites de télédétection à haute résolution devrait entraîner une augmentation de la fréquence temporelle des données satellitaires et de la précision de la classification. Des méthodes doivent être élaborées pour permettre d'exploiter des données SAR aux fins d'estimation de cultures autres que le riz.

Plusieurs opinions ont été exprimées sur les méthodes à utiliser pour estimer les surfaces, à savoir l'estimateur par régression ou le comptage de pixels. Au sujet de l'estimateur, Carfagna et Gallego (2005) soutiennent que les images satellitaires classifiées doivent être utilisées comme des variables auxiliaires dans un estimateur par régression ou dans les estimateurs basés sur les matrices de confusion. Ils considèrent aussi que les images classifiées ou photointerprétées ne doivent généralement pas être exploitées directement pour estimer les surfaces cultivées, car la proportion de pixels classifiés dans des cultures spécifiques est souvent fortement biaisée.

Toutefois, dans de nombreuses applications, les cartes classifiées des cultures sont essentielles en complément des estimations sur les surfaces cultivées, notamment pour planifier les récoltes de carrés de rendement. Par ailleurs, le personnel qualifié nécessaire pour renforcer l'utilisation de l'estimation par régression n'est pas toujours disponible, même dans de nombreuses organisations officielles (Gallego, 2006). Craig et Atkinson (2013) indiquent que la sous-estimation presque systématique des surfaces cultivées en cas de recours au système de comptage de pixels peut être améliorée par le recours à la régression. Il serait donc nécessaire d'adopter une approche complète incluant à la fois la constitution de bases de sondage, le comptage de pixels et la régression.

TABLEAU 6. TAILLE MOYENNE ET FRAGMENTATION DES EXPLOITATIONS AGRICOLES, 1995-2005.

Pays par continent (le nombre de pays couverts par l'étude est indiqué entre parenthèses)	Superficie moyenne par exploitation agricole (hectares)	Nombre moyen de parcelles par exploitation agricole
Total mondial (114)	5,5	3,5
Afrique (25)	11,5	3,0
Amérique du Nord et centrale (14)	117,8	1,2
Amérique du Sud (8)	74,4	1,2
Europe (29)	12,4	5,9
Asie (29)	1,0	3,2

Source : FAO, 2010.

5.9. RÉFÉRENCES BIBLIOGRAPHIQUES

Ahmad, I., Ghafoor, A., Bhatti, M.I., Akhtar, I.H., Ibrahim, M. et Rehman, O. 2015. « *Satellite remote sensing and GIS-based crops forecasting & estimation system in Pakistan* ». In : Srivastava, M.K. (dir.), *Crop Monitoring for Improved Food Security: Proceedings of the Expert Meeting. Vientiane, République démocratique populaire lao, 17 février 2014*. Publication RAP 2014/28. Publication conjointe de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) et de la Banque asiatique de développement (BASD) : Rome – Manille.

Asian Rice Crop Estimation and Monitoring Component (Asia-RiCE). 2016. *Asia-RiCE Phase 2 Work Plan*. Publication d'Asia-RiCE.

Bailey, J.T. et Boryan, C.G. 2010. *Remote Sensing Applications in Agriculture at the USDA National Agricultural Statistics Service*. Vol. 22030. Publication USDA/NASS : Fairfax, VA, États-Unis.

Bauer, M.E., Cary, T.K., Davis, B.J. et Swain, P.H. 1975. *Crop Identification Technology Assessment for Remote Sensing (CITARS): Results of CITARS Experiments Performed by LARS*. NASA CR 147389, Note d'information LARS 072175. Publication du Laboratory for Applications of Remote Sensing (LARS) : West Lafayette, IN, États-Unis.

Benedetti, R., Piersimoni, F. et Postiglione, P. 2015. *Sampling Spatial Units for Agricultural Surveys*. Springer-Verlag : Berlin.

Bhavsar, P.D. 1980. « Demonstrated Applications in India of Earth Resources Survey by Remote Sensing ». *Proceedings of the National Academy of Sciences, India*, vol. 46, A, n° 3, p. 275–85.

Boryan, C. 2012. *A Review of Four Area Sample Designs*. Document préparé pour le cours GGS796 de l'Université George Mason. Disponible auprès de la Division de recherche, USDA/NASS : Fairfax, Virginia, États-Unis.

Boryan, C., Yang, Z., Mueller, R. et Craig, M. 2011. « Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program ». *Geocarto International*, vol. 26, n° 5, p. 341–58.

Brisbane, J. et Mohl, C. 2014. « Utilisation possible de la télédétection pour produire des statistiques sur les grandes cultures à Statistique Canada ». In : *Au-delà des méthodes traditionnelles d'enquêtes : l'adaptation à un monde en évolution*. Recueil du Symposium 2014 de Statistique Canada.

Bryan, E.R. 1974. *The ERTS-I Investigation (ER-600). Volume I - ERTS-I Agricultural Analysis*. Mémoire technique X-58117 de la NASA.

Carfagna, E. 2001. *Cost-effectiveness of Remote Sensing in Agricultural and Environmental Statistics*. Communication sollicitée pour la conférence Caesar : *Conference on Agricultural and Environmental Statistical Applications in Rome*. 4–8 juin 2001. Rome, Istituto nazionale di statistica (Istat). Volume III, p. 617–27.

Carfagna, E. 2013. *Evaluating the Cost-Efficiency of Remote Sensing in Developing Countries*. Présentation élaborée pour le premier Comité consultatif scientifique de la Stratégie mondiale – Improving AG-Statistics. 18–19 juillet 2013. Rome, FAO.

Carfagna, E. et Gallego, F.J. 2005. « Using Remote Sensing for Agricultural Statistics ». *International Statistical Review*, vol. 73, n° 3, p. 389–404.

Chakraborty, M., Patnaik, C., Panigrahy, S. et Parihar, J.S. 2006. « Monitoring of wet season rice crop at state and national level in India using multi-date synthetic aperture radar data ». In : Kuligowski, R.J., Parihar, J.S. et Saito, G. (dir.), *Proceedings of SPIE Volume 6411, Agriculture and Hydrology Applications of Remote Sensing*. Publication de la SPIE : Goa, Inde.

Chakraborty, M., Panigrahy, S. et Sharma, S.A. 1997. « Discrimination of rice crop grown under different cultural practices using temporal ERS-1 Synthetic aperture radar data ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 52, n° 4, p. 183–91.

Choudhury, I. et Chakraborty, M. 2006. « SAR signature investigation of rice crop using RADARSAT data ». *International Journal of Remote Sensing*, vol. 27, n° 3, p. 519–34.

Clauss, K., Yan, H. et Kuenzer, C. 2016. « Mapping Paddy Rice in China in 2002, 2005, 2010 and 2014 with MODIS Time Series ». *Remote Sensing*, vol. 8, n° 5, p. 434.

Cochran, G.W. 1963. *Sampling Techniques*. John Wiley & Sons : New York, États-Unis.

Craig, M. et Atkinson, D. 2013. *A Literature Review of Crop Area Estimation*. Publication de la FAO : Rome.

Csornai, G., Wirnhardt, C., Suba, Z., Nádor, G., Tikász, L., Martinovich, L., Kocsis, A., Zelei, G., László, I., Bognár, E. 2006. « CROPMON: Hungarian Crop Production Forecast by Remote Sensing ». In : Baruth, B., Royer, A. et Genovese, G. (dir.), *ISPRS Archives Volume XXXVI-8/W48. Workshop proceedings: Remote sensing support to crop yield forecast and area estimates* (p. 65–70). Publication de l'ISPRS : Stresa, Italie.

Dadhwal, V.K., Singh, R.P., Dutta, S. et Parihar, J.S. 2002. « Remote sensing based crop inventory: A review of Indian experience ». *Tropical Ecology*, vol 43, n° 1, p. 107–22.

Del Blanco Medina, V. et Nafria García, D.A. 2015. *Mapa de cultivos y superficies naturales de Castilla y León*. Document élaboré en vue du XVI^e Congrès de l'Association espagnole de télédétection, 21–23 octobre 2015. Séville, Espagne.

Delincé, J. 2015. *Technical Report on Cost-Effectiveness of Remote Sensing for Agricultural Statistics in Developing and Emerging Economies*. Technical Report Series GO-09-2015. Rapport technique de la Stratégie mondiale : Rome.

Dhanju, M.S. et Shankaranarayana, H.S. 1978. *Agricultural Resources Inventory and Survey Experiment (ARISE)*. ISRO-SAC-TR-11-78. Publication de l'ISRO : Bangalore, Inde.

Eberhardt, I.D.R. et al. 2016. « Cloud Cover Assessment for Operational Crop Monitoring Systems in Tropical Areas ». *Remote Sensing*, vol. 8, n° 3, p. 219.

FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture). 2010. *Characterisation of Small Farmers in Asia and the Pacific*. Document élaboré en vue de la 23^e session de la Commission des statistiques agricoles pour l'Asie et le Pacifique (Point 10 à l'ordre du jour), 26–30 avril 2010. Siem Reap, Cambodge.

FAO. 2017. *Afghanistan: Monitoring of Rice Crop using Satellite Remote Sensing and GIS Technologies*. Publication de la FAO : Rome.

Ferreira, S.L., Newby, T. et du Preez, E. 2006. « Use of Remote Sensing In Support of Crop Area Estimates in South Africa ». In : Baruth, B., Royer, A. et Genovese, G. (dir.), *ISPRS Archives Volume XXXVI-8/W48. Workshop proceedings: Remote sensing support to crop yield forecast and area estimates* (p. 51–52). Publication de l'ISPRS : Stresa, Italie.

Fontana, D.C., Melo, R.W., Wagner, A.P.L., Weber, E. et Gusso, A. 2006. « Use of Remote Sensing For Crop Yield and Area Estimates in the Southern of Brazil ». In : Baruth, B., Royer, A. et Genovese, G. (dir.), *ISPRS Archives Volume XXXVI-8/W48. Workshop proceedings: Remote sensing support to crop yield forecast and area estimates* (p. 53–58). Publication de l'ISPRS : Stresa, Italie.

Gallego, F.J. 2000. *Statistical aspects of area estimates in MARS*. Document élaboré pour la MARS Conference: 10 years of demand-driven technical support, 22–23 avril 1999. Bruxelles.

Gallego, F.J. 2006. « Review of the Main Remote Sensing Methods for Crop Area Estimates ». In : Baruth, B., Royer, A. et Genovese, G. (dir.), *ISPRS Archives Volume XXXVI-8/W48. Workshop proceedings: Remote sensing support to crop yield forecast and area estimates* (p. 65–70). Publication de l'ISPRS : Stresa, Italie.

Gallego, F.J., Craig, M., Michaelsen, J., Bossyns, B. et Fritz S. (dir.). 2008. *Best practices for crop area estimation with Remote Sensing*. GEOSS Community of Practice Ag 0703a. Office des publications officielles des Communautés européennes : Luxembourg.

Gallego, F.J. et Delincé, J. 2010. « The European Land Use and Cover Area-frame statistical Survey (LUCAS) ». In : Benedetti, R., Bee, M., Espa, G. et Piersimoni, F. (dir.), *Agricultural Survey Methods* (ch. 10, p. 151–68). John Wiley & Sons : Chichester, Royaume-Uni.

Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A. et Kussul, O. 2014. « Efficiency assessment of using satellite data for crop area estimation in Ukraine ». *International Journal of Applied Earth Observation and Geoinformation*, vol. 29, p. 22–30.

Gallego, F.J., Peedell, S. et Carfagna, E. 1999. « The use of CORINE Land Cover to improve area frame survey estimates ». *Research in Official Statistics*, vol. 2, n° 2, p. 99–122.

Hanuschak, G.A., Allen, R.D. et Wigton, W.H. 1982. « Integration of Landsat data into the crop estimation program of USDA's statistical reporting service (1972-1982) ». In : *Proceedings of 8th International Symposium on Machine Processing of Remotely Sensed Data with Special Emphasis on Crop Inventory and Monitoring*. Publication conjointe de LARS et de la Purdue University : West Lafayette, IN, États-Unis.

Hoffman, S. 2016. *Preparation of the USDA World Ag Supply & Demand Estimates (WASDE) Report*. Article présenté lors du Colloque international sur les approches et méthodes de contrôle des cultures et de prévision de la production, 25–26 mai 2016. Dhaka.

Holmes, Q.A., Horvath, R., Cicone, R.C., Kauth, R.J. et Malila, W.A. 1979. *Development of Landsat-Based Technology for Crop Inventories*. Rapport technique SR-E9-00404 de la NASA. Publication de la NASA.

Husak, G.J., Marshall, M.T., Michaelsen, J., Pedreros, D., Funk, C. et Galu, G. 2008. « Crop area estimation using high and medium resolution satellite imagery in areas with complex topography ». *Journal of Geophysical Research*, vol. 113, n° D14.

Justice, C. et Becker-Reshef, I. 2007. *Developing a Strategy for Global Agricultural Monitoring in the Framework of Group on Earth Observations (GEO) Workshop Report*, 16–18 juillet 2007. Rome, FAO.

Li, Y., Zhu, X., Pan, Y., Gu, J., Zhao, A. et Liu, X. 2014. « A Comparison of Model-Assisted Estimators to Infer Land Cover/Use Class Area Using Satellite Imagery ». *Remote Sensing*, vol. 6, n° 9, p. 8904–22.

Lillesand, T.M., Kiefer, R.W. et Chipman, J. 2015. *Remote Sensing and Image Interpretation*. 7^e édition. John Wiley & Sons : New York.

MacDonald, R.B., Hall, F.G. et Erb, R.B. 1975. *The Use of LANDSAT Data in a Large Area Crop Inventory Experiment (LACIE)*. Document 46 du Colloque sur le traitement automatisé des données de télédétection, 3–5 juin 2017. West Lafayette, IN, États-Unis, Laboratory for Applications of Remote Sensing et Purdue University.

MacDonald, R.B. et Hall, F.G. 1980. « Global crop forecasting ». *Science*, vol. 208, n° 4445, p. 670–9.

MacDonald, R.B. 1984. « A summary of the history of the development of automated remote sensing for agricultural applications ». *IEEE Transaction on Geoscience & Remote Sensing*, vol. GE-22, n° 6, p. 473–81.

Navalgund, R.R. et Sahai, B. 1985. « Remote sensing applications in agriculture - Indian experience and plans ». In : *Proceedings of the Fourth Asian Agricultural Symposium* (p. 329–343), 28 février–1^{er} mars 1985. Kumamoto, Japon, Université Kyushu Tokai.

Nellis, M.D., Price, K.P. et Rundquist, D. 2009. *Remote Sensing of Cropland Agriculture*. Papers in Natural Resources. Article 217. Publication de l'Université du Nebraska–Lincoln : Lincoln, NE, États-Unis.

Nelson A. et al. 2014. « Towards an Operational SAR-Based Rice Monitoring System in Asia: Examples from 13 Demonstration Sites across Asia in the RIICE Project ». *Remote Sensing*, vol. 6, n° 11, p. 10773–812.

Nelson, A. et al. 2014. « Towards an operational SAR-based rice monitoring system in Asia: examples from 13 demonstration sites across Asia in the RIICE project ». *Remote Sensing*, Édition spéciale sur la télédétection au service de la production et de la sécurité alimentaires, vol. 6, n° 11, p. 10773–812.

Pan, Y., Zhang, J., Zhou, W., Zhao, J. et Yu, X. 2012. *Methodology of the Crops Acreage Estimation Using Remote Sensing and Survey Sampling in National statistical system of China*. Article rédigé pour la 1^{re} Conférence internationale sur l'agro-géo-informatique, 2–4 août 2017. Shanghai, Chine.

Panigrahy, S., Chakraborty, M., Manjunath, K.R., Kundu, N. et Parihar, J.S. 2000. « Evaluation of Radarsat ScanSAR synthetic aperture radar data for rice crop inventory and modeling ». *Journal of Indian Society of Remote Sensing*, vol. 28, p. 59–65.

Parihar, J.S. et Oza, S.R. 2006. « FASAL: An integrated approach for crop assessment and production forecasting ». In : Kuligowski, R.J., Parihar, J.S. et Saito, G. (dir.), *Proceedings of SPIE Volume 6411, Agriculture and Hydrology Applications of Remote Sensing*. Publication de la SPIE : Goa, Inde.

Patel, N.K., Medhavv, T.T., Patnaik, C. et Hussain, A. 1995. « Multi-temporal ERS-1 SAR data for identification of rice crop ». *Journal of Indian Society of Remote Sensing*, vol. 23, p. 33–9.

Potgieter, A.B., Apan, A., Dunn, P.K. et Hammer, G.L. 2007. « Estimating crop area using seasonal time series of Enhanced Vegetation Index from MODIS satellite imagery ». *Crop and Pasture Science*, vol. 58, n° 4, p. 316–25.

Prasad, S.V.S., Satya Savithri, T. et Murali Krishna, I.V. 2015. « Techniques in Image Classification; A Survey ». *Global Journal of Researches in Engineering: Electrical and Electronics Engineering*, vol. 15, n° 6, p. 17–32.

Radhakrishnan, K., Jayaraman, V. et Nageswara Rao, P.P. 1991. « The Economics of Remote Sensing ». *Current Science*. vol. 61, n° 3 et 4, p. 272–7.

Ray, S.S., Mamatha, S., Manjunath, K.R., Uday, R., Seshasai, M.V.R., Singh, K.K., Kimothi, M.M., Parihar, J.S. et Saxena, M. 2016. « CHAMAN: A National Level Programme for Horticultural Assessment & Development ». *Bulletin of the National Natural Resources Management System, NNRMS (B)*, vol. 40, p. 1–6.

Ray, S.S., Neetu, Mamatha, S. et Gupta, S. 2015. « Use of Remote Sensing in Crop Forecasting and Assessment of Impact of Natural Disasters: Operational Approaches in India ». In : Srivastava, M.K. (dir.), *Crop Monitoring for Improved Food Security: Proceedings of the Expert Meeting. Vientiane, République démocratique populaire lao, 17 février 2014* (p. 111–122). Publication RAP 2014/28. Publication conjointe de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) et de la Banque asiatique de développement (BAsD) : Rome – Manille.

Ray, S.S., Neetu, Manjunath, K.R. et Singh, K.K. 2016. *Crop Production Forecasting using Space, Agro-meteorology and Land based Observations: Indian Experience*. Article présenté lors du Colloque international sur les approches et méthodes de contrôle des cultures et de prévision de la production, 25–26 mai 2016. Dhaka.

Sahai, B. et Dadhwal, V.K. 1990. « Remote sensing in agriculture ». In : Verma, J.P. et Verma, A. (dir.), *Technology Blending and Agrarian Prosperity* (p. 83–98). Malhotra Publishing House : New Delhi.

Sahai, B., Chandrasekhar, S., Barde, N.K. et Nag Bhushna, S.R. 1977. « Agricultural resources inventory and surveys experiment ». In : Rycroft, M.J. et Stickland, A.C. (dir.), *COSPAR Space Research* (p. 3–8). Pergamon Press : Oxford, Royaume-Uni.

Sharman, M.J. 1993. « The agriculture project of the joint research centre: Operational use of remote sensing for agricultural statistics ». In : *Proceedings of the ITC International Symposium Operationalization of Remote Sensing* (p. 46–57), 19–23 avril 1993. Enschede, Pays-Bas.

Sharples, J.A. 1973. *The Corn Blight Watch Experiment: Economic Implications for Use of Remote Sensing for Collecting Data on Major Crops*. Rapport technique du LARS, article 121. Publication de la Purdue University : West Lafayette, IN, États-Unis.

Stehman, S.V. 2013. « Estimating area from an accuracy assessment error matrix ». *Remote Sensing of Environment*, vol. 132, p. 202–11.

Suga, Y. et Konishi, T. 2008. « Rice crop monitoring using X-, C- and L-band SAR data ». In : Neale, C.M.U., Owe, M. et d'Urso, G. (dir.), *Proceedings of SPIE Volume 7104, Remote Sensing for Agriculture, Ecosystems, and Hydrology X*. Publication de la SPIE : Cardiff, Pays de Galles, Royaume-Uni.

Temnikov, V. et Sergey, I. 2007. *Use of Geo-informational Systems for the Russian Agriculture Census*. Article rédigé pour la 4^e Conférence internationale sur les statistiques de l'agriculture (ICAS IV) : progresser en matière d'intégration et d'analyse statistique, 22–24 octobre 2007. Beijing, Chine.

Vogel, F. et Bange, G. 1999. *Understanding USDA Crop Forecasts*. Miscellaneous Publication No. 1554. Publication conjointe USDA/NASS et WOAB : Washington, D.C, États-Unis.

Whitworth, A. (dir.). 2013. *Evaluations and improvements in small area estimation methodologies*. Discussion Paper. National Centre for Research Methods (NCRM), Publication de l'Université de Sheffield : Sheffield, Royaume-Uni.

Wu, B. et Li, Q. 2012. « Crop planting and type proportion method for crop acreage estimation of complex agricultural landscapes ». *International Journal of Applied Earth Observation*, vol. 16, p. 101–12.

Wu, B., Gommès, R., Zhang, M., Zeng, H., Yan, N., Zou, W., Zheng, Y., Zhang, N., Chang, S., Qiang Xing, Q. et van Heijden, A. 2015. « Global Crop Monitoring: A Satellite-Based Hierarchical Approach ». *Remote Sensing*, vol. 7, n° 4, p. 3907–33.

Zhou, W. 2013. *Remote Sensing Application in China's Crop Acreage Estimation*. Article élaboré en vue du Congrès mondial 2013 sur les statistiques, 25–30 août 2013. Chine, RAS Hong Kong.

Zhou, W. 2016. *Small Area Estimation for Crop Acreage in Remote Sensing Assisted Crop Survey – A Case of Major Crop Acreage Estimation in Liaozhong County*. Article rédigé pour la 7^e Conférence internationale sur les statistiques de l'agriculture (ICAS VII), 26–28 octobre 2016. Rome.



6

Chapitre 6

Système d'Alerte Précoce et estimation des rendements agricoles

Oscar Rojas, spécialiste des ressources naturelles, Bureau sous-régional pour l'Amérique centrale, Division du climat et de l'environnement de la FAO

6.1. INTRODUCTION

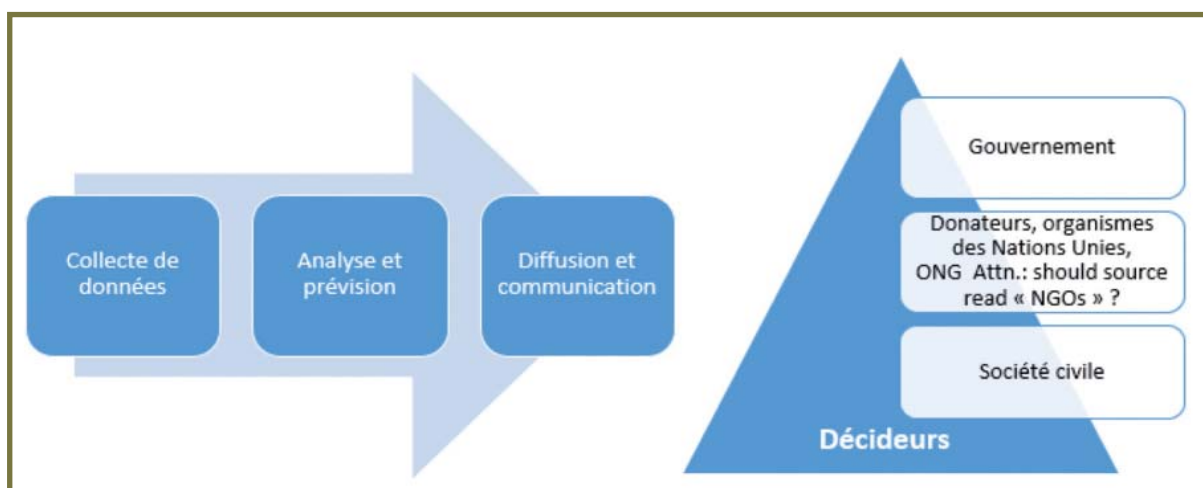
Ce chapitre passe en revue les concepts fondamentaux relatifs aux systèmes d'alerte précoce (SAP) et à la prévision des rendements agricoles, afin de mieux gérer les risques climatiques ayant une incidence sur la sécurité alimentaire. Le déclenchement des alertes précoces et la définition de la probabilité du risque à l'échelle du système sont des composantes fondamentales de l'information d'alerte précoce (IAP) ; des liaisons systématiques avec les solutions et les moyens d'alerte précoce permettraient de sauver bon nombre de vies et de moyens d'existence.

Il est essentiel que les systèmes d'alerte précoce visant à évaluer la disponibilité et la demande alimentaires d'une région ou d'un pays donnés puissent prévoir les rendements agricoles et la production globale. L'analyse méthodique des conditions de marché et le repérage des zones de surplus et de déficit d'un pays ou d'une région sont extrêmement utiles pour concevoir des politiques de lutte contre l'insécurité alimentaire adaptées. Des statistiques agricoles fiables et précises sont indispensables pour atteindre ces objectifs importants. Dans ce contexte, les données de télédétection sont essentielles à l'amélioration de la production de statistiques agricoles, car elles permettent d'introduire des mécanismes de vérification indépendants, notamment en cas de recours à des modèles d'échantillonnage utilisant une base aréolaire ou multiple. Les informations et données de télédétection peuvent être introduites au moment de la conception comme de l'estimation.

6.2. SYSTÈME D'ALERTE PRÉCOCE

Un système d'alerte précoce (SAP) se définit comme un système intégré de suivi, de collecte et d'analyse de données, et de communication permettant de prendre rapidement des décisions visant à protéger les populations et l'environnement (Davies *et al.*, 1991). Les SAP consacrés à l'agriculture et à la sécurité alimentaire¹ surveillent l'accès de la population aux aliments, dans le but d'émettre des avis en temps utile sur la survenue prochaine d'une crise alimentaire et déclencher ainsi une intervention appropriée. Leur réussite repose sur de nombreux facteurs, dont la plupart échappent à leur contrôle (Buchanan-Smith, 2000). Tout SAP consacré à la sécurité alimentaire doit en permanence surveiller la production, les stocks, les prix, le commerce et la consommation de denrées alimentaires, ainsi que tout phénomène extrême, y compris les bouleversements climatiques et les perturbations du marché pouvant affecter l'offre et la demande alimentaires. L'encadré 1 présente le contexte historique dans lequel s'inscrivent les systèmes d'information sur la sécurité alimentaire. La forme précise des SAP dépend de différents facteurs, tels que l'environnement des données, l'infrastructure de communication, le public visé et la nature du risque (Bailey, 2013). Ils partagent néanmoins tous certains points communs — qu'ils portent sur les pénuries alimentaires ou sur d'autres risques : ils collectent des données d'alerte précoce, qu'ils analysent pour produire des informations d'alerte précoce (IAP) qui seront par la suite communiquées aux décideurs (figure 1).

FIGURE 1. CARACTÉRISTIQUES DES SYSTÈME D'ALERTE PRÉCOCE.



La phase de **collecte de données** correspond au processus de rassemblement et de compilation des données d'alerte précoce, qui doivent être pertinentes, opportunes et fiables. La phase d'**analyse et de prévision** englobe les activités techniques de suivi et de production d'informations d'alerte précoce. Dans ce cadre, une importance particulière est accordée aux facteurs suivants : paramètres surveillés, méthode employée, prévision de certaines variables spécifiques et niveau de confiance accordé à tous ces éléments. La phase de **diffusion et de communication** désigne la transmission d'informations d'alerte précoce à certaines parties prenantes compétentes. Les principaux éléments à prendre en considération concernent ici les modes de diffusion des alertes, leurs destinataires et la forme que prennent les informations communiquées.

Rapidité et confiance sont fondamentalement antagonistes dans de nombreux SAP. La confiance accordée à la pertinence des informations communiquées s'accroît inévitablement avec le temps, à mesure que de nouvelles données sont collectées et analysées, ce qui réduit dans le même temps le délai d'alerte. La richesse des informations

1 « La sécurité alimentaire existe lorsque tous les êtres humains ont, à tout moment, la possibilité physique, sociale et économique de se procurer une nourriture suffisante, saine et nutritive leur permettant de satisfaire leurs besoins et préférences alimentaires pour mener une vie saine et active ». Sommet mondial de l'alimentation, 1996.

communiquées peut également être problématique. Plus les parties prenantes reçoivent d'informations d'alerte précoce, plus elles devraient être en mesure de prendre des décisions éclairées, ce qui est pourtant rarement le cas dans la pratique : elles peuvent en effet manquer de moyens leur permettant d'interpréter correctement les informations, le message fondamental étant alors perdu.

Idéalement, le SAP consacré aux pénuries alimentaires doit donc être en mesure d'anticiper à la fois les crises des moyens d'existence et les urgences humanitaires, et admettre qu'elles constituent deux phases d'un même processus. En surveillant certains facteurs de risque centraux tels que la météo, les récoltes et les marchés et en étudiant les données économiques des ménages, il doit pouvoir déceler les stress et les chocs liés aux moyens d'existence, et fournir ainsi des orientations aux interventions visant à consolider ces derniers et à éviter la mise en place de stratégies d'adaptation préjudiciables. Les SAP doivent pouvoir détecter le potentiel de transformation d'une crise en véritable urgence, permettant ainsi aux acteurs de se préparer à déployer une intervention humanitaire. À chaque étape du processus, l'action rapide doit chercher à empêcher la crise de dégénérer et à atténuer ses répercussions sur la vie et les moyens d'existence. Les calendriers des crises permettent de déterminer le moment opportun pour intervenir (Bailey, 2013).

Sivakumar (2014) appelle, dans le cadre des sécheresses actuelles, à la mise en place future d'actions proactives par les différents intervenants, afin qu'ils soient en mesure de faire face à leurs impératifs. Malgré l'apparition régulière de sécheresses tout au long de l'histoire de l'humanité et leurs effets à grande échelle sur plusieurs secteurs socio-économiques, aucun effort concerté n'a jusqu'à présent été déployé pour amorcer un dialogue sur la formulation et l'adoption de politiques nationales ciblées de lutte contre la sécheresse. L'absence de politiques de ce type implique que les autorités nationales, régionales et communautaires maintiendront le statu quo en se contentant de réagir, chacune de leur côté, aux répercussions de ces catastrophes. Selon Bailey (2013), la modification du statu quo ne peut se concrétiser que si les gouvernements savent qu'ils seront récompensés pour leur action en faveur de la réduction des risques de pénurie et sanctionnés pour leur immobilisme.

Une sécheresse (un risque) ne déclenche pas à elle (lui) seul(e) une situation d'urgence. Qu'il ou elle se transforme en situation d'urgence ou en catastrophe dépend de ses répercussions sur les communautés locales et l'environnement. Répercussions qui dépendent à leur tour de la vulnérabilité de ces derniers face à un « choc » de cette ampleur.

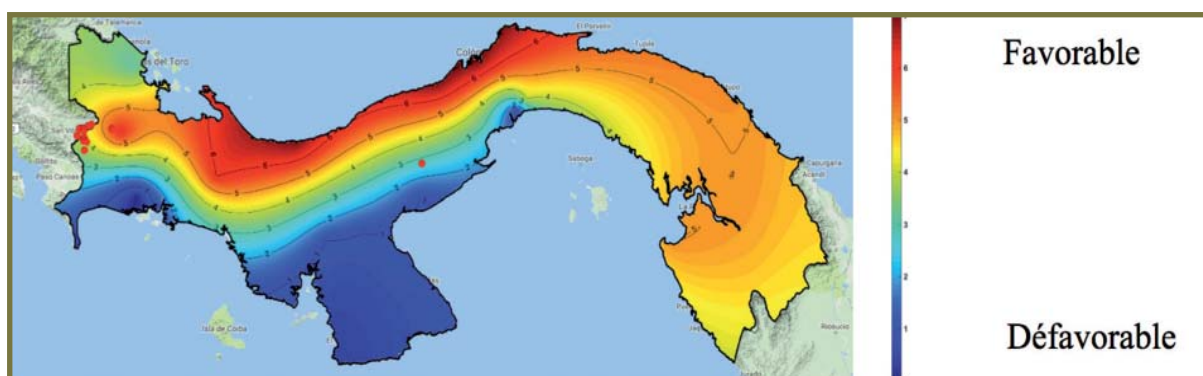
6.2.1. Ravageurs et maladies des plantes

Selon les estimations, les ravageurs causent 30 à 40 % de pertes de récoltes chaque année. Une baisse de 1 % seulement de ces pertes permettrait à des millions de personnes supplémentaires de se nourrir correctement. Plus d'un tiers de la population mondiale — qui devrait atteindre les 9 milliards d'ici à 2050 — se nourrit grâce à 500 millions de petits agriculteurs. Il est donc urgent, au nom de la sécurité alimentaire mondiale, d'aider ces derniers à lutter contre les ravageurs. En 1994, la FAO a créé l'Observatoire acridien², qui exploite les capacités des satellites actuels pour fournir continuellement un ensemble d'informations sur les conditions écologiques, telles que les précipitations et le développement de la végétation, qui sont des facteurs importants, utiles pour surveiller les habitats du criquet pèlerin et prévoir tout développement acridien. SATCAFE³, lui aussi développé récemment, est un système de surveillance des conditions climatiques pouvant favoriser la prolifération de la rouille du caféier (*Hemileia vastatrix* ; voir la figure 2). Il détecte, grâce aux technologies mobiles, le niveau d'infection dans les plantations de café. La rouille est la maladie du caféier qui a les plus importantes répercussions économiques au monde, tandis que le café est le produit agricole le plus important, en termes monétaires, des échanges internationaux. Une baisse, même minime, des récoltes de café ou une légère augmentation des coûts de production, causées par la rouille ont donc d'importantes conséquences sur les producteurs, les services d'appui connexes, voire les systèmes bancaires des pays tributaires des recettes provenant de l'exportation de café.

² <http://www.fao.org/ag/LOCUSTS/fr/activ/DLIS/satel/index.html>

³ <http://www.siatma.org/>

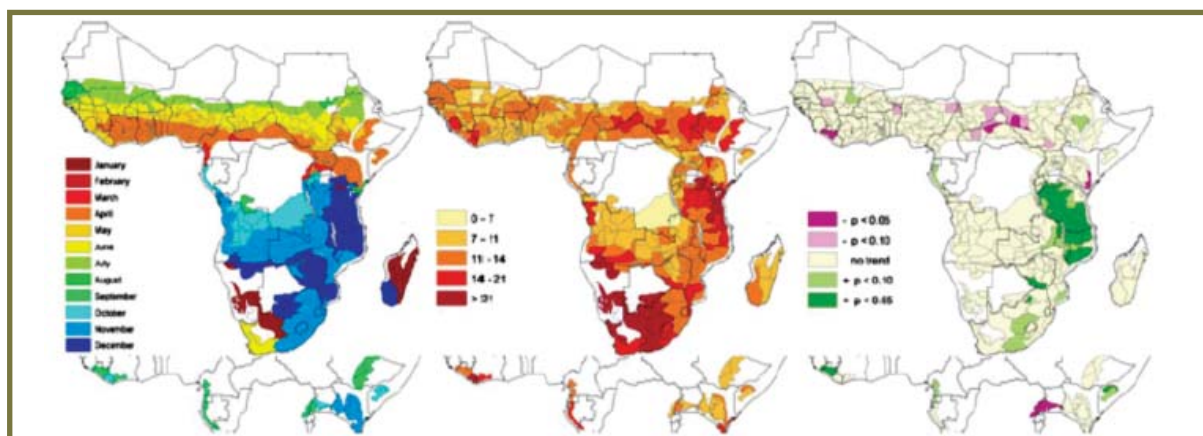
FIGURE 2. CONDITIONS CLIMATIQUES FAVORISANT OU NON LE DÉVELOPPEMENT DE LA ROUILLE DU CAFÉIER AU PANAMA.



6.2.2. Le modèle phénologique de surveillance de la végétation

La phénologie des végétaux (l'étude des cycles récurrents de végétation et leurs liens avec le climat) est une variable importante pour de nombreuses applications des sciences de la Terre et de l'atmosphère. À mesure que l'intérêt pour la recherche sur les changements à l'échelle du globe s'accroît, notamment, les modèles phénologiques faisant preuve de précision deviendront des outils de plus en plus incontournables, puisqu'ils permettent aux chercheurs de surveiller et d'anticiper les réactions des végétaux face aux variations climatiques d'une année sur l'autre. Aux débuts de la recherche sur la phénologie par satellite, Justice *et al.* (1985) ont évalué qualitativement la phénologie mondiale de nombreux types de couvertures du sol en s'appuyant sur l'indice différentiel normalisé de végétation (NDVI) (voir la section 6.4.2 ci-dessous). Goward *et al.* (1985) ont prouvé que les tendances observées du NDVI correspondaient à la variation saisonnière connue des végétaux aux États-Unis continentaux. Par la suite, les chercheurs ont utilisé des satellites pour interpréter la phénologie comme indicateur de changement de la couverture du sol en Amérique du Sud (Stone *et al.*, 1994) et déceler la dynamique phénologique des fruticées (Duncan *et al.*, 1993). Vrieling *et al.* (2011), dont les recherches portent sur les NDVI cumulés au cours de la saison (qui servent à calculer la productivité primaire nette), décrivent la phénologie des végétaux d'Afrique subsaharienne et évaluent la variabilité et les tendances des indicateurs phénologiques en se fondant sur des séries temporelles des NDVI couvrant la période 1982-2006 (figure 3).

FIGURE 3. MOYENNE EN DÉBUT LA SAISON (GAUCHE), ÉCART-TYPE TEMPOREL (σ_t) EN DÉBUT DE SAISON (AU MILIEU, EN JOURS) ET CORRÉLATION DE SPEARMAN EN DÉBUT DE SAISON (DROITE), FONDÉS SUR LA SÉRIE TEMPORELLE NDVI CALCULÉE PAR L'AVHRR. LA PARTIE INFÉRIEURE INDIQUE LES VALEURS RELATIVES À UNE ÉVENTUELLE DEUXIÈME SAISON, LE CAS ÉCHÉANT.

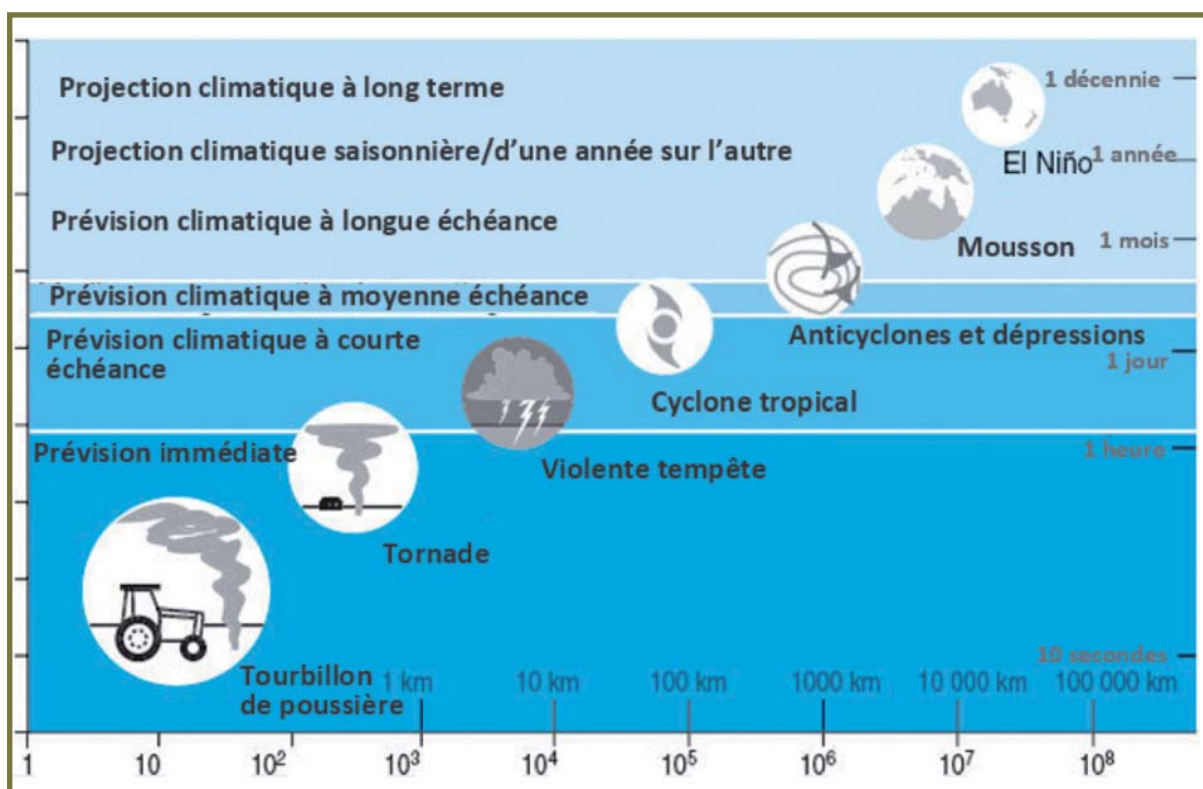


6.2.3. Quels délais pour une alerte précoce ?

La figure 4 indique les délais d'alerte relatifs aux risques climatiques. Elle montre que les délais d'alerte des SAP peuvent durer de quelques secondes pour les séismes à quelques mois pour les sécheresses, respectivement pour les risques les plus aux moins soudains. Plus spécifiquement, les SAP peuvent déclencher des alertes quelques dizaines de secondes avant la survenue d'un séisme, quelques jours ou heures avant une éruption volcanique, quelques heures avant un tsunami, et quelques minutes avant une tornade. Le délai d'alerte ouragan varie de quelques semaines à quelques heures. Le délai d'alerte peut durer plusieurs années, voire plusieurs décennies, pour les menaces à évolution lente (El Niño, le changement climatique, etc.). En ce qui concerne les sécheresses, le délai d'alerte varie de plusieurs mois à plusieurs semaines.

Les phénomènes à évolution lente peuvent gravement nuire à l'environnement et à la société, en l'absence d'adoption opportune de mesures préventives. Ces évolutions environnementales progressives et cumulatives nécessitent des technologies d'alerte précoce efficaces compte tenu de leurs retombées potentiellement importantes sur la société et l'environnement (PNUE, 2012).

FIGURE 4. QUELS DÉLAIS POUR UNE ALERTE RAPIDE? CE GRAPHIQUE MONTRE LE DÉLAI DE RÉACTION DES SAP ET LES SUPERFICIES D'IMPACT (EN PRÉCISANT LE DIAMÈTRE DE L'AIRE SPHÉRIQUE) POUR DIFFÉRENTS RISQUES HYDROMÉTÉOROLOGIQUES. SOURCE : GOLNARAGHI, 2005.



6.3. ALERTE ET INTERVENTION PRÉCOCES

La Croix-Rouge (2008) définit l'alerte et l'intervention précoces comme la mise en place d'actions humanitaires régulières, avant la survenue d'une catastrophe ou d'une situation d'urgence sanitaire, en mettant à profit toutes les informations scientifiques disponibles à court, moyen et long terme. L'évacuation de la population avant une inondation est toujours préférable aux actions de sauvetage pendant la catastrophe ou à la fourniture ultérieure d'une aide aux victimes. Il est également beaucoup plus efficace d'aider les agriculteurs à trouver d'autres moyens de subsistance en cas de mauvaise récolte que de leur procurer une aide alimentaire. Les incroyables avancées dans le domaine des sciences et technologies ont ouvert l'accès à un grand nombre de SAP. Les modèles de circulation générale (MCG) et les images satellite, les centres régionaux d'expertise, les services météorologiques nationaux et d'autres agences publiques, les rapports d'enquête locaux et les observations des communautés permettent tous de mieux comprendre les phénomènes et leurs possibles conséquences, ce qui permet de beaucoup mieux anticiper les menaces liées au climat. À court terme, les alertes de tempêtes imminentes peuvent aider les communautés à se préparer et à réagir immédiatement (p. ex., en évacuant les lieux) pour réduire les pertes humaines. À moyen terme, une prévision saisonnière fondée sur El Niño peut souligner la particulière intensité de la saison des tempêtes à venir, ou indiquer qu'une sécheresse persistante pourra entraîner une pénurie alimentaire. À plus longue échéance, les scénarios de l'évolution future du climat alertent la population à l'avance sur la multiplication des risques, fournissant ainsi une nouvelle analyse des risques tenant compte d'autres tendances telles que l'urbanisation et la croissance démographique.

À chaque extrémité de cette échelle temporelle, le risque de catastrophe correspond à la corrélation entre les dangers (cyclones, tempêtes, sécheresses, etc.) et le degré de vulnérabilité des communautés. Ces deux éléments sont en constante évolution. D'une part, la fréquence, l'intensité et la localisation des menaces fluctuent en raison du changement climatique. D'autre part, les phénomènes tels que l'urbanisation, la pauvreté, la croissance démographique ou les épidémies font sans cesse évoluer la nature de la vulnérabilité. La réduction des risques de catastrophe consiste à produire des cartes détaillées des risques, mais permet surtout d'observer sans cesse la nature évolutive des dangers et de la vulnérabilité et de prendre des mesures contre cette dernière et ses causes sous-jacentes.

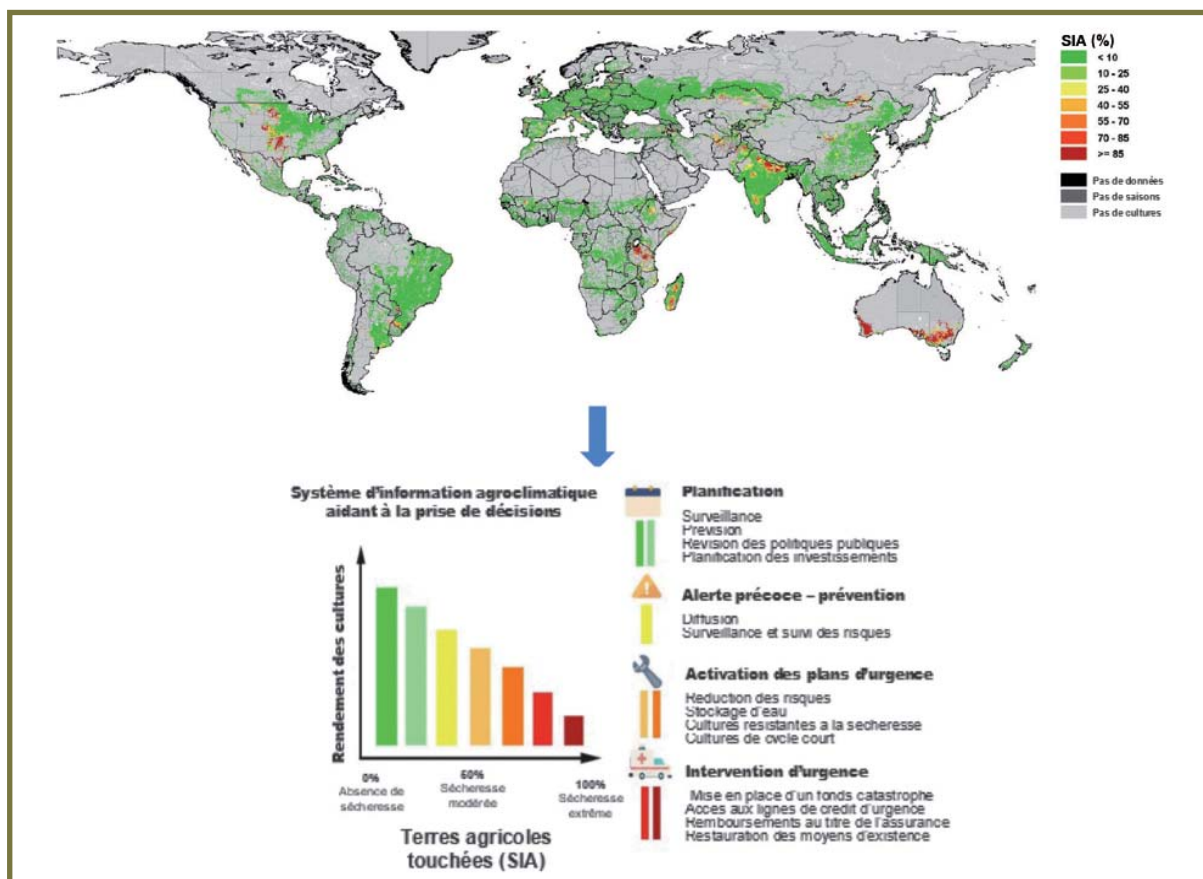
Les alertes précoces sont inutiles si elles ne sont pas suivies d'une intervention rapide. Le cyclone Nargis, l'ouragan Katrina et la crise alimentaire au Niger ne sont que quelques exemples parmi tant d'autres de catastrophes non évitées en raison d'un manque de réaction malgré l'existence d'informations fiables sur leur imminence. À l'échéance la plus courte, l'action pourrait prendre la forme d'une évacuation. À long terme, l'intervention rapide suppose de travailler en étroite collaboration avec les communautés locales afin d'évaluer les causes profondes des risques évolutifs auxquels elles sont confrontées et de les combattre. Parmi les innombrables mesures de réduction des risques qui existent, citons par exemple la construction de maisons sur pilotis, la plantation d'arbres pour prévenir les glissements de terrain, l'organisation de campagnes de sensibilisation et de prévention de la dengue, ou encore l'installation de systèmes de captage des eaux. La mise à jour des plans d'urgence et la mobilisation de bénévoles sont également des mesures relevant de l'intervention rapide. Sur le plan de la portée géographique, l'intervention rapide peut prendre différentes formes : à l'échelle locale, si les prévisions annoncent la survenue d'une inondation importante, la communauté devra en priorité chercher à protéger sa source d'eau principale de la contamination. À l'échelle nationale, le gouvernement peut actualiser ses plans d'urgence. Au niveau mondial, les institutions internationales peuvent mobiliser des ressources humaines et financières avant la catastrophe pour aider les pays exposés à réduire les conséquences de ces dangers, voire à empêcher les pertes humaines. Plus le nombre de mesures prises en réponse aux alertes à longue échéance est élevé — identification des communautés à risque, investissement dans la réduction des risques de catastrophe, amélioration de la capacité à intervenir — plus le nombre de vies et de moyens d'existence sauvés à courte échéance lors de l'inondation proprement dite sera important. De même, le renforcement des rapports entretenus avec les centres de connaissance mondiaux et régionaux et la mise en place de procédures normalisées de diffusion des informations auprès des acteurs appropriés amélioreront l'efficacité des interventions locales.

Il est généralement possible de déterminer avec précision, quelques heures à l'avance, où et quand aura lieu une grosse tempête. Mais pour faciliter les interventions dans ce cadre, il faut investir bien à l'avance dans la création d'un système global de gestion des urgences. Une période d'alerte de quelques jours permet la prise immédiate de mesures de préparation à la catastrophe : détermination des voies d'évacuation, mise en place de centres d'évacuation, protection des biens et mobilisation des organisateurs communautaires dans le cadre d'une intervention immédiate. Une alerte à plus longue échéance (plusieurs mois, voire plusieurs années à l'avance) sur la nature évolutive du risque de tempête permet, quant à elle, de multiplier les mesures de réduction des risques de catastrophe : aide à la plantation d'arbres par la communauté pour stabiliser les pentes, organisation communautaire pour réagir plus efficacement aux alertes, construction d'habitations anticycloniques, demande de construction d'abris, etc. La prise de conscience du risque accru suppose davantage d'investissements dans le renforcement des capacités à prendre des mesures précoces utiles et ce, quels que soient le moment et le lieu de survenue de la catastrophe.

L'exploitation de ces informations sur les risques peut également entraîner la prise de « mauvaises » décisions : par exemple, une prévision annonçant la probabilité à 80 % de rafales de vent de la force d'un ouragan à un moment et à un endroit précis. Même si les probabilités sont fortes, il n'y a aucune certitude que cet événement aura bien lieu — il a en effet 20 % de chances de ne pas se produire. Lors de la promotion d'actions rapides, ces doutes ne doivent pas être dissimulés : une description honnête des certitudes et des incertitudes vis-à-vis de l'avenir doit faire partie intégrante de toute communication à l'ensemble des parties prenantes, tout en occupant une place centrale dans l'évaluation et la prise en compte des risques.

Il est indispensable de transformer les informations scientifiques — qui se présentent souvent sous forme de cartes ou de pourcentages complexes — en messages simples et compréhensibles permettant aux personnes en danger de prendre des décisions judicieuses sur la meilleure façon de réagir à une menace imminente. Par exemple, Rojas (2016) propose d'utiliser le Système d'indice de stress agricole (ASIS, de l'anglais *Agriculture Stress Index System*) pour donner le coup d'envoi aux activités d'atténuation des effets de la sécheresse en milieu agricole (figure 5).

FIGURE 5. SYSTÈME D'ALERTE PRÉCOCE SUR LES SÉCHERESSES AGRICOLES ET INTERVENTION RAPIDE. LA CARTE REPRÉSENTE LE POURCENTAGE DE TERRES AGRICOLES TOUCHÉES PAR UNE SÉCHERESSE (SYSTÈME D'INFORMATION AGROCLIMATIQUE [SIA]) EN 2006.



Le graphique à barres associe le SIA aux activités d'atténuation des effets de la sécheresse. Source : Rojas, O., 2016.

Les progrès continus réalisés dans le domaine des technologies de télédétection, l'amélioration des prévisions et des modèles météorologiques, les nouvelles perspectives offertes par les technologies de l'information et de la communication, et les nombreuses possibilités d'étendre la couverture et la capacité des SAP dans les pays et les régions vulnérables signalent une capacité de prédiction toujours plus efficace. Les possibilités offertes aux gouvernements, aux organismes et aux prestataires de services d'alerte précoce pour favoriser ces avancées sont infinies. Cependant, en l'absence de réformes significatives dans d'autres domaines, les interventions rapides ne connaîtront pas d'améliorations similaires (Bailey, 2013).

6.4. PRÉVISION DES RENDEMENTS AGRICOLES

Les projections constituent la base de tout système d'alerte. Elles doivent être appliquées aux quatre domaines de la sécurité alimentaire (disponibilité, stabilité, accès et utilisation biologique) afin de donner assez de temps aux décideurs pour réagir en cas d'alerte, et être suffisamment fiables (en règle générale, plus elles portent sur le long terme, moins elles sont fiables) afin d'éviter les fausses alertes. Toutes les prévisions s'accompagnent d'une cote de probabilité (qui peut ou non être calculable) qui constitue un bon indicateur de leur fiabilité.

Passant en revue certaines des méthodes utilisées, Basso *et al.* (2015) ont constaté que la prévision des rendements agricoles devait impérativement passer par le recours à des modèles de simulation des cultures, c'est-à-dire des

représentations informatisées de la croissance des plantes, de leur développement et des récoltes correspondantes, simulées à l'aide d'équations mathématiques exprimant les fonctions de l'état du sol, des conditions météorologiques et des pratiques de gestion (Hoogenboom *et al.*, 2004). Ces outils sont divisés en modèles statistiques, mécaniques et fonctionnels. Les modèles de cultures ne constituent qu'une approximation du monde réel, et ils sont nombreux à ne pas tenir compte de facteurs importants tels que les adventices, les maladies, les insectes, les techniques de préparation du sol et les nutriments (Jones *et al.*, 2001). Ils ont toutefois joué un rôle important dans l'interprétation de résultats agronomiques, et les agriculteurs s'en servent de plus en plus fréquemment comme outils d'aide à la prise de décision. Il existe des modèles simples et des modèles complexes. Les premiers, qui recèlent peu de détails sur le système sol-plante, sont souvent utilisés pour estimer les rendements sur les grandes surfaces agricoles, à partir d'informations statistiques liées au climat et aux rendements obtenus dans le passé. Les mécanismes nationaux d'alerte précoce utilisent des modèles simples dans la plupart des cas.

L'analyse des données météorologiques et climatiques permet la communication quasiment en temps réel d'informations qualitatives et quantitatives sur l'état des cultures, offrant ainsi la possibilité d'émettre des alertes rapides en cas de situation préoccupante, afin que soient planifiées et organisées des interventions en temps opportun (Gommes *et al.*, 2006). La philosophie sous-tendant la prévision des récoltes se fonde sur différents types de données (p. ex., météorologiques, agrométéorologiques [phénologie, rendement], relatives aux sols [capacité de rétention d'eau], statistiques agricoles) collectées à partir de différentes sources (p. ex., la télédétection). Plusieurs indices, considérés comme des variables pertinentes pour la détermination des rendements agricoles, sont calculés à partir des données météorologiques et agronomiques : citons par exemple la satisfaction en eau des cultures, les surplus et excédents d'humidité, l'humidité moyenne des sols, etc.

Les résultats sont associés de façon empirique aux rendements agricoles à l'aide de techniques standard de régression, un processus connu sous le nom d'« étalonnage des modèles » qui donne lieu à une formule mathématique – la « fonction de rendement » – utilisée pour calculer les estimations de rendements sur la base des résultats du modèle. Les statistiques administratives disponibles relatives aux cultures sont utilisées pour étalonner les modèles de régression, dont la précision ultime dépend donc, dans une large mesure, de la qualité de ces statistiques d'entrée.

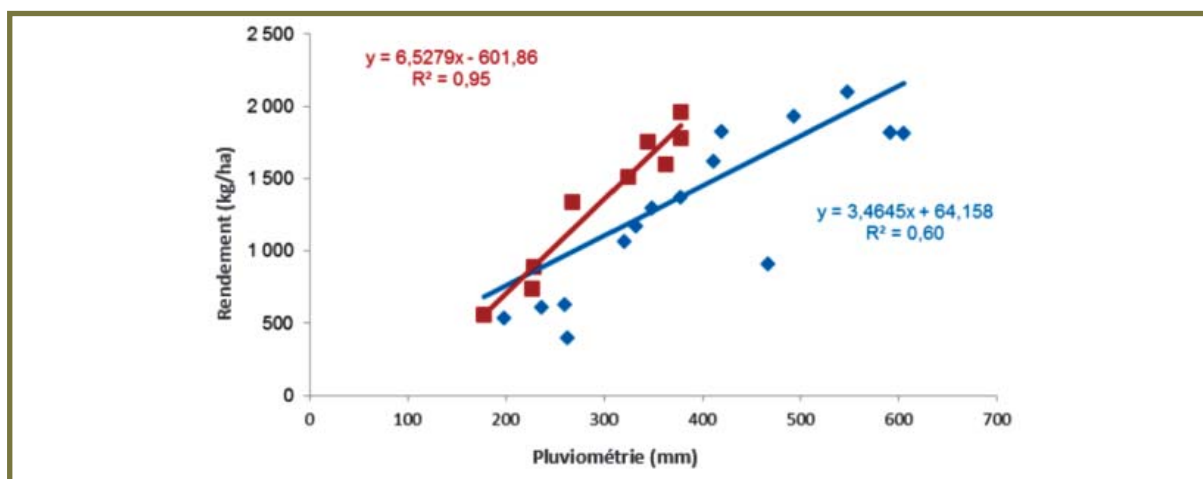
Comme l'a relevé Pulwarty (2007), le délai et la forme des données climatiques d'entrée (y compris les prévisions et projections), l'accès à des conseils avisés et la capacité à interpréter et à mettre en œuvre des informations et projections dans le cadre de processus décisionnels, sont tout aussi importants pour les utilisateurs individuels que l'amélioration des capacités de prédiction.

6.4.1. Le modèle agrométéorologique

Les statistiques agricoles s'appuient sur trois sources pour dégager les variabilités interannuelles des rendements agricoles : a) les tendances, b) les facteurs météorologiques directs et c) les effets indirects causés par les conditions météorologiques, les ravageurs, les maladies, les adventices, etc. Il convient toutefois également de prendre en compte le fait que la variabilité spatiale des rendements agricoles dépend fortement de la fertilité du sol. La question des répercussions des conditions météorologiques sur les disponibilités alimentaires se révèle encore plus compliquée en raison du rôle joué par les conditions météorologiques à différents niveaux de la chaîne alimentaire (Gommes, 2002). L'importance relative des sources recensées de variabilité de la production alimentaire dépend en grande partie du contexte socio-économique général. Dans de nombreux pays en développement, la composante technologique n'est pas particulièrement marquée, et les récoltes de certains des pays les plus pauvres du monde n'affichent aucune tendance. Dans ce cas, les conditions météorologiques peuvent avoir des conséquences gravissimes et menacer la sécurité alimentaire de millions de personnes. Lorsque les circonstances obligent progressivement les agriculteurs à passer à l'agriculture de marché, ceux-ci passeront par une phase de transition pendant laquelle leur vulnérabilité aux aléas climatiques s'accroîtra. Dans les pays ayant un secteur agricole plus développé (grâce à la mécanisation, à l'irrigation, à l'utilisation de variétés améliorées, à la prise de décision avancée sur le terrain, etc.), les tendances représentent une large proportion de la variabilité des récoltes (plus de 80 %). Si la tendance est retirée de la série temporelle, l'on pourra partir de l'hypothèse que la variabilité restante est en grande partie due aux conditions

météorologiques. Balaghi *et al.* (2013) proposent une fonction de rendement des cultures fondée sur des données météorologiques (pluviométrie), qui sert de variable explicative de la fluctuation des rendements agricoles au Maroc (figure 6). Les auteurs considèrent que la courbe de la droite de régression représente la productivité de l'eau de pluie (PEP, exprimée en g/l), qui se définit comme le rapport du rendement (kg/ha) sur la pluviométrie cumulée (mm) durant la campagne agricole (de septembre à avril). Dans les régions semi-arides, une variable météorologique simple telle que la pluviométrie permet d'expliquer la plupart des variations des rendements agricoles ; dans les pays tropicaux, en revanche, il conviendra d'intégrer d'autres facteurs météorologiques et liés à la teneur en humidité des sols. Afin de prévoir les rendements rizicoles au Bangladesh, Gommès (2001) s'appuie sur un modèle hydrique simple et l'analyse de ses principales composantes pour déterminer deux variables indépendantes, à savoir le taux d'évapotranspiration réelle et la température maximale au mois d'août (Tx8) (figure 7).

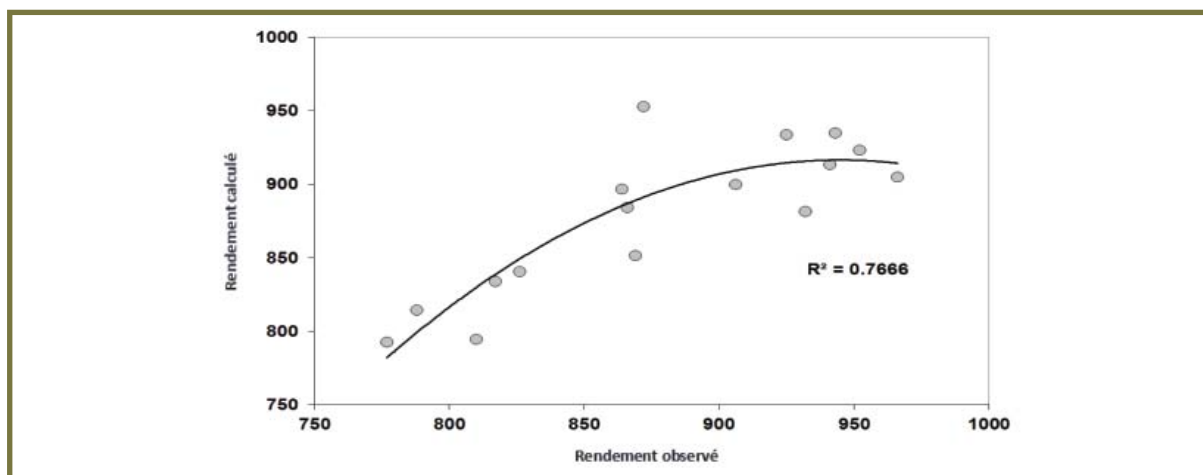
FIGURE 6. RELATION ENTRE LE RENDEMENT NATIONAL DU BLÉ DUR (KG/HA) ET LA PLUVIOMÉTRIE DE LA CAMPAGNE AGRICOLE (MM) (DONNÉES DE 1988 À 2011).



La pente de la droite de régression bleue indique la PEP du blé tendre au Maroc, qui est de 0,347 g/l. La pente de la droite de régression rouge indique la PEP maximale du blé dur à l'échelle nationale (0,653 g/l).

Source : Balaghi *et al.*, 2013.

FIGURE 7. RAPPORT ENTRE LES RENDEMENTS OBSERVÉS ET CALCULÉS POUR LES VARIÉTÉS DE RIZ À HAUT RENDEMENT AU RAJSHAHI, BANGLADESH (1983-1998).



Le modèle de régression est basé sur le taux d'évapotranspiration réelle et la température maximale au mois d'août (Tx8).
Source : Gommès, 2001.

La plupart des auteurs associent les variables météorologiques pour expliquer les fluctuations des rendements agricoles ; toutefois, si la superficie ensemencée demeure relativement constante dans un pays donné, les conditions météorologiques pourraient directement expliquer les variations de la production.

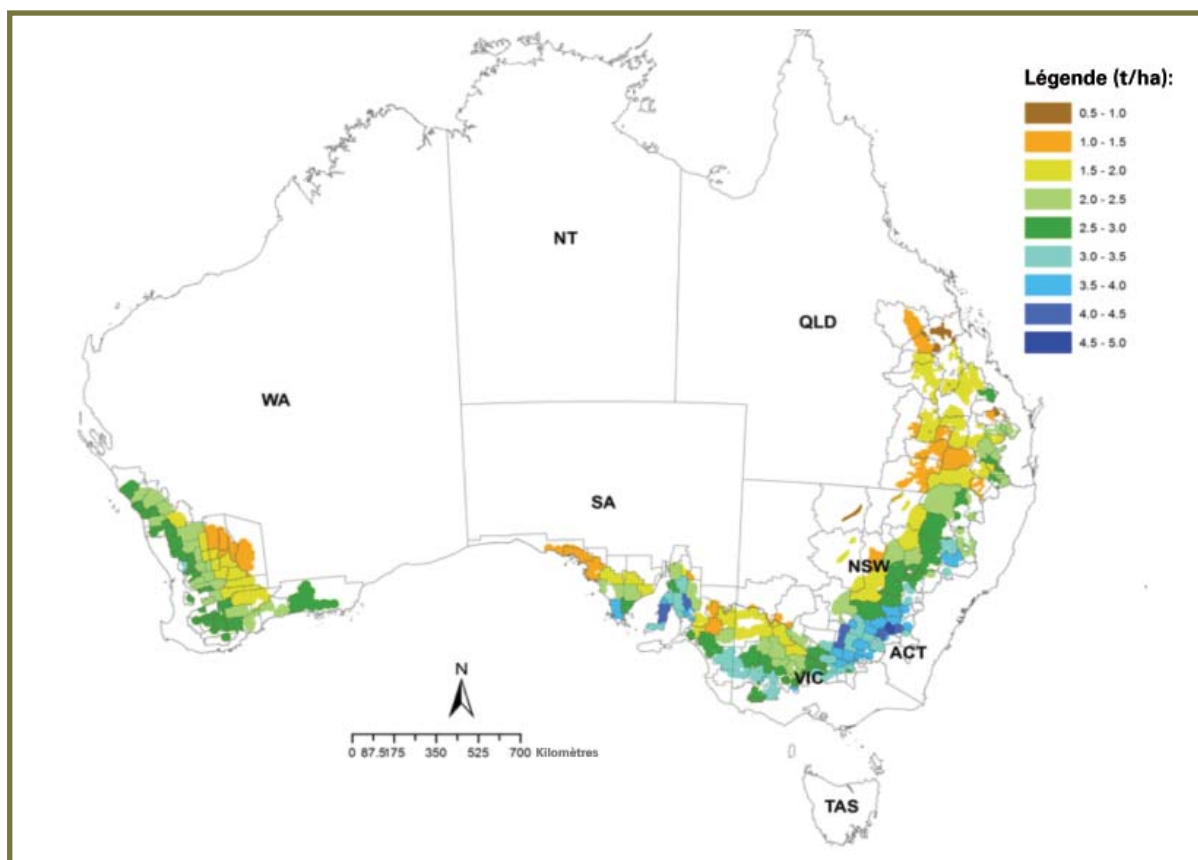
Le Centre commun de recherche (JRC) de la Commission européenne fournit quasiment en temps réel des informations relatives au suivi de la croissance des cultures et à la prévision des rendements pour l'Union européenne et les pays voisins. Il est actuellement en passe d'étendre ces activités aux principales régions productrices du monde. Il évalue également les effets du changement climatique sur le secteur agricole en procédant à des simulations d'impact de différents scénarios sur les modèles de cultures. La stratégie du JRC constitue un bon exemple d'utilisation opérationnelle de modèles complexes relatifs aux cultures ; le Centre s'appuie sur des modèles de simulation biophysique (établis grâce à son logiciel BioMA, de l'anglais *Biophysical Models Applications*) pour simuler les effets du changement climatique sur le secteur agricole et évaluer les stratégies d'adaptation envisageables. Un ensemble d'éléments de modélisation mis en œuvre dans ce cadre permet d'effectuer des simulations fondées sur des scénarios de changement climatique actuels et futurs pour différentes plantes cultivées dans divers systèmes agricoles.

Depuis 2012-2013, Statistique Canada travaille, en collaboration avec Agriculture et Agroalimentaire Canada et Environnement Canada, sur un modèle permettant d'estimer les rendements agricoles des principales cultures du pays. Statistique Canada a saisi l'occasion de produire de nouvelles estimations fondées sur ces rendements modélisés. Ces estimations pourraient, à terme, remplacer la collecte de données et éviter ainsi aux producteurs d'avoir à répondre à de trop nombreuses enquêtes⁴. Le modèle s'appuie sur des données tirées d'images satellite à basse résolution, des estimations effectuées à partir d'anciens levés de terrain et des informations agroclimatiques.

En Australie, Potgieter *et al.* (2014) proposent une méthode de calcul du rendement agrégé du blé à l'échelle des parcelles (avec un écart d'environ 2,6 %) fondée sur des informations de télédétection (figure 8). Le secteur de l'assurance pourrait utiliser les résultats de la première application de cette méthode dans le pays.

4 <http://www.23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5225>

FIGURE 8. SIMULATION À LONG TERME DU RENDEMENT MÉDIAN DU BLÉ À L'ÉCHELLE DES DISTRICTS (1901-2015).



Source : Potgieter *et al.*, 2014.

Grâce aux prévisions de rendements des cultures, les populations sont en mesure de mener des actions de réduction de la vulnérabilité aux risques climatiques, afin de se préparer aux conséquences d'éventuels déficits de production. Il s'agit donc d'un outil incontournable pour la prise de décision dans le domaine agricole, qui permet de prévoir des actions qui faciliteront le travail des compagnies d'assurance agricole (Balaghi *et al.*, 2013).

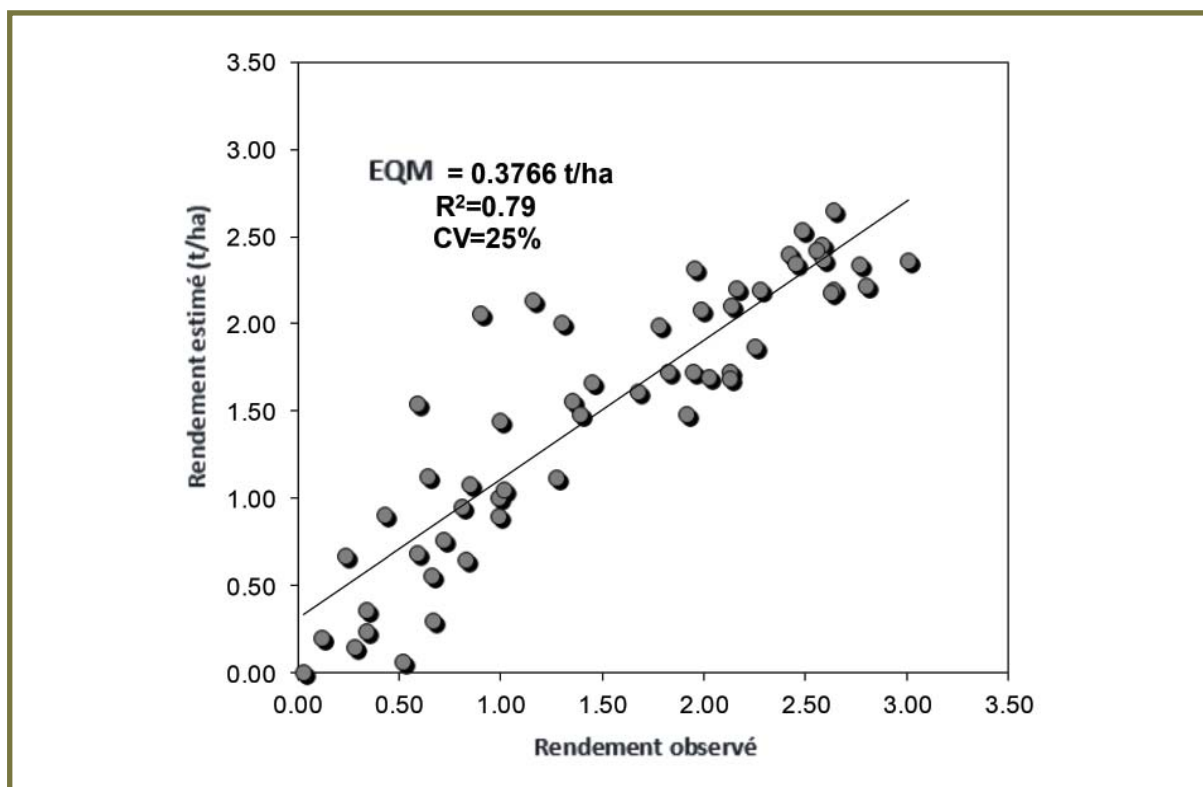
6.4.2. Le modèle fondé sur la télédétection

La télédétection se définit comme l'acquisition d'informations sur un objet par l'intermédiaire d'un appareil n'ayant pas de contact avec l'objet étudié (Lillesand et Keifer, 1994). Ces données peuvent être obtenues par différentes plateformes, telles que satellites, aéronefs, UAV et radiomètres portables, et collectées par différents appareils, tels que capteurs, caméras, appareils photo numériques ou magnétoscopes. Les instruments utilisés pour mesurer les rayonnements électromagnétiques sont appelés « capteurs ». Ceux-ci sont passifs lorsqu'ils ne produisent pas leur propre source d'énergie et ne réagissent qu'aux rayonnements d'origine naturelle ; on parle au contraire de capteurs actifs lorsqu'ils produisent leur propre source de rayonnement.

Les études menées sur le terrain et l'utilisation de scanners aéroportés (Tucker, 1979) prouvent l'utilité des propriétés de la réflectance spectrale du couvert végétal, en particulier l'association de la réflectance de la bande spectrale rouge et celle du proche infrarouge (également appelée « indices de végétation »), pour la surveillance des végétaux chlorophylliens. Ces indices de végétation se définissent comme des combinaisons mathématiques ou des ratios composés principalement de bandes spectrales rouges, vertes et infrarouges et conçus pour identifier les rapports fonctionnels entre les caractéristiques des cultures et les observations par télédétection (Wiegand *et al.*, 1990). Fortement modulés par les interactions entre le rayonnement solaire et les cultures au moment

de la photosynthèse, ces indices témoignent donc de la dynamique des propriétés biophysiques liées à la santé des cultures. Le tableau 1 recense les indices de végétation les plus utilisés dans le cadre du surveillance de la végétation. Le NDVI constitue l'un des indices fondés sur les deux bandes spectrales susmentionnées. Proposé par Deering (1978), il est encore aujourd'hui le plus utilisé pour étudier la santé des végétaux et la production végétale (MacDonald et Hall, 1980 ; Sellers, 1985). La recherche en matière de suivi des végétaux a montré que le NDVI était étroitement lié à l'indice foliaire et à l'activité photosynthétique des végétaux chlorophylliens. Il mesure indirectement la productivité primaire par le biais de son rapport quasi linéaire avec la Fraction de rayonnement photosynthétiquement actif absorbé (fAPAR) (Los, 1998 ; Prince, 1990). Il présente certaines limites bien connues, relatives par exemple aux effets de l'humidité du sol ou de l'anisotropie des surfaces. Les produits composites utilisés dans la plupart des applications ont tendance à restreindre ces effets, qui ne peuvent toutefois pas être totalement ignorés. Rojas (2007) propose une fonction de rendement basée sur la télédétection et les données agrométéorologiques pour estimer les rendements futurs du maïs au Kenya. Le NDVI du modèle et l'évapotranspiration réelle expliquent 79 % des fluctuations des rendements du maïs, avec une erreur quadratique moyenne s'élevant à 0,3766 t/ha (figure 9). Les modèles agrométéorologiques fournissent des informations sur le rayonnement solaire, la température, l'humidité de l'air et la quantité d'eau disponible dans le sol, la composante spectrale contenant quant à elle des informations sur la gestion des cultures, les variétés utilisées et les stress ignorés par les modèles agrométéorologiques (Rudorff et Batista, 1990).

FIGURE 9. RAPPORT ENTRE LES ESTIMATIONS DE RENDEMENT OBSERVÉS ET ESTIMÉS AU KENYA.



Modèle de regression basé sur des données ETA et NDVI pour les années 1985 à 2003.
Source : Rojas, 2007

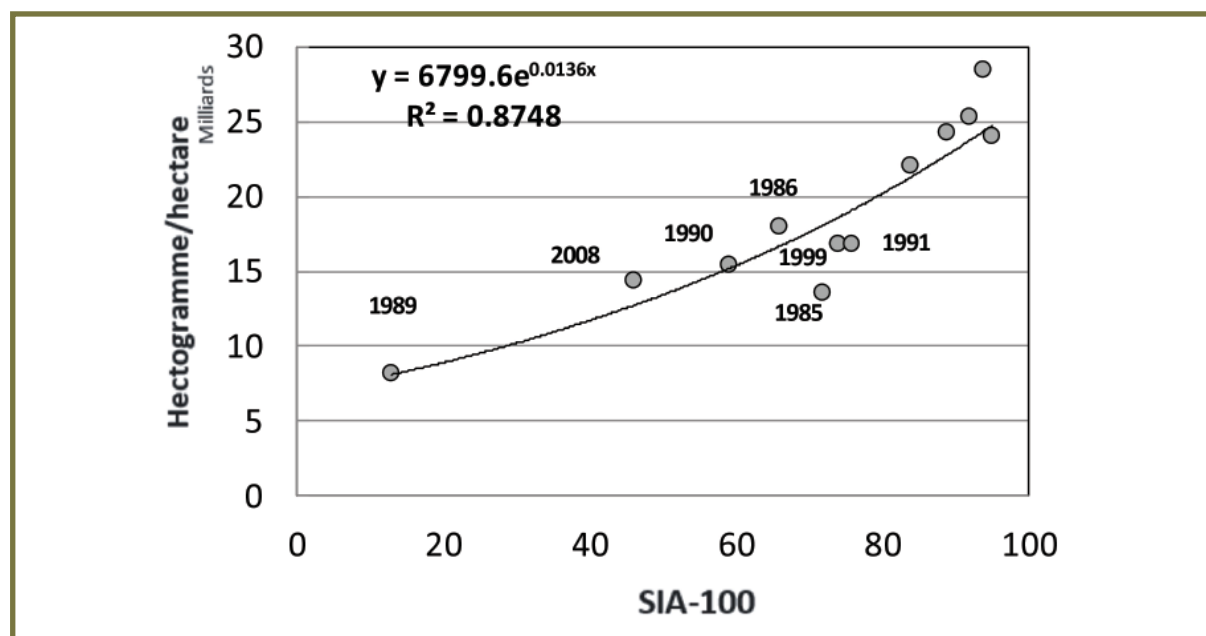
La FAO a mis sur pied un Système d'indice de stress agricole (ASIS, de l'anglais *Agriculture Stress Index System*), qui a pour objet de détecter les zones agricoles fortement susceptibles de subir un stress hydrique (sécheresse) à l'échelle mondiale. Celui-ci repose sur l'indicateur de la santé végétale (VHI, de l'anglais *Vegetation Health Index*), un indicateur composite englobant l'indicateur de l'état de la végétation (VCI, de l'anglais *Vegetation*

Condition Index) et l'indicateur des conditions de températures (TCI, de l'anglais *Vegetation Condition Index*). Le VCI (Kogan, 1994) provient du NDVI. L'algorithme du TCI est similaire à celui du VCI, mais se rapporte à la température de brillance T estimée à partir de la région infrarouge thermique du système AVHRR (canal 4). Kogan (1995) a proposé cet indice afin d'éliminer les effets de la nébulosité au moment de l'évaluation par satellite de l'état de la végétation, le canal 4 de l'AVHRR étant moins sensible à la vapeur d'eau dans l'atmosphère que le canal captant la lumière visible. Les températures de mi-saison élevées signalent des conditions défavorables ou de sécheresse, tandis que les températures basses témoignent principalement de conditions favorables (Kogan, 1995).

La première étape de l'ASIS consiste à calculer la moyenne temporelle du VHI afin d'évaluer l'intensité et la durée de la/des période(s) de sécheresse ayant lieu pendant le cycle de culture au niveau des pixels. L'ASIS se fonde sur des données satellite d'une résolution de 1 km enregistrées tous les dix jours (décadaires) par le capteur AVHRR du satellite METOP et portant sur la végétation et la température de la surface de la Terre.

La deuxième étape consiste à calculer le pourcentage de la superficie agricole touchée par la sécheresse (pixels présentant un VHI inférieur à 35 – une valeur considérée comme décisive par de précédentes études) afin d'évaluer l'étendue de la catastrophe. Enfin, l'ensemble de la zone administrative est classé en fonction du pourcentage de superficie touchée. L'ASIS évalue la gravité (intensité, durée et étendue) de la sécheresse et publie des résultats finaux au niveau administratif, afin d'en faciliter la comparaison avec les statistiques agricoles du pays (Rojas *et al.*, 2011 ; Van Hoolst *et al.*, 2016). Rojas (2016) propose une fonction de rendement basée sur l'indice de stress agricole (ISA) et les rendements du blé en Syrie (figure 10).

FIGURE 10. RAPPORT ENTRE L'INDICE DE STRESS AGRICOLE (ASI) ET LE RENDEMENT DU BLÉ EN SYRIE (1985-2012)



Source : FAO, GIEWS

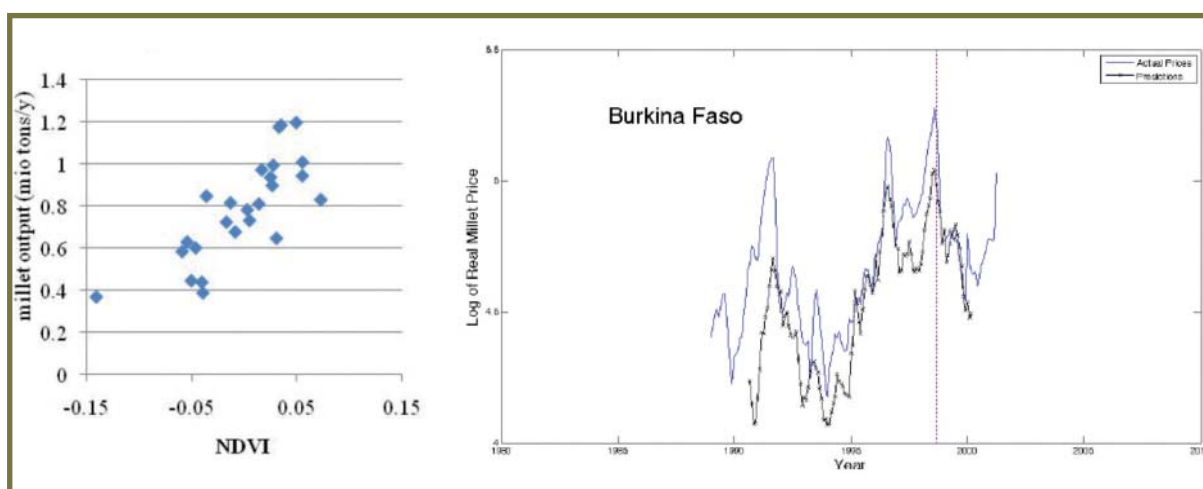
En Inde, le Centre national Mahalanobis de prévision des récoltes (NCFC, de l'anglais *Mahalanobis National Crop Forecast Centre*) utilise à la fois des modèles agrométéorologiques et des données de télédétection pour fournir différentes prévisions concernant la production de cultures en amont de la récolte, au niveau du pays, des États et des districts⁵. Le centre publie des prévisions de production à six reprises au cours du cycle de culture : avant la

⁵ <http://www.nrsc.gov.in/Agriculture>

saison, en début de saison, à la mi-saison, avant la récolte (au niveau des États et des districts), et dans le cadre d'une évaluation révisée des dommages. Il utilise des outils conventionnels, tels que l'économétrie et l'agrométéorologie (pour formuler ses prévisions d'avant-saison et de début de saison), des produits de télédétection à moyenne et haute résolution (pour ses prévisions de mi-saison) et l'imagerie haute résolution (pour ses prévisions d'avant-récolte à l'échelle des districts). La méthode du NCFC tire également sa force des expériences de coupe menées par télédétection à chaque campagne agricole pour entraîner les différents modèles agrométéorologiques et les produits de télédétection.

Brown *et al.* (2009) suggèrent d'utiliser le modèle de prévision des rendements agricoles pour prévoir le prix des aliments de base à l'aide des informations fournies par le NDVI. Ils se servent de l'indicateur comme d'un intermédiaire pour déterminer l'offre locale en millet, pour laquelle il n'existe pas suffisamment de données de marché. Le fait d'associer l'indicateur à des informations sur la superficie ensemencée augmentera vraisemblablement sa capacité de prédiction; toutefois, et c'est regrettable, l'accès à des informations exhaustives sur la superficie cultivée à l'échelle locale s'avère difficile en Afrique de l'Ouest. Le modèle explique 85 à 90 % des fluctuations de prix observées, et le taux d'erreur des prévisions à quatre mois varie en moyenne entre 13,4 % (au Niger) et 19,5 % (au Burkina Faso) dans le cadre d'observations de l'échantillon, ou entre 18,8 % (au Burkina Faso) et 21,9 % (au Mali) dans le cadre d'observations hors échantillon. La figure 11 illustre les rapports existant d'une part entre la production de millet et le NDVI, et d'autre part entre les cours réels de la céréale et les prédictions.

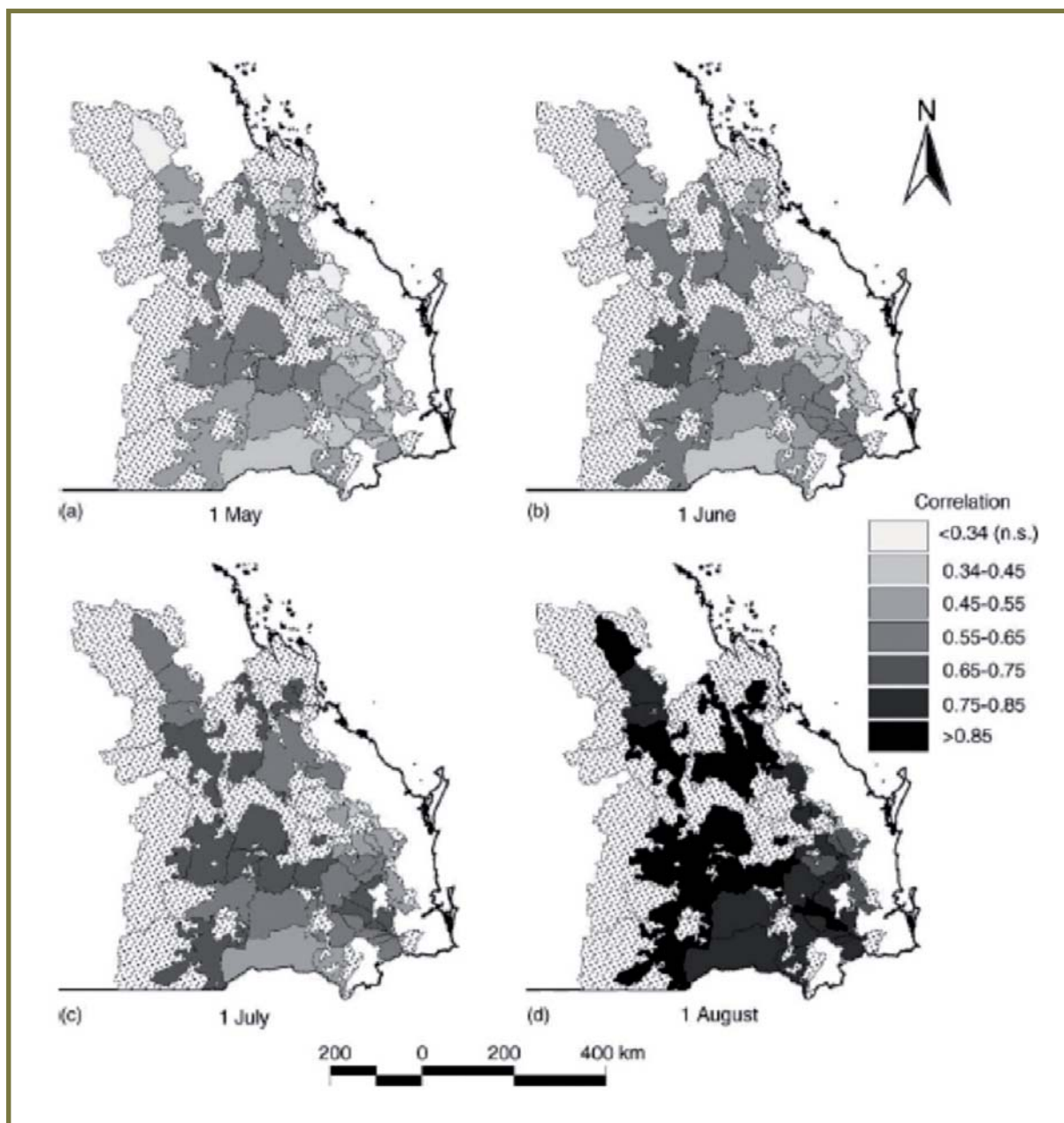
FIGURE 11. RAPPORT ENTRE LE NDVI ET LA PRODUCTION DE MILLET POUR LA PÉRIODE 1982-2006 (GAUCHE); PRIX MOYENS ET PRÉDICTIONS À QUATRE MOIS (DROITE), AU BURKINA FASO.



Source: Brown *et al.*, 2009.

Enfin, Hansen *et al.* (2004) ont recours à un MCG pour prévoir les rendements de blé à l'échelle régionale dans le nord-est de l'Australie. Dans ce cadre, l'exactitude des prédictions est généralement meilleure au niveau des États qu'à l'échelon inférieur (districts). Bien que le modèle de simulation de production du blé ait permis d'expliquer environ 75 % de la variance des rendements moyens à l'échelle de l'État en l'absence de toute tendance, les corrélations propres aux districts étaient généralement inférieures, justifiant en moyenne 58 % de la variance pondérée par la surface occupée par le blé dans chaque district. Les auteurs ont comparé les résultats du MCG aux prédictions fondées sur la seule climatologie, ainsi qu'aux phases d'El Niño-Oscillation australe (ENSO, de l'anglais *El Niño-Southern Oscillation*) et de l'indice d'oscillation australe (IOA). La figure 12 montre la corrélation entre les rendements de blé à l'échelle des districts observés dans le Queensland et les prévisions *a posteriori* fondées sur le MCG.

FIGURE 12. CORRÉLATION ENTRE LES RENDEMENTS DE BLÉ OBSERVÉS AU NIVEAU DES DISTRICTS DU QUEENSLAND ET LES PRÉVISIONS *A POSTERIORI* EFFECTUÉES À PARTIR DU MCG LES A) 1^{ER} MAI, B) 1^{ER} JUIN, C) 1^{ER} JUILLET ET D) 1^{ER} AOÛT, ENTRE 1975 ET 1993, AJUSTÉE AUX TENDANCES TECHNOLOGIQUES DE 2001.



Source : Hansen *et al.*, 2004.

Le Service national de statistiques agricoles du Département de l'agriculture des États-Unis (USDA/NASS) utilise les technologies de télédétection à des fins d'arpentage agricole et effectue des recherches dans ce domaine depuis le début des années 1970. Ses progrès considérables ces dernières années lui ont permis de faire évoluer le rôle joué par la télédétection, qui auparavant remplissait surtout une fonction de recherche. Elle fait désormais partie intégrante du programme d'estimation des surfaces agricoles de l'agence, qui porte sur l'ensemble des principales cultures produites dans les États très producteurs des États-Unis. L'agence est parvenue à ce résultat en particulier

grâce : 1) au renforcement de ses partenariats en matière de données ; 2) au perfectionnement de ses méthodes ; 3) à l'accessibilité accrue de logiciels commerciaux ; et 4) à l'amélioration de la qualité des données d'imagerie et auxiliaires. L'arpentage étant désormais fonctionnel, l'USDA concentre désormais ses efforts sur une autre transition entre recherche et contexte opérationnel : celle de l'estimation des rendements. À l'heure actuelle, il fournit au Comité de statistiques agricoles et aux bureaux de terrain du NASS des estimations des rendements du soja et du maïs pour dix grands États producteurs, pendant la saison. Bien que ces estimations soient déjà fiables, elles peuvent encore être améliorées grâce à d'autres expériences et recherches. Pour l'avenir, le NASS a commencé à mener des recherches afin de mesurer quantitativement l'évolution et l'état des cultures à l'aide de la télédétection. Il compte par ailleurs utiliser cette technologie pour cartographier l'humidité du sol et améliorer les évaluations et le suivi des catastrophes.

6.5. CONCLUSION

Ce chapitre a passé en revue le concept des systèmes d'alerte précoces et souligné leur importance, spécifiquement dans les domaines de la sécurité alimentaire et de la nutrition. Sa principale conclusion est la suivante : pour être efficace, un SAP doit être intégré aux plans nationaux d'atténuation et de réduction des risques de catastrophe liés au(x) danger(s) surveillé(s). Toute alerte peut aider les communautés à se préparer et à prendre des mesures immédiates ; elles ne peuvent, en revanche, atténuer les effets potentiels de la catastrophe à venir. Les concepts philosophiques très pertinents d'alerte et d'intervention précoces ont donc été introduits, en vue de leur mise en œuvre par les gouvernements et la société pour défendre les communautés, sauver des vies et protéger les ressources. En matière de production agricole, les techniques de prévision des rendements des cultures (qui évaluent essentiellement la disponibilité des quatre domaines de la sécurité alimentaire, mais qui produisent des effets indirects sur la stabilité et l'accès en agissant sur les marchés locaux) peuvent s'avérer utiles en permettant aux sociétés de se préparer aux conséquences d'une éventuelle pénurie en prenant des mesures de réduction de la vulnérabilité face aux risques climatiques. Plusieurs modèles plus ou moins complexes ont été proposés pour estimer les rendements des cultures à l'échelle régionale, de la simple régression entre variables climatiques aux fonctions de rendement plus sophistiquées, en passant par les produits de télédétection et les MCG. Le modèle choisi dépendra de l'accessibilité et de la qualité des données dans chaque pays et de l'objectif ultime des prévisions. La télédétection joue actuellement un rôle central en matière d'estimation des surfaces et des rendements des cultures dans les programmes statistiques. L'amélioration des capacités d'imagerie et des technologies de l'information a permis à la télédétection de quitter le domaine de la recherche pour intégrer un véritable processus de production. À l'avenir, le passage du domaine de la recherche à celui de la mise en œuvre opérationnelle de la télédétection devrait s'accélérer, réduisant ainsi la charge pesant sur les enquêtés, multipliant les données utiles sur le plan spatial et favorisant la réalisation d'économies en limitant la collecte de données dans le cadre d'enquêtes traditionnelles, dans le but ultime d'améliorer les statistiques agricoles.

ENCADRÉ 1. HISTORIQUE DES SYSTÈMES D'INFORMATION RELATIFS À LA SÉCURITÉ ALIMENTAIRE

- Jusqu'au 19^e siècle, les variations démographiques étaient dues à trois facteurs principaux (guerres, famines et épidémies), et souvent liées aux fluctuations de la production alimentaire. Le développement commercial et industriel a progressivement modifié cette tendance dans les pays industrialisés, la population dépendant de moins en moins de l'agriculture de subsistance. Jusqu'à là, le suivi des populations donnait une idée (*a posteriori*) de l'insécurité alimentaire.
- Jusqu'au début du 20^e siècle, l'agriculture de subsistance est restée la source principale d'alimentation dans les pays non industrialisés. Les vagues successives de colonisation et de décolonisation des pays les plus pauvres étaient fondées sur des politiques d'autosuffisance alimentaire visant à préserver l'indépendance politique. Les systèmes de surveillance de l'alimentation de l'époque reposaient essentiellement sur la connaissance de techniques élémentaires de production agricole.
- Les crises alimentaires majeures des années 1970 ont souligné l'importance de l'opinion publique mondiale et de la nécessité d'informer le grand public de la survenue d'une crise alimentaire. Le contrôle nutritionnel des enfants en bas âge a vu le jour à cette époque et constitué le point de départ de la collecte d'informations sur les plus démunis et de la fourniture d'une assistance à ces personnes.
- Plus de cent pays ont adopté la Déclaration universelle pour l'élimination définitive de la faim et de la malnutrition à l'occasion de la Conférence mondiale de l'alimentation organisée en 1974 à Rome. À l'époque, les analyses des causes de la faim foisonnaient. La notion moderne de « sécurité alimentaire » a été développée avec la création, en 1975, du Comité de la sécurité alimentaire mondiale (CSA). Bien que les gouvernements aient été exhortés à adopter des « politiques alimentaires nationales », leurs résultats n'ont pas été à la hauteur. L'accent a été mis sur la recherche d'un équilibre entre l'offre et la demande d'aliments de base, et les systèmes de surveillance de l'approvisionnement alimentaire se sont progressivement développés.
- Les agences publiques de mise en marché de nombreux pays (en particulier celles spécialisées dans les céréales) se sont vu confier de nouveaux pouvoirs de supervision des approvisionnements, et des efforts ont été déployés pour créer des stocks alimentaires de sécurité nationaux et régionaux (principalement dans les régions arides) destinés à être utilisés en cas de crise alimentaire. Dans ce contexte, les informations sur la production, l'approvisionnement national, les stocks et les importations pouvaient théoriquement être contrôlées par un seul organe national.
- Les systèmes statistiques et les comptes nationaux ont enregistré des demandes d'information toujours plus importantes, aux fins notamment de planification et de suivi économique. Les systèmes statistiques ont eu tendance à être fortement cloisonnés (confiés à un ministère ou à un organe national distinct) et axés sur leurs propres exigences internes. Leurs informations n'ont donc pas été particulièrement utiles au suivi de la sécurité alimentaire (p. ex., statistiques peu fiables ou publiées tardivement; zones couvertes incompatibles, rendant difficile la comparaison des données collectées à partir de systèmes différents). Certains partenaires de développement ont donc préféré établir — en Afrique subsaharienne notamment — des systèmes d'information parallèles qu'ils financent et supervisent eux-mêmes, et qui leur fournissent les informations dont ils ont besoin pour cibler leur aide et leur assistance en matière de sécurité alimentaire.

» »

- Plusieurs ajustements structurels ultérieurs et un processus de privatisation ont métamorphosé l'économie des pays en développement et leur trajectoire, produisant des effets majeurs sur la sécurité alimentaire de leur population. Les donateurs ont décidé de conjuguer leur aide économique à des politiques interventionnistes ciblées afin d'éviter la survenue de crises sociales ou alimentaires majeures. Il est devenu évident que le problème d'accès aux denrées de base (pauvreté) devait faire l'objet d'un examen plus approfondi et que l'aide sociale ne pouvait être organisée qu'en échange de l'obtention de certaines informations spécifiques.
- La baisse de l'excédent mondial de céréales et certaines mauvaises expériences en matière de distribution à grande échelle ont poussé la communauté internationale à réduire les aides et à remettre en question leur utilisation, compte tenu de leur effet négatif potentiel — lorsqu'elles sont inadaptées ou font l'objet d'une mauvaise gestion — sur la sécurité alimentaire. Les systèmes d'information existants ont ensuite été réorientés, afin de mieux cibler et surveiller la distribution de l'aide. Le développement et l'utilisation de l'imagerie satellite ont permis d'effectuer des estimations concernant la végétation — notamment les cultures pluviales — dans chaque pays, et donc de fournir des informations sur la production mondiale aux donateurs principaux.
- La mondialisation des échanges est désormais liée à la rapidité des communications modernes (puisque'ils peuvent être effectués sur Internet, par fax, par téléphone satellitaire, etc.) et des systèmes de gestion (ordinateurs). Il est beaucoup plus facile aujourd'hui de gérer des bases de données, de faire des pronostics et d'en diffuser les résultats. Les systèmes de suivi et de prévision de la sécurité alimentaire sont devenus plus faciles à gérer à l'échelle nationale — du moins en théorie, et les donateurs encouragent la formation des gestionnaires à l'utilisation de ces outils et médias modernes.
- La vitesse de l'urbanisation dans les pays en développement produit aussi des effets sur les données relatives à la sécurité alimentaire, puisqu'il est devenu essentiel de procéder à un suivi de la sécurité alimentaire et des groupes vulnérables à l'échelle urbaine et rurale.
- Parallèlement au développement de circuits d'information plus rapides, il est devenu nécessaire de décentraliser la prise de décision et de démocratiser la prise d'action concertée en y intégrant différents intervenants (publics, privés et de la société civile) de la sécurité alimentaire. Les systèmes d'information sur la sécurité alimentaire sont peu à peu devenus des centres nationaux d'échange d'informations à tous les niveaux.
- La nécessité de réduire le nombre de personnes sous-alimentées a souligné le besoin de décentraliser et de mener une action commune à tous les acteurs de la sécurité alimentaire. Cette exigence a été clairement mise en lumière par l'ensemble des participants au Sommet mondial de l'alimentation organisé en novembre 1996, à l'occasion duquel les représentants gouvernementaux ont décidé de prendre toutes les mesures nécessaires pour réduire de moitié le nombre de personnes sous-alimentées (estimé à 800 millions à l'époque) à l'horizon 2015. Depuis, des efforts ont été déployés pour compléter les bases de données des systèmes d'information en développant des indicateurs concrets pour le suivi des personnes sous-alimentées.
- Plus récemment, le 25 septembre 2015, les pays ont adopté un ensemble d'objectifs à atteindre pour mettre fin à la pauvreté, protéger la planète et assurer la prospérité de tous, dans le cadre d'un nouveau programme de développement durable. Chaque objectif s'accompagne de cibles spécifiques à atteindre au cours des 15 prochaines années. La proposition contient 17 objectifs assortis de 169 cibles qui couvrent un vaste ensemble de questions liées au développement durable — p. ex., éradiquer la pauvreté et la faim, améliorer la santé et l'éducation, améliorer la durabilité des villes, lutter contre le changement climatique et protéger les océans et les forêts.

TABEAU 1. LISTE DES INDICES DE VÉGÉTATION COURANTS, FORMULES MATHÉMATIQUES, ÉCHELLE DE DÉVELOPPEMENT ET PARAMÈTRES ESTIMÉS.

Indice	Formule	Référence	Échelle	Paramètre
NDVI	$\frac{(NIR - Red)}{(NIR + Red)}$	Rouse <i>et al.</i> , 1974	Canopée	Biomasse; fraction de végétation
Indice différentiel normalisé de végétation verte (GNDVI, de l'anglais <i>Green Normalized Difference Vegetation Index</i>)	$\frac{(NIR - Green)}{(NIR + Green)}$	Gitelson <i>et al.</i> , 1996	Canopée	Chlorophylle; fraction de végétation
Indice de réflectance photochimique (PRI, de l'anglais <i>Photochemical Reflectance Index</i>)	$\frac{(R570 - R531)}{(R570 + R531)}$	Gamon <i>et al.</i> , 1992	Canopée	Efficacité de la photosynthèse/ efficacité d'utilisation du rayonnement
Indice différentiel normalisé red edge (NDRE, de l'anglais <i>Normalized Difference Red Edge</i>)	$\frac{(R790 - R720)}{(R790 + R720)}$	Barnes <i>et al.</i> , 2000	Canopée	Chlorophylle/Azote
Indice de la teneur en chlorophylle de la canopée (CCCI, de l'anglais <i>Canopy Chlorophyll Content Index</i>)	$\frac{(NDRE - NDRE_{min})}{(NDRE_{max} - NDRE_{min})}$	Fitzgerald <i>et al.</i> , 2006	Canopée	Taux d'azote/chlorophylle
Indice de végétation par quotient (RVI, de l'anglais <i>Ratio Vegetation Index</i>)	$\frac{NIR}{Red}$	Jordan, 1969	Feuillage	Biomasse
Indice de végétation amélioré (EVI, de l'anglais <i>Enhanced Vegetation Index</i>)	$2.5 * \frac{(NIR - Red)}{(NIR + C1 * Red - C2 * Blue + L)}$ [C1 = 6; C2 = 7.5; L = 1]	Huete <i>et al.</i> , 2002	Canopée/ régional	Biomasse/couvert végétal
Indice de végétation amélioré 2 (EVI2, de l'anglais <i>Enhanced Vegetation Index 2</i>)	$G * \frac{NIR - Red}{NIR + \left(6 - \frac{7.5}{c}\right) * Red + 1}$ $Red = c * Blue$ $G = f(c)$	Jiang <i>et al.</i> , 2008	Canopée/ régional	Biomasse/couvert végétal
Indice de résistance atmosphérique visible (VAR green, de l'anglais <i>Visible Atmospherically Resistant Index</i>)	$\frac{(Green - Red)}{(Green + Red + Blue)}$	Gitelson <i>et al.</i> , 2002	Canopée/ régional	Fraction végétale/indice foliaire
Indice de résistance atmosphérique visible, 700 nm (VARI700)	$\frac{(R700 - 1.7 * Red - 0.7 * Blue)}{(R700 + 2.3 * Red - 1.3 * Blue)}$	Gitelson <i>et al.</i> , 2002	Canopée/ régional	Fraction végétale/indice foliaire
Indice triangulaire de végétation (TVI, de l'anglais <i>Triangular Vegetation Index</i>)	$0.5 * [20 * (R750 - R550) - 200 * (R670 - R550)]$	Brodge et Leblanc, 2000	Canopée	Chlorophylle
Indice triangulaire modifié de végétation 1 (MTVI 1)	$1.2 * [1.2 * (R800 - R550) - 2.5 * (R670 - R550)]$	Haboudane <i>et al.</i> , 2004	Canopée	Chlorophylle
Indice triangulaire modifié de végétation 2 (MTVI 2)	$1.5 * [1.2 * (R800 - R550) - 2.5 * (R670 - R550)]$ $\sqrt{(2 * R800 + 1)^2 - (6 * R800 - 5 * \sqrt{R670})} - 0.5$	Haboudane <i>et al.</i> , 2004	Canopée	Chlorophylle
Indice de chlorophylle terrestre MERIS (MTCI, de l'anglais <i>MERIS Terrestrial Chlorophyll Index</i>)	$\frac{R735.75 - R708.75}{R708.75 - R681.25}$	Dash et Curran, 2007	Canopée	Chlorophylle

Indice	Formule	Référence	Échelle	Paramètre
Réflectance de l'absorption de la chlorophylle (CAR, de l'anglais <i>Chlorophyll Absorption Reflectance</i>)	$\frac{[(a * R_{670} + R_{670} + b)]}{\sqrt{(a^2 + 1)}}$ $\left[\begin{array}{l} a = (R_{700} - R_{550}) / 150 \\ b = R_{550} - (a * 550) \end{array} \right]$	Kim <i>et al.</i> , 1994	Canopée	Chlorophylle
Indice de réflectance de l'absorption de la chlorophylle (CARI, de l'anglais <i>Chlorophyll Absorption Reflectance Index</i>)	$\frac{[(a * R_{670} + R_{670} + b)] * R_{700}}{\sqrt{(a^2 + 1)} * R_{670}}$	Kim <i>et al.</i> , 1994	Canopée	Chlorophylle
Indice modifié de réflectance de l'absorption de la chlorophylle (MCARI, de l'anglais <i>Modified Chlorophyll Absorption Reflectance Index</i>)	$[(R_{700} - R_{670}) - 0.2 * (R_{700} - R_{550})] * \frac{R_{700}}{R_{670}}$	Daughtry <i>et al.</i> , 2000	Feuillage/ canopée	Chlorophylle/indice foliaire/ réflectance du sol
MCARI 1	$1.2 * [2.5 * (R_{800} - R_{670}) - 1.3 * (R_{800} - R_{550})]$	Haboudane <i>et al.</i> , 2004	Canopée	Chlorophylle/indice foliaire/ réflectance du sol
MCARI 2	$\frac{1.5 * [2.5 * (R_{800} - R_{670}) - 1.3 * (R_{800} - R_{550})]}{\sqrt{(2 * R_{800} + 1)^2 - (6 * R_{800} - 5 * \sqrt{R_{680}}) - 0.5}}$	Haboudane <i>et al.</i> , 2004	Canopée	Chlorophylle/indice foliaire/ réflectance du sol
Indice transformé de réflectance de l'absorption de la chlorophylle (TCARI, de l'anglais <i>Transformed Chlorophyll Absorption Reflectance Index</i>)	$3 * [(R_{700} - R_{700}) - 0.2 * (R_{700} - R_{550}) * (R_{700} / R_{670})]$	Haboudane <i>et al.</i> , 2002	Canopée	Chlorophylle/indice foliaire/ réflectance du sol
Indice de végétation par différence pondérée (WDVI, de l'anglais <i>Weighted Difference Vegetation Index</i>)	$NIR - a * Red$	Clevers, 1989	Canopée	Indice foliaire/paramètres biophysiques
Indice perpendiculaire de végétation (PVI, de l'anglais <i>Perpendicular Vegetation Index</i>)	$\frac{1}{\sqrt{a^2 + 1 * (NIR - a * Red - b)}}$	Richardson et Wiegand, 1977	Canopée	Paramètres biophysiques de la canopée
Indice de végétation ajusté au sol (SAVI, de l'anglais <i>Soil-Adjusted Vegetation Index</i>)	$\frac{(1 + L^a) * (R_{800} - R_{670})}{(R_{800} + R_{670}) + L}$	Huete <i>et al.</i> , 1988	Canopée	Paramètres biophysiques de la canopée
Indice transformé de végétation ajusté au sol (TSAVI, de l'anglais <i>Transformed Soil-Adjusted Vegetation Index</i>)	$\frac{a^b * (R_{800} - a * R_{670} - b^c)}{[a * R_{800} + R_{670} - a * b]}$	Baret <i>et al.</i> , 1989	Canopée	Paramètres biophysiques de la canopée
Indice optimisé de végétation ajusté au sol (OSAVI, de l'anglais <i>Optimized Soil-Adjusted Vegetation Index</i>)	$\frac{(1 + 0.16) * (NIR - Red)}{(NIR + Red + 0.16)}$	Rondeaux <i>et al.</i> , 1996	Canopée	Paramètres biophysiques de la canopée
Indice modifié de végétation ajusté au sol (MSAVI, de l'anglais <i>Modified Soil-Adjusted Vegetation Index</i>) Indicateur de la santé végétale (VHI)	$\frac{(1 + L^c) * (R_{800} - R_{670})}{(R_{800} + R_{670}) + L}$ VHI= a * VCI + b * TCI	Qi <i>et al.</i> , 1994 Kogan, 1995	Canopée Canopée	Paramètres biophysiques de la canopée Biomasse, paramètres biophysiques

Source : adapté de Cammarano, 2010.

* L est un facteur d'ajustement au sol fixé à 0,5.

^b et ^c sont des coefficients du sol issus de l'équation suivante : $NIR_{90} = a * ROUGE_{90} + b$.

^d x est un facteur d'ajustement visant à minimiser les effets d'arrière-plan du sol. Il est fixé à 0,08.

^e L est un facteur d'auto-ajustement tiré de l'équation suivante : $L = 1 - 2 * x * NDVI * WDVI$.

6.6. RÉFÉRENCES BIBLIOGRAPHIQUES

Bailey, J.T. et Boryan, C.G. 2010. *Remote Sensing Applications in Agriculture at the USDA National Agricultural Statistics Service*. Vol. 22030. Publication USDA/NASS : Fairfax, VA, États-Unis.

Bailey, R. 2013. *Managing Famine Risk. Linking Early Warning to Early Action*. Rapport de Chatham House. Publication de Chatham House : Londres.

Balaghi, R., Jlibene, M., Tychon, B. et Eerens, H. 2013. *La prédiction agrométéorologique des rendements céréaliers au Maroc*. Publication de l'Institut national de recherche agronomique du Maroc : Rabat.

Basso, B., Cammarano, D. et Carfagna, E. 2013. *Review of Crop Yield Forecasting Methods and Early Warning Systems*. Publication de la FAO : Rome.

Brown, M., Higgins, N. et Hintermann, B. 2009. *A Model of West African millet prices in rural markets*. Document de travail n° 69 du CEPE. Publication du CEPE : Zurich.

Buchanan-Smith, M. 2000. «Early Warning Systems for Drought Preparedness and Drought Management». Actes de la réunion du groupe d'experts organisés du 5 au 7 septembre 2000 à Lisbonne, Portugal. Publication de l'Organisation météorologique mondiale (OMM) : Genève, Suisse.

Cammarano, D. 2010. *Spatial integration of remote sensing and crop simulation modelling for wheat nitrogen management*. Thèse de doctorat. Université de Melbourne, Australie.

Campbell, J.B. 1996. *Introduction to Remote Sensing*. Deuxième édition. Guilford Press : New York, États-Unis.

Davies, S., Buchanan-Smith, M. et Lambert, R. 1991. *Early warning in the Sahel and Horn of Africa: The state of the art. A review of the literature*. Vol. 1, Rapport de recherche n° 20. Publication de l'IDS : Brighton, Royaume-Uni.

Deering, D.W. 1978. *Rangeland reflectance characteristics measured by aircraft and spacecraft sensors*. Thèse de doctorat. Université A&M du Texas, College Station, États-Unis.

Duncan, J., Stow, D., Franklin, J. et Hope, A. 1993. «Assessing the relationship between spectral vegetation indices and shrub cover in the Iornada Basin, New Mexico». *International Journal of Remote Sensing*, vol. 14, n° 18, p. 3395–416.

FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture) et SMIAR (Système mondial d'information et d'alerte précoce sur l'alimentation et l'agriculture). 2014. «Syrian Arab Republic: Continued conflict and drought conditions worsen 2014 crop production prospects». GIEWS-Update. Publication de la FAO : Rome.

Fédération Internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge. 2008. *Early warning > Early action*. Publication de la Fédération Internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge : Genève, Suisse.

Golnaraghi M. 2005. «Système d'alerte précoce». Bibliothèque PNUE/GRID-Arendal des cartes et graphiques. http://maps.grida.no/go/graphic/early_warning_systems.

Gommes, R., Bernardi, M. et Petrassi, F. 2001. *Agrometeorological Crop Forecasting*. Publication de la FAO : Rome. Disponible à l'adresse : http://www.fao.org/nr/climpag/agromet/index_en.asp. Consulté le 10 juin 2017.

Gommes, R. 2002. *Agrometeorological models and remote sensing for crop monitoring and forecasting in Asia and the Pacific*. Publication de la FAO : Rome.

Gommes, R. 2001. *An introduction to the art of agrometeorological crop yield forecasting using multiple regressions*. Publication de la FAO : Rome.

Goward, S.N., Tucker, C.J. et Dye, D.G. 1985. «North American vegetation patterns observed with the NOAA advanced very high resolution radiometer». *Vegetation*, vol. 64, n° 1, p. 3–14.

Hansen, J.W., Potgieter, A. et Tippet, M.K. 2004. «Using a general circulation model to forecast regional wheat yields in northeast Australia». *Agricultural and Forest Meteorology*, vol. 127, n° 1-2, p. 77–92.

Hoogenboom, G., White, J. et Messina, C. 2004. «From genome to crop: integration through simulation modelling». *Field Crop Research*, vol. 90, n° 1, p. 145–63.

Jones, J., Keating, B. et Porter, C. 2001. «Approaches to modular model development». *Agricultural Systems*, vol. 70, n° 2-3, p. 421–43.

Justice, C.O., Townshend, J.R.G., Holben, B.N. et Tucker, C.J. 1985. «Analysis of the phenology of global vegetation using meteorological satellite data». *International Journal of Remote Sensing*, vol. 6, n° 8, p. 1271–318.

Kogan, F. 1995. « Application of vegetation index and brightness temperature for drought detection ». *Advances in Space Research*, vol. 15, n° 11, p. 91–100.

Kogan, F. 1994. « Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data ». *Bulletin of the American Meteorological Society*, vol. 76, n° 5, p. 655–68.

Lillesand, T.M et Kiefer, R.W. 1994. *Remote Sensing and Image Interpretation*. 3^e édition. John Wiley & Sons : New York, États-Unis.

Los, S.O. 1998. *Linkages between global vegetation and climate. An analysis based on NOAA Advanced Very High Resolution Radiometer data*. Thèse de doctorat. Vrije Universiteit, Amsterdam – Goddard Space Flight Centre, Greenbelt, Maryland.

MacDonald, R.B. et Hall, F.G. 1980. « Global crop forecasting ». *Science*, vol. 208, n° 4445, p. 670–9.

PNUE (Programme des Nations Unies pour l'environnement). 2012. *Early Warning Systems: A State of the Art Analysis and Future Directions*. Publication du PNUE : Nairobi.

Potgieter, A., Power, B., Mclean, J., Davis, P. et Rodríguez, D. 2014. «Spatial estimation of wheat yields from Landsat's visible, near infrared and thermal reflectance bands». *International Journal of Remote Sensing Applications*. vol. 4, n° 3, p. 134–43.

Prince, S.D. 1990. «High temporal frequency remote sensing of primary production using NOAA AVHRR». *Applications of Remote Sensing in Agriculture*, vol. 4, p. 169–83.

Pulwarty, R. 2007. « Communicating agroclimatological information, including forecasts for agricultural decisions ». In : Organisation météorologique mondiale (OMM), *Guide to Agrometeorological Practices* (chap. 17). Publication de l'OMM : Genève, Suisse. Disponible à l'adresse : <http://www.wmo.ch/web/wcp/agm/RevGAMP/>. Consulté le 10 juin 2017.

Rojas, O. Vrieling, A. et Rembold, F. 2011. «Assessing drought probability for agricultural areas in Africa with remote sensing». *Remote Sensing of Environment*, vol. 115, n° 2, p. 343–52.

Rojas, O. 2016. *Protocol for country-level ASIS: calibration and national adaptation process*. Publication de la FAO : Rome.

Rojas, O. 2007. « Operational maize yield model development and validation based on remote sensing and agrometeorological data in Kenya ». *International Journal of Remote Sensing*. vol. 28, n° 17, p. 3775–93.

Rudorff, B. et Batista, G.T. 1990. «Yield estimation of sugarcane based on agrometeorological-spectral models». *Remote Sensing of Environment*, vol. 33, n° 3, p. 183–92.

Sellers, P.J. 1985. «Canopy reflectance, photosynthesis and transpiration». *International Journal of Remote Sensing*, vol. 6, n° 8, p. 1335–72.

Sivakumar, M., Stefanski, R., Bazza M., Zelaya, S., Wilhite D. et Rocha, A. 2014. «High Level Meeting on National Drought Policy: Summary and Major Outcomes». *Weather and Climate Extremes*, vol. 3, p. 126–32.

Stone, T.A., Schlesinger, S., Houghton, R.A. et Woodwell, G.M. 1994. «A map of the vegetation of South America based on satellite imagery». *Photogrammetric Engineering & Remote Sensing*, vol. 60, n° 5, p. 541–51.

Tucker, C.J. 1979. « Red and photographic infrared linear combinations for monitoring vegetation ». *Remote Sensing of Environment*, vol. 8, n° 2, p. 127–50.

Van Hoolst, R., Eerens, H., Haesen, H., Royer, A., Bydekerke, L., Rojas, O., Li, Y. et Racionzer, P. 2016. « FAO's AVHRR-based Agricultural Stress Index System (ASIS) for global drought monitoring ». *International Journal of Remote Sensing*, vol. 37, n° 2, p. 418–39.

Vrieling, A., de Beurs, K. et Brown, M. 2011. «Variability of African farming systems from phenological analysis of NDVI time series». *Climatic Change*, vol. 109, n° 3-4, p. 455–77.

Wiegand, C.L., Gerbermann, A.H., Gallo, K.P., Blad, B.L. et Dusek, D. 1990. «Multisite analyses of spectral-biophysical data for corn». *Remote Sensing of Environment*, vol. 33, n° 1, p. 1–16.

Chapitre 7

Suivi du couvert forestier et de la déforestation

Frédéric Achard, Yeda Maria Malheiros de Oliveira et Danilo Mollicone

7.1. INTRODUCTION ET OBJECTIFS PRINCIPAUX

Les nombreux biens et services produits par la forêt contribuent à l'amélioration des moyens de subsistance et du bien-être des populations, et jouent un rôle essentiel à l'échelle des économies mondiales. Il est communément admis que l'obtention en temps utile de données fiables sur les ressources forestières est l'une des clés pour évaluer précisément les avantages des écosystèmes forestiers et faciliter le suivi et l'évaluation de l'efficacité des stratégies et des programmes en matière de sylviculture et d'autres utilisations des terres par les États et les autres acteurs concernés (MacDicken, 2015).

Par ailleurs, la demande mondiale de produits d'alimentation humaine et animale d'origine agricole et de biocarburants est en grande partie responsable de l'expansion des terres cultivées et des pâturages dans de nombreux pays en développement (DeFries *et al.*, 2010). Les conséquences environnementales de l'expansion des surfaces agricoles sont sensiblement différentes selon la nature du terrain sacrifié (forêt, forêt dégradée ou prairie). Dans les tropiques, plus de 55 % des terres agricoles apparues entre 1980 et 2000 ont vu le jour au détriment de forêts intactes et 28 % ont remplacé des forêts exploitées (Gibbs *et al.*, 2010). La disparition des forêts au profit de cultures commerciales s'est fortement accélérée au cours des dernières années, principalement en Amérique du Sud (de Sy *et al.*, 2015).

Plusieurs difficultés peuvent découler d'un manque d'informations et de statistiques relatives aux ressources forestières : connaissances insuffisantes ou imprécises quant à l'exploitation des ressources forestières nationales ; planification et stratégies inopérantes en matière de sylviculture et d'autres utilisations des terres ; mauvais ciblage des priorités et des projets par les donateurs ; et évaluation incorrecte des progrès réalisés dans le cadre des objectifs de développement, parmi lesquels la gestion durable des forêts (MacDicken *et al.*, 2015).

Le Programme de réduction des émissions liées à la déforestation et à la dégradation des forêts (REDD+), mené sous les auspices de la Convention-cadre des Nations Unies sur les changements climatiques (CCNUCC), est un système de paiement aux résultats octroyé aux pays en développement en échange de la réduction de leurs émissions de gaz à effet de serre (GES) d'origine forestière (CCNUCC, 2014). Ce système impose de définir des seuils de référence relatifs aux pertes de carbone forestier, afin d'apprécier leur évolution au fil du temps, ainsi que des méthodes fiables de suivi, de compte-rendu et de vérification de ces changements. La plupart des pays en développement n'ont pas encore mis en place le système national de suivi forestier exigé par la REDD+.

Bien que certains organismes nationaux (notamment au Brésil, en Inde et au Mexique) enregistrent d'excellents progrès à l'échelle nationale, plusieurs pays tropicaux n'ont établi aucun dispositif de suivi forestier, faute de capacités suffisantes. Des efforts ont été entrepris dans le but de renforcer les compétences techniques nécessaires à la mise en œuvre de systèmes nationaux de suivi forestier à l'échelle institutionnelle (Romijn *et al.*, 2015). Il est essentiel d'aider les pays en développement à consolider et améliorer leurs capacités à produire eux-mêmes des données statistiques sur les ressources forestières.

Au cours des dernières décennies, de grandes organisations internationales, telles que la FAO (voir FAO, 2015a) et l'Institut des ressources mondiales, ont considérablement amélioré la production et la diffusion de données sur les ressources du couvert forestier mondial. Ce chapitre propose des exemples robustes illustrant les progrès méthodologiques réalisés et leur application dans toute la ceinture tropicale, ainsi que les bonnes pratiques adoptées à l'échelle nationale.

Les techniques et méthodes de mesure fondées sur la télédétection ont enregistré des progrès remarquables (Achard et Hansen, 2012). Les données et les méthodes ne constituent plus un obstacle à la mise en œuvre du programme REDD+ au titre de l'Accord de Paris (CCNUCC, 2016). Par ailleurs, les professionnels de l'observation de la Terre et experts du carbone ont publié deux directives techniques visant à aplanir les difficultés méthodologiques liées à l'intégration des observations de télédétection et de terrain dans l'estimation des émissions et de l'élimination des GES forestiers : 1) le guide de référence sur la REDD+ de l'Observatoire mondial de la dynamique du couvert forestier et terrestre (GOFC-GOLD, de l'anglais *Global Observation of Forest and Land Cover Dynamics*) (GOFC-GOLD, 2016), et 2) le document d'orientation et de méthodologie de l'Initiative mondiale pour l'observation des forêts (GFOI, de l'anglais *Global Forest Observations Initiative*) (GFOI, 2014). Ces publications ont pour vocation d'aider les pays à repérer les données manquantes dans les inventaires forestiers nationaux et à mettre en place des systèmes nationaux de suivi forestier qui soient fonctionnels. Les pays sont invités à intégrer les normes internationales dans leurs programmes de suivi forestier, de manière à favoriser la comparabilité internationale.

Le renforcement des capacités de suivi national nécessaires à l'estimation des surfaces forestières bénéficie à terme aux décideurs, aux acteurs économiques, ainsi qu'aux populations subsistant grâce aux forêts. La disponibilité et la qualité des données sur les ressources forestières s'en trouvent améliorées, ce qui donne lieu à des stratégies et à des investissements plus judicieux.

Ce chapitre fournit des directives relatives à l'utilisation de la télédétection aux fins de statistiques forestières, et présente les méthodes de télédétection actuellement utilisées pour évaluer le couvert forestier et l'évolution des forêts à l'échelle mondiale et nationale. Ces travaux visent à soutenir la mise en place d'interventions de REDD+ et de dispositifs de suivi forestier dans les pays.

7.2. INFORMATIONS GÉNÉRALES SUR L'APPLICATION DE LA TÉLÉDÉTECTION AU SUIVI DU COUVERT FORESTIER

À partir des années 1990, les techniques de télédétection sont devenues suffisamment perfectionnées pour permettre l'observation de l'évolution des forêts. La faisabilité et la précision du suivi dépendent en grande partie de paramètres nationaux (notamment la disponibilité des données); en d'autres termes, les éventuels obstacles sont liés à des problèmes de définition, de ressources et d'indisponibilité des données plutôt qu'aux méthodologies proprement dites (GOFC-GOLD, 2016).

7.2.1. Définition des terres forestières, de la déforestation et de la dégradation

Plusieurs termes, définitions et autres éléments en rapport avec la REDD+ ne sont pas formellement arrêtés – notamment certains termes tels que «déforestation» et «dégradation des forêts». Les décisions concernant la REDD+ étant fondées sur les modalités en vigueur prévues par la CCNUCC et le Protocole de Kyoto, les définitions employées dans ces deux documents sont utilisées dans ce chapitre et présentées ci-dessous (voir GOFC-GOLD, 2016, pour plus de détails).

Terres forestières – Aux termes de la CCNUCC, «cette catégorie inclut toutes les terres à végétation ligneuse correspondant aux seuils utilisés dans la définition des terres forestières dans l'inventaire national des gaz à effet de serre. Elle inclut également les systèmes dont la structure végétale est actuellement inférieure aux seuils de la catégorie des terres forestières utilisés par le pays, mais qui pourrait potentiellement les dépasser». Les autres utilisations prédominantes éventuelles doivent également être prises en compte.

L'absence de définition unique des terres dites «forestières» et «non forestières» – dont les critères (taille des arbres, superficie et couvert arboré) sont actuellement très variables – empêche de mesurer correctement l'évolution de la déforestation. Bien qu'il existe d'innombrables définitions du terme «forêt», la plupart appliquent des paramètres communs pour établir certains seuils, concernant notamment la superficie, la hauteur et le couvert forestier minimums. Dans son *Évaluation des ressources forestières mondiales 2015* (FAO, 2015a), la FAO se base sur un couvert minimum de 10 %, une hauteur supérieure ou égale à 5 mètres et une superficie d'au moins 50 ares, ajoutant que l'utilisation prédominante du sol doit être forestière. En revanche, la plupart des études de télédétection se fondent sur une définition de la couverture du sol (Magdon *et al.*, 2014), l'utilisation des terres ne pouvant être déterminée uniquement par télédétection.

De leur côté, les États parties au Protocole de Kyoto ont fixé des seuils communs en matière de couvert boisé, de hauteur des arbres et de superficie pour définir les forêts présentes sur leur territoire national (CCNUCC, 2006). Les fourchettes suivantes sont appliquées, sachant que les jeunes peuplements n'ayant pas encore atteint le couvert ou la hauteur minimum répondent également à la définition du terme «forêt» :

- Superficie minimum : entre 0,05 et 1 hectare ;
- Possibilité d'atteindre une hauteur à maturité *in situ* comprise entre 2 et 5 mètres ;
- Couvert forestier minimum (ou matériel sur pied équivalent) : de 10 à 30 %.

Les pays jouissent donc d'une certaine marge pour établir leurs plans de suivi, l'analyse des données de télédétection pouvant s'adapter aux différents seuils de superficie et de couvert forestiers. Il est néanmoins indispensable d'harmoniser les classifications forestières appliquées à l'ensemble des activités de la REDD+, de manière à pouvoir associer les différents types d'informations, notamment les analyses de télédétection. La multiplicité des définitions a une incidence sur les paramètres techniques de l'observation de la Terre et peut affecter les coûts, la disponibilité des informations et la capacité d'intégration et de comparaison des données dans le temps.

Déforestation – La déforestation renvoie la plupart du temps à la conversion permanente ou à long terme des sols, dont l'utilisation forestière est remplacée par d'autres utilisations non forestières. Selon la décision 16/CMP.1 de la CCNUCC, la déforestation se définit comme « la conversion anthropique directe de terres forestières en terres non forestières ».

Dans la pratique, la déforestation telle que définie précédemment implique une réduction du couvert forestier, qui passe alors en deçà du seuil de définition de la forêt. Elle entraîne un changement de l'utilisation des terres, et généralement aussi de la couverture du sol. La plupart du temps, les forêts sont remplacées par des cultures annuelles ou permanentes (palmiers à huile, arbustes), des pâturages, des espaces urbains ou d'autres infrastructures humaines.

Dégradation des forêts – La dégradation des forêts est la conséquence de plusieurs phénomènes, notamment l'exploitation forestière non viable, l'agriculture itinérante et la collecte ou la combustion de bois de feu. Elle provoque le recul de la biomasse, l'éclaircissement des forêts et la mutation des structures forestières. En outre, elle modifie la composition des essences, ce qui perturbe les services écosystémiques des forêts, notamment leur potentiel futur en matière de piégeage et de stockage du carbone.

Un rapport du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC, 2003) propose cinq définitions potentielles de la dégradation, en précisant leurs avantages et inconvénients respectifs. La dégradation y est qualifiée comme suit :

« Une perte directe, à long terme, due aux activités humaines, persistant pendant X années ou plus, d'au moins Y % de stocks de carbone forestier [et de valeurs forestières], depuis un temps T, et non qualifiable de déforestation ».

Toutefois, les valeurs pour X, Y et T sont vraisemblablement difficiles à harmoniser dans la pratique. On pourrait alors peut-être renoncer à établir une définition spécifique et inclure toute « dégradation des forêts » comme une baisse nette du stock de carbone dans la catégorie « terres forestières restant terres forestières » à l'échelle nationale ou infranationale. Le document d'orientation et de méthodologie du GFOI (GFOI, 2014) n'offre aucune définition officielle de la dégradation, mais présente une procédure par étapes pour l'évaluer, en se fondant sur les méthodes du GIEC.

7.2.2. Caractéristiques du suivi de la déforestation par télédétection

La cartographie et le suivi des forêts tropicales sont un domaine d'application essentiel des technologies d'observation de la Terre. Des données récurrentes et fréquentes sont en effet nécessaires pour permettre l'évaluation annuelle du couvert forestier dans les zones humides et assujétis aux changements de saison, ainsi que le suivi régulier des processus de perturbation des écosystèmes forestiers. Le programme Landsat possède une bibliothèque d'images complète retraçant l'évolution des forêts sur le long terme, exploitée par divers programmes opérationnels et bien développés, tels que le projet de surveillance forestière Global Forest Watch (GFW)¹ de l'Institut des ressources mondiales, et le programme de suivi par satellite de la déforestation de l'Amazonie² (PRODES, de l'anglais *Program for Monitoring Deforestation of the Amazon by Satellite*) de l'Agence spatiale nationale brésilienne (INPE, du portugais *Instituto Nacional de Pesquisas Espaciais*). Les tentatives précédentes d'intégrer des images d'observation de la Terre à haute et moyenne résolutions dans la cartographie et le suivi de la dégradation des forêts ont échoué, en raison notamment de paramètres techniques inadaptés, de coûts élevés et d'incertitudes quant aux débouchés à long terme. Les programmes d'observation de la Terre actuels utilisent principalement des capteurs Landsat (30 m), dont les images couvrent l'ensemble des régions du monde et sont produites tous les ans. Ce type d'imagerie (environ 30 m) sert aujourd'hui principalement à calculer les pourcentages de couvert arboré ou à établir des cartogrammes binaires sur la répartition des zones forestières et non forestières, publiés sur une base annuelle (Hansen *et al.*, 2013).

¹ <http://www.globalforestwatch.org/>

² <http://www.obt.inpe.br/prodes/index.php>

Les techniques de télédétection appliquées au suivi de l'évolution des surfaces forestières (p. ex., la déforestation) permettent d'estimer les surfaces avec une grande précision et pourraient également servir à réaliser des cartographies spatiales des principaux écosystèmes forestiers (GOFC-GOLD, 2016). Pour suivre l'évolution du couvert forestier, il est conseillé d'utiliser au minimum des données de type Landsat (résolution de 30 m) ou supérieures (p. ex., Sentinel-2, qui présente une résolution de 10 m) et d'appliquer une unité cartographique minimale (UCM) située entre 1 et 5 hectares. On pourra ainsi évaluer les changements intervenus dans les zones forestières – et notamment calculer la superficie des zones déboisées et reboisées au cours de la période étudiée. Il est recommandé d'adopter une méthode hybride associant d'une part une segmentation et une classification numériques automatiques et d'autre part une interprétation et/ou une validation visuelle(s) des classes ou des polygones ainsi obtenus, pour des raisons de simplicité, de fiabilité et de rentabilité.

Différentes unités spatiales peuvent être appliquées pour identifier les forêts et détecter leur évolution. Voici quelques exemples d'UCM utilisées par les systèmes nationaux et régionaux actuels de suivi par télédétection : au Brésil, le système PRODES³ est passé d'une UCM initiale de 6,25 ha à une UCM de 1 ha aux fins de traitement numérique ; en Inde, le système national de suivi forestier applique une UCM de 1 ha ; le programme européen CORINE Land Cover, qui surveille l'évolution de la couverture du sol et de l'utilisation des terres, applique une UCM de 5 ha ; enfin, le programme de suivi de la déforestation du ministère de l'Environnement du Pérou et le dispositif de suivi de la déforestation de GFW ont adopté une UCM de 0,1 ha.

Il existe actuellement deux grands fournisseurs de données de télédétection de résolution moyenne (de 30 m × 30 m à 10 m × 10 m), gratuites et couvrant toutes les régions du monde : la NASA (satellites Landsat) depuis le début des années 1980, et l'ESA (satellites Sentinel du programme Copernicus) depuis le milieu des années 2010, bien que leur qualité laisse parfois à désirer en certains endroits des tropiques (en raison de la nébulosité, de la saisonnalité, etc.). Notons que l'intégralité des données d'archives Landsat des États-Unis (notamment celles de l'USGS) est disponible gratuitement, de même que les images satellite du programme d'imagerie spatiale sino-brésilien CBERS (de l'anglais *China-Brazil Earth Resources Satellite*). Le satellite CBERS-4 a été mis en orbite lors de la seconde phase de ce programme de coopération : ses images sont désormais utilisées dans le cadre de grands projets de contrôle de la déforestation et de surveillance de l'environnement en Amazonie. D'autres secteurs, parmi lesquels la surveillance des ressources en eau, l'urbanisation, l'occupation des sols et la formation à la pédologie, bénéficient également des images de ce satellite. Celles-ci sont d'ailleurs essentielles au fonctionnement de grands projets nationaux et stratégiques, au premier rang desquels le projet PRODES⁴ déjà cité et le programme CANASAT⁵ (suivi des plantations de canne à sucre au Brésil). La fusion des données de CBERS-4 et Sentinel-2 est d'ores et déjà à l'étude.

TABLEAU 1. CARACTÉRISTIQUES DES CAPTEURS LANDSAT 8 OLI ET SENTINEL-2

Pays	Satellite et capteur	Résolution et couverture	Coût des données d'archives	Caractéristiques
États-Unis	Landsat 8 OLI	15 m – 30 m 180 × 180 km ²	Données d'archives de l'USGS entièrement gratuites	Données systématiquement recueillies depuis juin 2013
UE	Sentinel-2	10 m – 20 m Fauchée de 290 km	Données d'archives de l'ESA entièrement gratuites	Données systématiquement recueillies depuis juillet 2016

3 Le projet PRODES de l'INPE publie un taux annuel brut de déforestation depuis 1988. Selon ses estimations, environ 750 000 km² de forêts ont disparu dans l'Amazonie brésilienne en 2010, soit à peu près 17 % de la surface originelle de la forêt.

4 <http://www.obt.inpe.br/prodes/>

5 <http://www.dsr.inpe.br/mapdsr/>

Le suivi de la déforestation s'appuie essentiellement sur des données optiques de résolution moyenne (p. ex., les données Landsat). D'autres types de capteurs plus récents, tels que les radars (ERS1/2 SAR, JERS-1, ENVISAT-ASAR et ALOS PALSAR 1/2) et la technologie lidar (détection et télémétrie par la lumière), pourraient eux aussi se prêter efficacement à une telle utilisation (De Sy *et al.*, 2012). Les radars, en particulier, s'accommodent relativement bien de la nébulosité persistante de certaines régions des tropiques, très contraignante pour l'imagerie optique. Bien que les lidars et les radars aient fait leurs preuves dans le cadre de projets d'étude, leur utilisation effective aux fins de suivi forestier à grande échelle ne s'est pas encore généralisée. L'utilisation des radars pourrait se populariser dans les années à venir en fonction de l'acquisition des données, de l'accès aux informations et des progrès scientifiques.

7.2.3. Caractéristiques du suivi de la dégradation des forêts par télédétection

Si la dégradation des forêts peut, la plupart du temps, être identifiée par télédétection, les méthodes optimales de suivi ne sont pas nécessairement les mêmes selon le type et le lieu de la dégradation, et le type de forêt concernée. Des approches robustes de suivi de la dégradation (et de la repousse) des forêts sont toujours en cours de développement. Comme l'indique le guide de référence sur la REDD+ du GOF-C-GOLD (2016), la dégradation ou la repousse des forêts, et les changements induits dans les stocks de carbone forestier, sont plus difficiles à quantifier que la déforestation. Ils sont en effet difficilement repérables par télédétection, sans compter qu'ils nécessitent des données de meilleure qualité, à des intervalles plus fréquents, et un traitement plus approfondi.

Certains obstacles techniques limitent la détection et l'enregistrement des évolutions de la canopée servant au suivi de la dégradation des forêts. Les changements modérés passent en effet généralement inaperçus, à moins de former un motif systématique sur les images satellite. Toutes les activités – par ailleurs nombreuses – qui contribuent à dégrader les stocks de carbone forestier ne peuvent pas être suivies efficacement et avec précision. Certaines doivent être observées avec d'autres moyens que la télédétection (Miettinen *et al.*, 2014). Avant de mettre en place un système de suivi de la dégradation des forêts, il est nécessaire d'évaluer les causes de la dégradation et ses conséquences probables sur les stocks de carbone forestier :

- Les zones forestières exploitées de manière sélective – où se trouvent des coupes, des routes et des dépôts de grumes – sont généralement visibles sur les images satellite, notamment le réseau routier et les dépôts. Les éclaircissements de la canopée dus à l'abattage des arbres sont visibles sur certaines images, notamment celles de Landsat, dont les techniques d'analyse plus avancées permettent de traiter des clichés recueillis à intervalle fréquent (Grecchi *et al.*, 2017).
- La dégradation des stocks de carbone forestier par les feux de forêt risque d'être plus difficile à surveiller avec les moyens d'imagerie satellite actuels (Miettinen *et al.*, 2016). Les feux de forêt tropicaux sont d'origine anthropique dans presque 100 % des cas.
- Les zones dégradées à cause de la surexploitation (bois de feu ou autre utilisation locale du bois) deviennent généralement des aires de pacage, ce qui empêche la régénération des peuplements – une situation plus fréquente dans les forêts plus sèches. Ces changements ne sont guère repérables sur les images satellite, sauf en cas de dégradation sévère provoquant une modification sensible de la canopée.

7.2.4. Disponibilité des données Landsat

En 1972, la NASA a mis en orbite le premier satellite Landsat doté d'un capteur de résolution moyenne capable de recueillir des données d'observation de la Terre à l'échelle du paysage. Six autres satellites Landsat d'observation de la Terre ont été lancés depuis (à raison d'un lancement tous les deux ou trois ans), qui ont permis au programme de recueillir des données en continu depuis 1972. Landsat 7 est toujours en service et couvre le même tracé avec une fréquence de revisite de 16 jours. Le dernier satellite en date, Landsat 8, a été mis en orbite en 2013 dans le cadre de la mission de continuité des données Landsat visant à poursuivre le programme de la NASA.

La quasi-totalité des données satellite mondiales produites par Landsat depuis le début des années 1990 peut être téléchargée gratuitement sur le portail en ligne de l'USGS⁶; elles contiennent notamment les corpus de données de l'enquête mondiale d'observation de la Terre (en anglais, *Global Land Survey [GLS]*). Ces derniers sont essentiels pour déterminer les taux historiques de déforestation, bien que la nébulosité persistante de certaines régions tropicales humides (en Afrique centrale, par exemple) en limite considérablement l'efficacité. Les archives complètes des missions Landsat 8 OLI (depuis juin 2013), Landsat 7 ETM+ (depuis 1999), Landsat 5 TM (depuis 1984), Landsat 4 TM (1982-1985) et Landsat 1-5 MSS (1972-1994) peuvent être commandées gratuitement auprès de l'USGS.

À ce jour, le programme Landsat reste la principale source d'images satellite de résolution moyenne (10-50 m) utilisées aux fins d'analyse, en raison de son faible coût et de sa licence d'utilisation libre. Deux facteurs principaux limitent l'utilisation des capteurs Landsat, à savoir la nature mixte du signal mesuré et le repérage laborieux des perturbations du couvert forestier. Ce deuxième aspect revêt une importance particulière dans les zones où se produisent de nombreux changements à petite échelle. Les autres sources de données satellite sont notamment ASTER, SPOT, IRS, CBERS, DMC, AVNIR-2 et Sentinel-2.

7.2.5. Disponibilité des données Sentinel-2

Le satellite Sentinel-2A (S2A) du programme Copernicus a été mis en orbite en 2015; depuis octobre 2015, il fournit des images optiques sur de larges périmètres avec des résolutions de 10 m (dans le visible et le proche infrarouge), 20 m (red-edge, proche infrarouge et infrarouge à ondes courtes) et 60 m (du visible à l'infrarouge à ondes courtes, à des fins de correction atmosphérique).

Il présente une large fauchée (290 km), une fréquence de revisite de 10 jours et assure une couverture globale (il inclut les masses terrestres). Le lancement de l'unité jumelle B du satellite S2, prévu pour 2017, permettra de faire tomber sa fréquence de revisite à cinq jours. Le lancement des versions C et D de ces satellites est d'ores et déjà envisagé, afin de garantir la disponibilité des données au moins jusqu'en 2027.

Les capacités propres aux capteurs de Sentinel-2 – et de Landsat 8 – pourront servir de base à un ensemble de services forestiers viables et harmonisés au niveau mondial, ouvrant ainsi la voie à diverses applications innovantes et ambitieuses, ainsi qu'à une refonte des systèmes de suivi, dans le but d'évaluer la dégradation des forêts avec une meilleure précision.

Le niveau de produit 2A de Sentinel-2 est le niveau standard des outils de traitement d'images contenus dans la boîte à outils de Sentinel-2 fournie par le programme Copernicus de l'ESA⁷. Les produits de niveau 1C appliquent des corrections radiométriques et géométriques (notamment l'orthorectification et le recalage spatial dans le cadre d'une géométrie cartographique fixe). Les produits de niveau 2A se situent quant à eux en bas de l'échelle de réflectance atmosphérique dans la géométrie cartographique, et nécessitent un traitement par l'utilisateur. Notons que les données satellite relatives aux pays membres de la REDD+ exigeraient un traitement supérieur de niveau 3. Cette procédure nécessite une mosaïque d'images acceptables (présentant une nébulosité minimale) provenant des séries temporelles réalisées par Sentinel-2 au-dessus de la ceinture tropicale, à des intervalles de 30 jours ou de trois mois. Les caractéristiques d'un produit principal standard de niveau 3 de Sentinel-2, mises au point par le programme Copernicus en 2017, autorisent sa mise à disposition systématique et gratuite sur une plateforme de distribution en accès libre.

La qualité technique des capteurs Sentinel facilite considérablement la distinction des classes de couverture du sol dans le cadre de l'utilisation des terres forestières, pour les forêts (détermination du type de forêt), mais aussi pour les paysages mixtes complexes mêlant terres agricoles et forestières (notamment les zones où se pratique l'agriculture itinérante). La résolution spatiale de 10 à 20 m de S2A (et S2B), associée à un temps de revisite de 10 jours (ou

⁶ <http://glovis.usgs.gov/>

⁷ http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Data_products

cinq jours en cas d'utilisation combinée des deux satellites), permettra d'évaluer sans difficulté l'état du couvert forestier et de localiser les petites zones de perturbation (à l'échelle des parcelles et des plants). L'excellent niveau de détail spatial et la possibilité de choisir les images saisonnières les plus utiles faciliteront le suivi des processus de transformation plus lents, notamment la disparition progressive du bois de feu ou la repousse des forêts due à l'abandon des terres agricoles. Le S2 se distingue également des autres capteurs par la complémentarité des bandes (visible, proche infrarouge et infrarouge à ondes courtes). Enfin, précisons que la compatibilité des bandes du S2 et de Landsat 8, de même que leurs niveaux très supérieurs de correction atmosphérique, augmenteront considérablement la convergence entre les capteurs et le potentiel de fusion des données.

La résolution spatiale (10 m) et la fréquence temporelle (temps de revisite de cinq ou dix jours) supérieures des images prises par Sentinel-2 permettront de repérer et de quantifier la dégradation des forêts dans les pays tropicaux de manière plus fiable et régulière que ne le permettent les images satellite actuelles de résolution moyenne. Les images de Sentinel-2 devraient donc permettre d'intensifier à brève échéance l'évaluation des conditions forestières.

L'arrivée de Sentinel-2 pourrait entraîner une transmission des capacités de suivi forestier aux autorités nationales et régionales au cours des cinq à dix prochaines années, sous forme, par exemple, d'un module rattaché ou directement intégré au système d'inventaire forestier national. Cette démarche nécessitera de renforcer considérablement les capacités; la priorité devra, dans la mesure du possible, être accordée à l'élaboration et l'utilisation d'un cadre méthodologique rigoureux. Ce dernier devra permettre de normaliser au maximum les estimations et les cartogrammes nationaux des surfaces forestières, et de garantir des niveaux de précision et de qualité homogènes, afin de pouvoir intégrer les données dans les applications régionales et mondiales.

En résumé, les données de type Landsat sont celles qui se prêtent le mieux à l'évaluation des taux historiques et des tendances liées à la déforestation. Le nombre de données Landsat 8 et Sentinel-2 accessibles librement et gratuitement est en augmentation depuis quelques années, ce qui devrait permettre d'évaluer plus en détail les images datant de moins de cinq ans dans plusieurs régions du monde.

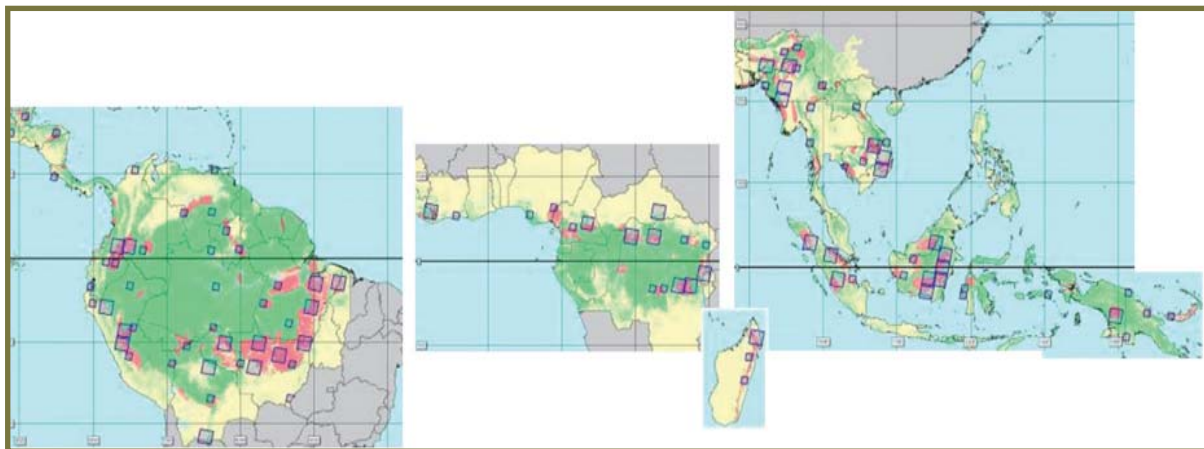
7.3. ÉVALUATION DES RESSOURCES FORESTIÈRES MONDIALES DE LA FAO : ENQUÊTE MONDIALE PAR TÉLÉDÉTECTION

7.3.1. Informations générales sur l'échantillonnage statistique conçu pour estimer la déforestation à partir de capteurs optiques de résolution spatiale moyenne

Dans l'idéal, il faudrait pouvoir analyser la totalité des surfaces forestières (couverture intégrale ou «de mur-à-mur») à l'aide de techniques d'imagerie de résolution spatiale moyenne (de type Landsat); toutefois, ce procédé risque de ne pas être adapté aux étendues vastes et hétérogènes. Il limiterait également d'autant les ressources disponibles pour l'analyse. Concernant l'analyse numérique à partir d'images satellite de résolution moyenne prises à l'échelle pantropicale ou continentale, plusieurs échantillonnages réussis ont été effectués à partir de la superficie forestière totale, afin de réduire les coûts et le temps d'analyse.

S'il reflète correctement les épisodes ponctuels de déforestation, l'échantillonnage permet de mettre en évidence les mouvements tendanciels de la déforestation (Achard *et al.*, 2002 ; Richards *et al.*, 2000). La répartition spatiale des épisodes de déforestation n'étant pas aléatoire, le modèle statistique doit être échantillonné dans les zones présentant un risque de déforestation (figure 1), en réalisant par exemple un échantillonnage systématique très forte densité (petit pas), si les ressources sont disponibles (Mayaux *et al.*, 2005).

FIGURE 1. RÉPARTITION DES UNITÉS D'ÉCHANTILLONNAGE DE L'ENQUÊTE TREES-II



Achard *et al.*, 2002 ; Richards *et al.*, 2000

Dans le cadre du programme d'évaluation des ressources forestières mondiales (FRA, de l'anglais *Forest Resources Assessment*) de 2010, la FAO a continué d'intensifier le suivi des changements du couvert forestier à l'échelle mondiale et continentale, en vue de compléter les rapports nationaux relatifs à cette question. Les progrès technologiques et l'amélioration de l'accessibilité des données de télédétection lui ont permis d'élargir la portée du FRA 2010 par rapport à celle du FRA 2000. Les conclusions de l'enquête par télédétection sur les forêts tropicales du FRA 2000 (figure 2) ont fait l'objet d'un chapitre dans le rapport principal y afférent (FAO, 2001) et d'un rapport par Drigo *et al.* (2009).

FIGURE 2. RÉPARTITION DES UNITÉS D'ÉCHANTILLONNAGE DE L'ÉVALUATION DES RESSOURCES FORESTIÈRES MONDIALES 2000



FAO, 2001 ; Drigo *et al.*, 2009

7.3.2. Méthode d'échantillonnage générale utilisée dans l'enquête mondiale par télédétection

Les enquêtes mondiales par télédétection des FRA 2010 et 2015 couvrent l'ensemble des régions (et pas uniquement les zones pantropicales). Elles avaient pour fonction d'évaluer les changements survenus dans les écosystèmes forestiers à partir d'un échantillon d'images satellite de résolution moyenne. Conçues de manière à produire des

estimations cohérentes et comparables de l'évolution sur vingt ans du couvert arboré et de l'utilisation des terres forestières, à l'échelle mondiale et régionale, elles avaient pour objectif de compléter les statistiques nationales extraites des enquêtes nationales par télédétection, de plus en plus abondantes dans les rapports principaux des FRA.

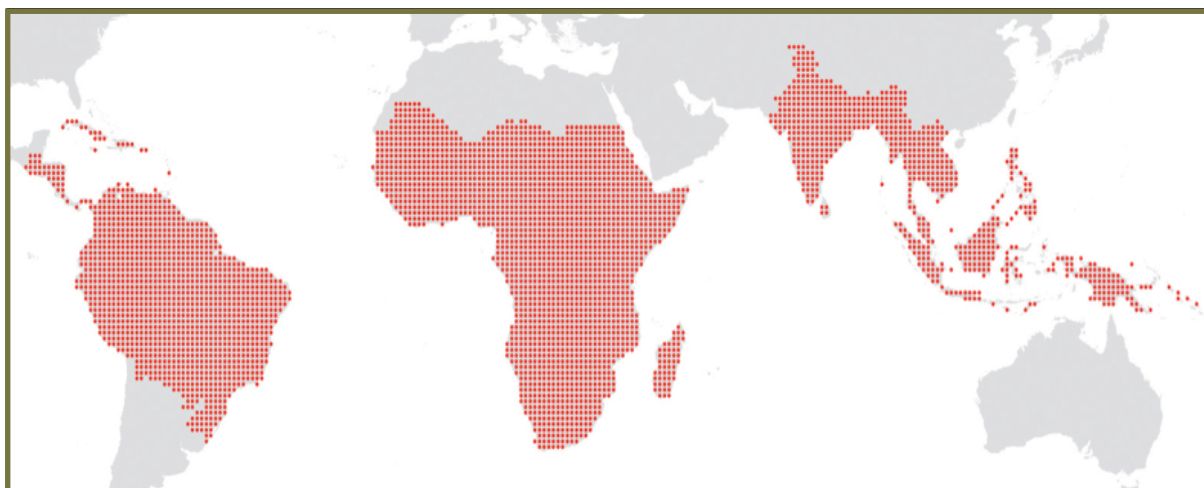
La FAO et le Centre commun de recherche (JRC) de la Commission européenne ont publié des estimations conjointes des changements de l'utilisation des terres forestières entre 1990 et 2005, pour l'enquête mondiale par télédétection de 2010 (FAO et JRC, 2012). Ce projet a ensuite été prolongé jusqu'en 2010 (pour couvrir la période 1990-2010) afin d'alimenter le FRA 2015 (Achard *et al.*, 2014 ; Keenan *et al.*, 2015).

L'enquête par télédétection du FRA 2010 se fonde sur un maillage de petites unités d'échantillonnage beaucoup plus serré que pour les rapports précédents, et une grille systématique : les unités correspondent à chaque point d'intersection des parallèles et des méridiens à raison d'un espacement de $1^{\circ} \times 1^{\circ}$. Ce procédé d'échantillonnage systématique général a été élaboré conjointement par la FAO et le JRC pour évaluer le taux de déforestation à l'échelle mondiale et continentale tous les cinq à dix ans (Mayaux *et al.*, 2005).

Chaque unité d'échantillonnage mesure $10 \text{ km} \times 10 \text{ km}$ et comporte une zone tampon de 5 km permettant d'obtenir des données contextuelles sur le couvert forestier (la taille complète des unités est donc de $20 \text{ km} \times 20 \text{ km}$ pour les informations relatives à la couverture du sol). Ces proportions s'expliquent par le fait qu'elles permettent un suivi spatialement explicite à une échelle permettant d'observer la gestion des terres.

Cet échantillonnage comprend donc environ 13 500 unités couvrant les zones terrestres du globe (environ 9 000 si l'on exclut les zones désertiques) et représente environ 1 % de la surface terrestre (0,8 % le long de l'équateur) où s'applique la grille géographique (figure 3).

FIGURE 3. RÉPARTITION DES UNITÉS D'ÉCHANTILLONNAGE DE L'ENQUÊTE MONDIALE PAR TÉLÉDETECTION SUR LES FORÊTS TROPICALES DU FRA 2010



7.3.3. Sélection et prétraitement des images satellite

L'enquête par télédétection du FRA 2010 de la FAO est une étude mondiale fondée sur des méthodes cohérentes et des séries temporelles régulières. Ces dernières peuvent être prolongées afin d'y intégrer les années plus récentes. Chaque point d'échantillonnage fait l'objet de séries temporelles constituées de données de télédétection de résolution moyenne, selon une procédure normalisée, décentralisée et dont la qualité est contrôlée. Les paragraphes suivants décrivent succinctement le jeu de données satellite et la procédure de traitement préliminaire utilisée par la FAO dans le cadre de l'enquête mondiale par télédétection sur les forêts tropicales du FRA 2010.

Pour chacune des unités d'échantillonnage, des images Landsat (E) TM orthorectifiées ont été recueillies gratuitement des archives du GLS, créées et mises à disposition par l'USGS (Gutman *et al.*, 2013). Quatre images ont ensuite été sélectionnées pour chaque unité : elles devaient présenter le moins de nébulosité possible et se rapprocher au maximum de la date cible du 30 juin pour les années 1990, 2000, 2005 et 2010. En cas de données GLS indisponibles, de mauvaise qualité (p. ex., les données Landsat 7 SLC-off) ou présentant une couverture nuageuse excessive sur certaines unités d'échantillonnage (Potapov *et al.*, 2010), d'autres scènes satellite tirées des archives Landsat de l'USGS ou d'autres agences spatiales (dont l'INPE) ont été utilisées (Beuchle *et al.*, 2011). Les images obtenues correspondaient aux périodes 1986–1993, 1999–2003, 2004–2007 et 2009–2011 pour les années 1990, 2000, 2005 et 2010 respectivement.

Les images sélectionnées ont subi de nombreux traitements préliminaires : contrôle de la géolocalisation, conversion en réflectance au sommet de l'atmosphère, masquage des nuages et des ombres et normalisation de l'image sur la base de paramètres pseudo-invariables (Bodart *et al.*, 2011). Les images multidates doivent impérativement présenter une bonne correspondance géométrique. Dans le cas qui nous occupe, la géolocalisation des images a parfois nécessité de recourir à certaines optimisations. L'image Landsat ETM (2000) a été choisie comme image de référence à cette fin ; les images secondaires (pour la plupart Landsat 5) ont été recalées pour s'aligner correctement sur l'image de référence.

7.3.4. Traitement et analyse des images satellite

Cette section présente le processus d'analyse des forêts tropicales évoqué précédemment.

Après le traitement préliminaire, les images satellite ont subi une segmentation automatique multidates, ce qui a permis de fragmenter l'unité d'échantillonnage (10 000 ha) en cellules délimitées (polygones) présentant des caractéristiques spectrales et structurelles similaires. L'UCM visée était de 5 ha. Les images segmentées ont ensuite été classées selon la couverture du sol, de manière automatisée et supervisée, puis classées en fonction de l'utilisation des terres, avec l'aide de spécialistes en photo-interprétation.

Les images prétraitées des quatre « époques » (soit les années 1990, 2000, 2005 et 2010) correspondant à chaque unité d'échantillonnage ont subi une segmentation en plusieurs étapes à l'aide du logiciel eCognition (Trimble©). Elle ont ensuite fait l'objet d'un processus de classification fondée sur les objets selon des fonctions d'appartenance, elles-mêmes définies à partir d'un ensemble de signatures spectrales prises le long de la ceinture tropicale (Raši *et al.*, 2011 et 2013). L'interprétation des images pour le repérage des changements du couvert forestier était basée sur des UCM de 5 ha (50 pixels d'une résolution de 30 m × 30 m). Une « échelle d'unité de détection » plus fine (environ 1 ha) a été utilisée durant la première phase de segmentation et d'étiquetage automatisés ; les éléments ont ensuite été regroupés en objets de 5 ha au cours de la phase d'interprétation.

Les objets ont été regroupés en cinq classes de couverture du sol : couvert arboré, mosaïque de couvert arboré, autres terres boisées, autre couverture du sol, et eau (voir le tableau 2 pour une description de chaque classe). La classe « couvert arboré » correspond à la définition de la forêt de la FAO (FAO, 2010).

TABLEAU 2. CLASSES DE COUVERTURE DU SOL UTILISÉES PAR LE JRC

Classe	Description
Couvert arboré	Objets couverts de 70 à 100 % d'arbres, lesquels se définissent comme des plantes d'une hauteur supérieure à 5 m dotées d'une tige ligneuse. La densité du couvert forestier est supérieure à 30 %.
Mosaïque de couvert arboré	Objets couverts de 30 à 70 % d'arbres.
Autres terres boisées	Objets couverts de plus de 50 % de plantes d'une hauteur inférieure à 5 m dotées ou non d'une ou de plusieurs tiges ligneuses.
Autre couverture du sol	Terres n'appartenant à aucune des catégories précédentes et constituées de prairies naturelles, de terres agricoles, de zones bâties, de sols nus et de roche.
Eau	Cours et étendues d'eau

Les informations relatives à la couverture du sol de chacun des objets ainsi classés (UCM de 5 ha) ont ensuite été corrigées en profondeur pour chaque année cible (Eva *et al.*, 2012).

Les chercheurs du JRC et de la FAO ont collaboré avec plus de 100 experts de la télédétection et de la sylviculture originaires de pays tropicaux, notamment de pays possédant de grandes superficies forestières tels que le Brésil, l'Inde, l'Indonésie et la République démocratique du Congo.

Notons que pour l'enquête par télédétection du FRA 2010 (FAO et JRC, 2012), la FAO ne s'est pas appuyée sur une classification de la couverture du sol comme celle du JRC (de Sy *et al.*, 2015) mais sur une classification de l'utilisation des terres (FAO, 2010b). Celle-ci comprend notamment une classe « forêt »⁸ et permet une meilleure estimation des facteurs. Dans l'enquête de la FAO, les jeunes plantations forestières sont considérées comme des « forêts » (si les arbres sont capables d'atteindre une hauteur supérieure à 5 m) ou comme une « autre couverture du sol » d'après la légende du JRC (si les arbres sont invisibles ou mesurent moins de 5 m).

7.3.5. Analyse statistique

Les données sur la nature et l'évolution de la couverture du sol de l'ensemble des unités d'échantillonnage permettent d'obtenir des estimations statistiques pour toute la zone d'étude. Étant donné le très petit nombre d'images satellite obtenues exactement à la même date en 1990, 2000, 2005 et 2010, les données sur (l'évolution de) la couverture du sol de chaque unité ont été dans un premier temps « normalisées » sur une base linéaire (selon une approximation la plus précise possible) à la date cible du 30 juin pour chaque année considérée, afin d'obtenir des statistiques sur la couverture du sol. Nous partons ici de l'hypothèse que les changements de la couverture du sol se sont opérés de façon linéaire dans le temps.

Les zones pour lesquelles aucune donnée n'était disponible en raison de la couverture nuageuse, d'une couverture satellitaire insuffisante ou de la faible qualité des images, quelle que soit l'année de référence, ont été considérées comme une perte de données non biaisée. Les chercheurs ont supposé en outre qu'elles comptaient les mêmes proportions de couverture du sol que les zones non nuageuses au même endroit. Ils ont pour cela modifié les matrices de changement de la couverture du sol des décennies 1990-2000 et 2000-2010, afin que leur taille soit relative à la superficie terrestre sans nuage totale des unités d'échantillonnage. Les unités manquantes (4, 39 et 3 pour la période 1990–2000 et 3, 39 et 3 pour la période 2000–2010, à partir de 1 230, 2 045 et 741 unités d'échantillonnage au total couvrant respectivement l'Amérique du Sud, l'Afrique et l'Asie du Sud-Est) ont été remplacées par les moyennes locales de sites d'échantillonnage environnants. Ces moyennes ont été obtenues en appliquant les pondérations suivantes ($\delta_{jj'}$) :

$$\delta_{jj'} = 1/d(j, j') = 1 / \left(\text{dif}(\text{lat})^4 + \text{dif}(\text{long})^4 \right) \quad [1]$$

où $d(j, j')$ correspond à la distance entre deux sites.

Pour la phase d'estimation statistique, les unités d'échantillonnage ont été pondérées en fonction de leur probabilité d'inclusion. En effet, la distance entre les unités varie selon qu'elles se situent sur un parallèle ou un méridien : toutes les unités ne présentent donc pas la même probabilité, même si le plan d'échantillonnage est systématique. Pour tenir compte de ces écarts de probabilité, chaque unité a reçu une pondération égale au cosinus de la latitude. L'effet de ces pondérations dans les zones tropicales reste toutefois modéré. Les unités sélectionnées contenant

8 Dans le FRA 2010, la classe « forêt » renvoie aux « terres occupant une superficie de plus de 0,5 hectare avec des arbres atteignant une hauteur supérieure à cinq mètres et un couvert arboré de plus de dix pour cent, ou avec des arbres capables d'atteindre ces seuils *in situ*. Sont exclues les terres à vocation agricole ou urbaine prédominante ».

une zone de mer ont été considérées comme des sites entiers et ont permis à ce titre de compenser les unités non sélectionnées contenant des surfaces terrestres (lorsque l'unité présentait une zone de mer en son milieu).

Les proportions des changements de la couverture du sol ont ensuite été extrapolées à la zone d'étude à l'aide de l'estimateur par expansion directe Horvitz-Thompson (Särndal *et al.*, 1992). L'estimateur de chaque changement de classe de couverture du sol correspond à la part moyenne de ce changement par unité d'échantillonnage, qui s'obtient par l'équation ci-dessous :

$$\bar{y}_c = \frac{1}{m} \sum_{i=1}^n w_i \cdot y_{ic} \quad [2]$$

où y_{ic} correspond à la part de changement de couverture du sol pour une classe de transition particulière dans la $i^{\text{ème}}$ unité d'échantillonnage. La pondération de l'unité d'échantillonnage est w_i tandis que m renvoie à la somme des échantillons pondérés.

En cas d'échantillonnage systématique, un biais positif est attribué à l'estimateur habituel «du sondage aléatoire simple» (Stehman *et al.*, 2011). Le recours à d'autres estimateurs fondés sur une estimation de la variance locale permet de réduire le biais, et donc de contourner en partie cet écueil. L'exemple ci-dessous utilise un estimateur de l'erreur type à partir d'une estimation de la variance locale :

$$s^2 = (1-f) \frac{\sum_{j \neq j'} w_{jj'} \delta_{jj'} (y_j - y_{j'})^2}{2 \sum_{j \neq j'} w_{jj'} \delta_{jj'}} \quad [3]$$

où f correspond au taux de sondage, la pondération $w_{jj'}$ est une moyenne des pondérations w_j et $w_{j'}$, et $\delta_{jj'}$ renvoie à une fonction décroissante [1] de la distance entre j et j' (à noter que s'il est déterminé que $\delta_{jj'} = 0$, on obtient l'estimateur de la variance habituel). L'erreur type (se) est ensuite calculée comme suit :

$$se = \frac{s}{\sqrt{n}} \quad [4]$$

où n correspond au nombre total de sites d'échantillonnage disponibles (à l'exclusion des sites manquants, même ceux pour lesquels on utilise une moyenne locale).

Les changements de couverture du sol ont été estimés en évaluant les matrices de transition (voir tableau 3) pour les décennies 1990-2000 et 2000-2010. Les objets étiquetés «mosaïque de couvert arboré» étaient considérés comme étant composés à 50 % de forêts, ce chiffre correspondant à la moyenne des limites supérieure et inférieure de pourcentage. La perte de couvert forestier a donc été calculée ainsi : 100 % du couvert arboré convertis en autres terres boisées, en autre couverture du sol ou en eau + 50 % du couvert arboré convertis en mosaïque de couvert arboré + 50 % de la mosaïque de couvert arboré convertis en autre couverture du sol, en autres terres boisées ou en eau.

7.3.6. Évaluation de la précision et de la cohérence des estimations des changements du couvert forestier

Les observations (données d'origine) à partir desquelles ont été obtenus les résultats présentés dans ce chapitre proviennent de l'interprétation d'images satellite. Ces alternatives aux observations de terrain présentent un risque d'erreur ou d'incertitude (biais) (Foody, 2010); ces difficultés n'ont toutefois pas été corrigées dans la présente évaluation. Le suivi de l'évolution des forêts tropicales doit souvent se satisfaire de données de substitution lorsqu'aucune observation de terrain n'est effectuée. Dans d'autres cas, seule l'interprétation des données satellite permet de surveiller, à un coût raisonnable, de grandes étendues forestières inaccessibles. Une évaluation indépendante a tout de même été réalisée sur 1 185, 1 552 et 830 points (soit un total de 3 567 points) systématiquement répartis dans un sous-échantillon aléatoire de 240, 338 et 166 unités couvrant respectivement l'Amérique du Sud, l'Afrique et l'Asie du Sud-Est (un point central auquel on ajoute quatre points dans les angles pour chaque unité d'échantillonnage). Tous les points considérés comme un changement de couverture du sol au cours de la décennie 1990-2000 ont également été sélectionnés, en se fondant sur une grille systématique de 9×9 (soit 81 points séparés de 1 km dans chaque unité d'échantillonnage). Cela représente respectivement 1 663, 1 194 et 1 425 points (soit 4 282 points au total) pour les trois sous-régions. Les polygones correspondants ont ensuite été soigneusement réinterprétés à vue par des experts indépendants en s'aidant de toutes les ressources secondaires disponibles (telles que les images de Google Earth, après vérification de la date de prise des clichés). Cette méthode a permis d'évaluer la «cohérence» des résultats de l'interprétation.

Afin de parachever cette évaluation, les résultats d'une sélection aléatoire de 34 des 411 unités d'échantillonnage couvrant l'Amazonie brésilienne (Eva *et al.*, 2012) ont été comparés aux interprétations de l'INPE pour la décennie 1990-2000 (INPE, 2013).

7.3.7. Résultats pour les tropiques

Les résultats des périodes 1990-2000 et 2000-2005 de l'enquête mondiale par télédétection du FRA 2010 ont été publiés à l'échelle mondiale (FAO et JRC, 2012), et ceux des décennies 1990-2000 et 2000-2010 à l'échelle des régions ou des pays tropicaux (Achard *et al.*, 2014; Beuchle *et al.*, 2015; Bodart *et al.*, 2013; Eva *et al.*, 2012; Mayaux *et al.*, 2013; Stibig *et al.*, 2014). Cette section présente brièvement les principaux résultats pour la région tropicale, afin d'illustrer les réalisations de l'enquête par télédétection du FRA 2010.

En 1990, on recensait 1 635 millions d'hectares (Mha) de forêt tropicale et 964 Mha d'autres terres boisées. En 2010, la superficie de la forêt tropicale n'était plus que de 1 514 Mha, soit une perte nette globale sur 20 ans de respectivement 56,9, 30,9 et 32,9 Mha en Amérique centrale/Amérique du Sud/Caraïbes, en Afrique subsaharienne et en Asie du Sud et du Sud-Est. À l'opposé, la superficie des autres terres boisées s'est étendue au cours de la même période pour atteindre 975 Mha, une hausse intervenue principalement en Asie du Sud-Est (18,6 Mha). En 2010, les forêts tropicales humides représentaient environ 64 % du couvert forestier tropical, soit 972 des 1 514 Mha répartis entre l'Amérique du Sud (599 Mha), l'Afrique (210 Mha) et l'Asie du Sud-Est (163 Mha).

La perte brute des forêts tropicales mondiales a atteint 8 Mha a-1 durant les années 1990 (soit 0,497 % par an) avant de fléchir légèrement à 7,6 Mha a-1 dans les années 2000 (0,494 % par an). Ce ralentissement est principalement dû à la baisse des taux de déforestation dans les forêts humides d'Afrique (de 0,70 à 0,36 Mha a-1) et d'Asie du Sud-Est (de 1,70 à 1,22 Mha a-1) (Achard *et al.*, 2014). De grandes étendues non forestières ont également été recolonisées par la forêt, soit 1,9 Mha a-1 dans les années 1990 et 1,6 Mha a-1 dans les années 2000.

7.3.8. Précision des estimations pour les tropiques

Les statistiques des changements des zones forestières (perte brute, gain brut et perte nette) contiennent peu d'erreurs types en raison de la taille importante des échantillons : de 4 à 10 % à l'échelle mondiale et de 11 à 19 % en moyenne à l'échelle régionale. La précision des cartes de couverture du sol des unités d'échantillonnage des régions tropicales a été évaluée pour la période 1990-2000. Le taux de similitude global entre les résultats de

l'enquête par télédétection et les réinterprétations considérées comme des données de référence s'élève à 92,9 % pour la superficie forestière et à 85,5 % pour les changements du couvert forestier. Le risque de biais (imputable à des erreurs d'interprétation) a été évalué en confrontant les estimations tirées des échantillons à celles tirées des données de référence. L'écart relatif ainsi obtenu est de -8,9 % pour l'estimation de la superficie forestière mondiale (les estimations de l'enquête par télédétection sont inférieures à celles découlant des données de référence) et de 11,2 % pour l'estimation de la déforestation brute mondiale (les estimations de l'enquête par télédétection sont supérieures). La comparaison d'une sélection aléatoire de 34 sites échantillonnés avec les interprétations de l'INPE pour la décennie 1990-2000 fait apparaître une bonne équivalence entre les résultats de l'INPE et ceux de l'enquête, tant pour la superficie forestière (1990) que pour la déforestation (1990-2000). Certaines courbes sont proches de 1 (1,017 et 1,008 respectivement) tout comme R^2 (0,986 et 0,978 respectivement) (Eva *et al.*, 2012).

7.3.9. Intensification du système d'échantillonnage pour l'estimation des changements du couvert forestier à l'échelle nationale

Le système mondial d'échantillonnage systématique décrit précédemment peut être intensifié afin d'obtenir des résultats à l'échelle nationale. Les estimations relatives à la déforestation découlant de deux niveaux d'intensité d'échantillonnage ont été comparées avec les estimations issues des inventaires officiels menés en Amazonie brésilienne et en Guyane française (Eva *et al.*, 2010).

Neuf lots de données échantillonnées ont été extraits de la carte intégrale officielle de la déforestation établie à partir des images satellite prises au-dessus de l'Amazonie brésilienne de 2002 à 2003 (INPE, 2016). À partir de ces informations, le système mondial d'échantillonnage systématique a pu estimer la superficie de la déforestation à 2,8 Mha – avec une erreur type de 0,1 Mha. Ce chiffre rejoint l'estimation du peuplement total réalisée à partir des interprétations «de mur-à-mur» de la déforestation, soit 2,7 Mha. L'écart relatif entre l'estimation moyenne découlant de la méthode d'échantillonnage et l'estimation du peuplement total s'établit à 3,1 %; l'erreur type représente 4 % de l'estimation du peuplement total. Lors des essais, le modèle d'échantillonnage systématique des forêts de l'Amazonie brésilienne a débouché sur une erreur type faible (moins de 5 % du taux de changement du couvert forestier).

En ce qui concerne la Guyane française, des images Landsat 5 TM ont été utilisées pour la période de référence historique (1990) et des images SPOT visibles à haute résolution (20 m × 20 m) pour l'année 2006, afin d'accroître le taux de sondage de l'étude. Les estimations ainsi obtenues du taux de déforestation en Guyane française sur la période 1990-2006 sont comparées avec celles publiées par l'autorité nationale compétente pour ses départements d'outre-mer en vertu des règles du Protocole de Kyoto (Stach *et al.*, 2009). Les statistiques officielles se fondent sur un système d'échantillonnage composé de près de 17 000 parcelles choisies selon les méthodes traditionnelles de l'inventaire forestier national (IFN) français et analysées à partir d'images satellite identiques obtenues entre 1990 et 2006. L'écart relatif entre les chiffres de l'IFN et ceux de l'étude intensifiée s'établit à 5,4 %. On constate donc, d'après ces résultats et ceux d'autres études (Steininger *et al.*, 2009), que l'intensification du système mondial d'échantillonnage permet d'obtenir des estimations du changement des superficies forestières proches de celles des IFN, avec un écart inférieur à 10 %. À noter toutefois que les erreurs découlant de l'utilisation de données de substitution ne sont pas prises en compte, et que seules les erreurs estimées de l'échantillonnage sont considérées.

7.3.10. L'avenir de l'évaluation des ressources forestières mondiales : objectif FRA 2020

Le processus du FRA est en constante amélioration : les évaluations sont optimisées d'une année sur l'autre du fait de nouveaux besoins d'informations, de la production de nouvelles données de meilleure qualité, et de l'introduction de méthodes et de technologies innovantes. Les récents événements, parmi lesquels la signature de l'Accord de Paris et le lancement des objectifs de développement durable (ODD), impliquent d'importants changements en matière de gestion forestière et de procédures : la portée et la fréquence de publication du FRA doivent donc être ajustées afin de répondre aux nouveaux besoins d'information.

Depuis 30 ans, le FRA bénéficie du soutien et des conseils techniques d'experts internationaux, par le biais de consultations régulières organisées par la FAO et la Commission économique des Nations Unies pour l'Europe (CENUE). La première consultation s'est tenue en 1987 ; les suivantes ont eu lieu en 1993, 1996, 2002 et 2006 (Kotka I-V) à Kotka (Finlande) et en 2012 à Ispra (Italie). La dernière consultation s'est tenue en juin 2017 à Joensuu, en Finlande.

La consultation vise à obtenir des recommandations quant à la portée de l'évaluation mondiale suivante (notamment la procédure de compte-rendu par les pays et la composante de télédétection) et à établir la fréquence de publication des rapports sur les principales variables et des rapports annuels sur les indicateurs des ODD.

7.4. AUTRES EXEMPLES D'ENQUÊTES PAR TÉLÉDÉTECTION AUX FINS DE STATISTIQUES FORESTIÈRES

7.4.1 Statistiques sur la déforestation du produit « couvert arboré mondial », université du Maryland

L'université du Maryland a récemment présenté une nouvelle méthode fondée sur les bonnes pratiques du GIEC et sur plusieurs ensembles de données de télédétection (De Sy *et al.*, 2012), visant à quantifier les pertes de carbone de surface dans les forêts tropicales sur la période 2000-2012 (Tyukavina *et al.*, 2015). Cette étude vient compléter les travaux déjà menés en République démocratique du Congo et au Pérou (Tyukavina *et al.*, 2013 ; Pelletier et Goetz, 2015).

Plus précisément, Tyukavina *et al.* (2015) appliquent une méthode associant un procédé d'échantillonnage et un ensemble de données « de mur-à-mur » sur la perte de couvert arboré (Hansen *et al.*, 2013), afin d'estimer le recul des forêts tropicales.

Le projet sur le « couvert arboré mondial » mené par l'université du Maryland (Hansen *et al.*, 2013) comporte des données intégrales remontant à 2000. À partir d'images Landsat, les chercheurs ont réalisé une cartographie complète, divisant les terres en quatre classes de « couvert arboré » (0–25 %, 26–50 %, 51–75 % et 76–100 %). Ils ont découvert que dans les tropiques, la superficie de la classe comprise entre 76 et 100 % (correspondant approximativement à la forêt tropicale humide) était de 1 324 Mha en 2000. En revanche, la surface dont le couvert arboré était supérieur à 25 % (soit 2 094 Mha) présentait un ordre de grandeur similaire aux estimations du FRA 2015 pour l'ensemble de la forêt tropicale, dont le seuil de couvert arboré était fixé à 10 % (Keenan *et al.*, 2015).

Tyukavina *et al.* (2015) ont réalisé une estimation non biaisée des pertes forestières à l'aide d'un échantillon aléatoire stratifié de 3 000 pixels (chaque pixel mesurant environ 0,1 ha) répartis dans les zones de forêt tropicale. Ils ont également fait la distinction entre « forêts naturelles » (forêts primaires et forêts secondaires matures, et forêts claires naturelles) et « forêts aménagées » (plantations, systèmes de sylviculture et zones d'agriculture de subsistance impliquant une rotation du couvert arboré). Leurs recherches confirment la valeur de la méthode d'échantillonnage, qui permet d'estimer les pertes de couvert forestier de manière plus fiable qu'une méthode de mur-à-mur, avec des résultats beaucoup plus élevés. Ceux-ci s'expliquent par les changements intervenant à petite échelle, qui ne figurent pas sur la carte complète de la perte de couvert forestier. La prise en compte adéquate de ces changements modestes peut s'avérer fort utile, puisqu'elle permet aux pays d'établir des seuils de référence précis.

Les estimations des pertes de couvert forestier peuvent varier très sensiblement suivant les définitions et les méthodes employées : pour Tyukavina *et al.*, les forêts sont des territoires présentant un couvert boisé supérieur à 25 %, tandis que la FAO s'appuie sur un seuil de 10 % et sur la définition de l'utilisation des terres. Par ailleurs, Tyukavina *et al.* se limitent aux pertes forestières brutes, tandis que la FAO rend compte des pertes forestières nettes (en intégrant le boisement et la repousse des forêts) (Keenan *et al.*, 2015).

Les travaux de Tyukavina *et al.* (2015) illustrent le potentiel actuel de l'échantillonnage des données satellite, dans l'optique de quantifier les pertes du couvert forestier tropical et les fuites de carbone associées.

7.4.2. Exemple au niveau national : unités paysagères de l'inventaire forestier national du Brésil

La superficie du Brésil est d'environ 8,5 millions de km², dont 4,9 millions sont recouverts de forêts (FAO, 2015b). Les forêts revêtent non seulement une importance environnementale et socio-économique capitale pour le pays, mais leurs services écosystémiques (conservation de la biodiversité, rétention du carbone) sont également essentiels pour l'ensemble de la planète. L'IFN du Brésil est assuré par le service forestier brésilien (SFB, du portugais *Serviço florestal brasileiro*)⁹, sous la tutelle du ministère de l'Environnement, en partenariat avec d'autres institutions telles que l'Entreprise brésilienne de recherche agricole (Embrapa, du portugais *Empresa Brasileira de Pesquisa Agropecuária*), des organismes publics spécialisés dans l'environnement, des universités, des instituts de recherche et des conservatoires botaniques. L'IFN est l'un des éléments phares du système national d'information sur les forêts (Freitas *et al.*, 2010) et constitue à ce titre l'une des principales étapes de la production de données fiables et régulières sur les ressources forestières (SFB, 2016).

L'IFN brésilien repose sur un cycle d'inventaire de cinq ans et vise essentiellement à recueillir des données sur les ressources forestières (forêts naturelles et plantées), afin d'orienter les politiques publiques relatives à l'exploitation et à la conservation de celles-ci. Quelques-uns des 27 États du pays ont déjà entamé le second cycle d'inventaire, mais la majorité n'a pas encore achevé le premier cycle.

Un plan d'échantillonnage systématique est appliqué à l'aide d'une grille nationale de points d'échantillonnage équidistants placés à un intervalle de 648 secondes, soit un périmètre par point d'environ 20 km × 20 km à l'équateur. Des grappes ou «unités d'échantillonnage parcellaires», composées chacune de quatre sous-unités de 20 m × 50 m, sont ensuite tirées sur la grille. Les données collectées sur le terrain comprennent des variables biophysiques portant sur l'état de la forêt et du milieu, ainsi que des variables socio-économiques (entretiens) pour comprendre l'utilisation et la perception des ressources forestières par les populations locales (Freitas *et al.*, 2010). Certains rapports préliminaires font également état des stocks forestiers et de la composition, de la santé et de la vitalité des écosystèmes forestiers. La comparaison des résultats des cycles d'inventaire successifs permet de dégager les caractéristiques des changements survenus dans le temps.

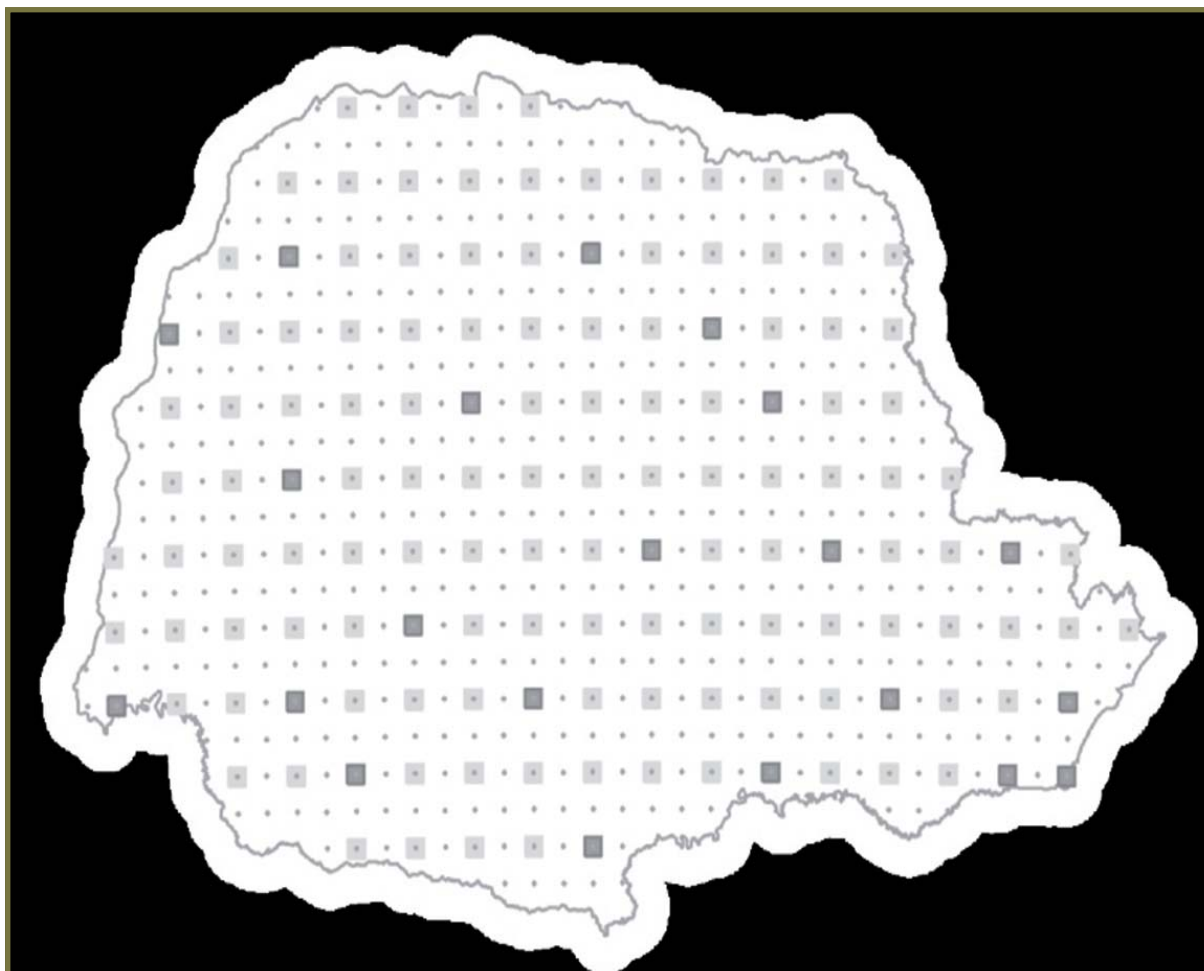
Outre les données collectées tous les 20 km × 20 km sur l'ensemble du territoire par les équipes de terrain, l'IFN renferme également une composante géospatiale qui fournit des données paysagères par le biais d'unités d'échantillonnage paysagères (UEP), d'une carte de l'utilisation des terres et de la couverture du sol (UTCS) et d'analyses spatiales (Luz *et al.*, 2015b). La méthodologie des UEP est le fruit d'une collaboration entre la FAO, le SFB¹⁰ et l'équipe de l'Embrapa chargée de la foresterie¹¹. Le plan d'échantillonnage utilisé pour collecter des données à l'échelle du paysage doit s'articuler sur une grille systématique identique à celle utilisée sur le terrain, mais avec des sous-échantillons systématiques de 40 km × 40 km (figure 4) (Freitas *et al.*, 2006). Chaque UEP mesure 10 km × 10 km (100 km²), son centre géométrique correspondant à l'emplacement d'une grappe de parcelles.

9 L'IFN brésilien est dirigé par Joberto Veloso de Freitas et Claudia Melo Rosa.

10 Naissa Batista da Luz et Jessica Maran.

11 Maria Augusta Doetzer Rosot, Marilice Cordeiro Garrastazú et Yeda Maria Malheiros de Oliveira.

FIGURE 4. RÉPARTITION DES UNITÉS D'ÉCHANTILLONNAGE PAYSAGÈRES DE L'IFN DU BRÉSIL DANS L'ÉTAT DU PARANÁ



Au Brésil, l'exploitation et la modification des ressources naturelles font depuis peu l'objet d'une réglementation intégrée dans le nouveau Code forestier, lequel prévoit notamment la mise en place d'un cadastre environnemental rural. Depuis 2011, le ministère de l'Environnement se procure chaque année des images satellite RapidEye couvrant l'ensemble du territoire, afin de faire appliquer cette réglementation. Ces données sont également mises à disposition des autres organismes publics. L'équipe de l'Embrapa chargée de la foresterie, le SFB et la FAO ont donc mené, dans le cadre de l'IFN, une étude paysagère pilote à l'aide des images RapidEye et Landsat 8. Une méthode d'analyse d'images fondée sur les objets a été élaborée à partir d'images RapidEye orthorectifiées (et mise en œuvre à l'aide du logiciel Definiens). Les polygones obtenus par segmentation des images RapidEye ont ensuite été classés à l'aide de plusieurs couches secondaires (indices de végétation augmentés, par exemple), de couches statistiques de séries temporelles (moyenne annuelle, valeur minimale/maximale, écart type) et de données du projet Global Forest Change (couvert arboré de 2013, par exemple). Le traitement a été effectué par l'éditeur de code de Google Earth. Une classification des images RapidEye et Landsat 8 non supervisée et fondée sur les pixels a également été effectuée à l'aide du logiciel IMPACT Toolbox (mis au point par le JRC). Ces images ont été intégrées comme données auxiliaires dans la classification des polygones tirés des images RapidEye (Luz *et al.*, 2015b).

Le paysage tel qu'appréhendé dans l'IFN brésilien forme un ensemble hétérogène d'écosystèmes, que l'on retrouve dans divers types d'UTCS en interaction les uns avec les autres (Luz *et al.*, 2015a). La mosaïque des classes d'UTCS – dont les composantes d'origine naturelle et anthropique influent sur la qualité des ressources forestières

disponibles – est définie de la manière suivante : (a) arbres/arbustes ; (b) forêt plantée ; (c) prairies naturelles ; (d) terres agricoles et pâturages ; (e) zones urbaines ; (f) sols nus ; (g) eau.

Une fois que l'UTCS a été cartographiée suivant les classes susmentionnées, la structure de chaque UEP est analysée. Cette méthodologie (spécialement adaptée aux UEP de l'IFN) aborde de manière innovante l'organisation spatiale du paysage, les mosaïques d'UTCS, ainsi que le morcellement, la connectivité et l'interface des écosystèmes. Outre les indicateurs traditionnels, tels que la composition du paysage, le modèle repose sur un procédé de fragmentation original : l'application d'un indice normalisé et comparable fondé sur la distance euclidienne globale de l'écosystème. Une autre méthode consiste à évaluer la qualité des zones ripariennes en étudiant leur connectivité structurelle (corridors biologiques) ainsi que le degré de pression anthropique, et en simulant des scénarios de protection. Ces zones sont également classées (Clerici et Vogt, 2012) en fonction de leur priorité de conservation. Cette méthode revêt d'autant plus d'importance que le Brésil vient de modifier, dans son Code forestier, la superficie de la végétation forestière devant être restaurée le long des berges. Les « arbres hors forêt » forment une rubrique spécifique de l'analyse des UEP ; plusieurs méthodes de classification et de différenciation de ces arbres ont été testées à l'aide d'images satellite RapidEye. Les définitions et critères pertinents ont été établis par la FAO en partenariat avec l'Institut de recherche pour le développement (IRD) (De Foresta *et al.*, 2013).

Les indicateurs paysagers (et leurs indices respectifs) suivants sont en cours de chiffrage :

- Composition du paysage (pourcentage de couvert d'arbres et d'arbustes, notamment la forêt naturelle, les autres terres boisées et les arbres hors forêt) et part des autres espaces naturels et/ou semi-naturels (notamment les prairies naturelles et les forêts plantées) ;
- Classification du paysage selon le degré (pourcentage) de présence de chaque classe d'UTCS ;
- Analyse morphologique du modèle spatial des habitats intégrée par Soile et Vogt (2008) dans le logiciel Guidos Toolbox (Vogt *et al.*, 2007 ; Saura *et al.*, 2011), qui permet de définir les catégories ou classes (principale, bordure, perforation, pont, boucle, branche et îlot) ;
- Mosaïque de paysage forestier, qui comprend plusieurs classes et indices (et permet d'étiqueter un lieu donné suivant l'étendue des zones d'agriculture intensive et des zones urbaines environnantes) et est intégrée dans le logiciel Guidos Toolbox ;
- Modèle d'interface en bordure, intégré dans le logiciel Guidos Toolbox, qui produit divers indices et évalue l'importance de la fragmentation liée au changement de l'utilisation des terres et de la couverture du sol en bordure des forêts ;
- Connectivité du paysage, qui englobe les connexions prioritaires du paysage et se fonde sur l'analyse morphologique du modèle spatial et le logiciel Conefor¹² (Saura et Torné, 2009) ; un classement des corridors structurels menacés présents dans le paysage est également proposé ;
- Fragmentation du paysage : cet indicateur aux concepts innovants vise à quantifier la fragmentation (dans le logiciel Guidos Toolbox) ; il permet de comparer les taux de fragmentation en différents points, et de mesurer et surveiller l'évolution du processus de morcellement ;
- Analyse des zones ripariennes fondée sur leur connectivité structurelle (corridors biologiques), le degré de pression anthropique et la simulation de scénarios de protection de ces milieux, selon les concepts exposés par Clerici *et al.* (2011) et Ivits *et al.* (2009).

Ces indicateurs, représentés par des groupes d'indices, permettent d'évaluer la qualité structurelle des UEP. La somme des indices sélectionnés sur une base linéaire pondérée permet d'obtenir un score unique par UEP et donc d'étudier et de comparer les unités entre elles, à des fins de restauration et de suivi de certains éléments paysagers.

Les travaux de l'IFN contribuent fortement à l'engagement du gouvernement brésilien en faveur du développement durable. En 2010, le pays s'est en effet volontairement engagé à réduire ses émissions de 80 % en Amazonie et de 40 % dans le Cerrado (région de Savannah) à l'horizon 2020. Il prévoit en outre d'intégrer les instruments existants

12 <http://www.conefor.org>

et d'encourager la coordination et les synergies entre eux, afin d'optimiser les résultats de la REDD+. Les outils de l'IFN peuvent soutenir ces engagements. Par ailleurs, le programme de l'IFN peut contribuer à faciliter la mise en œuvre et le suivi d'autres politiques nationales, notamment celles relatives aux forêts plantées et à l'intégration de l'agriculture, du bétail et de la sylviculture (agroforesterie iLPF, du portugais *integração lavoura pecuária floresta*).

La contribution prévue déterminée au niveau national (CPDN) du Brésil est jugée particulièrement ambitieuse. Dans les domaines de la sylviculture, de l'utilisation des terres et de la couverture du sol, le pays a fait les propositions suivantes : 1) améliorer le respect du nouveau Code forestier à tous les niveaux ; 2) renforcer les stratégies et les mesures dans le but d'éradiquer la déforestation illégale en Amazonie et de compenser les émissions de GES imputables au défrichement autorisé à l'horizon 2030 ; 3) restaurer et planter 12 Mha de forêt d'ici 2030 pour des usages multiples ; 4) diversifier les systèmes de gestion durable des forêts primaires grâce à l'utilisation de systèmes de géoréférencement et de traçabilité adaptés, afin de dissuader les auteurs de pratiques illégales et non durables. En outre, le pays s'est engagé à consolider son plan relatif à l'agriculture sobre en carbone (plan ABC, du portugais *Agricultura de Baixa Emissão de Carbono*) pour en faire le principal vecteur du développement durable dans le secteur agricole. Celui-ci prévoit notamment la restauration de 15 Mha de pâtures dégradées et une augmentation de 5 Mha des projets d'agroforesterie iLPF à l'horizon 2030. Les données de l'IFN sur les ressources forestières, l'utilisation des terres, la couverture du sol et la qualité des paysages s'avéreront donc fort utiles à la réalisation des objectifs et cibles de la CPDN.

L'analyse paysagère vient compléter la collecte de données sur le terrain et l'enquête socio-économique, les deux autres volets de l'IFN brésilien. Tous les volets de l'IFN ont été conçus de manière à pouvoir être intégrés et permettre une analyse conjointe. Les éléments physiques et biologiques recueillis lors de l'enquête de terrain (méthode fondée sur l'écosystème) et mélangés à des données spatialisées sur l'utilisation des terres, la couverture du sol et l'environnement socio-économique, peuvent être ainsi conformes à la méthode qui porte désormais le nom de « méthode fondée sur le paysage ».

7.4.3. Enquête mondiale de la FAO sur les forêts

Objectifs de l'enquête mondiale sur les forêts

L'enquête mondiale sur les forêts (EMF), qui, comme son nom l'indique, a une portée internationale, vise principalement à fournir des estimations mondiales et régionales des données des inventaires forestiers relatives à certains écosystèmes forestiers spécifiques. Ces données sont recueillies via un réseau mondial de parcelles.

L'EMF poursuit les objectifs spécifiques suivants :

- établir un réseau mondial d'échantillons de parcelles permanents, si possible sur des parcelles existantes, mais également sur de nouveaux terrains, le cas échéant ;
- produire des estimations mondiales géoréférencées et détaillées sur le carbone forestier, l'état de santé des forêts et d'autres éléments forestiers, par le biais des échantillons de parcelles et des données de télédétection ;
- créer un portail d'information et établir des règles de partage des données pour garantir l'accès libre et gratuit aux données et résultats ;
- mettre en évidence l'intérêt de la disponibilité d'une banque de données en ligne unique, permanente et librement accessible.

Les données devraient être recueillies en suivant un plan d'échantillonnage multi-échelles ; des protocoles de mesure seront mis au point pour évaluer les paramètres simples (p. ex., le couvert arboré) et complexes (p. ex., les types d'utilisation des terres) des forêts. Les données seront recueillies par les organisations partenaires, les autorités et les communautés locales et, le cas échéant, directement par la FAO. Elles seront librement accessibles sur un portail en ligne muni d'un système d'information géographique (SIG).

Première évaluation mondiale des zones arides

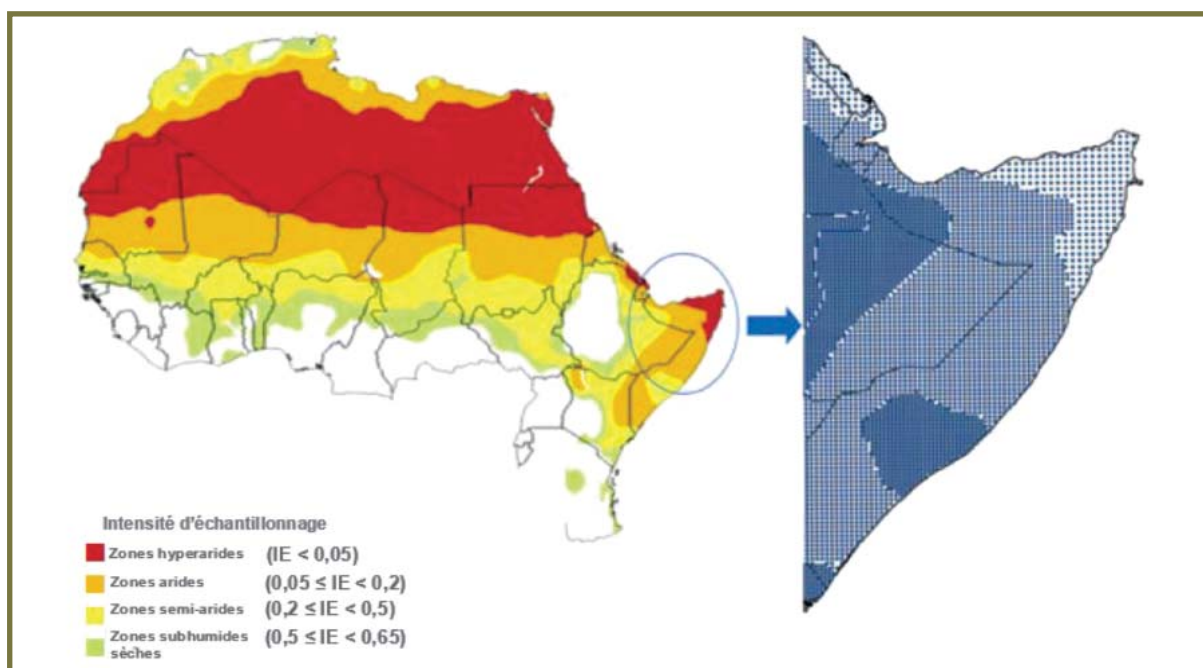
L'évaluation mondiale des zones arides (FAO, 2016) est une initiative pilote du projet de données libres sur les forêts mondiales (*World Forest Open Data*). Elle se concentre exclusivement sur les zones arides et n'utilise que

des images satellite accessibles au grand public, telles que celles fournies par Google Earth Engine et Bing Maps. La première édition de l'évaluation a réuni plus de 200 experts spécialisés dans la composition et l'utilisation des terres de régions arides spécifiques.

Les données de l'évaluation proviennent de 213 795 parcelles échantillonnées dans les différentes zones arides du globe. Chaque parcelle mesure 70 m × 70 m (environ 0,5 ha), soit la taille minimale requise pour qu'un terrain soit considéré comme une forêt suivant la définition du FRA (FAO, 2015a).

Chaque zone aride de la grille d'échantillonnage des parcelles a été traitée comme une strate indépendante. Le taux de sondage relative attribuée était de 0,5 pour les zones hyperarides, 1 pour les zones arides et 1,5 pour les zones semi-arides et subhumides sèches. Les données sont communiquées à l'échelle mondiale et régionale (FAO, 2016). L'erreur d'échantillonnage statistique de l'estimation de la superficie forestière totale dans l'ensemble des zones arides est de l'ordre de 1 %.

FIGURE 5. ILLUSTRATION DU TAUX DE SONDAGE DE LA PREMIÈRE ÉVALUATION MONDIALE DES ZONES ARIDES



FAO, 2016

Données et outils utilisés pour l'évaluation mondiale des zones arides

L'enquête a été préparée à l'aide du logiciel Collect (disponible dans le pack Open Foris¹³), puis intégrée dans l'outil Collect Earth, développé en collaboration avec Google Earth Outreach (Bey *et al.*, 2016).

Les données des parcelles échantillonnées ont été extraites de bibliothèques en ligne d'images satellite à l'aide de Collect Earth. Chaque parcelle s'étendait généralement sur plusieurs images obtenues à partir de Google Earth Engine et Bing Maps. Collect Earth est capable de visualiser les valeurs de réflectance et les indices définis par l'utilisateur (p. ex., l'indice différentiel normalisé de végétation) à partir des images Landsat et MODIS. Le travail

¹³ Open Foris est un ensemble d'outils mis au point par le Département des forêts de la FAO pour améliorer la flexibilité et l'efficacité de la collecte, l'analyse et la communication des données. Voir <http://www.openforis.org/>

d'interprétation visuelle a été facilité par les profils temporels des indices de végétation interannuels, calculés à partir de données satellite de faible résolution (résolution au sol de 30 à 250 m).

Des images Landsat (d'une résolution de 30 m) étaient disponibles pour chaque parcelle ; 89 % des terrains étaient également couverts par des images de plus haute résolution, dont plus de la moitié provenaient des satellites de DigitalGlobe et présentaient une résolution spatiale de moins d'un mètre. Le nombre de types d'images satellite était similaire pour toutes les classes d'utilisation des terres.

Dans Collect Earth, chaque parcelle échantillonnée se présente sous forme d'une grille de 49 points de contrôle permettant à l'utilisateur d'estimer avec précision la superficie des parcelles recouvertes d'arbres, d'arbustes ou d'autres éléments. Chaque photo-interprète mise sur ses connaissances du lieu et sur les données de télédétection pour faire progresser l'enquête.

Interprétation des images relatives à l'utilisation des terres

Pour chaque parcelle, des données portant sur plus de 70 variables ont été recueillies et enregistrées au point le plus récent photographié par les satellites. Ces variables permettent de définir les paramètres de la couverture du sol, de l'utilisation des terres, des changements intervenus dans l'utilisation des terres, et des autres processus clés (désertification, verdissement, etc.), en association avec les indicateurs biophysiques. L'année 2000, considérée comme la première année pour laquelle il existe une couverture mondiale cohérente par des données satellite (Landsat 7), a été choisie comme année de référence.

L'emploi simultané d'images satellite à faible et très haute résolutions permet de repérer plus facilement le type d'utilisation des terres et les changements survenus. Les données satellite et les connaissances locales se sont avérées insuffisantes dans certains cas, par exemple pour distinguer les arbres et les arbustes. Les chercheurs ont par conséquent décidé de se baser sur le diamètre de la cime : les sujets dont la cime mesurait plus de 3 m de diamètre étaient considérés comme des arbres, et les sujets plus petits comme des arbustes. Collect Earth ne permettant pas de mesurer directement la hauteur des arbres, c'est leur ombre (lorsqu'elle était visible) qui a été utilisée en complément du diamètre de la cime : les sujets d'au moins 5 m étaient considérés comme des arbres (même seuil que celui utilisé dans la définition de la forêt suivie par la FAO dans l'évaluation mondiale des zones arides).

L'utilisation des terres est évaluée d'après les six catégories définies par le GIEC (GIEC, 2006) : forêt, sol cultivé, prairie, zone humide, habitations, autres terres. Une utilisation prédominante est attribuée à chaque parcelle échantillonnée, à partir de la présence d'indicateurs clés de l'utilisation des terres, interprétés selon une règle hiérarchique (Martinez et Mollicone, 2012). Une parcelle au couvert arboré de plus de 10 % n'est donc pas considérée comme une forêt, sauf si son utilisation dominante correspond à la catégorie « forêt » du GIEC. Seule une utilisation des terres peut être attribuée à chaque parcelle.

Résultats et retours d'expérience de la première évaluation mondiale des zones arides

La FAO s'est appuyée sur cette méthode pour étudier environ 214 000 parcelles échantillonnées (Bastin *et al.*, 2017) dans le cadre de l'évaluation mondiale des zones arides (FAO, 2016). Il s'agit de la première évaluation statistique fondée sur un échantillonnage de l'utilisation des terres dans les zones arides, qui intègre notamment les forêts et le couvert arboré. Elle constitue ainsi une référence à partir de laquelle il est possible de suivre l'évolution des écosystèmes forestiers, du couvert arboré et de l'utilisation des terres au niveau mondial, régional et par zone aride.

Selon ce premier rapport, les zones arides du globe abritent 1,11 milliard d'hectares de forêts. Plus de la moitié (566 Mha) se trouvent dans la zone subhumide sèche, principalement au nord-est de l'Afrique australe et dans la zone pré-andine, à l'ouest de l'Amérique du Sud. Environ deux tiers des forêts des zones arides (742 Mha) présentent un couvert arboré supérieur à 40 % et sont donc considérées comme des forêts denses. Parmi elles, plus de la moitié (pour la plupart situées en Europe et en Amérique du Sud) possèdent un couvert d'au moins 90 %.

Les auteurs de l'évaluation se sont appuyés sur des bibliothèques en ligne d'images satellite gratuites et accessibles au grand public, un outil d'interprétation simple et le recueil systématique des données par des experts de l'utilisation des terres. Cette méthode présente les avantages suivants :

- une étude statistique d'un ensemble de variables de base peut être menée rapidement et à moindres frais en complément d'autres procédés, les parcelles échantillonnées étant analysées grâce à des images satellite et non sur le terrain ;
- il est possible de mobiliser de nombreux participants du fait de la simplicité du processus d'interprétation et de la gratuité des données et du logiciel.

La première évaluation mondiale des zones arides a été menée de manière participative et collaborative, et les premiers résultats (FAO, 2016; Bastin *et al.*, 2017) ont été obtenus moins d'un an après sa conception. Si elle a été l'occasion de tester la méthode à l'échelle mondiale, celle-ci doit encore être adaptée aux zones non arides, notamment afin de réduire le risque d'interprétation erronée.

Le manque de cohérence constitue une source d'erreur potentielle : plus de 200 photo-interprètes ont été sollicités, et toutes les images des 213 795 parcelles évaluées ne provenaient pas de la même source. Pour pallier ce risque, tous les photo-interprètes ont été dotés des mêmes outils et modules de formation. Des mesures supplémentaires visant à réduire les incohérences et les erreurs d'interprétation seront introduites dans le cadre d'une évaluation pilote portant sur tous les types de terres menée à l'échelle de la planète.

Cette méthode peut être adaptée pour cibler les besoins spécifiques de pays et d'autres utilisateurs souhaitant mener un échantillonnage plus intensif au niveau régional, national et paysager. Un exemple d'utilisation régionale est celui de l'évaluation de référence de la « Grande muraille verte » menée dans plus de 20 pays au nord et au sud du Sahara, dont les données viennent enrichir celles déjà recueillies en Afrique du Nord, dans le Sahel et dans la Corne de l'Afrique.

Prochaines étapes

Les résultats de l'évaluation mondiale des zones arides ont fait l'objet d'un rapport publié début 2017 (Bastin *et al.*, 2017), après des levés et analyses de terrain supplémentaires. Cette évaluation devrait avoir lieu tous les deux ans.

L'utilisation du logiciel Collect Earth et des autres outils utiles au suivi et aux évaluations de référence doit être encouragée par le biais de manifestations et d'ateliers de renforcement des capacités au niveau régional et national. Ces outils permettent en effet d'évaluer de manière innovante et économique les arbres, les forêts, la nature et les changements de l'utilisation des terres quel que soit le milieu, comme en témoigne cette première évaluation mondiale des zones arides.

La FAO prévoit d'appliquer cette méthode à une étude pilote mondiale portant sur tous les types de terres, en échantillonnant environ 250 000 parcelles supplémentaires qui seront photo-interprétées à l'aide de Collect Earth. Des statistiques forestières mondiales seront dressées pour un total d'environ 500 000 parcelles mesurant approximativement 0,5 ha, puis intégrées dans la base de données de l'enquête mondiale sur les forêts.

7.5. COMPLÉMENTARITÉ DES ESTIMATIONS DES CHANGEMENTS DANS LES DOMAINES FORESTIER ET AGRICOLE

La complémentarité des estimations des changements dans les domaines forestier et agricole est mise en évidence dans l'étude financée par la Commission européenne (CE) et menée par VITO, l'Institut international pour l'analyse des systèmes appliqués (IIASA), l'Institut de recherche sur le travail et la société (HIVA) et le Comité national des Pays-Bas de l'Union internationale pour la conservation de la nature (UICN), sur l'impact de la consommation européenne sur la déforestation (ci-après le « rapport technique de la CE ») (Commission européenne, 2013).

Déforestation mondiale brute pour la période 1990-2008 (FRA 2010)

Selon le FRA 2010, le volume mondial brut de la déforestation pour la période 1990-2008 se chiffre à 239 Mha, soit environ 13 Mha en moyenne par an, avec des écarts importants entre les régions. Au cours de la même période, la déforestation brute a été en partie compensée par le boisement et l'expansion naturelle des forêts (115 Mha), ce qui donne lieu à une déforestation nette de 124 Mha (FAO, 2010a).

Selon le rapport technique de la CE (2013), qui s'appuie sur les bases de données du FRA 2010 (FAO, 2010a) et du domaine de l'utilisation des terres de FAOSTAT (FAOSTAT, 2011), la déforestation brute est imputable à cinq facteurs principaux : l'expansion des terres agricoles ; l'exploitation forestière (qui précède la mise en culture) ; l'urbanisation ; les risques naturels (p. ex., les incendies) ; et les facteurs inexpliqués. Les changements d'utilisation des terres liés à ces facteurs, enregistrés dans FAOSTAT à l'aide d'un modèle de transition, correspondent aux zones de déforestation identifiées dans le FRA 2010. Notons qu'une petite part de la déforestation due à l'expansion des terres agricoles est imputée à la catégorie « exploitation forestière pour la production industrielle de bois rond » afin de tenir compte du défrichement précédant la mise en culture des terres forestières.

Déforestation inexpliquée

Les cas de déforestation ne pouvant être attribués ni à l'agriculture, ni à l'exploitation forestière, ni à l'urbanisation, ni aux risques naturels sont regroupés sous la rubrique « facteurs inexpliqués » (Commission européenne, 2013). Sur les 239 Mha de déforestation brute à l'échelle mondiale, environ 58 Mha (24 %) ne peuvent être imputés avec certitude à la conversion des forêts à des fins manifestes de consommation, ni à d'autres facteurs connus. Les données relatives à la déforestation constituent la plus grande source d'incertitude. La FAO estime pourtant que la différence de superficie de 9 % entre le rapport principal et l'enquête par télédétection du FRA 2010 est tout à fait honorable, si l'on considère la diversité des méthodes utilisées (FAO et JRC, 2012). Par ailleurs, ni l'évaluation au niveau des pays ni l'enquête par télédétection du FRA 2010 ne réussissent à mesurer précisément le couvert arboré dans les zones plus clairsemées (entre 10 et 30 %), ce qui laisse planer des incertitudes notables en ce qui concerne les zones arides et les forêts dégradées. Contrairement au continent africain, les taux de déforestation en Amérique du Sud rapportés dans l'enquête par télédétection et le rapport principal du FRA 2010 sont relativement cohérents. Enfin, les statistiques sur la production agricole et les données sur le commerce ont été jugées plutôt fiables.

La déforestation inexpliquée peut être en partie attribuée à des erreurs statistiques sur la déforestation (données en trop) et la superficie occupée par les zones agricoles au niveau national (données insuffisantes), comme ce fut le cas dans l'enquête par télédétection du FRA 2010. Elle peut aussi correspondre, dans une certaine mesure, aux effets de la dégradation à long terme provoquée par certaines pratiques informelles, telles que l'exploitation forestière illégale et la collecte non durable du bois de feu. On suppose que la déforestation liée à ces pratiques est prise en compte dans cette catégorie, faute de statistiques mondiales complètes sur ce phénomène et de corrélation avérée entre ces pratiques et leurs effets. Enfin, il peut arriver que la transformation des forêts en terres agricoles soit en réalité plus importante que ne le laisse supposer la surface agricole productive (lorsqu'elle est communiquée).

Facteurs de la déforestation

Sur les 182 Mha déboisés restants, environ 41 Mha (17 %) sont le fait de risques naturels (incendies d'origine naturelle ou humaine principalement) qui n'ont pas été inclus dans la catégorie d'expansion des terres agricoles (Commission européenne, 2013). L'Indonésie a perdu à elle seule 9 Mha de forêts en raison de l'ENSO, en 1997 et 1998. Près de 9 Mha (4 %) supplémentaires ont disparu à cause de l'urbanisation. Les 132 Mha restants (55 % de la déforestation mondiale brute) sont attribuables de toute évidence à la mise en culture des terres forestières, à l'élevage de ruminants et à la production industrielle de bois rond (bûcheronnage).

Sur les 132 Mha (55 %) de forêts déboisées en faveur de la production mondiale de produits agricoles et forestiers, seuls 4,5 Mha (2 %) ont été attribués à l'exploitation forestière : ils représentent donc uniquement l'abattage effectué en amont de la transformation en terres agricoles (Commission européenne, 2013).

Enfin, sur la déforestation totale imputable à l'agriculture (128 Mha ou 53 %), 69 Mha (29 %) ont été défrichés directement ou indirectement pour répondre à la demande mondiale en matière d'alimentation humaine et animale, de biocarburants et de fibres. Environ 58 Mha (24 %) de forêts ont été mis en herbage pour le bétail.

7.6. RÉFÉRENCES BIBLIOGRAPHIQUES

Achard, F., Beuchle, R., Mayaux, P., Stibig, H.-J. et al. 2014. « Determination of tropical deforestation rates and related carbon losses from 1990 to 2010 ». *Global Change Biology*, vol. 20, n° 8, p. 2540–54.

Achard, F. et Hansen, M.C. (dir.). 2012. *Global Forest Monitoring from Earth Observation*. CRC Press, Taylor & Francis Group : Boca Raton, FL, États-Unis.

Achard, F., Eva, H.D., Stibig, H.-J., Mayaux, P., Gallego, J., Richards, T. et Malingreau, J.P. 2002. « Determination of Deforestation Rates of the World's Humid Tropical Forests ». *Science*, vol. 297, n° 5583, p. 999–1002.

Bartalev, S.S., Kissiyar, O., Achard, F., Bartalev, S.A. et Simonetti, D. 2014. « Assessment of forest cover of Russia by combining a wall to wall coarse resolution land cover map with a sample of 30m resolution forest maps ». *International Journal of Remote Sensing*, vol. 35, n° 7, p. 2671–92.

Bastin, J.F., Berrahmouni, N., Grainger, A. et al. 2017. « The extent of forest in dryland biomes ». *Science*, vol. 356, n° 6338, p. 635–8.

Beuchle, R., Grecchi, R.C., Shimabukuro, Y.E., Seliger, R., Eva, H.D., Sano, E. et Achard, F. 2015. « Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach ». *Applied Geography*, vol. 58, p. 116–27.

Beuchle, R., Eva, H.D., Stibig, H.-J. et al. 2011. « A Satellite Data set for Tropical Forest Change Assessment ». *International Journal of Remote Sensing*, vol. 32, n° 22, p. 7009–31.

Bodart, C., Eva, H.D., Beuchle, R. et al. 2011. « Pre-processing of a Sample of Multi-scene and Multi-date Landsat Imagery used to Monitor Forest Cover Changes over the Tropics ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, n° 5, p. 555–63.

Bodart, C., Brink, A., Donnay, F., Lupi, A., Mayaux, P. et Achard, F. 2013. « Continental estimates of forest cover and forest cover changes in the dry ecosystems of Africa for the period 1990-2000 ». *Journal of Biogeography*, vol. 40, n° 6, p. 1036–47.

Bey, A., Sánchez-Paus Díaz, A., Maniatis, D. et al. 2016. « Collect Earth: Land Use and Land Cover Assessment through Augmented Visual Interpretation ». *Remote Sensing*, vol. 8, n° 10, p. 807.

Brink, A.B. et Eva, H.D. 2009. « Monitoring 25 years of land cover change dynamics in Africa: a sample based remote sensing approach ». *Applied Geography*, vol. 29, n° 4, p. 501–12.

CCNUCC (Convention-cadre des Nations Unies sur les changements climatiques). 2016. *Rapport de la Conférence des Parties sur sa vingt et unième session, Décision 1/CP.21 : Adoption de l'Accord de Paris*. Publication de la CCNUCC : Bonn, Allemagne.

CCNUCC. 2014. *Key decisions relevant for reducing emissions from Deforestation and forest degradation in developing countries (REDD+)*. Decision Booklet REDD+. Publication de la CCNUCC : Bonn, Allemagne. Disponible à l'adresse : http://unfccc.int/files/land_use_and_climate_change/redd/application/pdf/compilation_redd_decision_booklet_v1.1.pdf. Consulté le 10 juin 2017.

CCNUCC. 2006. *Rapport de la première session de la Conférence des Parties agissant comme réunion des Parties au Protocole de Kyoto, Décision 16/CMP.1 : Utilisation des terres, changement d'affectation des terres et foresterie*. Publication de la CCNUCC : Bonn, Allemagne.

Clerici, N., Weissteiner, C.J., Paracchini, M.L. et Strobl, P. 2011. *Riparian zones: where green and blue networks meet—Pan-European zonation modelling based on remote sensing and GIS*. JRC Scientific and Technical Report. Office des publications de l'Union européenne : Luxembourg.

Commission européenne. 2013. *The impact of EU consumption on deforestation: Comprehensive analysis of the impact of EU consumption on deforestation*. Rapport technique 2013-063. Office des publications de l'Union européenne : Luxembourg.

De Foresta, H., Somarriba, E., Temu, A., Boulanger, D., Feuilly, H. et Gauthier, M. 2013. *Towards the assessment of trees outside of forests. A thematic report prepared in the framework of the global forest resources assessment*. Document de travail de l'Évaluation des ressources forestières n° 183. Publication de la FAO : Rome.

DeFries, R.S., Rudel, T.K., Uriarte, M. et Hansen, M.C. 2010 « Deforestation driven by urban population growth and agricultural trade in the twenty-first century ». *Nature Geoscience*, vol. 3, p. 178–81.

De Sy, V., Herold, M., Achard, F., Asner, G.P., Held, A., Kelldorfer, J. et Verbesselt, J. 2012. « Synergies of multiple remote sensing data sources for REDD+ monitoring ». *Current Opinion in Environmental Sustainability*, vol. 4, n° 6, p. 696–706.

De Sy, V., Herold, M., Achard, F., Beuchle, R., Clevers, J.G.P.W., Lindquist, E. et Verchot, L. 2015. « Land use patterns and related carbon losses following deforestation in South America ». *Environmental Research Letters*, vol. 10, n° 12.

Drigo, R., Lasserre, B. et Marchetti, M. 2009. « Patterns and trends in tropical forest cover ». *Plant Biosystems*, vol. 143, n° 2, p. 311–27.

Eva, H.D., Achard, F., Beuchle, R., de Miranda, E.E., Carboni, S., Seliger, R., Vollmar, M., Holler, W., Oshiro, O.T. et Barrena Arroyo, V. 2012. « Forest cover changes in tropical South and Central America from 1990 to 2005 and related carbon Emissions Removals ». *Remote Sensing*, vol. 4, n° 5, p. 1369–91.

Eva, H.D., Carboni, S., Achard, F., Stach, N., Durieux, L., Faure, J.-F. et Mollicone, D. 2010. « Monitoring forest areas from continental to territorial levels using a sample of medium spatial resolution satellite imagery ». *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, n° 2, p. 191–7.

FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture) et JRC (Centre commun de recherche de la Commission européenne). 2012. *Changement d'utilisation des terres forestières mondiales 1990–2005*. Étude FAO : Forêts n° 169. Publication de la FAO : Rome. Disponible à l'adresse : <http://www.fao.org/forestry/fra/remotesensingsurvey/fr/>. Consulté le 10 juin 2017.

FAO. 2016. *Arbres, forêts et utilisation des terres dans les zones arides – Première évaluation mondiale*. Publication de la FAO : Rome.

FAO. 2015a. *Évaluation des ressources forestières mondiales 2015 : comment les forêts de la planète changent-elles ?* Rapport de synthèse FRA-2015. Publication de la FAO : Rome.

FAO. 2015b. *Country Report Brazil – Global Forest Resource Assessment 2015*. Disponible à l'adresse : <http://www.fao.org/documents/card/en/c/6261857f-c0da-4f72-98fd-a18e9ca50509/>. Publication de la FAO : Rome. Consulté le 10 juin 2017.

FAO. 2010a. *Évaluation des ressources forestières mondiales 2010. Rapport principal*. Publication de la FAO : Rome.

FAO. 2010b. *Termes et définitions utilisés dans FRA 2010*. Document de travail 144/E de la FAO. Publication de la FAO : Rome.

FAO. 2001. *Évaluation des ressources forestières mondiales 2000*. Étude FAO : Forêts n° 140. Publication de la FAO : Rome.

FAOSTAT. 2011. *Séries chronologiques et données croisées relatives à l'alimentation et l'agriculture*. Publication de la FAO : Rome. Disponible à l'adresse : <http://www.fao.org/faostat/fr/#home>. Consulté en 2011 pour le rapport de la Commission européenne (2013).

Foody, G.M. 2010. « Assessing the accuracy of land cover change with imperfect ground reference data ». *Remote Sensing of Environment*, vol. 114, n° 10, p. 2271–85.

Freitas, J.V., Oliveira, Y.M.M., Brena, D.A., Gomide, G.L.A., Silva, J.A., Collares, J.E., Mattos, P.P., Rosot, M.A.D., Sanquetta, C.R., Vencatto, M.F., Barros, P.L.C., Santos, J.R., Ponzoni, F.L. et Shimabukuro, Y.E. 2008. « The New Brazilian National Forest Inventory ». In : McRoberts, R.E., Reams, G.A., Van Deusen, P.C. et McWilliams, W.H. (dir.), *Proceedings of the Eight Annual Forest Inventory and Analysis Symposium, Monterey, CA, October 16-19, 2006* (p. 9–12) Publication de l'USDA Forest Service : Washington, D.C., États-Unis.

Freitas, J.V., Oliveira, Y.M.M., Rosot, M.A.D., Gomide, G.L.A. et Mattos, P.P. 2010. « Brazil National Forest Inventory Report ». In : Tomppo, E., Gschwantner, T., Lawrence, M. et McRoberts, R.E. (dir.), *National forest inventories: pathways for common reporting* (p. 89–95). Springer-Verlag : Heidelberg, Allemagne.

Gibbs, H.K., Ruesch, A.S., Achard, F., Clayton, M., Holmgren, P., Ramankutty, N. et Foley, J.A. 2010. « Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s ». *Proceedings of the National Academy of Sciences USA*, vol. 107, n° 38, p. 16732–7.

GOFC-GOLD. 2016. *A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals associated with deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation*. Version COP22-1 du rapport GOFC-GOLD. Publication du GOFC-GOLD Land Cover Project Office – Université de Wageningen : Wageningen, Pays-Bas. Disponible à l'adresse : <http://www.gofcgold.wur.nl/redd/index.php>. Consulté le 10 juin 2017.

Grassi, G., Federici, S. et Achard, F. 2013. « Implementing conservativeness in REDD+ is realistic and useful to address the most uncertain estimates ». *Climatic Change*, vol. 119, n° 2, p. 269–75.

Grecchi, R., Beuchle, R., Shimabukuro, Y.E., Aragão, L.E, Arai, E., Simonetti, D. et Achard, F. 2017. An integrated remote sensing and GIS approach for monitoring areas affected by selective logging: a case study in northern Mato Grosso, Brazilian Amazon. *International Journal of Applied Earth Observation and Geoinformation*, vol. 61, p. 70–80.

Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC). 2006. Représentation cohérente des terres. In : *Lignes directrices 2006 du GIEC pour les inventaires nationaux de gaz à effet de serre. Volume 4 : Agriculture, foresterie et autres affectations des terres* (chapitre 3). Publication du GIEC : Kanagawa, Japon.

Gutman, G., Huang, C., Chander, G., Noojipady, P. et Masek J.G. 2013. « Assessment of the NASA/USGS global land survey (GLS) datasets ». *Remote Sensing of Environment*, vol. 134, p. 249–65.

Hansen, M.C., Potapov, P.V., Moore, R. et al. 2013. « High-resolution global maps of 21st-century forest cover change ». *Science*, vol. 342, n° 6160, p. 850–3.

Hansen, M.C., Stehman, S.V. et Potapov, P.V. 2010. « Quantification of global gross forest cover loss ». *Proceedings of the National Academy of Sciences USA*, vol. 107, n° 19, p. 8650–5.

Initiative mondiale pour l'observation des forêts (GFOI). 2014. *Integrating remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and Guidance from the Global Forest Observations Initiative*. Publication du Group on Earth Observations : Genève, Suisse. Disponible à l'adresse : <http://www.gfoi.org/methods-guidance/>. Consulté le 10 juin 2017.

Instituto Nacional de Pesquisas Espaciais (INPE). 2016. *Monitoramento da Floresta Amazônica Brasileira por Satélite*. Publication de l'INPE : São José dos Campos, Brésil. Disponible à l'adresse : <http://www.obt.inpe.br/prodes/index.html>. Consulté le 10 juin 2017.

Ivits, E., Cherlet, M., Mehl, W. et Sommer, S. 2009. « Estimating the ecological status and change of riparian zones in Andalusia assessed by multi-temporal AVHRR datasets ». *Ecological Indicators*, vol. 9, n° 3, p. 422–31.

Keenan, R.J., Reams, G.A., Achard, F., de Freitas, J., Grainger, A. et Lindquist, E. 2015. « Dynamics of global forest area: results from the FAO Global Forest Resources Assessment 2015 ». *Forest Ecology and Management*, vol. 352, p. 9–20.

Luz, N.B., Oliveira, Y.M.M., Rosot, M.A.D., Garrastazú, M.C., Franciscon, L., Mesquita, H.N. Jr et Freitas, J.V. 2015a. *Classificação híbrida de imagens Landsat-8 e RapidEye para o mapeamento do uso e cobertura da terra nas Unidades Amostrais de Paisagem do Inventário Florestal Nacional do Brasil*. Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR, João Pessoa-PB, Brazil, 25–29 avril 2015. Publication de l'INPE : São José dos Campos, Brésil. Disponible à l'adresse : <http://www.dsr.inpe.br/sbsr2015/files/p1606.pdf>. Consulté le 10 juin 2017.

Luz, N.B., Oliveira, Y.M.M., Rosot, M.A.D., Garrastazú, M.C., Franciscon, L., Mesquita, H.N. Jr, Freitas, J.V. et Costa, C.R. 2015b. « Developments in forest monitoring under the Brazilian National Forest Inventory: multi-source and hybrid image classification approaches ». Document élaboré pour le XIV^e Congrès forestier mondial, 7–11 septembre 2015. Durban, Afrique du Sud. Disponible à l'adresse : <http://foris.fao.org/wfc2015/api/file/5547d6d115ae74130ace6a2f/contents/e09e6ff0-e135-48d6-82e7-28466b86f63d.pdf>. Consulté le 10 juin 2017.

MacDicken, K.G., Sola, P., Hall, J.E., Sabogal, C., Tadoum, M. et de Wasseige, C. 2015. « Global progress toward sustainable forest management ». *Forest Ecology and Management*, vol. 352, p. 47–56.

MacDicken, K.G. 2015. « Global Forest Resources Assessment 2015: What, why and how? » *Forest Ecology and Management*, vol. 352, p. 3–8.

Magdon, P., Fischer, C., Fuchs, H. et Kleinn, C. 2014. Translating criteria of international forest definitions into remote sensing image analysis. *Remote Sensing of Environment*, vol. 149, p. 252–62.

Martinez, S. et Mollicone, D. 2012. From land cover to land use: a methodology to assess land use from remote sensing data. *Remote Sensing*, vol. 4, p. 1024–45.

Mayaux, P., Pekel, J.-F., Desclée, B. et al. 2013. « State and evolution of the African rainforests between 1990 and 2010 ». *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 368, n° 1625, doi : 10.1098/rstb.2012.0300.

Mayaux, P., Holmgren, P., Achard, F., Eva, H.D., Stibig, H.-J. et Branthomme, A. 2005. « Tropical forest cover change in the 1990's and options for future monitoring ». *Philosophical Transactions of The Royal Society B*, vol. 360, n° 1454, p. 373–84.

Miettinen, J., Shimabukuro, Y.E., Beuchle, R., Grecchi, R.C., Velasco-Gomez, M., Simonetti, D. et Achard, F. 2016. « On the extent of fire-induced forest degradation in Mato Grosso, Brazilian Amazon, in 2000, 2005 and 2010 ». *International Journal of Wildland Fire*, vol. 25, n° 2, p. 129–36.

Miettinen, J., Stibig, H.-J. et Achard, F. 2014. « Remote sensing of forest degradation in Southeast Asia - aiming for a regional view through 5-30 m satellite data ». *Global Ecology and Conservation*, vol. 2, p. 24–36.

Pelletier, J. et Goetz, S.J. 2015. « Baseline data on forest loss and associated uncertainty: advances in national forest monitoring ». *Environmental Research Letters*, vol. 10, n° 2.

Penman, J., Gytarsky, M., Krug, T. et al. (dir.). 2003. *Definitions and Methodological Options to Inventory Emissions from Direct Human induced Degradation of Forests and Devegetation of Other Vegetation Types*. Publication du GIEC : Kanagawa, Japon.

Potapov, P., Hansen, M.C., Gerrand, A.M. et al. 2010. « The global Landsat imagery database for the FAO FRA remote sensing survey ». *International Journal of Digital Earth*, vol. 4, n° 1, p. 2–21.

Raši, R., Bodart, C., Stibig, H.-J. et al. 2011. « An automated approach for segmenting and classifying a large sample of multi-date Landsat imagery for pan-tropical forest monitoring ». *Remote Sensing of Environment*, vol. 115, n° 2, p. 3659–69.

Richards, T.S., Gallego, J. et Achard, F. 2000. « Sampling for forest cover change assessment at the pan-tropical scale ». *International Journal of Remote Sensing*, vol. 21, n° 6-7, p. 1473–90.

Romijn, E., Herold, M., Kooistra, L., Murdiyarso, D. et Verchot, L. 2012. « Assessing capacities of non-annex I countries for national forest monitoring in the context of REDD+ ». *Environmental Science Policy*, vol. 19-20, p. 33–48.

Romijn, E., Lantican, C.B., Herold, M., Lindquist, E., Ochieng, R., Wijaya, A., Murdiyarso, D. et Verchot, L. 2015. « Assessing change in national forest monitoring capacities of 99 tropical countries ». *Forest Ecology and Management*, vol. 352, p. 109–23.

Särndal, C.E., Swensson, B. et Wretman, J. 1992. *Model Assisted Survey Sampling*. Springer-Verlag : New York, États-Unis.

Saura, S. et Torné, J. 2009. « Conefor Sensinode 2.2: a software package for quantifying the importance of habitat patches for landscape connectivity ». *Environmental Modelling & Software*, vol. 24, n° 1, p. 135–9.

Serviço florestal brasileiro (SFB). 2016. *Inventaire forestier national du Brésil*. Disponible à l'adresse : www.florestal.gov.br. Consulté le 21 juin 2017.

Stach, N., Salvado, A., Petit, M., Faure, J.F., Durieux, L., Corbane, C., Joubert, P., Lasselin, D. et Deshayes, M. 2009. « Land use monitoring by remote sensing in tropical forest areas in support of the Kyoto Protocol: the case of French Guiana ». *International Journal of Remote Sensing*, vol. 30, n° 19, p. 5133–49.

Stehman, S.V., Hansen, M.C., Broich et M., Potapov, P.V. 2011. « Adapting a global stratified random sample for regional estimation of forest cover change derived from satellite imagery ». *Remote Sensing of Environment*, vol. 115, n° 2, p. 650–8.

Steininger, M.K., Godoy, F. et Harper, G. 2009. « Effects of systematic sampling on satellite estimates of deforestation rates ». *Environmental Research Letters*, vol. 4, n° 3.

Stibig, H.-J., Achard, F., Carboni, S., Raši, R. et Miettinen, J. 2014. « Changes in tropical forest cover of Southeast Asia from 1990 to 2010 ». *Biogeosciences*, vol. 11, p. 247–58.

Tyukavina, A., Baccini, A., Hansen, M.C., Potapov, P.V., Stehman, S.V., Houghton, R.A., Krylov, A.M., Turubanova, S. et Goetz, S.J. 2015. « Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012 ». *Environmental Research Letters*, vol. 10, 074002.

Tyukavina, A., Stehman, S.V., Potapov, P.V., Turubanova, S.A., Baccini, A., Goetz, S.J., Laporte, N.T., Houghton, R.A. et Hansen, M.C. 2013. « National-scale estimation of gross forest aboveground carbon loss: a case study of the Democratic Republic of the Congo ». *Environmental Research Letters*, vol. 8, 044039.

Verhegghen, A., Eva, H.D., Desclée, B. et Achard, F. 2016. « Review of recent forest cover and forest cover change assessments in Cameroon ». *International Forestry Review*, vol. 18, suppl. 1, p. 14–25.

Chapitre 8

Organisations, ressources et compétences nécessaires à l'adoption de la télédétection dans le domaine des statistiques agricoles

John Latham

8.1. CONTEXTE

Il est important de disposer de données précises et à jour sur les cultures (informations temporelles durant la période de végétation et estimations au moment de la récolte) afin d'obtenir des estimations fiables à des fins de sécurité alimentaire, de planification, de formulation de politiques et de stratégies de commercialisation agricole. On sait d'expérience que la qualité des données agricoles disponibles (recensements agricoles, suivi des cultures et estimations du rendement) est généralement inconsistante et inégale. Il est par conséquent nécessaire d'améliorer les techniques de collecte et de diffusion des données. Grâce aux récents progrès en matière de traitement des données géospatiales et de télédétection, alliés à la réduction des coûts et à l'augmentation de la couverture, de la fréquence et de la résolution des images satellite, on dispose de nouvelles données et méthodes fiables pour renforcer les systèmes de données (Bell et Dalton, 2007 ; Murray, 2010). Les systèmes de données de télédétection ont connu une transformation radicale : ils ne font plus office de jeux de données purement accessoires destinés à la validation, mais proposent désormais des premières estimations fiables de la production agricole.

Les méthodes traditionnelles de prédiction des rendements agricoles sont axées sur des modèles intégrant le climat, les sols et les engrais, l'eau d'irrigation et l'agrométéorologie en tant que fonctions de réponse pour décrire le rendement agricole et évaluer la production (Wiegand et Richardson, 1990). Ces techniques reposent généralement sur de solides concepts physiologiques et physiques. Cependant, leur capacité à fournir des résultats fiables s'avère limitée lorsqu'elles sont confrontées à une importante variabilité spatiale des sols, à des facteurs de stress ou à de

mauvaises pratiques de gestion (Wiegand et Richardson, 1984 et 1990). Les technologies géospatiales (télédétection et SIG) sont néanmoins présentées comme des outils potentiellement utiles pour le suivi agricole en raison de leur couverture synoptique et de leur capacité à assurer un suivi dans le temps (Hinzman *et al.*, 1986 ; Quarmby *et al.*, 1993). Une grande importance a été accordée à la mise au point de procédures rapides et fiables permettant de fournir des prévisions rapide des récoltes dans la saison, ainsi que des estimations à la fin de la saison.

Les recensements agricoles recourent de plus en plus aux technologies géospatiales et à la télédétection par satellite, ou encore à l'utilisation de cartes de couverture du sol créées à d'autres fins de façon à privilégier les méthodes de sondage fondées sur une base de sondage aréolaire (BSA) plutôt que sur une base de type liste. Les jeux de données issus de la télédétection et autres SIG contribuent également à l'élaboration et au maintien des bases de sondage principales (Carfagna, 2013) intégrant plusieurs couches de données géospatiales (districts de recensement, limites des parcelles et zones cultivées, par exemple), et qui permettent une réutilisation efficace de la base de sondage (Carfagna et Gallego, 2005). Bien que les techniques de télédétection n'aient pas éliminé la nécessité des recensements agricoles, l'intégration de la télédétection et des SIG dans les procédures d'enquête et d'échantillonnage vise à accroître la précision, l'efficacité pratique, la reproductibilité et la rentabilité en réduisant le temps nécessaire à la création des bases de sondage. À l'heure actuelle, l'utilisation conjointe de GPS, de SIG et de techniques d'imagerie permet de renforcer l'efficacité des enquêtes et des recensements agricoles sur le terrain (Carfagna *et al.*, 2013).

Toutefois, l'utilisation de la télédétection dans le cadre du suivi, des estimations et de la diffusion des données agricoles nécessite des données et des ressources techniques suffisantes pour procéder à l'analyse des mégadonnées, un domaine qui devra être bien maîtrisé par la nouvelle génération de spécialistes de la modélisation géospatiale. Dans ces circonstances, il est recommandé d'utiliser la télédétection par satellite pour assurer la surveillance des cultures et des systèmes agricoles, et de recourir au renforcement des compétences et à la formation concernant le traitement des images et leur utilisation lors des opérations sur le terrain. Malgré des difficultés inhérentes, ces méthodes ont été largement développées dans d'autres secteurs ; lorsque l'on souhaite utiliser la télédétection dans le domaine de l'agriculture, il est donc possible d'exploiter des données acquises et traitées à d'autres fins et de bénéficier de possibilités de partage institutionnel intéressantes.

Les sections suivantes détaillent les exigences techniques et pratiques relatives à l'utilisation de la télédétection et des SIG dans le cadre des statistiques agricoles et fournissent des exemples d'intégration de la télédétection dans les statistiques et la diffusion de données agricoles.

8.2. ORGANISATIONS

Les organisations du secteur public sont généralement les principales productrices de statistiques agricoles. Toutefois, depuis quelques années, ce travail est de plus en plus souvent réalisé par des organisations du secteur privé. Les entreprises commerciales, et en particulier certaines industries (raffineries de sucre, fabricants d'aliments pour animaux et producteurs de boissons, notamment), externalisent généralement la collecte des informations temporelles dont elles ont besoin pour améliorer leur efficacité de fonctionnement et gérer leurs activités industrielles. Le système conventionnel employait essentiellement des méthodes manuelles et les données n'étaient disponibles qu'à la fin de la saison ; en outre, il n'était pas possible d'obtenir une perspective temporelle et globale à un moment précis.

Les organisations qui utilisent des données satellite doivent impérativement être multidisciplinaires afin de pouvoir combiner simultanément les informations issues de la télédétection par satellite, les SIG, les statistiques, l'agronomie, l'agrométéorologie, l'économie et le développement de logiciels. Des spécialistes d'autres domaines peuvent être intégrés aux programmes afin de répondre aux besoins spécifiques d'une industrie donnée ou d'étudier

les questions relatives aux modes de consommation, à la sécurité alimentaire et aux régimes d'importation et d'exportation. Les organisations doivent également recruter du personnel technique afin d'assurer la saisie des données, les travaux de maintenance essentiels au développement des données et à la conduite des enquêtes sur le terrain, même si la nature et les modalités de ce type de tâches ont également été profondément modifiées par la production participative. Les défis actuels consistent à préserver la représentativité des données obtenues grâce à la production participative et à veiller à ce qu'elles ne soient pas biaisées.

8.3. RESSOURCES

Les ressources dont les organisations doivent disposer pour se mettre à produire des statistiques agricoles peuvent globalement être réparties dans les catégories suivantes :

- personnel qualifié ;
- laboratoires (matériel informatique et logiciels ; ces besoins sont toutefois de plus en plus limités en raison du recours à l'informatique dématérialisé pour le stockage le traitement et les analyses) ;
- détermination des données d'entrée (les données multicateurs sont souvent privilégiées) ;
- planification du travail ;
- formation ; et
- financements.

Les statistiques agricoles peuvent tirer parti de l'utilisation de données de télédétection et du traitement géospatial de différentes manières, selon un certain nombre de critères et de besoins en ressources qui varient en fonction des objectifs et de la finalité des enquêtes. Ces critères sont les suivants :

- i. disponibilité de données de référence permettant d'identifier les caractéristiques les plus appropriées de la BSA, compte tenu des particularités du territoire ;
- ii. élaboration et maintenance de la BSA ;
- iii. stratification de la BSA (en utilisant par exemple les données de couverture du sol sur la zone cultivée ou sur des cultures spécifiques ou saisonnières) ;
- iv. cartographie géospatiale de la répartition des cultures spécialisées ou mineures et des grandes et moyennes exploitations commerciales afin d'établir des bases de sondage sélectives ;
- v. appui aux procédures et à la logistique des enquêtes agricoles sur le terrain ;
- vi. mise à disposition de variables supplémentaires (estimateur par régression ou par calage et estimation sur petits domaines) ;
- vii. suivi et prévision du rendement agricole ; et
- viii. cartographie des facteurs circonstanciels (ravageurs et maladies, par exemple) et des catastrophes naturelles et indices dérivés de la productivité des cultures saisonnières (NDVI, notamment).

Outre son application directe aux statistiques agricoles, la télédétection peut également s'avérer utile dans des domaines connexes. Ainsi, l'imagerie satellite est couramment employée dans l'évaluation des pâturages et terres de parcours destinés au bétail, la gestion des ressources naturelles, le suivi de l'agriculture durable, du développement rural et de la dégradation des terres, les évaluations de l'impact des activités humaines et la cartographie des dangers et vulnérabilités. Ces différents usages pourront avoir une influence sur la sélection des ressources et caractéristiques d'images (saisonnalité, date, résolution, intervalle de revisite ou encore bandes spectrales), qui devront donc être adaptées à une telle exploitation multi-usage. Ils nécessiteront également différents niveaux de données géospatiales auxiliaires, de soutien financier, d'équipement et de formation.

8.3.1. Personnel qualifié

Les statistiques relatives à la production végétale reposent essentiellement sur deux éléments, à savoir l'estimation de la surface cultivée et la prévision du rendement. La surface peut être estimée à l'aide d'une base de sondage aréolaire ou d'une classification d'image portant sur la totalité d'une région. Les analystes spécialisés en télédétection et en SIG sont traditionnellement chargés du traitement des images et de l'élaboration des bases aréolaires, tandis que les statisticiens sont responsables de la conception des échantillons, de l'extrapolation et des estimations, ainsi que du calcul du coefficient de variation de ces dernières. Les analystes d'images calculent souvent la superficie à l'aide d'une classification par pixels ou par objets. Le personnel de terrain est chargé de mener les enquêtes sur le terrain et de collecter les signatures des cultures. On observe parfois un chevauchement des compétences requises, notamment lorsque certains processus d'échantillonnage et de listage nécessitent à la fois des ressources géospatiales et des procédures d'échantillonnage statistique (échantillonnage « en serpent »), par exemple).

L'amélioration des statistiques agricoles repose sur l'association et la collaboration d'un large éventail de disciplines (tableau 1).

TABLEAU 1. SYNTHÈSE DES DISCIPLINES REQUISES POUR UTILISER LA TÉLÉDÉTECTION DANS LE DOMAINE DES STATISTIQUES ET DE LA DIFFUSION DES DONNÉES AGRICOLES.

Domaines thématiques et expertise	Description des besoins
Analyste spécialisé en télédétection et en SIG	Produits obtenus à partir d'images satellite (notamment NDVI et indice de teneur en eau par différence normalisée [NDWI, de l'anglais <i>Normalised Difference Water Index</i>]). Jeux de données relatifs à la couverture du sol et à l'utilisation des terres, traitement d'image (amélioration d'image, classification d'image par pixels et par objets, interprétation et extraction des caractéristiques, production de cartes thématiques, signatures spectrales des cultures, calendriers phénologiques et cartes des cultures obtenus par télédétection), utilisation des ressources géospatiales pour l'échantillonnage et l'élaboration de bases aréolaires
Statisticien	Conception de bases de sondage, élaboration de bases aréolaires, sélection de segments aléatoires, conception de questionnaires destinés à la collecte de données de terrain, correction des données de terrain, estimation des surfaces cultivées, analyse statistique destinée à la prévision du rendement agricole
Agronome	Calendrier des cultures, calendrier d'acquisition des images satellite, évaluation de l'état des cultures à l'aide des données satellite et des informations recueillies sur le terrain
Agrométéorologue	Collecte et mise à jour des paramètres météorologiques
Développeur de logiciels	Mise au point d'applications bureautiques ou Web destinées à la collecte de données de terrain et à l'automatisation du traitement des données
Personnel de terrain	Collecte des données de terrain à l'aide de systèmes GPS, validation, listage et échantillonnage agricole ou coupe échantillon
Personnel d'appui	Saisie des données, numérisation, impression

8.3.2. Laboratoires : configuration matérielle et logicielle requise

Les laboratoires doivent disposer du personnel et des équipements nécessaires. L'espace affecté au laboratoire pourra dépendre de l'ampleur des activités réalisées par l'organisation concernée. D'importants systèmes d'analyse peuvent désormais présenter des dimensions réduites grâce à la disponibilité de l'informatique dématérialisée et les technologies analytiques.

Le matériel informatique nécessaire à un laboratoire de statistiques agricoles géospatiales est présenté dans le tableau 2.

TABLEAU 2. SYNTHÈSE DES BESOINS D'UN LABORATOIRE DE STATISTIQUES AGRICOLES GÉOSPATIALES (MATÉRIEL INFORMATIQUE).

Élément	Description/finalité
Poste de travail	Processeur : core i7, Xeon RAM : > 32 Go de RAM Carte graphique : 2 à 6 Go Capacité du disque dur : 2 à 4 To Pour le traitement des images
Ordinateur portable/tablette	Collecte de données sur le terrain
GPS	Enregistrement des coordonnées des champs
Smartphone/tablette	Collecte de données sur le terrain et transmission en temps quasi réel
Appareil photo	Appareils photo numériques pour la validation et la vérification des classifications
Imprimante/traceur	Imprimante couleur A4 et A3 et traceur A0 pour l'impression des rapports et des cartes
Scanner	Numérisation des cartes historiques en format A3 ou supérieur
Serveur de stockage/sauvegarde	Capacité de stockage de données : 16 To
Réseau	Pour la connexion des postes de travail

Les appareils mobiles modernes peuvent intégrer la géolocalisation GPS des photographies, la communication et le stockage des données pour les applications hors ligne, réduisant ainsi le nombre de systèmes utilisés sur le terrain.

Les logiciels nécessaires à un laboratoire de statistiques agricoles géospaciales sont présentés dans le tableau 3.

TABLEAU 3. SYNTHÈSE DES BESOINS D'UN LABORATOIRE DE STATISTIQUES AGRICOLES GÉOSPATIALES (LOGICIELS).

Discipline	Logiciel
Statistiques	SPSS : analyse statistique
	Excel Stat : calcul de superficie et de rendement
	R (logiciel libre) : calcul statistique
SIG/télédétection	ArcGIS 10.x : développement de SIG, requêtes et analyse spatiales
	QGIS (logiciel libre) : développement de SIG, requêtes et analyse spatiales
	ERDAS Imagine 9.x : traitement d'image, classification
	ENVI : traitement d'image, classification
	E-Cognition : traitement d'image, classification par objets
	FME : conversion, transformation et automatisation des tâches dans le domaine géospatial
Applications mobiles d'enquêtes et entretien individuel assisté par ordinateur (CAPI, de l'anglais <i>Computer-Assisted Personal Interviewing</i>)	ODK (Open Data Kit) etc./applications Android personnalisées d'enquête géospatiale sur le terrain (p. ex. MAGIS)
	Logiciels CAPI, p. ex. Survey System (Banque mondiale), Area Frame Survey System (AFSS), Collect Mobile (FAO)
Diffusion des données/métadonnées	GeoNetwork : métadonnées géospaciales, téléchargement et affichage des données Applications de cartographie Web : diffusion des informations (MapStore, GeoServer, Leaflet, Arc Server, etc.)

8.3.3. Données d'entrée

Les données d'entrée issues de l'observation de la Terre et les données géospatiales auxiliaires nécessaires à la production des statistiques agricoles constituent un élément essentiel du processus de mise en œuvre. L'adéquation des différentes sources d'images est déterminée par un certain nombre de facteurs (résolution spatiale et temporelle, fréquence de couverture, résolution spectrale et type de capteur), avec les implications que cela suppose sur l'obtention de ces images et les coûts d'acquisition et de traitement. La résolution des images est un paramètre décisif dans les critères de sélection ; elle est toutefois mise en balance avec divers facteurs connexes tels que la superficie du terrain, la fréquence de revisite, ainsi que des facteurs pratiques liés notamment aux coûts et au traitement. La disponibilité croissante de données haute résolution (résolution supérieure à 1 m et inférieure à 5 m) et très haute résolution (résolution inférieure à 1 m), mais également de largeurs de fauchée plus importantes (le satellite Sentinel-2, par exemple, fournit gratuitement des images d'une résolution comprise entre 9 et 10 m pour une fauchée de 290 km) modifie la disponibilité et la pertinence des différentes résolutions d'images satellite. Delincé (2017) a étudié l'adéquation des images en fonction de la taille des champs et des parcelles et mis en évidence la prédominance de petits et très petits champs dans l'agriculture, en particulier en Afrique et en Asie, où plus de 80 % des champs relèvent de ces catégories selon la carte mondiale des terres cultivées d'une résolution de 1 km² réalisée par l'Institut international pour l'analyse des systèmes appliqués (IIASA) (Fritz, 2015). Du fait de cette prépondérance de petites exploitations, il est préférable de privilégier des données de meilleure résolution dans ces environnements agricoles afin d'accroître le nombre de pixels compris dans chaque parcelle (autrement dit, de réduire le nombre de pixels mixtes). Cependant, cette situation évolue, et l'apparition de nouveaux capteurs disposant d'une fauchée plus large et d'une meilleure résolution laisse entrevoir des solutions supplémentaires. Les nouvelles constellations de satellites et l'utilisation de capteurs mixtes (Deimos-1 et 2/UrtheDaily, OptiSar, PanGeo Alliance et Planet Labs, par exemple) offrent des capacités, des fréquences et des résolutions de données inédites qui ouvrent de nouvelles possibilités d'application.

Les données d'entrée englobent également les données de validation de la télédétection et celles des enquêtes de terrain. L'imagerie satellite est parfois employée à des fins de pseudo-vérification sur le terrain : des images et photographies aériennes de meilleure résolution remplacent alors les enquêtes sur site. Bien que les images THR simplifient l'identification des objets au sol, il est important de comprendre pour quels besoins et applications spécifiques ces images sont proposées. En règle générale, plus la résolution spatiale des images est élevée, plus le coût d'acquisition augmente, de même que les capacités nécessaires au stockage, au traitement et à l'analyse de ces images. Toutefois, la possibilité d'analyser des mégadonnées grâce à l'informatique dématérialisée minimise de plus en plus ces considérations.

8.3.4. Planification du travail

L'intégration des données de télédétection et des données de couverture du sol qui en découlent dans les procédures d'estimation des surfaces cultivées et du rendement agricole se déroule selon les étapes suivantes :

- i. sélection et acquisition des images satellite ou de données déjà classées ;
- ii. prétraitement et traitement des images ;
- iii. prise en compte des données auxiliaires et des données de validation sur le terrain afin de générer des classifications pouvant être intégrées dans la base de sondage aréolaire ; et
- iv. élaboration de la BSA, stratification et calcul d'estimation des surfaces cultivées.

La période d'acquisition des images satellite dépend des stades phénologiques des cultures ; la sélection des images peut toutefois dépendre de l'objectif spécifique de l'enquête (tableau 4).

TABEAU 4. EXEMPLE DE CALENDRIER DES CULTURES EN ASIE DESTINÉ À IDENTIFIER LES PÉRIODES D'ACQUISITION DES IMAGES ET À ATTEINDRE LES OBJECTIFS D'ÉCHANTILLONNAGE.

Tâche	Mois											
	Jan.	Fév.	Mars	Avr.	Mai	Juin	Jui.	Août	Sept.	Oct.	Nov.	Déc.
Acquisition												
Prétraitement												
Enquête de terrain												
Classification des images												
Validation												
Estimation des cultures												

 Cultures d'hiver (*rabi*)
  Cultures d'été (*kharif*)

Il est toujours utile d'organiser une évaluation par des tiers afin de repérer les éventuels défauts et lacunes des systèmes proposés ou approuvés. Il convient d'engager des experts réputés appartenant à une organisation ou source indépendante afin de valider les procédures et les conclusions. Un audit technique des données peut également contribuer à améliorer la qualité de ces dernières.

8.3.5. Formation requise

Compte tenu des exigences techniques liées à l'utilisation de la télédétection et des SIG, il est indispensable d'investir dans la formation et le renforcement des capacités afin de faciliter la mise en œuvre. Il convient donc de proposer des formations, des supports d'apprentissage en ligne et un soutien appropriés concernant l'application de l'estimation des surfaces cultivées et du rendement agricole. Des formations régulières dispensées par des organismes nationaux et internationaux spécialisés sont nécessaires afin de mieux accompagner l'utilisation des nouvelles technologies, même si l'essentiel du traitement spécifique des images de télédétection et des analyses géospatiales peut être effectué au niveau national par des capteurs plus spécialisés (SAR, notamment) pouvant exiger un renforcement des capacités supplémentaire.

Les programmes de formation envisagés pour l'intégration et l'application de la télédétection et des SIG sont les suivants :

- i. notions de base dans les domaines suivants :
 - a. télédétection,
 - b. SIG,
 - c. statistiques,
 - d. agronomie ;
- ii. télédétection et prétraitement, traitement, classification, analyse et communication des images ;
- iii. création et mise en forme de données SIG, conception de bases de données géoréférencées et outils d'analyse géospatiale ;
- iv. méthodes de classification de la couverture du sol, segmentation d'image, photo-interprétation, classification automatisée et manuelle, vérification sur le terrain et évaluation de la précision (à l'aide notamment de la pseudo-vérification) ;
- v. prise en compte des paramètres agro-environnementaux déduits des données de télédétection (calendriers des cultures, phénologie et réaction des plantes) et des informations recueillies sur le terrain dans la prévision des rendements ;

- vi. application de la télédétection et des SIG aux BSA, à la stratification, à la sélection des échantillons, à la conception de questionnaires (à l'aide des informations issues des données de télédétection) et à l'échantillonnage probabiliste ;
- vii. application des logiciels de SIG à l'élaboration de plans d'échantillonnage aréolaires, à la stratification, à l'échantillonnage probabiliste, à la conception de questionnaires, à l'échantillonnage de contrôle et aux erreurs non liées à l'échantillonnage ;
- viii. utilisation de GPS et collecte de données sur le terrain (pour les agents recenseurs) ;
- ix. techniques d'échantillonnage statistique appliquées à l'estimation ponctuelle et par intervalle des paramètres de chaque strate et à l'évaluation de la précision de ces estimations.

Le niveau, le périmètre et la périodicité des formations seront déterminés en fonction des besoins et/ou de la composition de l'équipe.

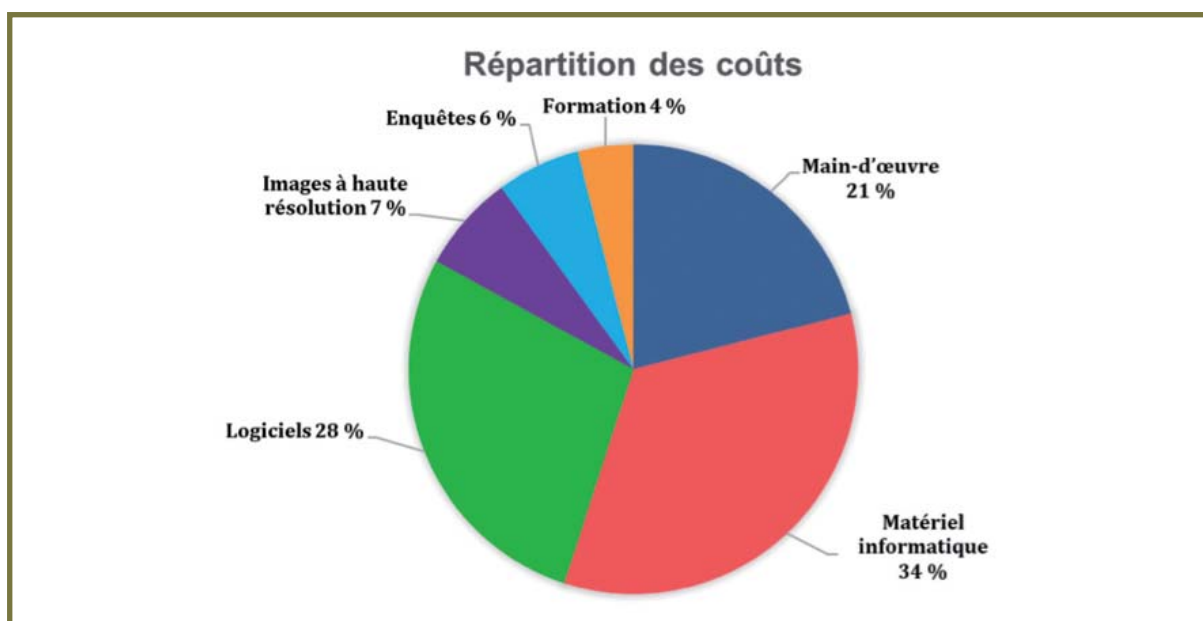
8.3.6. Financement

L'intégration de la télédétection dans le domaine des statistiques agricoles nécessite l'affectation de ressources financières suffisantes. Il est possible de couvrir les coûts relatifs à l'établissement de statistiques nationales en utilisant l'imagerie satellite pour de multiples applications ; il est donc important de veiller à obtenir des licences autorisant plusieurs utilisations. Toutefois, la disponibilité des images à résolution moyenne fournies par les satellites Sentinel-1, Sentinel-2 et Landsat 8 réduit la part des coûts liée à l'acquisition de données, ce qui permet de privilégier les activités de traitement, de collecte de données sur le terrain et de validation. Ces coûts sont en outre compensés par l'efficacité accrue des statistiques ainsi obtenues, les gains d'efficacité enregistrés lors de l'exécution pratique et logistique des opérations de terrain et la possibilité de mutualiser les coûts avec d'autres applications.

Les vérifications réalisées à l'aide de données haute résolution et THR peuvent réduire le coût des activités de validation sur le terrain ; elles ne permettent pas d'éliminer la phase d'enquête sur le terrain, mais peuvent les rendre plus efficaces en supprimant la nécessité d'effectuer des mesures au sol lorsque le dénombrement sur le terrain s'appuie sur des SIG et des images satellite contemporaines pour mesurer la superficie des parcelles.

Le matériel informatique et les logiciels engendrent essentiellement des coûts ponctuels. La répartition des coûts liés à l'utilisation d'un système de suivi agricole faisant appel à la télédétection ou à un SIG est illustrée à la figure 1.

FIGURE 1. RÉPARTITION DES COÛTS D'UN SYSTÈME DE SUIVI AGRICOLE FAISANT APPEL À LA TÉLÉDÉTECTION OU À UN SIG.



Cette répartition peut varier selon les pays, mais également selon les cultures, les saisons et les zones concernées. Les Nations Unies ont conclu une série de contrats d'achat avec les principaux fournisseurs d'images, notamment MacDonald Dettwiler (MDA, responsable des satellites QuickBird, IKONOS, WorldView-1, WorldView-2, GeoEye-1, WorldView-3, KOMPSAT-2, KOMPSAT-3, ZY-3 et RADARSAT-2) et Airbus DS Geo (dont les images sont issues des satellites TerraSAR-X, SPOT-6/7 et Pléiades). Les sources d'images peuvent également être mises à disposition dans le cadre d'interventions humanitaires, conformément à la Charte internationale « Espace et catastrophes majeures »¹. L'intégration d'un suivi agricole national dans les missions d'évaluation des récoltes et de la sécurité alimentaire de la FAO et du Programme alimentaire mondial (PAM) permet d'assurer une assistance technique à la réalisation de prévisions sur la production végétale.

8.4. MISE EN ŒUVRE DU PROGRAMME : ÉTUDES DE CAS

Les programmes de soutien à la cartographie de la couverture du sol mis en place par la FAO illustrent bien l'application des données de télédétection aux statistiques agricoles à différentes échelles. En Éthiopie, au Pakistan, en Afghanistan et au Bangladesh, ces programmes ont permis de créer des bases de données nationales de la couverture du sol à partir d'images à haute résolution et favorisé l'intégration des produits cartographiques dans les enquêtes agricoles, le renforcement des capacités, la formation et la mise au point de systèmes informatiques. Au Rwanda, des orthophotographies aériennes THR ont servi de point de départ à des enquêtes à base multiple. Les cartes de couverture du sol normalisées suivant le système LCCS de la FAO réalisées récemment au Pakistan (en particulier dans les provinces du Sind, du Pendjab, du Khyber Pakhtunkhwa, du Baloutchistan et dans les zones tribales sous administration fédérale) et en Afghanistan constituent des ressources pertinentes pour l'élaboration des bases de sondage aréolaires. Il est toutefois nécessaire d'évaluer la capacité des données à contribuer à la stratification (classes fondées sur le pourcentage de terres cultivées, par exemple), même avec les données de couverture du sol existantes. En effet, ces dernières ne pourront se prêter à la stratification qu'à condition de présenter une résolution appropriée, de couvrir certaines classes et d'être relativement récentes eu égard au rythme de l'évolution du paysage.

8.4.1. Exemple 1 : Éthiopie – application de la stratification d'une base de sondage aréolaire

L'une des principales utilisations des données géospatiales et de télédétection dans le domaine des statistiques agricoles intervient au niveau de la conception : base de sondage aréolaire, stratification et sélection des échantillons (Carfagna, 2013). En Éthiopie, l'Agence centrale de statistique (CSA, de l'anglais *Central Statistical Agency*) a eu recours à un échantillonnage traditionnel de type liste (fondé sur la liste des districts de recensement) pour élaborer des enquêtes agricoles réalisées sur le terrain et auprès des ménages. Il a été reconnu que cette méthode présentait certaines limites : les facteurs d'extrapolation reposent sur le nombre de ménages plutôt que sur les surfaces cultivées, et certains champs peuvent être omis en raison d'erreurs non liées à l'échantillonnage. La CSA a procédé à une évaluation comparative des méthodes fondées sur les bases de sondage aréolaire et par liste lors de la saison de *meher* 2008-2009 (saison des pluies qui s'étale de juin à octobre). La base aréolaire, qui utilisait les districts de recensement (DR) comme unités primaires d'échantillonnage (UPE), était identique à celles employées pour les enquêtes de recensement de la population menées auprès d'environ 150 à 200 ménages, et numérisées dans un SIG. La stratification des UPE a été entreprise d'après le pourcentage de surfaces cultivées calculé lors de l'interprétation manuelle des données de télédétection (Programmes nationaux de méthodes statistiques pour l'agriculture en Éthiopie, 2008).

Les DR ont été répartis en cinq classes : une pour les DR ne comportant aucune surface cultivée, puis quatre classes à intervalles égaux pour atteindre 100 % de surface cultivée. La strate qui ne contenait aucune surface cultivée a été

¹ <https://disasterscharter.org/fr/web/guest/home>

exclue de l'enquête, bien que cela puisse fausser les résultats en cas de production marginale (peuplement forestier ou agriculture urbaine, notamment). La première évaluation (réalisée dans le Choa occidental) s'est fondée sur l'interprétation manuelle des images afin de calculer la superficie cultivée dans chaque DR. Si ce processus n'a pas initialement eu recours à la segmentation d'image et à une cartographie complète de la couverture du sol, les essais ultérieurs (effectués dans la région d'Oromia) ont employé un système de classification de la couverture du sol reposant sur les méthodes LCCS de la FAO (Di Gregorio *et al.*, 2005). Les méthodes fondées sur le système LCCS ont l'avantage d'être reproductibles et d'offrir un certain niveau d'automatisation et d'homogénéité grâce à l'adoption de normes ISO (19144-2).

Des unités secondaires d'échantillonnage (USE) ont été définies à l'aide de technologies SIG : les DR ont ainsi été divisés en plusieurs unités de terre de 40 ha, suivant l'interprétation des délimitations visibles sur les images, de façon à obtenir le nombre d'USE préétabli.

La sélection des échantillons s'est déroulée selon un processus en deux étapes : les UPE (équivalents aux DR) ont d'abord été affectés aux différentes strates de manière proportionnelle, puis des DR ont été sélectionnés dans chaque strate en fonction de la probabilité proportionnelle à la taille (autrement dit, selon le nombre d'USE affecté au DR) afin de constituer l'échantillon. Les USE ont uniquement été générées au sein des UPE sélectionnées. Pour l'enquête 2009, deux « segments » ont été sélectionnés par échantillonnage aléatoire au sein du DR choisi. Des iconocartes des segments sélectionnés ont été réalisées afin de faciliter l'enquête de terrain conduite par les agents recenseurs.

8.4.2. Exemple 2 : Pakistan – Service d'information sur les cultures

Au Pakistan, le Service d'information sur les cultures (CRS, de l'anglais *Crop Reporting Service*) est un service global placé sous l'autorité du secrétaire à l'Agriculture et chargé de fournir des informations sur les cultures au gouvernement fédéral pakistanais, au Bureau fédéral de statistiques, aux provinces, au Bureau de statistiques du Pendjab, aux universités, aux chercheurs en agriculture, au Département de l'alimentation et de la vulgarisation agricole, entre autres organismes. Si ce service avait initialement recours à des enquêtes sur les revenus, ses méthodes ont évolué, et il réalise actuellement des enquêtes à base de sondage aréolaire en s'appuyant sur le traitement géospatial et la télédétection. Le CRS, organisé au niveau provincial, bénéficie du concours de responsables des statistiques au niveau des *tehsils*, mais également de statisticiens assistants et de correspondants chargés des cultures qui travaillent dans les villages couverts par les bases de sondage aréolaires. Il emploie 1 611 professionnels diplômés en statistiques, en économie et en mathématiques et exerce ses activités dans les provinces du Pendjab, du Sind, du Khyber Pakhtunkhwa et du Baloutchistan. Il reçoit un financement provincial normatif, proportionnel à la superficie de chaque province dans laquelle il opère ; ces subventions sont complétées par des activités réalisées dans le cadre des projets.

Le programme prévoit des mesures de renforcement des capacités : des formations de remise à niveau sont ainsi proposées par l'ancien ministère de l'Alimentation et de l'Agriculture ou par des bureaux de statistiques fédéraux et provinciaux au début des saisons de *rabi* et de *kharif*.

Les méthodes, outils et techniques ont évolué depuis les premières enquêtes d'opinion menées par le Département du revenu jusqu'au milieu des années 1950. Par la suite, l'estimation des rendements agricoles s'est appuyée sur des coupes échantillons. Pour estimer les surfaces cultivées et le rendement agricole, le ministère de l'Alimentation et de l'Agriculture et le Bureau fédéral de statistiques ont mis au point une base de sondage statistique à partir des registres relatifs aux revenus fonciers 1973-1974, obtenus auprès du Département des revenus fonciers en 1978. Ces données ont servi de base à la stratification et les UPE, délimitées suivant les frontières administratives, ont été stratifiées par classes en fonction de la taille des villages. Dans chaque district des régions productrices de blé, 20 à 40 villages ont été sélectionnés de façon à établir la base aréolaire correspondant à la culture du blé. Cette méthode a ensuite été élargie à d'autres zones cultivées afin de constituer la base d'échantillonnage principale des villages. La structure d'échantillonnage a été mise au point par le Bureau fédéral de statistiques et les enquêtes ultérieures ont été menées dans la province par le CRS.

Ce système a été appliqué à la culture du coton (550 villages), de la canne à sucre (450 villages) et du riz (550 villages). Un correspondant chargé des cultures (personnel de première ligne) a été affecté dans un ou plusieurs villages de la base de sondage aréolaire afin de conduire les enquêtes saisonnières sur le terrain et de rendre compte des surfaces semées et cultivées grâce à ces enquêtes. Le rendement agricole a été estimé à partir de trois échantillonnages aléatoires, dont chacun a été répété deux fois. Ces estimations ont servi à déterminer la superficie cultivée et le rendement agricole au niveau du district. La superficie utilisée pour la prévision initiale de la production végétale provenait du CRS. Pour l'estimation finale, en revanche, la surface cultivée a été obtenue auprès du Département du revenu, et le rendement agricole auprès du CRS.

À l'heure actuelle, le système global utilise à la fois des techniques objectives et subjectives. La base d'échantillonnage élaborée en 2004 par le Bureau fédéral de statistiques repose sur une base de sondage aréolaire couvrant plusieurs cultures, notamment le blé, le coton, le riz, la canne à sucre et le maïs. Une base d'échantillonnage avait également été définie pour les petites cultures (mangue, agrumes, pomme de terre, pois chiches, haricot mungo et haricot urd, notamment) en collaboration avec l'ancien ministère fédéral de l'Alimentation et de l'Agriculture à la fin des années 1990.

TABLEAU 5. TECHNIQUES D'IDENTIFICATION DE LA COUVERTURE CULTIVÉE.

Type de technique	Type de technique	Type de technique
Objective	<i>Rabi</i>	Blé, pois chiches, pomme de terre, lentilles, oignon, graines oléagineuses, maïs de printemps, agrumes
	<i>Kharif</i>	Coton, riz, canne à sucre, maïs d'automne, sésame, graines de guar, haricot mungo, haricot urd, arachides, mangue
Subjective	<i>Rabi</i>	Orge, légumes secs, tabac, ail, piments, tomate, navet, banane, goyave, dattes, raisin, cultures fourragères <i>rabi</i> , condiments <i>rabi</i> et autres fruits, légumes et légumineuses <i>rabi</i>
	<i>Kharif</i>	Millet indien, sorgho sucré, curcuma, graines oléagineuses, autres légumineuses <i>kharif</i> , gombo, cultures fourragères <i>kharif</i> , autres fruits et légumes <i>kharif</i>

8.4.3. Exemple 3 : Pakistan – Unité géospatiale opérationnelle SUPARCO/FAO

La FAO et la SUPARCO ont collaboré à la mise en place au niveau fédéral d'un nouveau Système d'information agricole intégré (AIS, de l'anglais *Agricultural Information System*) qui fait appel à la télédétection, aux technologies géospatiales et à différents outils de diffusion. L'objectif consiste à proposer une méthode unifiée, fiable et rapide permettant de fournir des statistiques agricoles officielles et harmonisées. Ce programme a mis au point un solide système d'estimation des surfaces cultivées, d'élaboration de bases de sondages aréolaires et de procédures de diffusion des statistiques agricoles auprès des parties prenantes et des décideurs.

L'estimation des surfaces cultivées s'appuie sur le traitement des données satellite acquises pour une période donnée, la conduite d'enquêtes visant à vérifier la réalité de terrain durant la période de récolte, la collecte des signatures des cultures, le traitement effectué en laboratoire, ainsi qu'une évaluation de la précision. Ces activités sont réalisées avec le concours des infrastructures de la station réceptrice de la SUPARCO. L'acquisition des images satellite au niveau national s'effectue en deux fois pour les cultures *rabi* (printemps) et *kharif* (automne), selon les étapes suivantes :

- une première fois, quatre semaines après l'achèvement des semis (juin-juillet pour les cultures *kharif* et décembre-janvier pour les cultures *rabi*) ;
- une deuxième fois, huit semaines après l'achèvement des semis (août pour les cultures *kharif* et février pour les cultures *rabi*).

La navigation GPS en temps réel a permis de procéder à des vérifications approfondies sur le terrain. Les classifications d'échantillons de terrain ont servi à contrôler la classification par maximum de vraisemblance des images satellite multitudes réalisée à l'aide du logiciel de traitement d'image ERDAS.

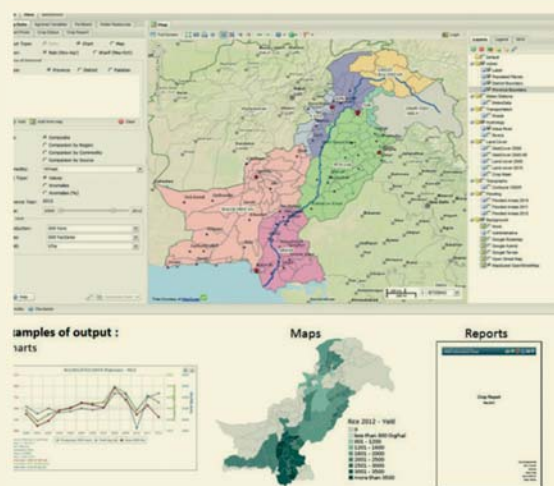
L'élaboration de la base de sondage aréolaire s'est appuyée sur la télédétection et le traitement SIG des images acquises en février et en septembre, et a ensuite été utilisée pour définir les strates en fonction de l'intensité de la végétation et de l'interprétation visuelle de la couverture du sol (dix classes différentes). Le Pakistan a été divisé en dix zones, dont quatre au Pendjab, deux dans les provinces du Sind et du Khyber Pakhtunkhwa et une au Baloutchistan. La stratégie d'échantillonnage a permis de définir les UPE et USE. Ces dernières, numérotées par ordre serpentin, ont été utilisées pour garantir une répartition spatiale homogène dans les zones identifiées. Vingt à trente segments d'environ 30 ha ont été sélectionnés dans chaque strate, pour un total de 379 segments dans chaque région.

Suivant la conception de l'échantillonnage aréolaire, des facteurs d'extrapolation ont été définis afin d'estimer la superficie semée dans chaque strate de chaque zone au moyen d'un modèle statistique, et ce pour différentes cultures. Les données générées ont fait l'objet d'un examen critique par une équipe d'experts en agronomie, en télédétection et en statistiques, l'objectif étant de normaliser la technique en faisant appel à la classification d'image et aux courbes de tendance historiques. La diffusion des informations constitue une composante essentielle du système de suivi des cultures par satellite mis au point par la SUPARCO et la FAO : des communiqués mensuels sur les cultures fournissent des informations sur les indices de végétation (indicateur du rythme de croissance des cultures), la disponibilité en engrais et en eau d'irrigation, la situation agrométéorologique durant les périodes de récolte, ainsi que les prévisions ou estimations du rendement et de la production agricoles pour différentes saisons. Ces communiqués sont diffusés en version papier et en ligne². Dans le cadre de ce programme, quatre mécanismes et outils de diffusion ont été mis en place : le Portail d'information sur les cultures, le Système de suivi satellitaire de l'agriculture mondiale pour le Pakistan (GLAM, de l'anglais *Global Agricultural Monitoring*), le Système mobile d'information et de géolocalisation agricoles (MAGIS, de l'anglais *Mobile Agricultural Geo-Tagging Information System*) et le Système d'échantillonnage aréolaire (AFSS, de l'anglais *Area Frame Sampling System*).

PORTAIL D'INFORMATION SUR LES CULTURES

Le Portail d'information sur les cultures est une plateforme Web en libre accès, mise en place pour faciliter l'analyse et la diffusion des données sur les cultures du Pakistan et des variables climatiques, agronomiques, hydrologiques et économiques dans ce domaine.

Il propose un filtrage détaillé des archives de données à partir d'une combinaison de paramètres spatiotemporels définis par les utilisateurs, qui peuvent ainsi s'intéresser à des cultures et des facteurs agronomiques ou environnementaux spécifiques, ces derniers étant stockés dans le système de façon à pouvoir générer des produits normalisés (tableaux récapitulatifs, cartes, graphiques ou encore rapports définis par les utilisateurs). Le portail intègre également des modules d'information sur l'eau, les engrais, les conditions agrométéorologiques, les prix du marché et l'état des cultures, qui viennent compléter le système d'information agricole général.



² <http://suparco.gov.pk/webroot/pages/pak-scms.asp>

GLAM

Le système GLAM est un portail Web utilisant des données MODIS, conçu et mis en place au Pakistan par l'université du Maryland (États-Unis) afin d'offrir un accès direct à des analyses Web de l'état et des anomalies de la végétation à différentes échelles spatiales et temporelles. Il a pour objectif d'aider le CRS à l'échelle provinciale à publier en temps voulu des prévisions améliorées sur l'état des cultures et la production végétale.

Ce système propose la définition de différents paramètres d'entrée, la visualisation des images et l'interprétation graphique de l'état de la végétation. Il utilise comme paramètre d'entrée un produit normalisé composite, obtenu à partir d'images MODIS enregistrées sous différents angles (8 jours et 16 jours). Ce portail intègre également des couches de données vectorielles qui permettent un suivi de l'état de la végétation à l'échelle des provinces, des divisions et des districts. Il fournit par ailleurs une grille d'analyse rectangulaire fondée sur l'indice d'orientation agricole, destinée à étudier les anomalies de la végétation allant de 0,25 km² à 4 900 km². Enfin, le système GLAM se caractérise par sa capacité à générer automatiquement des profils graphiques des indices NDVI et NDWI à partir d'images MODIS et d'une analyse d'histogramme.



MAGIS

L'application mobile MAGIS a été développée au Pakistan en vue d'aider le CRS et les correspondants chargés des cultures à recueillir de façon plus efficace les données relatives à la base de sondage de type liste.

Elle permet de collecter des données numériques et géoréférencées sur le terrain, de les transmettre au serveur central MAGIS situé au siège du CRS via Internet ou le réseau de téléphonie mobile, et de les mettre à la disposition des statisticiens afin qu'ils puissent les analyser et établir des comptes rendus.



AFSS

Le CRS recueille des informations sur les cultures et d'autres paramètres connexes à l'intérieur des segments définis pour les périodes de récolte *rabi* et *kharif*. L'application mobile AFSS, dotée d'un système GPS, remplace les anciens formulaires papier. Elle permet de valider les enregistrements sur le terrain, de réaliser des photographies sur le terrain et de les combiner avec les images de télédétection pour la saisie des données spatiales des différentes zones. Sa fonctionnalité de saisie des données de terrain, expérimentée lors de la conception de l'échantillonnage aréolaire des provinces du Pendjab et du Sind, améliore réellement l'efficacité et la précision de la saisie des données, ainsi que la synchronisation avec les bases de données principales, ce qui assure une gestion efficace des enregistrements.



Le recours à un système d'estimation des cultures par télédétection offre les avantages suivants :

- i. création de méthodes et d'outils d'enquêtes agricoles et rurales rigoureux, qui renforcent la qualité (rapidité, précision, exactitude, fiabilité et rentabilité) des efforts d'échantillonnage et de leurs résultats à l'échelle des provinces pakistanaises et améliorent l'ensemble des estimations nationales des surfaces cultivées, du rendement et de la production agricole ;
- ii. renforcement des ressources humaines et des capacités techniques dans quatre antennes régionales du CRS et deux grandes universités agricoles concernant l'utilisation des nouvelles technologies géospatiales et de télédétection et les méthodes statistiques de collecte de données, de production de statistiques agricoles et de diffusion des informations sur les cultures. Les antennes provinciales du CRS sont en mesure d'assurer des services efficaces de suivi des cultures ;
- iii. intégration d'applications mobiles dans le processus de suivi des cultures, qui offre de nouvelles perspectives intéressantes en ce qui concerne l'amélioration des mesures et de la saisie des données sur le terrain, la validation des données et les possibilités d'entretiens individuels assistés par ordinateur. La synchronisation finale des données autorise les opérations à distance ainsi qu'une validation des enregistrements et un post-traitement rapides ;
- iv. mise en place d'un système Web personnalisé d'analyse des informations et de transmission des données, et création d'une base de données relationnelle unique pour l'évaluation des cultures, destinée aux analystes des cultures du CRS ;
- v. démonstration des améliorations apportées en matière d'estimation et de prévision des récoltes par l'utilisation concertée et le partage des données de terrain, des images satellite et des données dérivées. Cette base de connaissances pourra servir à d'autres organismes du Pakistan qui envisagent d'intégrer la télédétection, les SIG, les bases de sondage aréolaires et autres pratiques connexes dans leurs statistiques agricoles ;
- vi. élaboration de protocoles et de procédures visant à anticiper l'éventuelle mise hors service future ou l'apparition de nouveaux satellites, mais également les modifications de configuration des données, de façon à garantir la continuité des opérations ou à assurer une transition ou une migration dans des délais raisonnables.

8.4.4. Exemple 4 : Bangladesh – unité géospatiale opérationnelle

Cet exemple illustre les multiples utilisations possibles des données produites à d'autres fins que le suivi des cultures, et montre comment on peut, grâce à des actions collaboratives, proposer une carte nationale de couverture du sol créée à partir d'images satellite et adaptée à l'élaboration de bases de sondage aréolaires et à l'établissement de comptes rendus sur les cultures. Les initiatives de ce type accroissent la rentabilité des activités liées aux rapports sur les cultures.

La FAO, en collaboration avec des organisations privées, des institutions gouvernementales et des universités du Bangladesh, élabore actuellement un jeu de données national de couverture du sol destiné à servir de base de référence pour plusieurs applications (suivi environnemental et évaluation du potentiel de productivité agricole dans le cadre d'une étude sur l'impact de divers scénarios climatiques, notamment).

La création d'un jeu de données de couverture du sol implique diverses étapes de développement des ressources techniques et humaines :

- Renforcement des capacités
- Définition des équipes et de leur rôle
- Définition des légendes
- Sélection et acquisition des images satellite
- Prétraitement des images
- Segmentation par objets

- Interprétation
- Contrôle qualité et publication finale

8.4.4.1. Formation

Les activités de formation et de renforcement des capacités, soutenues par la FAO, comprenaient notamment l'élaboration d'une nouvelle légende de couverture du sol à l'aide de la norme ISO relative au cadre LCML et du système LCCS, mais également l'organisation d'un atelier visant à mettre au point un Système de référence national permettant de définir l'ensemble des classes et attributs de référence nécessaires à l'identification de la légende de couverture du sol utilisée pour le Bangladesh. Des professionnels appartenant à diverses institutions locales ont participé à cet atelier, notamment des représentants de l'Université d'ingénierie et de technologie du Bangladesh, du Centre pour les services d'information géographique et environnementale, du service du Département des forêts du Bangladesh chargé de l'identification et de la gestion des ressources, de la Société bangladaise de géoinformatique et de l'Organisation de recherche spatiale et de télédétection du Bangladesh.

8.4.4.2. Définition des équipes

La réalisation de cartes nationales de couverture du sol représente une tâche complexe, qui suppose la mise en place d'un certain nombre de groupes dotés de compétences et d'aptitudes spécifiques. On a ainsi identifié la nécessité de disposer des services suivants :

- un service chargé des données géospatiales (SIG) ;
- des photo-interprètes ;
- un service d'appui technique ; et
- un service de contrôle de la qualité.

Le **service SIG** est responsable de toutes les activités relatives à la gestion, au stockage, à l'analyse et à la diffusion des informations géospatiales sollicitées et produites durant l'élaboration de la carte de couverture du sol. Il aide les photo-interprètes à assurer l'édition et autres tâches en rapport avec les SIG et à gérer les informations auxiliaires requises pour l'analyse des images (soutien concernant notamment les images satellite disponibles).

Les **photo-interprètes** réalisent l'essentiel de la cartographie de la couverture du sol. Sous la supervision d'un responsable technique, ils sont chargés de la classification visuelle ou semi-automatique des images satellite, afin d'affecter les codes correspondants aux polygones de couverture du sol générés lors du processus de segmentation.

L'**équipe d'appui technique** est chargée de dispenser des formations et d'apporter un soutien au superviseur technique et aux photo-interprètes. Elle fournit des solutions techniques et améliore la productivité opérationnelle.

L'**équipe de contrôle de la qualité** est chargée d'examiner de façon critique l'interprétation de la couverture du sol, qu'elle a le pouvoir de commenter ou de rejeter.

Outre ces services, une **assistance informatique** est nécessaire afin de gérer le matériel informatique, les logiciels, les sauvegardes, les licences et toute autre question connexe.

8.4.4.3. Définition des légendes

La définition des légendes est un élément essentiel du processus et repose sur un certain nombre de critères :

- la disponibilité et la qualité des images satellite fournies, et la capacité des photo-interprètes à distinguer les différentes classes ;
- la disponibilité d'informations auxiliaires (enquêtes sur le terrain et jeux de données issus d'autres projets de cartographie, par exemple) ;
- la représentativité des différentes classes sur le territoire national ; et
- les exigences liées à des classes spécifiques afin de répondre aux grands objectifs du produit cartographique et de permettre les applications prévues.

Les classes de couverture du sol ont été identifiées à partir d'une analyse du diagramme du Système de référence national. Elles ont ensuite été décrites à l'aide de la norme ISO relative au cadre LCML.

8.4.4.4. Sélection et acquisition des images satellite

La sélection des images satellite destinées à la réalisation de cartes de couverture du sol doit tenir compte d'un certain nombre de facteurs : le budget disponible, la qualité des images tout au long de l'année et les classes de couverture du sol à identifier. Au Bangladesh, la principale période de récolte se situe durant la mousson, une saison fortement marquée par la présence de nuages. Procéder à l'interprétation lors de la saison sèche facilite l'identification de la végétation ligneuse naturelle et cultivée, qui se distingue alors clairement des terres nues non végétalisées. C'est pourquoi on a sélectionné à la fois des images optiques à haute résolution (images SPOT6/7 d'une résolution spatiale de 6 m) enregistrées lors de la saison sèche et des données SAR issues d'images radar multitudes (Sentinel-1) pour la saison des pluies, ces dernières n'étant pas affectées par les conditions météorologiques.

8.4.4.5. Prétraitement des images

Les images obtenues ont dû être soumises à un prétraitement en vue d'une analyse plus approfondie. Le processus de prétraitement comprend différentes tâches : composition de bandes, mosaïquage des images, reprojection, modification des formats et modification de la profondeur de pixels. Il englobe également des techniques d'amélioration d'image et de suppression des nuages et du bruit. Les images SAR ont nécessité un traitement différent de celui requis pour les images issues de capteurs optiques, ce qui a donc exigé des compétences supplémentaires en matière de traitement d'image.

8.4.4.6. Segmentation par objets

À l'issue du prétraitement, les images ont été segmentées à l'aide d'un algorithme multirésolution de façon à générer une couche composée de polygones vectoriels non classés. À ce stade, le facteur le plus déterminant a été la sélection des paramètres de segmentation utilisés par l'algorithme pour répartir et segmenter les images d'entrée. Aucun paramétrage spécifique n'a été suggéré, car le paramétrage dépend de nombreux facteurs et il est nécessaire d'effectuer plusieurs essais avant de segmenter les objets à identifier, de façon à obtenir le niveau de résolution et de discrimination correspondant à la légende de classification.

8.4.4.7. Interprétation

L'interprétation correspond à la phase durant laquelle les photo-interprètes affectent des codes de couverture du sol aux différents polygones. Ce travail de longue haleine a fait l'objet d'une planification et d'un suivi rigoureux afin de normaliser l'interprétation lorsque celle-ci fait appel à des techniques manuelles. Il est possible de recourir à une interprétation automatique préliminaire fondée sur les caractéristiques spectrales et géométriques des objets et sur les jeux de données d'apprentissage des zones classées pour effectuer des classifications dirigées. Une analyse multitemporelle a permis d'extraire les valeurs moyennes de chaque objet (valeurs NDVI, notamment) pouvant être utilisées dans les techniques de regroupement statistique afin de sélectionner les configurations végétales similaires.

8.4.4.8. Contrôle qualité et publication finale

L'évaluation de la précision s'est appuyée sur une série de points de contrôle : les procédures de classification et de contrôle qualité ont ainsi été vérifiées pour l'ensemble du processus d'interprétation, de façon à ce que le jeu de données de couverture du sol réponde aux attentes en matière de qualité des résultats et de cohérence. L'équipe s'est également chargée de la structuration finale du jeu de données (métadonnées et documentation des données, notamment).

8.4.5. Exemple 5 : Rwanda – enquêtes par sondage probabiliste à objectifs multiples

L'Institut national de statistiques du Rwanda s'est appuyé sur des orthophotographies THR (d'une résolution de 0,25 m) pour concevoir et mettre en œuvre des bases de sondage multiples pour le pays (d'une superficie de 26 000 km²), en associant une base de sondage aréolaire et une base de sondage de type liste afin d'améliorer les statistiques agricoles nationales. Cette base de sondage multiple couvre 560 exploitations agricoles spécialisées dans la culture de produits de base ayant un poids important dans les estimations économiques et agricoles nationales. Ces exploitations spécialisées sont toutes recensées dans une base de sondage de type liste, ce qui évite les erreurs d'échantillonnage ; elles sont également prises en compte dans la base de sondage aréolaire. Soulignons qu'il est important d'utiliser des termes normalisés (notamment « exploitation agricole », « parcelle agricole », « champ », « agriculteur », « régisseur salarié » et « répondant ») afin d'assurer la cohérence des enquêtes agricoles. Ces termes s'inspirent des ontologies de la FAO (FAO, 2015).

La première étape de la conception d'une enquête agricole consiste à définir ses objectifs, de façon à pouvoir atteindre le niveau de précision requis. En règle générale, la conception de l'enquête repose sur les questions suivantes :

- i. Quels sont les objectifs et les attentes de cette enquête ?
- ii. Quelle est la définition retenue pour le terme « exploitation agricole » ?
- iii. Quelles sont les variables agricoles à étudier, sur quelle échelle de temps et à quelle périodicité (saisonnalité, par exemple) ?
- iv. Quelles enquêtes relatives aux produits agricoles de base sont jugées inappropriées, et de ce fait, sur quoi doit porter cette nouvelle enquête ?
- v. Quel est le niveau de précision requis ?
- vi. À quelle échelle les données seront-elles synthétisées : territoire national, État, région, bassin hydrographique, etc. ?

Cette initiative rwandaise visait à proposer une enquête agricole nationale, saisonnière et à objectifs multiples (cultures, élevage, forêts et produits agricoles), portant sur diverses variables pour chaque exploitation agricole :

- i. superficie plantée et récoltée, superficie destinée à la plantation, rendement agricole potentiel et effectif pour chaque culture ou variété de culture, production végétale et nombre d'arbres ;
- ii. inventaire du bétail et des volailles (nombre, type, âge, sexe, race et utilisation, par exemple) ;
- iii. production de lait, d'œufs, de miel et de semences ;
- iv. nombre et types de méthodes de culture et d'intrants agricoles (main-d'œuvre, type et quantité de semences, d'engrais et de pesticides, source de l'eau d'irrigation, drainage, importance de l'agriculture itinérante, stocks, machines, matériel et bâtiments agricoles) ;
- v. nombre et types d'exploitations agricoles (situation géographique, statut juridique et régime foncier, par exemple) ;
- vi. coûts de production et valeur des ventes ;
- vii. population travaillant dans le secteur de l'agriculture (caractéristiques démographiques élémentaires concernant les agriculteurs, les membres de leur famille et les régisseurs salariés travaillant sur l'exploitation et les jours de travail, notamment).

La base de sondage aréolaire s'est appuyée sur les strates d'utilisation des terres dans les 30 districts ; ces strates ont été divisées en « segments » ou unités d'échantillonnage non chevauchantes, dont chacune a ensuite été subdivisée en lopins non chevauchants. Chaque lopin correspond soit à la surface d'une exploitation agricole à l'intérieur d'un segment, soit à une surface de terrain n'appartenant à aucune exploitation. La zone totale de l'enquête a été entièrement subdivisée en lopins non chevauchants. Les segments et lopins ont été délimités en fonction des frontières physiques ou des champs reconnaissables afin de pouvoir être identifiés sans équivoque par les agents recenseurs. Les lopins, qu'ils correspondent ou non à une exploitation agricole, ont fait office

d'unités déclarantes pour chaque variable de l'enquête. L'enquête par sondage probabiliste menée au Rwanda s'est basée sur une enquête par sondage en grappes stratifié ; elle s'est servie d'un échantillon systématique constitué à partir des UPE en associant une base de sondage aréolaire et une base de sondage de type liste recensant les exploitations spécialisées (base de sondage multiple). Le recensement a été effectué sous forme d'entretiens à l'aide de formulaires papier afin de recueillir les variables susmentionnées. Les entretiens portent sur les lopins sélectionnés, sans tenir compte de ceux qui correspondent à des terres non agricoles (friches ou surfaces en eau, par exemple) ; il est prévu d'utiliser également des outils d'enquête CAPI et des GPS afin de faciliter la collecte de données.

L'élaboration de la base de sondage aréolaire peut s'appuyer sur différentes strates d'utilisation des terres (proportion de terres cultivées, pratiques agricoles spécifiques, prédominance de certaines cultures et taille moyenne des champs cultivés) dont les délimitations peuvent être localisées sur le terrain. Au sein de chaque strate, les UPE sont classées en fonction de leurs similitudes et sélectionnées suivant la probabilité proportionnelle à la taille ; dans l'UPE, les segments sont sélectionnés selon une probabilité égale. La surface de chaque UPE a été mesurée afin de pouvoir déterminer le nombre total d'USE dans chaque strate et dans la totalité de la base d'échantillonnage sans devoir cartographier chaque segment. Des photographies orthorectifiées (imprimées à une échelle de 1:1 000) de chaque segment sélectionné ont été utilisées par les agents recenseurs afin d'identifier et de mesurer les surfaces agricoles ; toutefois, ces photographies ne sont actuellement pas utilisées en format numérique sur le terrain. Dans le cadre des enquêtes par questionnaire, la surface des champs peut être fournie par les agriculteurs lors des entretiens ou calculée d'après les photographies ; il est cependant possible d'obtenir des évaluations plus précises en cartographiant les champs concernés par rapport aux orthophotographies et en faisant appel aux technologies SIG mobiles. La supervision de l'enquête a été facilitée par les calculs de surface réalisés à l'aide des images satellite.

La conception de la base de sondage aréolaire a fourni des formules permettant de calculer la variance et l'extrapolation directe pour l'estimation des segments fermés, pondérés et ouverts. Les concepteurs ont veillé à ce que le nombre total de segments soit connu pour chaque strate et chaque UPE et à ce que la probabilité de sélection de chaque segment soit égale à sa probabilité conditionnelle lors de l'échantillonnage à deux degrés. L'estimation par extrapolation directe de la valeur totale de chaque variable de l'enquête a été calculée selon la formule suivante :

$$Y_c = \sum_{h \in S} \sum_{j \in B_h} \sum_{k \in G_{hj}} e_{hjk} \sum_{m \in T_{hjk}} t_{hjk m}$$

où :

Y_c	= estimation par échantillon de la valeur totale de la variable d'enquête y
S	= ensemble des strates d'utilisation des terres
B_h	= ensemble des sous-strates comprises dans la strate h
G_{hj}	= ensemble des segments compris dans la sous- strate j de la strate d'utilisation des terres h
T_{hjk}	= ensemble des lopins compris dans le segment k de la sous- strate j de la strate d'utilisation des terres h
e_{hjk}	= facteur d'extrapolation pour l'ensemble des lopins compris dans le segment k
$t_{hjk m}$	= valeur de la variable y à l'intérieur du lopin m

8.5. CONCLUSION

Les applications de la télédétection et leur intégration dans les SIG ont rendu les opérations de suivi des cultures plus simples, plus rapides et plus précises. La prise en compte de ces techniques et de ces données dans les procédures d'élaboration de bases de sondage aréolaires, de stratification et de collecte de données de terrain modifie peu à peu la manière dont se déroulent les opérations d'estimation relative aux cultures. Les multiples usages des images de télédétection et la réalisation de produits dérivés dans le cadre des statistiques agricoles (cartes de couverture du sol, notamment) permettent d'intégrer les données satellite dans les estimations de superficies cultivées et de rendement agricole de façon efficace et rentable. L'utilisation de ce type de données dans les statistiques agricoles est soumise à une série de conditions fondamentales concernant les aspects organisationnels, les ressources, les opérations de traitement d'images thématiques, les activités d'échantillonnage, la collecte de données de terrain et les compétences géospatiales et statistiques. Elle repose également sur la collaboration entre les organismes responsables des différents éléments essentiels du suivi et des comptes rendus agricoles : services statistiques, agences de cartographie, agences spatiales ou encore agences thématiques utilisatrices. Les programmes intégrés tels que le Système d'information agricole pakistanais et d'autres évolutions dans ce domaine illustrent la coordination entre le traitement d'image, l'appui à l'échantillonnage et aux enquêtes de terrain et la diffusion des informations aux utilisateurs et à des interfaces complètes destinées aux décideurs. Les programmes collaboratifs à objectifs multiples favorisent eux aussi l'utilisation des images satellite et l'application potentielle de l'analyse d'image à des zones plus étendues, qui permettrait de couvrir des phénomènes comme les catastrophes naturelles (inondations, sécheresses et glissements de terrain, notamment) et leur impact sur la production végétale.

8.6. RÉFÉRENCES BIBLIOGRAPHIQUES

Bell, K.P. et Dalton, T.J. 2007. « Spatial Economic Analysis in Data-Rich Environments ». *Journal of Agricultural Economics*, vol. 58, n° 3, p. 487–501.

Carfagna, E. 2013. *Using satellite imagery and geo-referencing technology for building a master sampling frame*. Communication sollicitée pour la session IPS110 du 59^e Congrès mondial de la statistique sur « L'élaboration de base de sondage principale pour les enquêtes intégrées », p. 1078–83, Hong Kong, 25–30 août 2013. Institut international de statistique, La Haye, Pays-Bas. Disponible à l'adresse : <http://2013.isiproceedings.org/Files/IPS110-P1-S.pdf>.

Carfagna, E. et Gallego, F.J. 2005. « Using Remote Sensing for Agricultural Statistics ». *International Statistical Review*, vol. 73, n° 3, p. 389–404.

Carfagna E., Pratesi, M. et Carfagna, A. 2013. *Methodological developments for improving the reliability and cost-effectiveness of agricultural statistics in developing countries*. Session STS043 du 59^e Congrès mondial de la statistique sur « Le recours aux informations géo-spatiales dans les sondages aréolaires et les estimations effectuées dans le cadre d'enquêtes agricoles et environnementales », p. 1930–5, Hong Kong, 25–30 août 2013. Disponible à l'adresse : <http://2013.isiproceedings.org/Files/STS043-P1-S.pdf>.

Delincé, J. 2017. « Rentabilité de l'application de la télédétection aux statistiques agricoles ». In : Delincé, J. (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 9), Manuel de la Stratégie mondiale : Rome.

FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture). 2015. *Mise en œuvre de la stratégie globale au niveau pays : expérience pilote de pays échantillon – Rwanda*. RAF/AFCAS/15–5d.

Murray, H., Lucieer, A. et William, R. 2010. « Texture-based Classification of sub-Antarctic Vegetation Communities on Heard Island ». *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, n° 3, p. 138–49.

Quarmby, N.A., Milnes, M. et Hindle, T.T. 1993. « The use of multi-temporal NDVI measurements from AVHRR for crop estimation and prediction ». *International Journal of Remote Sensing*, vol. 14, n° 2, p. 199–210.

Wiegand, C.L. et Richardson, A.J. 1984. « Leaf Area, Light Interception, and Yield Estimates from Spectral Components Analysis ». *Agronomy Journal*, vol. 76, n° 4, p. 543–8.

Wiegand, C.L. et Richardson, A.J. 1990. « Use of Spectral Vegetation Indices to Infer Leaf Area, Evapotranspiration and Yield: II Results ». *Agronomy Journal*, vol. 82, p. 630–6.

Chapitre 9

Rentabilité de la télédétection aux fins de statistiques agricoles

Jacques Delincé

La meilleure méthode pour évaluer la rentabilité de la télédétection aux fins de statistiques agricoles consiste à comparer les gains obtenus (qui s'expriment généralement sous forme de réduction de la variance d'échantillonnage) aux coûts supplémentaires engagés (imagerie, analyse des données, formation du personnel et investissements dans les logiciels et équipements). Carfagna (2001, 2013), Nelson *et al.* (2007), Tenkorang et Lowenberg-DeBoer (2008), Miller *et al.* (2012), Gallego (2014) et Delincé (2015) proposent ainsi plusieurs solutions intéressantes.

9.1. LA QUESTION DES COÛTS

Les coûts associés à l'utilisation de la télédétection peuvent être globalement classés en deux catégories: le coût d'achat des images d'une part, et le coût du traitement des données d'autre part (achat et entretien des équipements et logiciels, recrutement du personnel, formation). La baisse considérable du prix des équipements et logiciels ces dernières années est imputable à la multiplication récente des logiciels en libre accès. L'analyse d'images à partir de l'informatique dématérialisée étant devenue une pratique courante, le stockage sur disque et les ordinateurs bon marché permettent d'analyser d'importantes quantités d'images. Il convient cependant d'accorder une attention particulière à la disponibilité et aux compétences du personnel (Latham, 2017). Il sera en effet nécessaire de s'entourer d'une équipe de spécialistes possédant les qualifications, la formation et l'expérience nécessaires dans les domaines suivants : SIG, analyse d'images, statistiques, modélisation des récoltes, agrométéorologie, pédologie et phytotechnie. Les coûts engendrés seront en majorité consacrés à ces ressources humaines.

Actuellement, des images à haute, moyenne et basse résolution sont disponibles gratuitement au format brut et sous forme de produits dérivés, tels qu'images corrigées géométriquement (moyenne quadratique de 1,5 pixel)

et radiométriquement (sommet de l'atmosphère), indices de végétation, mosaïques régionales ou nationales et couverture périodique par ciel dégagé. Grâce aux initiatives de l'Agence américaine d'observation océanique et atmosphérique (NOAA), de l'Institut d'études géologiques des États-Unis (USGS) et de l'Agence spatiale européenne (ESA), d'importants sites de stockage accessibles gratuitement et en temps réel permettent de télécharger ou de traiter en ligne ce que la Commission de statistique des Nations Unies considère comme le défi des mégadonnées (2015). Les principaux sites Internet concernés sont les suivants :

- Google Earth Engine (GEE), disponible à l'adresse : <https://earthengine.google.com/>, qui fournit des données en temps réel tirées de MODIS, Landsat et Sentinel-1, -2 et -3, ainsi que des images d'archive très haute résolution de Globeview/Airbus ;
- Earth Explorer (USGS), disponible à l'adresse : <http://earthexplorer.usgs.gov/>, qui fournit des données Landsat/MODIS ;
- WELD (USGS), disponible à l'adresse : <https://landsat.usgs.gov/WELD.php>, qui fournit des données d'archive Landsat 7 et 5 ;
- l'Australian Geoscience Datacube (AGDC), disponible à l'adresse : <http://www.datacube.org.au/>, qui ne couvre que l'Australie ;
- la plateforme de données de l'ESA, disponible à l'adresse : <https://scihub.copernicus.eu/>, à partir de laquelle peuvent être téléchargées des données de Sentinel-1, -2 et -3 ;
- la base de données SPIRITS du JRC, disponible à l'adresse : <http://spirits.jrc.ec.europa.eu/download/downloaddata/>, qui renferme des données agrométéorologiques mondiales sur dix jours provenant du Centre européen pour les prévisions météorologiques à moyen terme (CEPMMT) et des bases de données CHIRPS (*Climate Hazards Group Infrared Precipitation with Station*) et TAMSAT (*Tropical Applications of Meteorology using SATellite data and ground-based observations*) ;
- les indicateurs FAO-SMIAR tirés du capteur AVHRR du satellite METOP (en particulier l'indice de stress agricole [ISA]), disponibles à l'adresse : http://www.fao.org/giews/earthobservation/asis/index_1.jsp?lang=fr.

Il n'est nécessaire d'acquérir que des images THR présentant une distance d'échantillonnage au sol inférieure à 5 m. Les offres des trois prestataires suivants sont fondées sur une constellation de satellites : RapidEye (quatre bandes, résolution de 5 m, 1 dollar US/km²), AIRBUS-SPOT (cinq bandes, résolution comprise entre 2 et 8 m, 3,5 dollars US/km²) et WorldView (huit bandes, résolution de 31 cm, 20 dollars US/km²). Par ailleurs, les nouveaux acteurs qui font leur apparition sur le marché depuis 2010 (p. ex., Terra Bella, Planet, Satellogic, Blacksky Global) proposent des flottes de « microsattelites » d'une résolution multibandes de 1 à 2 m ; ils n'ont toutefois pas réussi à réduire suffisamment leurs prix pour être compétitifs.

Le lecteur souhaitant une description plus détaillée des caractéristiques et des prix des capteurs peut consulter l'analyse complète proposée par Delincé (2015, chapitre 6).

9.2. DOMAINES D'APPLICATION ET GAINS RELATIFS

La télédétection pouvant contribuer de différentes façons aux statistiques agricoles, nous tâcherons dans un premier temps de décrire brièvement les différents cas de figure.

9.2.1. Optimisation des plans d'échantillonnage

La télédétection sert en premier lieu à améliorer le plan d'échantillonnage des recensements ou enquêtes agricoles. Cette utilisation est également la moins onéreuse. La démarche intuitive consiste dans ce cas à utiliser des images récentes (Stratégie Mondiale, 2015), voire des cartes de couverture du sol tirées d'images (Waldner *et al.*, 2015). Au début, les données obtenues par télédétection permettront au minimum d'élaborer des cartes de référence à partir desquelles les frontières administratives pourront être précisément localisées.

- Dans le cadre d'un recensement, le travail commence généralement avec l'élaboration d'une carte de prérecensement (les districts de recensement [DR] constituant ainsi les unités primaires d'échantillonnage [UPE]) et l'identification, sur le terrain, du nombre de ménages ou d'exploitations par DR échantillonné (ceux-ci représentant alors les unités secondaires d'échantillonnage [USE]). Le recours à l'imagerie (d'une résolution de 0,5 à 2 m) dans un environnement SIG permet de définir efficacement les DR : le territoire est ainsi sous-divisé en espaces délimités par des frontières physiques correspondant à 50-100 exploitations, de telle façon qu'un seul agent recenseur pourra collecter les données relatives à plusieurs DR pendant la période de recensement (Geospace, 2007). Des projets menés au Lesotho, en Namibie, en République-Unie de Tanzanie et aux Seychelles (Loots, 2015) ont montré que, les gains importants réalisés — en matière de temps et de travail — compensaient largement les coûts liés à l'adoption des technologies requises (imagerie, SIG, consultation d'experts et formation). La qualité globale du recensement s'en est trouvée également renforcée, grâce à une meilleure planification et à la transparence et la traçabilité accrues du travail effectué ; toute l'infrastructure ainsi créée peut par ailleurs facilement être réutilisée à un stade ultérieur, ce qui est particulièrement intéressant dans le cas du recours aux des approches basées sur les Bases de Sondage Principales (Stratégie Mondiale, 2015b).
- Deux principales possibilités s'offrent aux agents effectuant une enquête régionale : 1) définir quatre ou cinq strates aussi homogènes que possible au sein de chaque groupe (homogénéité intra), mais aussi différentes que possible les unes des autres (hétérogénéité inter) ; ou 2) définir des UPE dans un plan à deux degrés qui soient aussi similaires que possible les unes des autres, mais qui présentent une homogénéité intra minimale. La complexité peut être renforcée grâce à un échantillonnage avec probabilité proportionnelle à la taille (PPT) ou à l'ajout d'étapes intermédiaires. Quelle que soit la base de sondage principale sélectionnée (liste, district, point ou multiple), l'imagerie peut largement contribuer à atteindre les objectifs visés, puisqu'elle fournit des informations détaillées et à jour adaptées à un traitement numérique moderne. En outre, même une classification approximative des images permettra d'obtenir un substitut pour calculer l'intensité des cultures, qui pourra servir à optimiser la taille de l'unité d'échantillonnage (à partir de la corrélation spatiale) et des taux de sondage variables (à partir du ratio entre la variance de l'échantillonnage et l'intensité agricole ; voir Benedetti *et al.*, 2015).
- Examinons tout d'abord la base de sondage aréolaire (BSA) élaborée dans le cadre de l'enquête agricole de juin (JAS, de l'anglais *June Agricultural Survey*) du NASS de l'USDA. Elle se compose traditionnellement d'UPE stratifiées, qui comportent elles-mêmes des USE appelées « segments » (Cotter *et al.*, 2010). Dans un premier temps, l'ensemble du territoire est divisé en UPE délimitées par des frontières physiques grâce à une interprétation visuelle des images satellite (chaque UPE étant elle-même composée de six à huit segments). Les UPE sont ensuite stratifiées par intensité des classes agricoles à l'aide de la photo-interprétation. Enfin, une fois que des échantillons en ont été extraits, les UPE qui ont été sélectionnées sont sous-divisées en segments (de taille variant entre 2,5 et 21 km²). Un segment est sélectionné de façon aléatoire par UPE échantillonnée. L'échantillon final du JAS comprend donc environ 11 000 segments (le taux de sondage étant de 0,2 %). Cette approche, baptisée « traditionnelle » par Boryan *et al.* (2016), nécessite au total 30 mois-personne par État. Le NASS a récemment adopté une méthode automatisée fondée sur la délimitation et la stratification des UPE par classification automatique des couches de données sur les cultures de l'année précédente (méthode CDL, de l'anglais *Crop Data Layers*). Cette approche n'exige donc plus que 12 mois-personne par État, réduisant ainsi les coûts par 2,5. Partant de l'enquête de 2013 sur le Dakota du Sud (qui portait sur 578 segments couvrant une superficie totale de 199 730 km²), Boryan *et al.* ont comparé la base de sondage traditionnelle du JAS à la base automatisée, dont les données ont été simulées de deux façons différentes : en extrayant l'utilisation des terres du segment de la base de données sur les unités de terre communes déclarées auprès de l'Agence de services aux exploitations agricoles (UTC-FSA), ou en déterminant l'utilisation des terres à partir des CDL 2013 (UTC-CDL).

TABLEAU 1. COMPARAISON DES ESTIMATIONS DES CULTURES PAR STRATIFICATION TRADITIONNELLE (ST) ET STRATIFICATION AUTOMATIQUE (SA), DAKOTA DU SUD, 2013.

Culture	Estimation du NASS (en millions d'hectares)	Différence de superficie entre ST et SA de l'UTC-FSA (%)	CV relatif à la ST du JAS (%)	CV relatif à la SA de l'UTC-FSA (%)	CV relatif à la SA de l'UTC-CDL (%)
Maïs	6 200 000	6,74	4,9	4,3	3,9
Soja	4 600 000	19,03	5	6,0	5,2
Blé de printemps	1 190 000	47,10	17,6	14,3	13,3
Blé d'hiver	1 400 000	46,34	17,9	16,8	19

Boryan *et al.*, 2016

Dans les deux cas, la variance d'échantillonnage était comparable. Rappelons cependant que le coût d'élaboration de la base de sondage a été divisé par 2,5.

- Comparons ces résultats aux coûts engendrés dans les pays en développement lorsqu'ils utilisent la méthodologie traditionnelle du NASS. Par exemple, le Maroc a renouvelé sa base aréolaire sur 66 900 km² à l'aide d'images SPOT-5 d'une résolution de 2,5 m. Un échantillon de 810 segments mesurant entre 4 et 30 ha (pour un taux de sondage de 0,1 %) répartis dans 18 provinces (d'une superficie totale de 49 000 km²) était disponible, menant à une précision relative de 1,8 % pour les céréales (1 950 000 ha), de 4,1 % pour les terres en jachère (650 000 ha) et de 16,3 % pour les pâturages (120 000 ha). Relevons tout d'abord que, compte tenu de son taux de sondage plus faible, l'enquête menée au Maroc présente des coefficients de variation deux fois plus faibles que ceux du NASS pour la même superficie cultivée. La résolution des images utilisées par le NASS varie entre 20 et 56 m, contre 2,5 m au Maroc. L'efficacité totale de la stratification pour les céréales s'élève à 3 au Maroc. Le type d'imagerie choisi peut donc être un facteur important à prendre en compte lors de l'élaboration d'une base de sondage. En ce qui concerne les coûts, la réalisation d'une base de sondage stratifiée avec création de cartes sur l'utilisation des terres a coûté 7 dollars US/km² (50 % du coût étant consacrés à l'achat d'images SPOT, 15 % à la photo-interprétation et 35 % à la validation sur le terrain). Les coûts d'extraction d'échantillons étaient quant à eux mineurs, le logiciel SIG nécessaire coûtant environ 20 000 dollars US et le temps requis s'élevant à 10 jours par province.
- Procédons à une autre comparaison, cette fois avec la base de sondage utilisée dans dix provinces chinoises (soit 1 652 083 km²) par le Bureau national des statistiques de Chine, à l'aide d'un plan d'échantillonnage stratifié à deux degrés avec probabilité proportionnelle à la taille (PPT) dans un premier temps (la taille de chaque UPE variant entre 1 et 5 km²) et sélection aléatoire dans un deuxième temps (chaque USE mesurant entre 2 et 5 ha et le taux de sondage total s'élevant à environ 0,2 %). Pour la province d'Anhui (139 400 km²), un échantillon constitué de 6 000 segments entraîne un coefficient de variation de 1,3 % pour le blé (2 200 000 ha), de 0,9 % pour le riz cultivé pendant la campagne intermédiaire (1 900 000 ha) et de 3 % pour le maïs (1 000 000 ha). Le bon niveau de précision obtenu résulte de la stratification et de l'échantillonnage avec PPT, fondé sur la classification des images des satellites chinois GF1 et ZY3 (résolution de 2 m). Les coûts associés s'élèvent à 75 000 dollars US par province, soit environ 0,5 dollar US/km².
- La stratification peut également s'appuyer sur des bases de sondage par points, élaborées à partir des images satellite disponibles :
 - L'enquête aréolaire sur l'utilisation et l'occupation des sols (LUCAS), qui existe depuis 2001 et couvre l'ensemble du territoire européen, permet d'évaluer, en comparant le plan systématique d'origine au plan stratifié utilisé en 2006, les avantages de la stratification effectuée à partir d'images satellite (Gallego, 2007). En ce qui concerne les céréales produites dans l'UE, les gains relatifs en matière de rentabilité (résultant de la stratification elle-même et des différents taux de sondage) s'élevaient à 1,76 ; pour le blé commun et l'orge, à 1,64 ; pour le maïs, à 1,70 ; pour la pomme de terre, à 1,44 ; pour la betterave à sucre, à 1,60 ; et pour le tournesol, à 2,01. Sachant que la stratification était effectuée

à partir d'images disponibles gratuitement (orthophotographies aériennes créées pour le Système d'identification des parcelles agricoles [SIPA] du Système intégré de gestion et de contrôle [SIGC]), et que le photo-interprète pouvait analyser 500 points par jour environ (1 million pour l'ensemble du territoire de l'UE, qui comptait à l'époque 25 États membres), le coût de la stratification s'élevait donc aux alentours de 200 000 dollars US pour une enquête, pour un coût annuel d'environ 3,5 millions d'euros. Même si l'exploitation pluriannuelle de la base de sondage n'est pas envisagée, les 5 % de dépenses du projet consacrées à l'analyse des images sont largement compensés par la réduction des variances, qui vont de 30 % pour la pomme de terre à 50 % pour le tournesol.

- Haïti constitue une autre étude de cas pour laquelle il a été possible d'évaluer les coûts et avantages liés à l'utilisation de photographies aériennes d'archives d'une résolution de 50 cm pour élaborer une base de sondage constituée de points. Près d'1,7 million de points situés sur une grille de 125 m ont été photo-interprétés (en l'espace de 3 400 jours ouvrables), ce qui a débouché sur la création d'un échantillon stratifié de 25 000 points (représentant chacun près de 1 km²). Cette base de sondage étant conçue pour être utilisée pendant 10 ans, le coût supplémentaire lié à l'utilisation d'une stratification fondée sur l'imagerie s'élevait à 2 %. Cette méthode a toutefois favorisé une baisse de la variance de l'ordre de 8 à 38 % pour les grandes cultures, illustrant ainsi la rentabilité de la télédétection pour la stratification en Haïti.

9.2.2. Couches de données sur les cultures (CDL)

Les CDL ont été créées par le NASS de l'USDA il y a 20 ans. Décrites comme coûtant 75 000 dollars US par État en 1997, elles se sont énormément améliorées avec le temps, aussi bien en ce qui concerne le nombre de catégories de cultures que la précision des cartes, les outils de diffusion et les coûts d'élaboration. Muller et Harris (2013) décrivent en détail différents aspects du portail CropScape de 2001 : couverture (48 États), aspects techniques (imagerie et données de référence sur le terrain), liste des produits proposés et communauté d'utilisateurs. Environ 5 000 scènes Landsat, Resourcesat et DMC (disponibles gratuitement ou à moindre coût – moins de 0,10 dollar US/km²) sont classées chaque année, puis organisées sous forme de CDL (échelle de 1:100 000) et publiées de façon provisoire en août et définitive en février, après la récolte. Les images sont traitées et intégrées à partir des logiciels ERDAS, See5, ArcGIS et SAS. CropSpace a été développé à l'extérieur (Yang, 2014) en moins de 15 mois-personne. Il s'exécute sur deux serveurs (respectivement 32 GB et 1 TB d'espace disque) et fonctionne sur les logiciels Apache, Tomcat, Mapserver, GDAL et PostgreSQL. Bien que le NASS ne communique que peu d'informations sur les coûts d'exploitation de son projet, il précise tout de même que ce bien public satisfait les besoins du secteur agroalimentaire privé (et plus particulièrement ceux liés aux services financiers, d'assurance et d'aide à la prise de décision).

9.2.3. Estimateurs améliorés

En ce qui concerne les estimations, la fusion des données issues des enquêtes de terrain et des satellites est généralement effectuée par le biais d'estimateurs de régression ou de calage (Stratégie mondiale, 2015a). Gallego *et al.* (2014) présentent des résultats détaillés portant sur une région du nord de l'Ukraine (78 500 km²) comptant 2,45 millions d'hectares de terres cultivées. Quatre-vingt-dix segments carrés de 4 km x 4 km ont fait l'objet d'un lever de terrain en 2010 (taux de sondage de 1,8 %, taille des parcelles allant jusqu'à 250 ha). La région tout entière a par la suite été couverte par les satellites MODIS, Landsat 5, AWiFS, LISSIII et RapidEye. La classification des images a été entraînée sur des données collectées le long d'une route, indépendamment des segments aréolaires. La rentabilité moyenne des grandes cultures (blé, orge, maïs et soja) s'élevait respectivement à 1,59, 1,54, 1,48 et 1,50, la performance des capteurs étant par conséquent relativement similaire. La comparaison des coûts du lever de terrain et de l'imagerie (le coût de la classification des images est si négligeable à l'heure actuelle qu'il peut être écarté) donne lieu à une situation tout autre, la rentabilité n'étant assurée qu'avec les deux capteurs gratuits (MODIS et Landsat TM), le prix d'achat des trois autres étant trop élevé (AWiFS 0,92, LISSIII 0,43 et RapidEye 0,18). Replaçons toutefois l'étude dans son contexte. Premièrement, les parcelles en Ukraine sont généralement étendues ; les capteurs présentant une résolution grossière permettent donc d'effectuer une classification aussi précise que les capteurs présentant une résolution plus fine. Ce résultat serait impossible dans la plupart des pays africains ou asiatiques, où les parcelles sont généralement plus petites. Deuxièmement, cette étude était uniquement fondée sur la méthode

de la classification par maximum de vraisemblance. L'USDA a aujourd'hui recours à la méthode de classification par arbre de décision (See5); l'ESA, après avoir testé plusieurs algorithmes (machines à vecteurs de support, arbres de décision, boosting d'arbres et forêts aléatoires) sur 12 sites partout dans le monde (Inglada *et al.*, 2015) compte appliquer un algorithme de forêt aléatoire à ses services de transmission de données Sentinel-2. Enfin, de plus en plus d'images sont disponibles en accès libre : non seulement les images tirées des satellites MODIS ou Landsat 8 sont des produits rectifiés ou classés (résolution de 250 à 500 m et de 15 à 30 m, respectivement) téléchargeables gratuitement en temps quasi réel, mais les produits Sentinel-1 (SAR-GRD, résolution de 9 m), Sentinel-2 (résolution de 10 m) et Sentinel-3 (résolution de 300 m) sont désormais accessibles sur la plateforme de l'ESA et Google Earth Engine.

9.2.4. Suivi des cultures et prévision des récoltes

Le suivi des cultures et la prévision des récoltes sont généralement effectués dans le cadre d'un système connu de modélisation des cultures, tel que DSSAT¹, BIOMA², APSIM³, ORYZA⁴, STICS⁵, CERES⁶, CROPSYST⁷ ou EPIC⁸, qui s'inscrivent pour la plupart dans le projet AGMIP⁹ (Rosenzweig *et al.*, 2013). Le choix du modèle dépend principalement des cultures étudiées, de la région concernée et de la nature du modèle – en libre accès ou non. Tous les modèles dépendent dans tous les cas des informations suivantes :

- Frontières administratives (GADM, <http://www.gadm.org/>);
- Zonage agro-écologique (GAEZ, <http://www.fao.org/nr/gaez/fr/>);
- Masques des terres cultivées (base de données sur la sécurité alimentaire mondiale, résolution de 1 km : <http://e4ftl01.cr.usgs.gov/provisional/MEaSURES/GFSAD/>; couche de données unifiée sur les terres cultivées, résolution de 250 m : <http://maps.elie.ucl.ac.be/geoportail/>);
- Paramètres des cultures (ECOCROP, <http://ecocrop.fao.org/ecocrop/srv/en/home>);
- Base de données sur les sols (<http://www.fao.org/soils-portal/soil-survey/cartes-historiques-et-bases-de-donnees-des-sols/base-harmonisee-mondiale-de-donnees-sur-les-sols-version-12/fr/>);
- Bases de données agrométéorologiques en temps réel et archives (températures [GSOD] : <https://data.noaa.gov/dataset/dataset/global-surface-summary-of-the-day-gsod>; rayonnement [MERRA2] : http://disc.sci.gsfc.nasa.gov/datareleases/merra_2_data_release; précipitations : [CHIRPS] <http://chg.geog.ucsb.edu/data/chirps/>, [SPIRITS] <http://spirits.jrc.ec.europa.eu/download/downloaddata/>, [FEWS] <http://earlywarning.usgs.gov/fews> ou <http://www.isac.cnr.it/~ipwg/data/datasets.html>);
- Indices de végétation en temps réel et archives (METOP, <http://www.eumetsat.int/website/home/Data/index.html>; MODIS, <http://modis.gsfc.nasa.gov/data/dataproduct/mod13.php>; ou SEN3/VEG, <http://land.copernicus.eu/global/products/ndvi>); et
- Pratiques de gestion locales (voir le calendrier cultural de l'AMIS, <http://www.amis-outlook.org/>, ou <http://www.usda.gov/oce/weather/pubs/Other/MWCACP/>).

Étant donné que la plupart des produits spatiaux utilisés pour le suivi des récoltes sont disponibles gratuitement, les frais engagés découleront principalement des coûts d'exploitation du système de suivi lui-même et de l'estimation de la superficie cultivée.

Prenons pour exemple le Centre national Mahalanobis de prévision des récoltes de l'Inde (<http://www.ncfc.gov.in/>) : la publication de prévisions périodiques portant sur huit cultures nécessite un budget annuel de 1,7 million de dollars US, qui couvre les coûts afférents aux bureaux, les salaires (de 31 employés), les levers de terrain (10 % du

1 Modèle « Decision Support System for Agrotechnology Transfer ».

2 Modèle « Biophysical Models Applications ».

3 Modèle « Agricultural Production Systems sIMulator ».

4 Plus d'informations à l'adresse : <http://oryza.com/>.

5 Simulateur mulTidisciplinaire pour les Cultures Standard.

6 Modèle « Crop Environment Resource Synthesis ».

7 Modèle « Cropping Systems Simulation ».

8 Modèle « Environmental Policy Integrated Climate ».

9 Projet « Agricultural Model Intercomparison and Improvement ».

budget total), l'imagerie (20 % du budget total), les équipements (19 stations de travail) et les logiciels (licences ERDAS, ARCGIS, GEOMATICA STAT).

Autre exemple : le budget annuel consacré par l'Institut de télédétection et Terre numérique (RADI) de Chine au système CROPWATCH (<http://www.cropwatch.com/cn/htm/en/index.shtml>), dédié au suivi des cultures dans les différentes régions chinoises et dans les principales zones de production du monde (31 pays assurant 80 % de la production mondiale de maïs, de blé, de riz et de soja), s'élève à 1,5 million de dollars US. Ce budget est ventilé de la façon suivante : ressources humaines (soit 15 personnes) : environ 35 % ; collecte de données d'imagerie : 20 % seulement, grâce au partage interinstitutionnel des données ; location des ordinateurs, réseaux et logiciels : à peu près 10 % ; missions sur le terrain, recherche, coûts des réunions et déplacements, logistique, etc. : 35 % environ.

9.3. PERTINENCE DES CAPTEURS

Plusieurs exemples permettent de confirmer la rentabilité de la télédétection aux fins de statistiques agricoles. Il convient de vérifier si les satellites disponibles sont adaptés à la taille la plus fréquente des parcelles agricoles dans les différentes régions du monde. Pour répondre à cette question, nous avons retravaillé les données publiées par Fritz (2015). Ces données ont été tirées d'une carte des terres cultivées d'une résolution de 1 km² (terres arables, cultures permanentes et prairies permanentes) divisée en cinq catégories : absence de terres cultivées, terres cultivées sur des parcelles très petites, petites, moyennes et grandes. Pour les terres cultivées, la base de données IIASA, téléchargeable gratuitement, présente une carte interpolée de la taille des parcelles d'une résolution de 1 km². Il est donc possible, à l'aide d'un masque des pays, de créer une base de données indiquant le nombre total de pixels par pays (superficie des pays en km²), la superficie cultivée totale et la superficie occupée par chacune des quatre classes de taille des parcelles. Ces catégories respectant la recommandation de GEOGLAM relative au suivi satellitaire de l'agriculture (voir Waldner *et al.*, 2015, tableau 4), il est facile de calculer la superficie agricole de chaque pays pouvant être suivie par satellite, à une distance d'échantillonnage au sol de 100 à 500 m (MODIS, Sentinel-3), de 20 à 100 m (Landsat 5, AWiFS, DMC, DEIMOS), de 5 à 20 m (Sentinel-1 et 2), ou de moins de 5 m (SPOT, RapidEye, LISS, WorldView).

Le suivi agricole présente des facettes variées (superficie cultivée, superficie par type de culture, ou suivi des récoltes à l'échelle des régions ou des parcelles). Cette étude portera plus spécifiquement sur la superficie régionale occupée par les principales catégories de cultures (telles que les céréales), d'autres terres arables, les terres cultivées permanentes et les prairies permanentes.

TABEAU 2. SUPERFICIE (EN MILLION D'HECTARES) OCCUPÉE PAR LES DIFFÉRENTES CLASSES DE TAILLE DES PARCELLES, VENTILÉE PAR RÉGION.

Région	Terres cultivées	Très petite parcelle	Petite parcelle	Parcelle de taille moyenne	Grande parcelle
Afrique	773,0	242,6	394,1	110,4	25,9
Moyen-Orient	107,8	9,0	70,0	26,6	2,2
Asie	1 411,6	472,5	673,1	179,3	86,7
Amérique centrale et du Sud	665,8	21,7	154,5	295,9	193,7
Europe	1 165,9	14,0	281,2	532,8	337,9
Amérique du Nord	856,7	1,9	68,0	454,5	332,3
Océanie	130,5	0,0	9,8	34,3	86,4
Monde	5 111,3	761,6	1 650,8	1 633,9	1 065,0

Comme le montre le tableau 2, la grande majorité des terres agricoles relève des catégories «petite parcelle» et «parcelle de taille moyenne». La catégorie «très petite parcelle» ne représente que 15 % des terres cultivées à l'échelle mondiale ; l'Asie et l'Afrique sont moins favorisées de ce point de vue, un tiers des surfaces agricoles de ces deux régions entrant dans cette catégorie.

On peut déduire des informations ci-dessus que les capteurs MODIS et Sentinel-3 sont adaptés au suivi des surfaces de 21 % des terres agricoles de la planète, en particulier en Océanie et en Amérique du Nord. Ils sont en revanche inutilisables en Afrique et en Asie. Les satellites à résolution moyenne, tels que Landsat 8, permettent de procéder au suivi de la moitié des surfaces agricoles du monde ; l'Asie et l'Afrique sont encore une fois désavantagées, puisque moins de 20 % de leurs zones agricoles peuvent être couvertes par ces capteurs. L'arrivée des satellites Sentinel-1 (radar) et -2 (optique) a considérablement amélioré la situation. Leurs capteurs sont adaptés à plus de 90 % pour cinq des régions du monde, et à près de 70 % pour l'Afrique et l'Asie. Les satellites THR se montrent quant à eux totalement adaptés, mais ils sont beaucoup trop onéreux pour l'ensemble des systèmes de statistique examinés.

TABLEAU 3. RÉOLUTION DES SATELLITES ET POURCENTAGE RELATIF DE COMPATIBILITÉ AVEC LES SUPERFICIES CULTIVÉES.

Région	Spot/RapidEye	Sentinel-1 et 2	Landsat 8/AWiFS	MODIS/Sentinel-3
Afrique	100	69	18	3
Moyen-Orient	100	92	27	2
Asie	100	67	19	6
Amérique centrale et du Sud	100	97	74	29
Europe	100	99	75	29
Amérique du Nord	100	100	92	39
Océanie	100	100	92	66
Monde	100	85	53	21

Il est intéressant de souligner que la résolution et le coût des images ne sont pas les seuls facteurs restrictifs. Dans les zones tropicales, la couverture nuageuse peut sérieusement compromettre les taux ci-dessus, sauf dans les cas de Sentinel-1 (haute résolution) et de RISAT 1 (résolution moyenne). Whitcraft *et al.* (2015) fournissent une analyse détaillée des tendances de la couverture nuageuse pendant la période de végétation agricole, et montrent que la probabilité d'obtention d'images dégagées pose encore problème aujourd'hui, particulièrement en Inde et en Asie du Sud-Est.

En outre, les satellites THR n'ont pas vocation à couvrir l'ensemble du territoire national, et les « mosaïques de timbres-poste » qu'ils produisent sont très coûteuses et leur traitement très chronophage. Outre sa pertinence, Sentinel-2 se caractérise par une fauchée de 290 km et une fréquence de revisite de cinq jours, contre 170 km et 16 jours pour Landsat 8 ; en d'autres termes, ce dernier suppose de mosaïquer trois fois plus d'images que Sentinel-2.

Le tableau 4 classe les pays d'Afrique et d'Asie par niveau de difficulté d'utilisation de l'imagerie satellite aux fins de statistiques agricoles, en fonction du pourcentage de terres cultivées par classe de taille de parcelle.

TABLEAU 4. CLASSIFICATION DES PAYS D'AFRIQUE ET D'ASIE PAR NIVEAU DE DIFFICULTÉ D'UTILISATION DE L'IMAGERIE SATELLITE AUX FINS DE STATISTIQUES AGRICOLES, EN FONCTION DU POURCENTAGE DE TERRES CULTIVÉES, PAR CLASSE DE TAILLE DE PARCELLE.

AFRIQUE					ASIE				
Liste des pays se prêtant le mieux à l'utilisation de l'imagerie satellite aux fins de statistiques agricoles					Liste des pays se prêtant le mieux à l'utilisation de l'imagerie satellite aux fins de statistiques agricoles				
	Très petite parcelle	Petite parcelle	Parcelle de taille moyenne	Grande parcelle		Très petite parcelle	Petite parcelle	Parcelle de taille moyenne	Grande parcelle
Burundi	99,8	0,2	0,0	0,0	Sri Lanka	99,8	0,2	0,0	0,0
Madagascar	86,3	13,4	0,3	0,0	Népal	86,3	13,4	0,3	0,0
Érythrée	79,6	18,2	2,2	0,0	Bangladesh	79,6	18,2	2,2	0,0
Togo	79,2	19,9	0,8	0,0	Inde	79,2	19,9	0,8	0,0
Rwanda	76,3	21,4	2,2	0,0	République de Corée	76,3	21,4	2,2	0,0
Guinée équatoriale	73,1	26,9	0,0	0,0	Pakistan	73,1	26,9	0,0	0,0
Mauritanie	70,9	28,4	0,7	0,0	Rép. dém. populaire lao	70,9	28,4	0,7	0,0
Éthiopie	69,7	29,2	1,0	0,1	Chine	69,7	29,2	1,0	0,1
Guinée-Bissau	65,5	34,5	0,0	0,0	Myanmar	65,5	34,5	0,0	0,0
Ouganda	65,1	32,8	2,1	0,0	Rép. pop. dém. de Corée	65,1	32,8	2,1	0,0
Liste des pays se prêtant le moins bien à l'utilisation de l'imagerie satellite aux fins de statistiques agricoles					Liste des pays se prêtant le moins bien à l'utilisation de l'imagerie satellite aux fins de statistiques agricoles				
	Très petite parcelle	Petite parcelle	Parcelle de taille moyenne	Grande parcelle		Très petite parcelle	Petite parcelle	Parcelle de taille moyenne	Grande parcelle
Gambie	5,1	91,9	2,4	0,7	Taïwan	10,1	89,9	0,0	0,0
Sénégal	16,7	82,6	0,7	0,0	Arménie	7,5	88,8	3,7	0,0
Somalie	7,1	79,3	13,5	0,0	Bhoutan	14,0	86,0	0,0	0,0
Angola	10,3	78,5	10,4	0,8	Timor-este	17,6	82,4	0,0	0,0
Egypte	20,3	77,6	2,1	0,0	Cambodge	15,8	80,8	3,4	0,0
Zimbabwe	10,2	76,8	12,7	0,3	Thaïlande	12,2	79,2	8,5	0,0
Mali	21,0	76,5	2,5	0,0	Jammu-et-Cachemire	19,9	78,7	1,4	0,0
Guinée	23,8	74,9	1,3	0,0	Viet Nam	16,2	76,6	6,6	0,5
Sierra Léone	24,3	74,2	1,6	0,0	Indonésie	19,2	69,6	10,9	0,3
Lesotho	9,8	73,9	16,4	0,0	Philippines	14,4	66,0	17,6	2,0

9.4. CONCLUSION

Trois facteurs principaux garantissent la rentabilité de la télédétection aux fins de statistiques agricoles : la baisse du prix des images due à la mise en place de systèmes gratuits à long terme par la NASA et l'ESA, présentant les résolutions nécessaires au suivi du rendement des cultures (METOP, MODIS, Sentinel-3) et à l'estimation de la superficie (Landsat 8, Sentinel-1 et 2) ; la hausse de la qualité en matière de disponibilité (garantie à long terme), de résolution de l'image (jusqu'à 10 m), de taille des unités composant la base de sondage (jusqu'à 290 m x 290 m), de temps de revisite (jusqu'à 5 jours) et de nombre de canaux radiométriques (plus de 10) ; et les applications en libre accès — SIG, systèmes d'analyse des images et outils dématérialisés (cloud computing) de traitement à distance (équipements et logiciels, tels que Google Earth Engine) — sont aujourd'hui devenues la norme.

Le meilleur moyen de procéder au suivi des rendements des cultures, essentiel à la sécurité alimentaire et à la gestion des marchés, passe par des systèmes de télédétection qui communiquent régulièrement et en temps voulu sur l'évolution des récoltes, tout en réduisant au maximum la charge de travail sur le terrain, qui coûte cher. Plusieurs systèmes existent de longue date à l'échelle nationale (CONAB au Brésil, NCFC en Inde, DMN au Maroc, DSNM au Mozambique, SUPARCO au Pakistan, VEGA-PRO en Russie, CSE au Sénégal et CNT en Tunisie) et mondiale (CropWatch, CropExplorer, JRC-MARS, FEWS-NET, FAO-SMIAR, PNUE-DEWA, PAM-VAM, AMIS et GEOGLAM).

L'estimation des superficies cultivées à partir de la télédétection se révèle également rentable. CropScape (États-Unis), AGRIFOOD (Canada), le NCFC (Inde) et SUPARCO (Pakistan) utilisent des images de la campagne en cours. Si la taille des parcelles demeure un facteur restrictif dans 70 pays, au moins 125 pays ont la possibilité de se tourner vers la télédétection pour estimer les superficies cultivées au cours de la campagne. Par ailleurs, la plupart des bureaux nationaux de statistique utilisent des images d'archive pour optimiser leurs plans d'échantillonnage, même lorsqu'ils ont recours à une approche par liste (pour les recensements et les enquêtes).

9.5. RÉFÉRENCES BIBLIOGRAPHIQUES

Benedetti, R., Piersimoni, F. et Postiglione, P. 2015. *Sampling Spatial Units for Agricultural Surveys*. Springer-Verlag : Berlin et Heidelberg.

Boryan, C.G., Yang, Z. et Seffrin, R. 2016. «Post stratification assessment of the NASS automated stratification method based on the Cropland Data Layer». In : Institute of Electrical and Electronics Engineers, Inc. (dir.), *Geoscience and Remote Sensing (IGARSS), 2016 IEEE International Symposium* (p. 5933–6). Publication de l'IEEE.

Carfagna, E. 2001. *Cost-effectiveness of Remote Sensing in Agricultural and Environmental Statistics*. Communication sollicitée pour la conférence *Caesar : Conference on Agricultural and Environmental Statistical Applications in Rome*. 4–8 juin 2001. Rome, Istituto nazionale di statistica (Istat). Volume III, p. 617–627.

Carfagna, E. 2013. *Evaluating the Cost-Efficiency of Remote Sensing in Developing Countries*. Présentation élaborée pour le premier Comité consultatif scientifique de la Stratégie mondiale – Improving AG-Statistics. 18–19 juillet 2013. Rome, FAO.

Commission de statistique des Nations Unies. 2015. *Rapport du Groupe de travail mondial des Nations Unies sur l'utilisation des mégadonnées en statistique officielle*. Rapport E/CN.3/2015/4. Publication du Conseil économique et social des Nations Unies : New York, États-Unis.

Cotter, J., Davies, C., Nealon, J. et Roberts, R. 2010. «Area Frame Design for Agricultural Surveys». In : Benedetti, R., Bee, M., Espa, G. et Piersimoni, F. (dir.), *Agricultural Survey Methods* (ch. 11, p. 169-92). John Wiley & Sons : Chichester, Royaume-Uni.

Delincé, J. 2015. *Technical Report on Cost-Effectiveness of Remote Sensing for Agricultural Statistics in Developing and Emerging Economies*. Technical Report Series GO-09-2015. Rapport technique de la Stratégie mondiale : Rome.

Fritz, F., See, L. et McCallum, I., You, L., Bun, A., Moltchanova, E. et al. 2015. « Mapping global cropland and field size ». *Global Change Biology*, vol. 21, n° 5, p. 1980–92.

Gallego, F.J. 2007. «*Sampling efficiency of the EU point survey LUCAS 2006*». Article présenté lors de la 56^e session de l'Institut international de statistique, 22-29 septembre 2007. Lisbonne.

Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A. et Kussul, O. 2014. «Efficiency assessment of using satellite data for crop area estimation in Ukraine». *International Journal of Applied Earth Observation and Geoinformation*, vol. 29, p. 22–30.

GeoSpace International. 2007. *Census mapping methodology using remote sensing and GIS technology*. Rapport élaboré pour le Bureau national de statistique des Seychelles. Publication de Geospace International : Pretoria, Afrique du Sud.

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre Canto, G., Bontemps, S., Defourny, P. et Koetz, B. 2015. «Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery ». *Remote Sensing*, vol. 7, n° 9, p. 12356–79.

Latham, J. 2017. «Organisation, ressources et compétences nécessaires à l'adoption de la télédétection dans le domaine des statistiques agricoles». In : J. Delincé (dir.), *Manuel sur l'application de la télédétection aux statistiques agricoles* (chapitre 8). Manuel de la Stratégie mondiale : Rome.

Loots, H. 2015. «The use of Hexagon's Smart Client for Census software for the demarcation of census enumeration areas for the 2016 Population and housing census in Lesotho ». Article préparé en vue de la conférence *Geomatics Indaba: Conference and exhibition of surveying, geospatial information, GIS, mapping, remote sensing and location-based business*, 11–13 août 2015. Gauteng, Afrique du Sud.

Miller, H.M., Richardson, L., Koontz, S.R., Loomis, J. et Koontz, L. 2012. *Users, uses and value of Landsat satellite imagery: results from the 2012 Survey of Users*. USGS Open-File Report 2013-1269. Publication de l'USGS : Washington.

Muller, R. et Harris, M. 2013. «Reported uses of CropScape and the National Cropland Data Layer Program». Article préparé pour la sixième Conférence internationale sur les statistiques de l'agriculture (ICAS VI), 23–25 octobre 2013. Rio de Janeiro, Brésil.

Nelson, G.C., Schimmelpfennig, D., Sumner, D. et Buck, J. 2007. «Can satellite-based land imaging data be made more valuable for agriculture?» Rapport destiné à l'USDA.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G. et Winter, J.M. 2013. «The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. » *Agricultural and Forest Meteorology*, vol. 170, p. 166–82.

Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS). 2015a. *Spatial Disaggregation and Small-Area Estimation Methods for Agricultural Surveys: Solutions and Perspectives*. Rapport technique GO-07-2015 de la GSARS : Rome.

Stratégie mondiale pour l'amélioration des statistiques agricoles et rurales (GSARS). 2015b. *Manuel sur les bases de sondage principales pour les statistiques agricoles – développement de la base de sondage, conception du plan d'échantillonnage et estimation*. Manuel de la GSARS : Rome.

Tenkorang, F. et Lowenberg-DeBoer, J. 2008. «On-farm profitability of remote sensing in agriculture.» *Journal of Terrestrial Observation*, vol. 1, n° 1, p. 50–9.

Waldner, F., Fritz, S., Di Gregorio, A. et Defourny, P. 2015. «Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps». *Remote Sensing*, vol. 7, n° 6, p. 7959–86.

Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I. et Justice, C.O. 2015. «Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations». *Remote Sensing of Environment*, vol. 156, p. 438–47.



Annexe

Estimation et correction des biais liés au comptage de pixels

Javier Gallego

A1. INTRODUCTION

Cette annexe s'adresse aux lecteurs cherchant à obtenir des informations plus détaillées sur les caractéristiques des principaux estimateurs présentés dans le chapitre 5. Elle propose également des références bibliographiques complémentaires, principalement destinées aux institutions envisageant de mettre en œuvre un système d'estimation non biaisée des surfaces cultivées fondé sur la télédétection.

La télédétection peut être utilisée à différentes fins dans le cadre des statistiques agricoles (Carfagna et Gallego, 2005). Nous n'étudierons ici son utilisation qu'au cours des dernières étapes de l'estimation des surfaces cultivées, en laissant notamment de côté son rôle dans l'élaboration de bases de sondage.

Lorsque la télédétection a fait son apparition, beaucoup croyaient que le fait de compter les pixels contenus dans les images satellite classées suffisait à obtenir des estimations précises des surfaces cultivées (McDonald et Hall, 1980). Le programme LACIE de l'USDA était initialement fondé sur le comptage de pixels contenus dans un échantillon de morceaux d'images de 6 x 5 milles nautiques. Suite à une analyse plus approfondie, cette méthode a été remplacée par la mise en place d'un estimateur par régression associant deux covariables, à savoir des données sur le terrain portant sur un échantillon de segments et des images satellite classées proposant une couverture complète (Hanuschak *et al.*, 1980). La rentabilité de la télédétection est devenue un objectif majeur, qui semblait hors de portée dans les années 1980 (Allen, 1990), mais fut finalement atteint pour de bon dans les années 1990 (Hanuschak *et al.*, 2001).

Le projet MARS de l'Union européenne, plus récent, est parvenu à la même conclusion en suivant une autre voie. En particulier, ses « estimations de modification rapide des surfaces cultivées » (répondant au nom d'Activité B ou d'Action 4) ont constitué une tentative de production d'estimations de la modification des surfaces cultivées à partir du comptage de pixels, sur un ensemble de 60 sites mesurant 40 × 40 km chacun. L'évaluation ultérieure de la méthode a précisé que la marge de subjectivité pouvait varier entre ± 10 % et ± 30 % pour les grandes cultures (Gallego, 2006).

Plusieurs articles traitant de la cartographie des cultures ou de la couverture du sol partent encore aujourd'hui de l'hypothèse selon laquelle les surfaces peuvent être estimées par simple comptage des pixels, sans analyse spécifique des propriétés de l'estimateur. L'utilisation d'images satellite classées aux fins d'estimation directe induit deux sources d'erreur principales : la présence de pixels mixtes (délimitations), et la classification erronée de pixels purs. Une erreur d'échantillonnage peut également se produire en cas de couverture non intégrale de la région étudiée (recours à un échantillon d'images).

Les effets des pixels mixtes sur les images à faible résolution ont fait l'objet de nombreuses études (Mayaux et Lambin, 1995 ; Verbeiren *et al.*, 2008 ; Wu, 2009 ; Atzberger *et al.*, 2013). Certaines recherches se sont également intéressées aux images à haute résolution (Chhikara, 1984). Moody et Woodcock (1994, 1996) ont examiné l'impact de l'échelle (taille des pixels) sur les estimateurs de surface par comptage de pixels, en simulant des classifications d'images à faible résolution à partir de classifications Landsat TM. Ils ont ainsi remarqué que le comptage de pixels avait tendance à surestimer les classes dominantes et à sous-estimer les classes secondaires, un effet qui peut être toutefois plus ou moins important en fonction des paysages ou algorithmes de classification utilisés. Waldner et Defourny (2017) ont quant à eux principalement axé leurs recherches sur le biais créé par les pixels mixtes, dû à la résolution de l'image et à certaines caractéristiques des capteurs, ainsi que sur l'indice de Matheron relatif à la complexité paysagère. Selon leur étude, le biais induit par l'utilisation d'images Landsat 8 peut toucher plus de 10 % de la surface cultivée pour une grande partie des paysages d'Afrique du Sud, même en supposant que la précision de la classification équivaut à 100 % pour les pixels purs. Cette situation s'améliore nettement avec les images du capteur Sentinel-2.

Les utilisateurs des technologies de télédétection peuvent croire dans un premier temps que plus la résolution des images est fine, plus la classification des images sera précise. Le problème est toutefois plus complexe : plusieurs études suggèrent en effet qu'une résolution élevée accentue la variabilité de la réponse radiométrique des pixels purs au sein d'une même classe, détériorant ainsi la séparabilité entre les classes, et donc la précision de la classification des pixels purs. À l'opposé, une résolution spatiale plus faible augmente la proportion de pixels mixtes. Le choix d'une résolution optimale constitue un problème délicat qui suppose de faire des compromis (Latty *et al.*, 1985 ; Cushnie, 1987 ; Treitz *et al.*, 1992 ; Atkinson et Curran, 1997 ; Hsieh *et al.*, 2001 ; Boschetti *et al.*, 2004 ; Duveiller et Defourny, 2010 ; Löw et Duveiller, 2014). Il pourrait sembler raisonnable de choisir une résolution présentant une grande majorité de pixels purs (90 %, par exemple), sans essayer d'atteindre une très grande majorité (99 %, par exemple). La validité de cette hypothèse n'a toutefois pas encore été confirmée.

A2. BIAIS LIÉ AU COMPTAGE DE PIXELS

Nous nous intéressons ici à l'impact des classifications erronées sur les estimations fondées sur le comptage de pixels, en partant de l'hypothèse que l'incidence des pixels mixtes et des données manquantes est dérisoire. Le problème du biais lié au comptage de pixels est bien connu dans le monde de la télédétection depuis la fin des années 1970 (Bauer *et al.*, 1978 ; Card, 1982 ; Hay, 1988 ; Czaplewski, 1992). Plusieurs méthodes ont été mises au point pour le corriger. Nous décrivons ici une expression du biais prenant la forme d'une matrice d'erreur (ou matrice de confusion), en insistant sur la nécessité de faire preuve de prudence au moment de la correction.

Le rapport qu'entretiennent une image classée et la couverture du sol réelle peut être exprimé sous forme de matrice de confusion. Considérons la matrice de confusion A pour l'ensemble des pixels. En pratique, A est inconnue et doit être estimée. Chaque cellule A_{gc} correspond à la superficie de la classe g (terrain, ou *ground* en anglais) classée dans la catégorie c , à savoir la part occupée par cette classe par rapport à la superficie totale de la région. A_{cc} correspond à la surface correctement classée dans la catégorie c . A_{+c} correspond à la surface totale classée dans la catégorie c , et A_{c+} à la surface totale inconnue appartenant à la catégorie c . L'estimateur par comptage de pixels utilise A_{+c} pour estimer A_{c+} . Il ne présente aucune variance, sauf dans les cas où il est calculé à partir d'un échantillon d'images. Le biais correspond à la différence entre les erreurs de commission et d'omission.

$$B = A_{+c} - A_{c+} = \sum_{c' \neq c} A_{c'c} - \sum_{c' \neq c} A_{cc'} = \Phi_c - \Psi_c \quad (2)$$

Le tableau A1 illustre les sommes à effectuer pour calculer l'erreur de commission Φ_c et l'erreur d'omission Ψ_c dans une matrice de confusion, exprimées sous forme de surfaces totales. La matrice de confusion est souvent exprimée, dans la littérature relative à la télédétection, sous forme de nombre d'unités d'échantillonnage ; ce choix s'avère bien souvent erroné, puisqu'il peut être fortement biaisé s'il est considéré comme un estimateur de A .

TABLEAU A1. MATRICE DE CONFUSION CONTENANT UNE CLASSE M. JAUNE : ERREUR DE COMMISSION POUR LA CLASSE 1 (BLÉ). ROSE : ERREUR D'OMISSION POUR LA MÊME CLASSE.

	Classification				Total
	1	2	M	
1 blé	A_{11}	A_{12}	A_{1M}	A_{1+}
2 maïs	A_{21}	A_{22}	A_{2M}	A_{2+}
....
M absence de culture	A_{M1}	A_{M2}	A_{MM}	A_{M+}
Total	A_{+1}	A_{+2}	A_{+M}	A_{++}

Il n'existe malheureusement aucun théorème relatif à un algorithme de classification permettant d'assurer que les erreurs de commission et d'omission tendront à se compenser. À noter que ces erreurs sont généralement exprimées sous forme de ratios, les erreurs Φ_c et Ψ_c étant respectivement divisées par A_{+c} et A_{c+} (Congalton et Green, 1999). Il devient alors légèrement plus compliqué d'exprimer le biais (Gallego, 2004). Il est préférable d'exprimer Φ_c et Ψ_c sous forme de superficies totales, afin de simplifier l'expression du biais.

En pratique, la matrice de confusion A est inconnue. Les totaux A_{+c} des colonnes sont connus, et il est donc possible d'estimer le reste de la matrice et les indicateurs en découlant, y compris Φ_c et Ψ_c . Si les estimateurs non biaisés Φ_c et ψ_c peuvent être calculés, il sera nécessaire d'obtenir un estimateur non biaisé du biais, qui pourra être corrigé.

Citons quelques règles de base devant être respectées à cette fin (pour plus de détails, voir par exemple Stehman [2009], Foody [2002] et McRoberts et Walters [2012]).

- Les données de validation utilisées pour calculer la matrice de confusion doivent être choisies à l'aide d'une méthode statistique (échantillonnage aléatoire stratifié, échantillonnage systématique avec origine aléatoire, etc.). Cette règle n'est pas toujours respectée dans le domaine de la télédétection en raison du coût élevé de la collecte de données sur le terrain ; elle est pourtant nécessaire au calcul des pondérations utilisées aux fins d'extrapolation, au risque d'obtenir une correction du biais totalement erronée (Gallego *et al.*, 2016).
- L'échantillon de validation ne doit pas comprendre les données servant à entraîner l'algorithme de classification. Le coût de la collecte de données sur le terrain incite souvent les praticiens à les y inclure, mais cette pratique est déconseillée.
- L'échantillon de validation et le jeu de données d'entraînement ne doivent pas être spatialement corrélés (Congalton et Green, 1999 ; Hammond et Verbyla, 1996 ; Zhen *et al.*, 2013). Par exemple, l'utilisation d'un système d'échantillonnage à deux degrés et la sélection de jeux de données d'entraînement et de validation dans une même UPE peuvent entraîner un biais optimiste important, notamment si l'algorithme de classification comprend un grand nombre de paramètres (Gallego et Rueda, 1993). En ce qui concerne les classifieurs par réseaux de neurones, le problème du surapprentissage à l'origine du biais optimiste a fait l'objet de nombreuses études dans d'autres domaines (Tetko *et al.*, 1995 ; Hawkins, 2004 ; Piotrowski et Napiorkowski, 2013 ; Lawrence *et al.*, 1997). Le degré de prise en compte de cette problématique dans le domaine de la télédétection reste toutefois incertain.
- Les pixels frontière ne doivent pas être exclus. Cette condition n'est pas obligatoire dans les cas où seule une mesure de la précision de la classification est nécessaire. Les professionnels de la télédétection ne prennent souvent en compte que les pixels purs (situés loin des délimitations) afin d'éviter les biais pessimistes pouvant découler d'une erreur de cadrage (Verbyla et Hammond, 1995 ; Foody, 2008). Toutefois, si la correction des biais relatifs au comptage de pixels se fait au moyen d'une matrice de confusion, les pixels mixtes ou frontière doivent eux aussi être pris en considération, ce qui suppose d'étendre la matrice aux sous-pixels ou classifications douces (Pontius et Cheuk, 2006 ; Silván-Cárdenas et Wang, 2008).

En principe, le biais est estimé à partir d'un échantillon de données de validation, et sa variance doit être calculée en tenant compte du plan d'échantillonnage relatif aux dites données. Si la part de pixels mixtes est dérisoire (p. ex., dans les paysages composés de très grandes parcelles homogènes), la distribution binomiale dans chacune des strates utilisées aux fins de sélection de l'échantillon peut servir à estimer les variances de Φ_c et de ψ_c . Rien ne garantit que leurs distributions seront indépendantes, mais l'on pourra tout de même raisonnablement additionner les deux variances pour obtenir la variance du biais. La variance du biais estimé correspondra alors à la variance de l'estimation par comptage de pixels corrigée du biais.

A2.1. Estimateurs par calage

La correction du biais lié au comptage de pixels s'effectue le plus fréquemment au moyen d'estimateurs par calage. L'emploi de ce terme par la littérature relative à la télédétection est lié, mais pas tout à fait identique, à son utilisation par la littérature statistique (Brown, 1982 ; Deville et Särndal, 1992).

Partons de l'hypothèse suivante :

A : matrice de confusion relative à la population totale. Inconnue.

T : vecteur colonne présentant les totaux issus des données de terrain $T_g = A_{g+}$ (T signifie « réalité de terrain », de l'anglais *truth*). Inconnu.

R : vecteur colonne présentant $R_c = A_{+c}$, soit la surface totale de pixels classés dans chaque catégorie. Connu.

a : matrice de confusion connue, établie à partir d'un échantillon de pixels tests a_{gc} = surface estimée de la classe g classée dans la catégorie c .

t : vecteur colonne présentant les totaux calculés à partir des données de terrain sur l'échantillon $t_g = a_{g+}$.

r : vecteur colonne présentant $r_c = \sum_g a_{gc}$, le nombre de pixels classés dans chaque catégorie de l'échantillon.

D_T , D_R , D_t , et D_r sont les matrices diagonales composées des éléments T_g , R_c , t_g et r_c .

Le vecteur R fournit des informations biaisées mais complètes sur la superficie cultivée. L'objectif est d'estimer T à partir des éléments connus R et de la matrice de confusion a .

$E_c = A' D_t^{-1}$ et $E_g = A D_R^{-1}$ et sont des matrices d'erreur relative (inconnues) portant sur la population. E_c indique la probabilité qu'un pixel soit classé dans la catégorie c , sachant que la réalité de terrain correspond à g , et E_g indique la probabilité selon laquelle la réalité de terrain correspond à g , sachant que le pixel est classé dans la catégorie c .

Des matrices similaires, $e_c = a' D_t^{-1}$ et $e_g = a D_r^{-1}$, peuvent être calculées à partir de l'échantillon. Il s'agit d'estimateurs approximativement sans biais de E_c et E_g , à condition que les données de l'échantillon soient extrapolées à l'aide de pondérations correctes inversement proportionnelles à la probabilité d'échantillonnage.

Les formules suivantes sont calculées directement à partir des définitions ci-dessus :

$$R = A' D_T^{-1} T = E_c T \quad T = A D_R^{-1} R = E_g R \quad T_g = R_g + N \left(\frac{T_g - R_g}{N} \right) \quad (2)$$

si E_c est carré et régulier, alors $T = E_c^{-1} R$

Ce calcul est à l'origine de trois estimateurs de calage :

Étalonnage direct :

$$\hat{T}_{dir} = e_g R \quad (3)$$

qui est un estimateur par maximum de vraisemblance utilisé dans le cadre d'une distribution multinomiale (Tenenbein, 1972). La matrice e_g contient des estimations des probabilités conditionnelles $p(g/c)$ de chaque classe de terrain g classée dans la catégorie c .

Étalonnage inverse :

$$\hat{T}_{inv} = e_c^{-1} R \quad (4)$$

qui correspond à un estimateur par maximum de vraisemblance présentant des propriétés asymptotiques (Grassia et Sundberg, 1982). La matrice e_c contient des estimations des probabilités conditionnelles $p(c/g)$ de chaque classe d'image c classée dans la catégorie g sur le terrain.

Bauer *et al.* (1978), Priesley et Smith (1987) et Hay (1988) présentent cet estimateur de façon légèrement différente. Dans un article commentant la publication de Hay, Jupp (1989) analyse le risque d'instabilité lorsque est presque singulier, c'est-à-dire lorsque certaines racines latentes sont très faibles.

Correction additive

$$T_{ad}(g) = R_g + n \left(\frac{t_g - r_g}{n} \right) \quad (5)$$

proposé par Dymond (1992).

Woodcock et Gopal (2000) proposent une version élargie de l'étalonnage direct pour les classifications floues, de façon à pouvoir calculer des estimations pour tous les niveaux d'adhésion confondus.

A2.2. Calcul de la variance de certains estimateurs

Pour évaluer les estimateurs associant données de terrain et images classées, il est important de calculer la variance du type d'estimateur choisi. Ce calcul peut s'avérer complexe dans de nombreux cas. Nous n'étudierons ici que les situations les plus simples, et en particulier le cas de l'étalonnage direct

$\hat{T}_{dir} = e_g R$. Pour une classe de terrain donnée g ,

$$\hat{T}_{dir}(g) = \sum_c \frac{n_{gc}}{n_c} R_c \quad (6)$$

c'est-à-dire la superficie estimée à partir d'un échantillon stratifié dans lequel les classes c correspondent aux strates de taille R_c connues à partir des images classées.

Si l'échantillonnage est aléatoire dans chaque classe c et que la taille de l'échantillon n_c a été préalablement déterminée au sein de chaque classe, la variance peut être calculée à partir de la formule binomiale habituelle suivante :

$$V_{dir}(g) = \sum_c \frac{1-f_c}{n_c} \frac{n_{gc}}{n_c} \left(1 - \frac{n_{gc}}{n_c}\right) R_c^2 \quad (7)$$

où le facteur de correction de population finie f est généralement insignifiant dans les applications relatives à la télédétection.

Une taille fixe d'échantillon n_c suggère que la classification des images est connue avant l'échantillonnage. C'est généralement le cas pour l'estimation de la superficie de la couverture du sol (ou du changement de couverture du sol) à des fins de suivi environnemental ou forestier. Si la stratification a déjà été effectuée, les estimateurs de la superficie et de la variance sont calculés séparément au sein de chaque strate h puis additionnés à un stade ultérieur.

Si l'échantillon doit être préparé avant la phase de classification des images, par exemple en cas d'estimation de la superficie occupée par des cultures annuelles, il ne pourra être de taille fixe n_c . La taille de l'échantillon n_c correspondant à différentes classes sera répartie sur une base multinomiale en cas d'échantillonnage aléatoire simple. Nous nous retrouverons alors dans une situation de post-stratification définie par la classification, qui nécessitera un terme supplémentaire dans l'équation permettant de calculer la variance (Cochran, 1977, p. 135).

Pour l'estimateur par régression présenté dans le chapitre 5 qui se fonde sur le lien existant entre la part d'un champ occupée par une culture y et la proportion x d'une classe connexe (le sous-indice de stratification est omis ici aussi) :

$$y_{Reg} = y + \hat{b} (\bar{X} - x) \quad (8)$$

l'approximation habituelle permettant de calculer la variance est la suivante :

$$Var(y_{Reg}) = Var(y) (1 - r_{xy}^2) \quad (9)$$

Lorsque l'échantillon est de taille restreinte, ou lorsqu'une ou plusieurs observations très influentes ont été émises, il pourra être nécessaire de procéder à une approximation plus complexe (Gallego *et al.*, 2014). Soulignons que l'application raisonnable de l'estimateur par régression nécessite des répartitions à peu près continues de x et y comprises entre 0 et 1 : les unités d'échantillonnage ne doivent donc pas prendre la forme de points (ce qui concentrerait les deux répartitions sur 0 et 1), mais plutôt de segments aréolaires suffisamment importants pour englober plusieurs parcelles cultivées ou paysagères.

A2.3. Matrices de confusion exprimées sous forme de nombres de points

Les formules présentées ci-dessus renvoient à l'application d'estimateurs par calage (étalonnage direct et inverse) à partir de matrices de confusion extrapolées, en supposant que ces dernières ont été élaborées sur la base d'un échantillon présentant une probabilité d'échantillonnage p_i connue et extrapolé à l'aide de la pondération suivante : $1/p_i$. L'absence de biais n'est approximative que lorsque le dénominateur se compose d'estimateurs non biaisés ;

dans tous les cas, les distorsions seront mineures si l'échantillon est plus important que le nombre de cellules composant la matrice de confusion.

Dans la littérature relative à la télédétection, les matrices de confusion sont souvent exprimées sous la forme suivante : a_{gc} , c'est-à-dire le nombre de pixels appartenant à la classe de terrain g ayant été classés dans la catégorie c . À ce stade, il convient de distinguer trois situations différentes en fonction du plan d'échantillonnage utilisé pour sélectionner les échantillons sur le terrain.

1. Si la probabilité d'échantillonnage est uniforme pour toute la région, la matrice de confusion a devient un estimateur non biaisé de A par la simple application d'un facteur d'extrapolation constante, et peut être utilisée à des fins d'étalonnage direct et inverse. C'est souvent le cas si l'échantillon est sélectionné avant l'acquisition et la classification des images. Grâce à cette stratégie d'échantillonnage, la matrice de confusion a exprimée sous forme de nombre de points peut être utilisée à des fins d'étalonnage direct comme inverse.
2. Si l'échantillon de points utilisés aux fins d'observation sur le terrain est sélectionné après classification des images, les classes c peuvent servir de strates pour assurer l'échantillonnage desdits points. Cette technique est habituellement recommandée lorsque l'échantillon de terrain sert à valider les cartes de couverture du sol obtenues à partir d'images et à élaborer des estimations des surfaces associant observations sur le terrain et images classées (Stehman, 2009). La taille de l'échantillon n_c est déterminée par l'application d'un système d'attribution (Wagner et Stehman, 2015) ou de règles générales telles que «au moins 50 points par strate». Cette stratégie d'échantillonnage permet d'utiliser la technique d'estimation par étalonnage direct avec a sans extrapolation à A . Les erreurs de commission sont ici corrigées, puisque la probabilité d'échantillonnage est homogène dans chacune des classes c ; les erreurs d'omission peuvent quant à elles être fortement biaisées lorsque la probabilité d'échantillonnage varie considérablement d'une classe à l'autre (Gallego, 2016). L'étalonnage inverse ne peut être utilisé sans extrapolation correcte à A .
3. Le nombre de points n_g à observer au sein de chaque classe de terrain g est déterminé en amont. Cette méthode est parfois appliquée à l'aide d'un échantillonnage dirigé que l'on pourra dénommer «échantillonnage par quotas» : les enquêteurs sont déployés sur le terrain avec pour mission de collecter un nombre donné de points. Leur objectif peut être, par exemple, d'enregistrer les coordonnées de 50 points situés dans des champs de maïs, de 50 points situés dans des champs de blé et de 50 autres points dans d'autres classes de couverture du sol. Les enquêteurs circulent au sein du périmètre étudié et choisissent des points qu'ils considèrent comme étant «représentatifs», le plus souvent au milieu de grandes parcelles. Cette méthode est pratique pour collecter à moindres frais des données de terrain, dont l'impact sur d'éventuels biais liés aux estimateurs demeure toutefois méconnu. Des méthodes d'échantillonnage probabiliste dans le cadre desquelles n_g est fixé au préalable sont envisageables, bien que les probabilités d'échantillonnage risquent d'être complexes. Quoi qu'il en soit, à la connaissance de l'auteur de cette annexe, cette méthode n'a jamais été appliquée dans le cadre d'enquêtes agricoles. Si l'on part du principe que cette stratégie d'échantillonnage peut être appliquée avec une probabilité uniforme pour chaque classe de terrain g , il est possible d'appliquer la méthode d'étalonnage inverse à a , et les erreurs d'omission seront corrigées, alors que les erreurs de commission seront biaisées. La méthode d'étalonnage direct ne peut être utilisée sans extrapolation à A . Si l'analyste est certain que la probabilité conditionnelle $p(c/g)$ est calculée à partir d'un échantillon par quotas, le recours à un estimateur par étalonnage inverse fondé sur la matrice de confusion a exprimée en nombre de points ou de pixels peut s'avérer une solution bon marché. L'extrapolation à A à partir de l'échantillon par quotas est impossible, tout comme l'utilisation de l'estimateur par étalonnage direct. Les erreurs de commission et d'omission calculées à partir de a ne peuvent être comparées pour déterminer si le comptage direct des pixels à partir des images classées surestime ou sous-estime chaque classe c .

L'échantillonnage probabiliste est généralement recommandé pour valider les cartes de couverture du sol et estimer des surfaces à partir de données de terrain et d'images satellite classées (Congalton et Green, 1999 ; Foody, 2002 ; Stehman *et al.*, 2003, 2005 ; Wagner et Stehman, 2015 ; Olofsson *et al.*, 2013, 2014 ; Li *et al.*, 2014). Il est cependant tout à fait légitime de chercher des systèmes moins coûteux si une enquête fondée sur cette méthode coûte trop cher. Le recours à un estimateur par étalonnage inverse fondé sur un échantillonnage par quotas sur le terrain peut s'avérer adapté, mais sa pertinence risque de dépendre de la répartition des pixels purs et mixtes.

A2.4. Quelle est la meilleure méthode pour corriger le biais ?

Tous les estimateurs de calage ont des difficultés à gérer les pixels mixtes, l'unité d'échantillonnage étant un point ou un pixel, et la précision de positionnement étant généralement du même ordre de grandeur qu'un pixel. On compare ainsi un pixel de l'image aux données de terrain correspondant à un pixel voisin. Lorsqu'une parcelle est de taille réduite ou contient des cultures mixtes, comme c'est souvent le cas dans les secteurs agricoles moins développés, la classification des images n'est pas très précise et les estimateurs de calage sont encore moins efficaces en raison de l'imprécision du copositionnement. Par exemple, Ceccarelli *et al.* (2016) montrent que les estimateurs par calage présentent une efficacité relative s'approchant de 1 dans les tests menés au Sénégal et au Kenya, malgré l'utilisation d'images d'une résolution d'environ 5 m captées par RapidEye. En d'autres termes, pour cet exemple précis, l'utilisation ultérieure d'images classées n'a pas vraiment amélioré la précision des enquêtes sur le terrain. Cette même étude reconnaît toutefois l'utilité des images pour l'élaboration de plans d'échantillonnage de meilleure qualité.

En cas d'utilisation d'un estimateur par étalonnage inverse dans le cadre d'un échantillonnage par quotas de points situés au milieu de champs relativement vastes, le problème de recalage (imprécision du copositionnement) ne se pose plus. L'impact de l'échantillonnage non probabiliste sur les estimations demeure méconnu, même si l'on sait que cette méthode introduit un biais optimiste dans la matrice de confusion (Hammond et Verbyla, 1996).

Le problème des pixels mixtes peut être formellement réglé en tenant compte des parts de couverture du sol occupées par les sous-pixels ; la réaction des estimateurs de calage en présence de classificateurs de sous-pixels n'est toutefois pas encore suffisamment connue.

Après comparaison de \hat{T}_{dir} et \hat{T}_{inv} dans différents jeux de données, Czaplewski et Catts (1992) et Walsh et Burk (1993) jugent que l'estimateur direct satisfait davantage aux différents critères étudiés : faisabilité, biais et variance. Yuan (1997) conclut que \hat{T}_{inv} présente une variance légèrement inférieure à celle de \hat{T}_{dir} et de moins bons résultats dans le cadre de la correction additive de Dymond ; d'après ses résultats, il est préférable d'utiliser \hat{T}_{inv} du point de vue du biais si l'échantillon est restreint par rapport au nombre de cellules comprises dans la matrice de confusion.

L'estimateur par régression est généralement moins sensible aux imprécisions de positionnement. Il est approprié lorsque les unités d'échantillonnage correspondent à des segments ou à des grappes de points contenant un nombre relativement important de pixels. Il peut également être considéré comme une méthode de correction du biais, et n'exclut pas les pixels mixtes qui sont naturellement inclus dans les unités d'échantillonnage. Si les unités d'échantillonnage correspondent à des segments dont l'ordre de grandeur est au moins une fois supérieur à l'imprécision du recalage, l'incertitude du positionnement n'aura qu'un impact mineur. Par exemple, si les images présentent une résolution de 20 m, l'imprécision du positionnement est susceptible d'être de l'ordre de 20 m. Si les unités d'échantillonnage sont des segments carrés de 300 m x 300 m ou de 500 m x 500 m, elles correspondront plus ou moins à une même parcelle dans les données de terrain et sur l'image, y compris les pixels mixtes. De ce point de vue, l'estimateur par régression apparaît probablement comme la méthode de correction de biais la plus sûre pour l'estimation des surfaces cultivées, même s'il est appliqué moins fréquemment dans la littérature sur la télédétection que les estimateurs par calage.

Kerdiles *et al.* (2013) et Liu *et al.* (2014) enregistrent une efficacité relative de l'estimateur par régression comprise entre 1,6 et 2,7 dans des zones tests de Chine composées de paysages complexes caractérisés par de fines bandes.

En comparant différents estimateurs dans trois régions pilotes de Chine, Li *et al.* (2014) montrent que les estimateurs par régression obtiennent de meilleurs résultats que les estimateurs par calage. Ils observent par ailleurs une dégradation plus importante de l'étalonnage inverse en cas de baisse de la précision de la classification des images.

A3. LE PROBLÈME DE SUBJECTIVITÉ DE L'ESTIMATION PAR COMPTAGE DE PIXELS

Le comptage de pixels peut poser un dilemme si un avis est forgé *a priori* à partir d'informations externes sur la surface occupée par une culture donnée c . Supposons, par exemple, que l'analyste s'attend à une surface d'environ 1 million d'hectares. Il calculera alors une première estimation de la surface à l'aide d'un système de classification donné. Ce dernier englobe le choix du type d'algorithme, les règles permettant d'éliminer les valeurs aberrantes ou de combler les données manquantes, et le réglage de certains paramètres spécifiques, tels que la probabilité préalable contenue dans le classifieur par maximum de vraisemblance.

Si la surface classée c est loin du chiffre escompté de 1 million d'hectares, un analyste consciencieux tentera probablement d'améliorer la classification en réexaminant le système initial jusqu'à obtenir des paramètres plus adaptés. Cette méthode d'amélioration de la classification est certainement acceptable, conçue pour répondre à la question «où la culture c est-elle produite?». Des doutes quant à l'objectivité de la surface estimée émergent toutefois en cas de recours au comptage direct des pixels. C'est pourquoi il convient de souligner l'importance d'une correction de biais rigoureuse dans le cadre de l'estimation de surfaces à partir d'images classées.

A4. RÉFÉRENCES BIBLIOGRAPHIQUES

- Allen, J.D.** 1990. «A Look at the Remote Sensing Applications Program of the National Agricultural Statistics Service». *Journal of Official Statistics*, vol. 6, n° 4, p. 393–409.
- Atkinson, P.M. et Curran, P.J.** 1997. «Choosing an appropriate spatial resolution for remote sensing investigations». *Photogrammetric Engineering and Remote Sensing*, vol. 63, n° 12, p. 1345–51.
- Atzberger, C. et Rembold, F.** 2013. «Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets». *Remote Sensing*, vol. 5, n° 3, p. 1335–54.
- Bauer, M.E., Hixson, M.M., Davis, B.J. et Etheridge J.B.** 1978. «Area estimation of crops by digital analysis of Landsat data». *Photogrammetric Engineering and Remote Sensing*, vol. 44, p. 1033–43.
- Boschetti, L., Flasse, S.P. et Brivio, P.A.** 2004. «Analysis of the Conflict between Omission and Commission in Low Spatial Resolution Dichotomic Thematic Products: The Pareto Boundary». *Remote Sensing of Environment*, vol. 91, n° 3-4, p. 280-92.
- Brown, P.J.** 1982. «Multivariate calibration», *Journal of the Royal Statistical Society, B*, vol. 44, n° 3, p. 287–321.
- Card, D.** 1982. «Using Known Map Category Marginal Frequencies to Improve Estimates of Thematic Map Accuracy». *Photogrammetric Engineering and Remote Sensing*, vol. 48, p. 431–9.
- Carfagna, E. et Gallego, F.J.** 2005. «Using Remote Sensing for Agricultural Statistics». *International Statistical Review*, vol. 73, n° 3, p. 389–404.
- Ceccarelli, T., Downie, M. et Remotti, D.** 2016. *Role of Earth Observations for Crop Area Estimates in Africa. Experiences from the AGRICAB Project*. Article rédigé pour la 7e Conférence internationale sur les statistiques de l'agriculture (ICAS VII), Rome. 26–28 octobre 2016.
- Chhikara, R.S.** 1984. «Effect of mixed (boundary) pixels in crop proportion estimation». *Remote Sensing of Environment*, vol. 14, p. 207–18.
- Congalton, R.G. et Green, K.** 1999. *Assessing the accuracy of remotely sensed data: principles and practices*, Lewis publishers : Boca Raton, États-Unis.
- Cushnie, J.L.** 1987. «The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies». *International Journal of Remote Sensing*, vol. 8, n° 1, p. 15–29.
- Czaplewski, R.L.** 1992. «Misclassification bias in areal estimates». *Photogrammetric Engineering and Remote Sensing*, vol. 58, n° 2, p. 189–92.
- Deville, J.C. et Särndal, C.E.** 1992. «Calibration estimators in survey sampling». *Journal of the American Statistical Association*, vol. 87, n° 418, p. 376–82.
- Duveiller, G. et Defourny, P.** 2010. «A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing». *Remote Sensing of Environment*, vol. 114, n° 11, p. 2637-50.

- Foody, G.** 2008. « Harshness in image classification accuracy assessment ». *International Journal of Remote Sensing*, vol. 29, n° 11, p. 3137–58.
- Foody, G.M.** 2002. « Status of land cover classification accuracy assessment ». *Remote Sensing of Environment*, vol. 80, n° 1, p. 185–201.
- Gallego, F.J.** 2006. « Review of the Main Remote Sensing Methods for Crop Area Estimates ». In : *Workshop Proceedings: Remote sensing support to crop yield forecast and area estimates*. Archives internationales de photogrammétrie, de télédétection et de sciences de l'information spatiale, XXXVI, 8/W48, 65–70. Disponible à l'adresse : http://www.isprs.org/publications/PDF/ISPRS_Archives_WorkshopStresa2006.pdf.
- Gallego, F.J., Kussul, N., Skakun, S., Kravchenko, O., Shelestov, A. et Kussul, O.** 2014. « Efficiency assessment of using satellite data for crop area estimation in Ukraine ». *International Journal of Applied Earth Observation and Geoinformation*, vol. 29, p. 22–30.
- Gallego, F.J.** 2004. « Remote sensing and land cover area estimation ». *International Journal of Remote Sensing*, vol. 25, n° 15, p. 3019–47.
- Gallego, F.J. et Rueda C.** 1993, « Balanced a Priori Probabilities for Maximum Likelihood Classification of Satellite Images ». *Bulletin de la 49^e réunion de l'Institut international de statistique*. Contributions, livre 1, p. 467–8.
- Gallego, J., Sannier, C. et Pennec, A.** 2016. « Validation of Copernicus Land Monitoring Services and Area Estimation ». Article présenté lors de la septième Conférence internationale sur les statistiques de l'agriculture (ICAS VII), 26–28 octobre 2016. Rome.
- Hammond, T.O. et Verbyla, D.L.** 1996. « Optimistic bias in classification accuracy assessment ». *International Journal of Remote Sensing*, vol. 17, n° 6, p. 1261–6.
- Hanuschak, G., Hale, R., Craig, M., Mueller, R. et Hart, G.** 2001. « The new economics of remote sensing for agricultural statistics in the United States ». In : *Proceedings of the Conference on Agricultural and Environmental Statistical Applications in Rome (CAESAR), June 5-7*, vol. 2, p. 427-37.
- Hanuschak, G.A., Sigman, R., Craig, M.E., Ozga, M., Luebbe, R.C., Cook, P.W., Kleweno, D.D. et Miller, C.E.** 1980, « Crop-Area Estimates from Landsat. Transition from research and development to timely results ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-18, n° 2, p. 160–6.
- Hawkins, D.M.** 2004. « The problem of overfitting ». *Journal of Chemical Information and Computer Sciences*, vol. 44, n° 1, p. 1–12.
- Hay, A.M.** 1988. « The derivation of global estimates from a confusion matrix ». *International Journal of Remote Sensing*, vol. 9, n° 8, p. 1395–8.
- Hsieh, P., Lee, L. C. et Chen, N.** 2001. « Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing ». *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, n° 12, p. 2657–63.
- Kerdiles, H., Dong, Q., Spyratos, S. et Gallego F.J.** 2013. « Use of high resolution imagery and ground survey data for estimating crop areas in Mengcheng county, China ». In : *Proceedings of 35th International Symposium on Remote Sensing of Environment (ISRSE35)*.

- Latty, R.S., Nelson, R., Markham, B., Williams, D., Toll, D. et Irons, J.** 1985. «Performance comparisons between information extraction techniques using variable spatial resolution data». *Photogrammetric Engineering & Remote Sensing*, vol. 51, n° 9, p. 1459–70.
- Lawrence, S., Giles, C.L. et Tsoi, A.C.** 1997. «Lessons in neural network training: Overfitting may be harder than expected». In : *Proceedings of the National Conference on Artificial Intelligence*, p. 540–5.
- Li, Y., Zhu, X., Pan, Y., Gu, J., Zhao, A. et Liu, X.** 2014. «A Comparison of Model-Assisted Estimators to Infer Land Cover/Use Class Area Using Satellite Imagery». *Remote Sensing*, vol. 6, n° 9, p. 8904–22.
- Liu, J., Chen, Z., Wang, L., Wang, X., Dong, Q. et Gallego, J.F.** 2014. «Assessing the crop acreage at county level on the North China Plain using an adapted regression estimator method». Article présenté lors de Troisième conférence internationale sur l’agro-géo-informatique, 11–14 août 2014. Beijing, Chine.
- Löw, F. et Duveiller, G.** 2014. «Defining the spatial resolution requirements for crop identification using optical remote sensing». *Remote Sensing*, vol. 6, n° 9, p. 9034–63.
- MacDonald, R.B. et Hall, F.G.** 1980. «Global crop forecasting». *Science*, vol. 208, n° 4445, p. 670–9.
- Mayaux, P. et Lambin, E.F.** 1995. «Estimation of tropical forest area from coarse spatial resolution data: a two-step correction function for proportional errors due to spatial aggregation». *Remote Sensing of Environment*, vol. 53, n° 1, p. 1–15.
- McRoberts, R.E. et Walters, B.F.** 2012. «Statistical inference for remote sensing-based estimates of net deforestation». *Remote Sensing of Environment*, vol. 124, p. 394–401.
- Moody, A. et Woodcock, C.** 1994. «Scale-dependent errors in the estimation of land cover proportions: implications for global land-cover datasets». *Photogrammetric Engineering and Remote Sensing*, vol. 60, n° 5, p. 585–94.
- Moody, A. et Woodcock, C.E.** 1996. «Calibration-based models for correction of area estimates from coarse resolution land-cover data». *Remote Sensing of Environment*, vol. 58, n° 3, p. 225–241.
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E. et Wulder, M.A.** 2014. «Good practices for estimating area and assessing accuracy of land change». *Remote Sensing of Environment*, vol. 148, p. 42–57.
- Piotrowski, A.P. et Napiorkowski, J.J.** 2013. «A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling». *Journal of Hydrology*, vol. 476, p. 97–111.
- Pontius Jr., R.G. et Cheuk, M.L.** 2006. «A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions». *International Journal of Geographical Information Science*, vol. 20, n° 1, p. 1–30.
- Silván-Cárdenas, J.L. et Wang, L.** 2008. «Sub-pixel confusion-uncertainty matrix for assessing soft classifications», *Remote Sensing of Environment*, vol. 112, n° 3, p. 1081–95.
- Stehman, S.V., Sohl, T.L. et Loveland, T.R.** 2003. «Statistical sampling to characterize recent United States Land-cover change», *Remote Sensing of Environment*, vol. 86, n° 3, p. 517–29.

- Stehman, S.V.** 2009. «Sampling designs for accuracy assessment of land cover». *International Journal of Remote Sensing*, vol. 30, n° 20, p. 5243–72.
- Stehman, S.V. et Foody, G.M.** 2008. «Accuracy assessment». In : *The SAGE handbook of remote sensing*, p. 297–314. SAGE Publications Ltd. : Londres.
- Stehman, S.V., Sohl, T.L. et Loveland, T.R.** 2005. «An evaluation of sampling strategies to improve precision of estimates of gross change in land use and land cover». *International Journal of Remote Sensing*, vol. 26, n° 22, p. 4941–57.
- Stehman, S.V., Wickham J.D., Smith J.H. et Yang L.** 2003. «Thematic accuracy of the 1992 national land-cover data for the Eastern United States: statistical methodology and regional results». *Remote Sensing of Environment*, vol. 86, n° 4, p. 500–16.
- Tetko, I.V., Livingstone, D. J. et Luik, A.I.** 1995. «Neural network studies. 1. Comparison of overfitting and overtraining». *Journal of Chemical Information and Computer Sciences*, vol. 35, n° 5, p. 826–33.
- Treitz, P. M., Howarth, P. J., Suffling, R. C. et Smith, P.** 1992. «Application of detailed ground information to vegetation mapping with high spatial resolution digital imagery». *Remote Sensing of Environment*, vol. 42, n° 1, p. 65–82.
- Verbeiren, S., Eerens, H., Piccard, I., Bauwens, I. et Van Orshoven, J.** 2008. «Sub-pixel classification of SPOT-VEGETATION time series for the assessment of regional crop areas in Belgium». *International Journal of Applied Earth Observation and Geoinformation*, vol. 10, n° 4, p. 486–97.
- Verbyla, D.L. et Hammond, T.O.** 1995. «Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids». *International Journal of Remote Sensing*, vol. 16, n° 3, p. 581–7.
- Wagner, J.E. et Stehman, S.V.** 2015. «Optimizing sample size allocation to strata for estimating area and map accuracy». *Remote Sensing of Environment*, vol. 168, p. 126–33.
- Waldner, F. et Defourny, P.** 2017. «Where can pixel counting area estimates meet user-defined accuracy requirements?». *International Journal of Applied Earth Observation and Geoinformation*, vol. 60, p. 1–10.
- Wu, C.** 2009. «Quantifying high-resolution impervious surfaces using spectral mixture analysis». *International Journal of Remote Sensing*, vol. 30, n° 11, p. 2915–32.
- Yuan, D.** 1997. «A simulation comparison of three marginal area estimators for image classification». *Photogrammetric Engineering and Remote Sensing*, vol. 63, n° 4, p. 385–92.
- Zhen, Z., Quackenbush, L.J., Stehman, S.V. et Zhang, L.** 2013. «Impact of training and validation sample selection on classification accuracy and accuracy assessment when using reference polygons in object-based classification». *International Journal of Remote Sensing*, vol. 34, n° 19, p. 6914–30.

