



Valorisation de l'information géographique en agriculture de précision

Jérémy Lherbier

► To cite this version:

Jérémy Lherbier. Valorisation de l'information géographique en agriculture de précision. Géographie. Université Joseph-Fourier - Grenoble I, 2005. Français. NNT : . tel-00010902

HAL Id: tel-00010902

<https://theses.hal.science/tel-00010902>

Submitted on 8 Nov 2005

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

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

THESE

Présentée par

Jérémy LHERBIER

pour obtenir le grade de

Docteur de l'Université Joseph Fourier – Grenoble 1

Discipline : Géographie

**Valorisation de l'information géographique
en agriculture de précision.**

Soutenue publiquement le 26 octobre 2005 devant le jury composé de :

Pierre DUMOLARD, directeur de thèse,

Joël CHARRE, rapporteur,

Michel BUSSI, rapporteur,

Maryse Guigo, examinatrice.

Thèse préparée au sein du laboratoire S.E.I.G.A.D., Institut de Géographie Alpine, U.J.F.

Remerciements

Cette thèse est le fruit d'une collaboration, dans le cadre d'une convention C.I.F.R.E., entre la société GEOSYS® de Toulouse et le laboratoire S.E.I.G.A.D. (Systèmes Environnementaux, Information Géographique et Aide à la Décision) de l'Institut de Géographie Alpine (Université Joseph Fourier). Un troisième partenaire a également contribué à ce travail : l'association Loi 1901 Arvalis Institut du Végétal, basée à Boigneville.

Ayant rédigé le projet de thèse chez Arvalis Institut du Végétal, je commencerais par remercier chaleureusement M. Denis Boigontier, initiateur du projet, pour la confiance qu'il m'a accordée. Je tiens également à saluer les autres membres de la N.T.I. : Corinne Deplanque, Rani Elmeouche, David Decluseau, Claudine Albertalli et Vincent Ayala (la vieille garde...) mais aussi Xavier Mérour, Florent Gasc, Benoît de Solan, Erwan Lefeuvre, Laëtitia Renoult, Françoise Font, Béatrice Ariaux et Francis Hénault. Et d'autres personnes appartenant à d'autres services ou d'autres stations expérimentales : François Piraux, Jean-Paul Nicolletti, Jean-Paul Gillet et Anne-Monique Bodilis. Merci également à M. Thierry Dufour, exploitant agricole ainsi qu'à M. Dominique Rouiller et Mme Martine Chalat, d'Epis-Centre.

Une fois la thèse engagée, c'est au sein de l'entreprise GEOSYS que j'ai passé la majeure partie de mon temps de travail. Je remercie M. Damien Lepoutre et toute l'équipe de direction pour leur accueil et pour m'avoir donné les moyens matériels et humains de réaliser ce doctorat. Je remercie également l'ensemble des salariés et ex-salariés pour leur bonne humeur et leurs innombrables coups de main dans des domaines aussi variés que la géomatique, l'agronomie, le guidage par G.P.S. dans les labyrinthes administratifs et la découverte des produits régionaux : Florent Magot, Philippe Lonvaud, Perrine Claux, Céline Fabre, La CORINE Team (Charles Drieu La Rochelle, Béatrice Pintat-Faracet, Stéphane Nivelet, Jean-Luc Amiel, Vincent Léger, Jean-Fabrice Bonnefoux), Philippe Doussat, Vincent Lelandais, Michel Guillermin, Marie Argentier, Marc Jacquin, Olivier François, Sylvie Carlevatto, Dominique Vormus, Huguette Wurtz, Matthieu Hyrien, Frédéric Biard, Fanny Rudelle, Charlotte Diagoras, Grégoire Darras, Chris Ovejero et Julien Orensanz. Enfin, il est une personne qui devrait figurer dans la liste ci-dessus mais, étant en plus mon parrain au sein de GEOSYS, je tiens tout particulièrement à signifier à Fabien Lahoche ma plus sincère gratitude.

Au sein du laboratoire S.E.I.G.A.D., où j'ai passé l'autre partie de mon temps de travail, je remercie très sincèrement M. Pierre Dumolard, mon directeur de thèse, pour son humanité, pour son soutien et pour sa disponibilité indéfectible. Toute ma gratitude et mon amitié aux thésards du labo : Nicolas Perrin, Jean-Christophe Loubier, Laure Charleux-Handsone, Rémi Lhotellier, Olivier Bouhet, Hayan Saffour et Ludovic Dupuis (ainsi que leurs conjoints respectifs) pour avoir cultivé nos différences (pas seulement gastronomiques) et pour les pauses café qui se terminaient facilement en discussions acharnées et passionnantes sur nos sujets de travail.

L'étude des configurations de données n'aurait pas pu être réalisée au sein du S.E.I.G.A.D sans l'aide précieuse de Céline Antunes et de Cécile Cotterot. Et les innombrables simulations ont pu être opérées grâce à la puissante machine de calcul parallèle (S.G.I. Altix 350) de la Maison des Sciences de l'Homme – Alpes à laquelle M. Bernard Bouhet et M. Bernard Rapacchi m'ont aimablement donné accès. A ce sujet, je remercie François Bouhet pour m'avoir aidé à maîtriser Juliette.

Un grand merci à Mme Christine Margetic, professeur auprès de l'Institut de Géographie et d'Aménagement Régional de l'Université de Nantes, pour son aide.

Cette thèse étant une conséquence indirecte de ce qu'ils m'ont apporté, peut être sans en avoir vraiment conscience, je tiens ici à remercier M. Jean-Luc Marcy, Conservateur-Archéologue départemental du Musée Départemental de Préhistoire d'Arras ; et M. Philippe Martin, Maître de Conférence auprès de l'U.F.R. d'Histoire-Géographie du Pôle d'Arras de l'Université d'Artois.

Sur un tout autre plan, il me reste à remercier, dans une famille décidément très présente, l'amatrice éclairée de Bêtises qu'est Mme Françoise Bouhet.

Enfin, il est de nombreuses personnes, dans ma famille comme dans ma belle-famille, qui, sans avoir participé directement au travail, n'en ont pas moins joué un rôle dans son accomplissement. Un grand merci à mes parents et à mes beaux-parents, à ch'tiote Marcelle et au Pèp' à r'ssorts, à mon frère, mes sœurs, mes beaux-frères et mes belles-sœurs (merci, Julie, pour la traduction) ainsi qu'à mes amis nordistes pour les récréations.

La même chose, tout le reste et bien plus encore à ma Milie et à ma Scroufipougnette...

A ceux qui montrent la voie
sans ouvrir le chemin.

Résumé

Dans le domaine agricole, l'application du principe du développement durable s'est faite en même temps que la prise de conscience des limites de l'agriculture productiviste pour donner naissance à l'agriculture durable. Cette dernière englobe toute une série d'approche de la conduite des cultures dont les plus connues sont l'agriculture biologique et l'agriculture raisonnée.

Une autre approche, l'agriculture de précision, s'appuie sur de l'information géographique à très grande échelle et des outils adaptés (Systèmes d'Information Géographique et analyse spatiale) pour gérer la variabilité intraparcellaire afin d'optimiser les coûts d'exploitation et de limiter les impacts environnementaux des opérations culturales. La pratique de l'agriculture de précision, en particulier la mise en oeuvre opérationnelle de la modulation des intrants, nécessite le découpage de la parcelle en quelques grandes entités homogènes.

Une méthode de zonage a donc été développée. Dans un premier temps, des tests ont été effectués sur la base de l'Aide MultiCritère à la Décision. Mais cette approche a été abandonnée pour des raisons d'applicabilité. La méthode développée au final s'appuie sur une classification mixte intégrant plusieurs contraintes spatiales et permettant la prise en compte du savoir de l'exploitant agricole. Elle permet d'obtenir rapidement et simplement des zonages satisfaisants tout en offrant quelques indicateurs de validité.

Une étude visant à modéliser la structure spatiale de données onéreuses en utilisant d'autres l'étant nettement moins a été menée. Nous avons montré qu'avec notre méthode de zonage, il était possible d'approcher correctement la structure spatiale de plusieurs variables agronomiques pertinentes en se basant sur une combinaison de cartes de résistivité électrique, de rendement et d'images satellitaires.

MOTS-CLÉS : Analyse spatiale, zonage, agriculture durable, agriculture de précision, entité homogène.

Abstract

In the agricultural field, the principle of sustainable development began to be applied when productivist agriculture showed its limits. This gave rise to sustainable agriculture which includes a whole series of approach to control the cultures among which biological agriculture and reasoned agriculture are the most known.

Another approach, precision agriculture, relies on very large scale geographical information and adapted tools (Geographical Information Systems and spatial analysis) to manage in-field variability in order to optimize exploitation costs and to limit the environmental impacts of the farming operations. The practice of precision agriculture, in particular the operational implementation of the modulation of inputs, requires the division of the parcel into a few big management zones.

A zoning method thus has been developed. Initially, tests were carried out on the basis of MultiCriteria Decision Making. But this approach was abandoned for reasons of applicability. The method finally developed is based on a hybrid clustering integrating several spatial constraints and allowing to take into account the farmer's knowledge. It makes it possible to obtain quickly and simply satisfactory zonings while offering some validity indicators.

A study aiming at modelling the spatial structure of expensive data while using others cheaper was conducted. We have demonstrated that with our zoning method, it was possible to correctly approach the spatial structure of several relevant agronomic variables using a combination of electrical resistivity, yield and satellite imagery maps.

KEYWORDS : Spatial analysis, zoning, sustainable agriculture, precision agriculture, management zone.

Table des matières

RESUME - ABSTRACT	4
TABLE DES MATIERES	6
TABLE DES FIGURES	11
TABLE DES TABLEAUX	14
 INTRODUCTION GENERALE	 16
 PREMIERE PARTIE : L'AGRICULTURE DE PRECISION	 18
1. DE L'AGRICULTURE A L'AGRICULTURE DE PRECISION	19
1.1. INTRODUCTION	19
1.1.1. Une activité très ancienne...	19
1.1.2. Qui influence les espaces...	19
1.1.3. ...Et les sociétés	20
1.2. L'EMERGENCE ET LES LIMITES DE L'AGRICULTURE PRODUCTIVISTE	20
1.2.1. La première révolution agricole	20
1.2.2. La seconde révolution agricole	21
1.2.3. Les limites de l'agriculture productiviste	22
1.3. L'AGRICULTURE DURABLE	23
1.3.1. Définition et objectifs	23
1.3.2. Deux exemples de modes de production relevant de l'agriculture durable	24
2. L'AGRICULTURE DE PRECISION, UNE APPROCHE DE L'AGRICULTURE DURABLE	28
2.1. L'EMERGENCE DE L'AGRICULTURE DE PRECISION	28
2.2. DEFINITION DE L'AGRICULTURE DE PRECISION	29
2.3. LES ENJEUX DE L'AGRICULTURE DE PRECISION	30
2.3.1. Un gain économique	30
2.3.2. La préservation de l'environnement	31
2.3.3. D'autres objectifs connexes	31
2.4. UNE APPROCHE DE L'AGRICULTURE DURABLE	31
2.5. UN EXEMPLE DE MISE EN ŒUVRE DE L'AGRICULTURE DE PRECISION	32
3. TYPES DE MODULATION ET ETAPES DE LA PRATIQUE DE L'AGRICULTURE DE PRECISION	33
3.1. LA MODULATION DIRECTE	33
3.2. LA MODULATION INDIRECTE	35
3.2.1. La mise en évidence et la caractérisation de la variabilité intraparcellaire	35
3.2.2. La prise de décision	36
3.2.3. La modulation effective	39
4. LES VERROUS AU DEVELOPPEMENT DE L'AGRICULTURE DE PRECISION	39
4.1. LES VERROUS SOCIO-ECONOMIQUES	40
4.2. LES VERROUS TECHNOLOGIQUES	40
4.3. LES VERROUS LIES A LA GESTION AGRONOMIQUE	42
 SECONDE PARTIE : GEOMATIQUE ET AGRICULTURE DE PRECISION	 44
5. INFORMATION GEOGRAPHIQUE, S.I.G., ANALYSE SPATIALE ET AGRICULTURE DE PRECISION	45
5.1. INFORMATION GEOGRAPHIQUE ET AGRICULTURE DE PRECISION	45
5.1.1. Les analyses de sol	46

5.1.2. Les cartes de rendement	48
5.1.3. Les images satellitaires et aériennes issues de la télédétection	50
5.1.4. La conductivité et la résistivité électrique du sol	53
5.1.5. Les croisements d'informations	55
5.2. S.I.G. ET AGRICULTURE DE PRECISION	55
5.2.1. Définition des S.I.G.	55
5.2.2. Modes de représentation de l'information géographique des S.I.G.	56
5.2.3. L'utilité des S.I.G. en agriculture de précision	58
5.2.4. Les logiciels dédiés à l'agriculture de précision	59
5.3. ANALYSE SPATIALE ET AGRICULTURE DE PRECISION	59
5.3.1. Définition de l'analyse spatiale	59
5.3.2. Les principales méthodes d'analyse spatiale des S.I.G.	61
5.4. LIMITES ET LACUNES DES CAPACITES D'ANALYSE SPATIALE DES S.I.G.	61
5.4.1. Les limites	62
5.4.2. Les lacunes	63
6. LE DEVELOPPEMENT D'UNE METHODE DE ZONAGE	65
6.1. LES CONTRAINTES DE TRAVAIL	65
6.2. LES METHODES DE ZONAGE EXISTANTES	66
6.2.1. Méthodes sans contraintes spatiales	66
6.2.2. Méthodes avec contraintes spatiales	67
6.3. APPROCHES NON RETENUES	69
6.3.1. Les classifications floues	70
6.3.2. Les réseaux neuronaux	70
6.3.3. La segmentation d'images	72
TROISIEME PARTIE : APPROCHE DE LA MISE AU POINT D'UNE METHODE DE ZONAGE PAR AIDE MULTICRITERE A LA DECISION	74
7. L'AIDE MULTICRITERE A LA DECISION	75
7.1. VOCABULAIRE	75
7.2. LA CRITIQUE DE LA RECHERCHE OPERATIONNELLE	76
7.2.1. La recherche opérationnelle : réussites et échecs	76
7.2.2. Le postulat de la recherche d'un optimum	76
7.2.3. Conclusion sur les domaines d'échecs de la recherche opérationnelle	77
7.3. L'INTERET DE L'APPROCHE PAR AIDE MULTICRITERE A LA DECISION	78
7.3.1. La prise en compte de la multiplicité des critères	78
7.3.2. La prise en compte de la multiplicité des acteurs	78
7.3.3. La non commensurabilité des critères	78
7.3.4. Les oppositions entre critères	78
7.3.5. Conclusion sur l'intérêt de l'A.M.C.D.	78
7.4. LES ETAPES D'UNE ETUDE PAR A.M.C.D.	79
7.4.1. La définition des actions potentielles	79
7.4.2. La définition de la famille des critères	79
7.4.3. La restitution des échelles de préférence	79
7.4.4. Le choix des pondérations	80
7.4.5. La saisie de la matrice des jugements	81
7.4.6. L'agrégation	81
8. L'INTERET D'UNE INTEGRATION S.I.G. - A.M.C.D.	83
8.1. LES CARACTERISTIQUES DES PROBLEMES DECISIONNELS A REFERENCE SPATIALE	83
8.1.1. Le besoin en données à référence spatiale	83
8.1.2. Des décisions qui concernent le territoire	83
8.1.3. La multiplicité des intervenants	84
8.1.4. La nécessité de tenir compte de l'existant	84
8.2. LES PROBLEMES DECISIONNELS A REFERENCE SPATIALE SONT SOUVENT DES PROBLEMES MULTICRITERES	84
8.3. LA COMPLEMENTARITE S.I.G. - A.M.C.D.	85
8.3.1. Une puissante interface graphique	85
8.3.2. Un système de gestion de base de données	85
8.3.3. La technique de superposition de couches, les contraintes et les critères	85

8.3.4. La comparaison des solutions	86
8.3.5. La prise en compte de l'espace	86
8.4. L'ADAPTATION DE LA PROCEDURE D'AGREGATION MULTICRITERE AU PROBLEME DECISIONNEL A REFERENCE SPATIALE	87
8.5. VERS DES SYSTEMES D'AIDE A LA DECISION A REFERENCE SPATIALE	87
8.6. POURQUOI L'INTEGRATION S.I.G. - A.M.C.D. NE S'EST PAS FAITE AVANT ?	88
9. EXEMPLES D'APPLICATION	89
9.1. LA DEFINITION D'ENTITES HOMOGENES	89
9.1.1. Objectifs	89
9.1.2. Méthode d'A.M.C.D. retenue	89
9.1.3. Renseignement des paramètres d'E.L.E.C.T.R.E. TRI	90
9.1.4. Application de la méthode	92
9.1.5. Conclusion	93
9.2. LE PRE-DIAGNOSTIQUE D'UN RISQUE DE POLLUTION	93
9.2.1. Objectif	94
9.2.2. Etude initiale GEOSYS S.A. – Arvalis Institut du Végétal	94
9.2.3. Approche par A.M.C.D.	96
9.2.4. Conclusion	99
QUATRIEME PARTIE : APPROCHE DE LA MISE AU POINT D'UNE METHODE DE ZONAGE PAR LA STATISTIQUE MULTIVARIEE	101
10. NOTRE METHODE DE ZONAGE	102
10.1. FONCTIONNEMENT GLOBAL DE NOTRE METHODE DE ZONAGE	102
10.1.1. Une méthode fondée sur une classification mixte...	102
10.1.2. Appuyée par une analyse factorielle préalable...	103
10.1.3. Et soumise à des contraintes spatiales...	103
10.1.4. Synthèse et fonctionnement global	103
10.2. FONCTIONNEMENT DETAILLE DE NOTRE METHODE DE ZONAGE	105
10.2.1. Mise en forme des cartes	105
10.2.2. Le choix des cartes initiales et leur pondération	108
10.2.3. L'Analyse en Composantes Principales	109
10.2.4. La classification par k-moyennes	110
10.2.5. La morphologie mathématique	112
10.2.6. La suppression des points isolés restants	114
10.2.7. La différenciation des entités statistiques en entités spatiales	115
10.2.8. La Classification Ascendante Hiérarchique avec contrainte de contiguïté	115
10.2.9. Le choix des cartes de zonage	117
10.2.10. Calculs d'indicateurs de crédibilité	118
11. L'ANALYSE DES DONNEES D'ENTREE	118
11.1. L'ANALYSE EN COMPOSANTES PRINCIPALES	118
11.1.1. Principes	118
11.1.2. Exemple d'application	118
11.2. L'ANALYSE DES CORRESPONDANCES MULTIPLES	119
11.2.1. Principes	119
11.2.2. Exemple d'application	120
11.3. L'AUTOCORRELATION SPATIALE	121
11.3.1. Principes	121
11.3.2. Exemple d'application	122
12. LA VALIDATION DE LA METHODE DE ZONAGE	123
12.1. LA VALIDATION STATISTIQUE SPATIALE DES CARTES DE ZONAGE	123
12.1.1. Calcul d'indicateurs locaux pour chaque carte initiale	124
12.1.2. Calcul des Indices d'Hétérogénéité	126
12.1.3. Autres informations fournies	128
12.2. L'ANALYSE DISCRIMINANTE, UNE AUTRE METHODE DE VALIDATION ENVISAGEABLE	129
13. EXEMPLE D'APPLICATION	130

13.1. OBJECTIFS	130
13.2. RENSEIGNEMENT DES PARAMETRES DE NOTRE METHODE DE ZONAGE	131
13.3. APPLICATION DE LA METHODE	131
13.4. CONCLUSION	133
CINQUIEME PARTIE : ETUDE DES CONFIGURATIONS DE DONNEES	134
14. CONTEXTE DE L'ETUDE	135
14.1. OBJECTIF DE L'ETUDE	135
14.2. LES DONNEES DISPONIBLES	135
14.2.1. Les données tests	135
14.2.2. Les données de référence	137
14.3. RESULTATS ATTENDUS	137
14.4. PARCELLES ETUDIEES	138
14.4.1. Parcelle "La Jaillière"	138
14.4.2. Parcelle "Imbault"	141
14.4.3. Parcelle "Magneraud"	142
14.4.4. Parcelle "Ouzouer"	142
14.5. ETUDES SIMILAIRES DEJA ENTREPRISES	143
14.5.1. Par Dampney, King, Lark, Wheeler, Bradley et Mayr	143
14.5.2. Par Domsch, Kaiser, Witzke et Zauer	143
14.5.3. Par Frogbook, Oliver et Derricourt	144
15. MODE OPERATOIRE DE L'ETUDE	144
15.1. ANALYSE DES DONNEES	145
15.2. APPLICATION DE LA METHODE DE ZONAGE	145
15.2.1. Mise en forme préalable des données	145
15.2.2. Choix des cartes initiales et pondération	146
15.3. VALIDATION SPATIALE DES RESULTATS	147
15.4. MISE EN FORME DES RESULTATS ET DEFINITION DES CONFIGURATIONS DE DONNEES	147
16. PRESENTATION DES RESULTATS	148
16.1. ANALYSE DES DONNEES	148
16.1.1. Considération sur l'hétérogénéité des parcelles étudiées	148
16.1.2. Autocorrélation spatiale des parcelles	150
16.1.3. Les Analyses en Composantes Principales (A.C.P.)	151
16.1.4. Les corrélations	152
16.2. PRESENTATION DES CONFIGURATIONS DE DONNEES	157
16.2.1. Pour la référence "pourcentage d'argile"	157
16.2.2. Pour la référence "teneur en P2O5"	161
16.2.3. Pour la référence "pH H2O"	165
16.2.4. Résultats obtenus pour les autres données de référence	168
17. CONCLUSIONS DE L'ETUDE	175
17.1. CONCERNANT LES GROUPES DE DONNEES DE REFERENCE	175
17.1.1. La granulométrie	175
17.1.2. Les teneurs physico-chimiques	178
17.1.3. Les pH	178
17.1.4. La profondeur de sol	178
17.2. CONCERNANT LES GROUPES DE DONNEES TESTS	179
17.2.1. La résistivité électrique	179
17.2.2. Le rendement	179
17.2.3. L'image satellitaire	180
17.3. CONCERNANT LE NOMBRE DE ZONES DES CARTES DE ZONAGE	180
17.3.1. Nombre de zones et qualité de la carte de zonage	180
17.3.2. Nombre de zones et adéquation des données tests au sujet de cette étude	181
17.4. CONCERNANT LES CONFIGURATIONS DE DONNEES	182
17.5. POUR UNE EVENTUELLE POURSUITE DE L'ETUDE	182
18. COMPARAISON DES ZONAGES OBTENUS AVEC LE LOGICIEL ECOGNITION ET NOTRE METHODE DE ZONAGE	183

<i>18.1.1. Modalités des comparaisons</i>	183
<i>18.1.2. Comparaison des cartes de zonage issues de l'image satellitaire</i>	184
<i>18.1.3. Comparaison des cartes de zonage issues des cartes de rendement</i>	185
<i>18.1.4. Comparaison des cartes de zonage issues des cartes de résistivité électrique</i>	185
<i>18.1.5. Comparaison des cartes de zonage issues de l'ensemble des données</i>	186
<i>18.1.6. Conclusion</i>	187
 CONCLUSION GENERALE	 189
 ANNEXES	 191
19. VOCABULAIRE COMMUN EN A.M.C.D.	191
19.1. ACTEUR	191
19.2. ACTION	191
19.3. AIDE A LA DECISION	191
19.4. CONSEQUENCE	191
19.5. CRITERE	191
19.6. ECHELLE DE PREFERENCE	192
19.7. TABLEAU DE PERFORMANCES	192
20. PRISE EN COMPTE DES INFORMATIONS INITIALES DANS UNE STRATEGIE DE DEFINITION D'ENTITES HOMOGENES	192
21. MISE EN FORME DES TABLEAUX D'I.G.V.S.	194
 BIBLIOGRAPHIE	 200

Table des figures

Figure 1 : Exemple de "marée verte" près de la Pointe de Guettes, en Bretagne	22
Figure 2 : Evolution des surfaces en agriculture biologique entre 1995 et 1998	25
Figure 3 : Principe de l'agriculture raisonnée	27
Figure 4 : Contexte de l'agriculture de précision	32
Figure 5 : Fonctionnement du système N-Sensor	34
Figure 6 : Localisation des mesures de réflectance avec le système N-Sensor	34
Figure 7 : Les trois étapes de l'agriculture de précision	35
Figure 8 : Exemple de localisation d'analyses de sol sur la parcelle Imbault	47
Figure 9 : Exemple de carte de rendement corrigée sur la parcelle Imbault	50
Figure 10 : Exemple d'image satellitaire sur la parcelle Imbault	51
Figure 11 : Le drone Pixy au décollage	52
Figure 12 : Exemple de photo réalisée avec le Pixy	53
Figure 13 : Méthode de mesure de la résistivité électrique	54
Figure 14 : Exemple de carte de résistivité électrique sur la parcelle Imbault	55
Figure 15 : Exemples de représentation de données en mode image	57
Figure 16 : Figurés élémentaires permettant la représentation vectorielle des objets géographiques en mode vecteur	58
Figure 17 : Vue simplifiée d'un réseau de neurones artificiel	71
Figure 18 : Restitution des sens de préférence pour les critères "habitabilité" et "puissance"	80
Figure 19 : Cartes de K ₂ O, de CaO et de pourcentage d'argile sur la parcelle "Baziège"	89
Figure 20 : Les trois interrogations permettant d'estimer la pondération des cartes initiales	91
Figure 21 : Carte de zonage obtenue sur la parcelle "Baziège"	92
Figure 22 : Cartes correspondantes à la contrainte et aux critères retenus pour l'approche par A.M.C.D.	98
Figure 23 : Cartographie des niveaux de risque de pollution diffuse sur le bassin versant de la Nonette	98
Figure 24 : Fonctionnement global de la méthode de zonage	104
Figure 25 : Création d'un fichier "A.S.C.I.I. Raster" à partir d'une carte de semis de points	106
Figure 26 : Création d'une carte de zonage en mode vecteur à partir d'un fichier "A.S.C.I.I. Raster"	107
Figure 27 : Région "utile" et régions de "no data" d'une carte en mode image	107
Figure 28 : Les trois interrogations permettant d'estimer la pondération des cartes initiales	108
Figure 29 : Principe du changement de système d'axes de l'A.C.P.	109
Figure 30 : Exemple de classification par <i>k</i> -moyennes sur deux classes dans un espace à deux variables en trois itérations	111
Figure 31 : Eléments structurants à quatre et huit connecteurs permettant de considérer respectivement le 4-voisinage et le 8-voisinage du point central	112
Figure 32 : Exemple de dilatation d'un objet avec les éléments structurants à 4 et 8 connecteurs	113
Figure 33 : Exemple d'applications successives du filtre	114
Figure 34 : Exemple d'application de la fermeture et du lissage à la carte de zonage issue des <i>k</i> -moyennes	115
Figure 35 : Représentation de deux itérations successives lors de la C.A.H.	116
Figure 36 : Historique des distances de fusion de la C.A.H. avec contrainte de contiguïté et de taille	117
Figure 37 : Analyse en Composantes Principales de données tests sur la parcelle Imbault	119

Figure 38 : Diagramme des colonnes (composantes 1 et 2) issu de l'A.C.M. sur les données de profondeur de sol et les trois voies de résistivité électrique	120
Figure 39 : Echantillon du modèle numérique de terrain de la forêt de Harvard	122
Figure 40 : Carte générée avec une fonction aléatoire	123
Figure 41 : Calcul des différents coefficients de variation	125
Figure 42 : Calcul des quotients de coefficients de variation	126
Figure 43 : Calcul des indicateurs zonaux de validité spatiale	128
Figure 44 : Exemple de contenu du fichier d'informations fourni avec la carte de zonage	129
Figure 45 : Cartes de K ₂ O, de CaO et de pourcentage d'argile sur la parcelle "Baziège"	131
Figure 46 : Carte de zonage obtenue sur la parcelle "Baziège" avec notre méthode de zonage	132
Figure 47 : Carte de zonage obtenue sur la parcelle "Baziège" avec la méthode E.L.E.C.T.R.E. TRI	132
Figure 48 : Cartes des données tests de résistivité électrique sur la parcelle "La Jaillière"	138
Figure 49 : Cartes des données tests de rendement et rendement moyenné sur la parcelle "La Jaillière"	139
Figure 50 : Cartes des données tests des canaux de l'image satellitaire sur la parcelle "La Jaillière"	139
Figure 51 : Composition colorée correspondant aux trois canaux de l'image Landsat sur la parcelle "La Jaillière"	140
Figure 52 : Cartes des données de référence concernant la granulométrie sur la parcelle "La Jaillière"	140
Figure 53 : Cartes des données de référence concernant les teneurs physico-chimiques sur la parcelle "La Jaillière"	141
Figure 54 : Cartes des données de référence concernant les pH sur la parcelle "La Jaillière"	141
Figure 55 : Cartes des données de référence concernant la profondeur de sol sur la parcelle "La Jaillière"	141
Figure 56 : Méthodologie de configuration et de validation des simulations	145
Figure 57 : Analyse en Composantes Principales des données tests et des données de référence sur la parcelle "La Jaillière"	152
Figure 58 : Zonage correspondant à la meilleure configuration de données pour la référence "pourcentage d'argile" pour un petit nombre de zones sur la parcelle "La Jaillière"	159
Figure 59 : Zonage correspondant à la meilleure configuration de données pour la référence "pourcentage d'argile" pour un grand nombre de zones sur la parcelle "La Jaillière"	160
Figure 60 : Zonage correspondant à la meilleure configuration de données pour la référence "teneur en P ₂ O ₅ " pour un petit nombre de zones sur la parcelle "La Jaillière"	163
Figure 61 : Zonage correspondant à la meilleure configuration de données pour la référence "teneur en P ₂ O ₅ " pour un grand nombre de zones sur la parcelle "La Jaillière"	164
Figure 62 : Zonage correspondant à la meilleure configuration de données pour la référence "pH H ₂ O" pour un petit nombre de zones sur la parcelle "La Jaillière"	166
Figure 63 : Zonage correspondant à la meilleure configuration de données pour la référence "pH H ₂ O" pour un grand nombre de zones sur la parcelle "La Jaillière"	167
Figure 64 : Triangle de définition de la texture d'un sol à partir des différents pourcentages granulométriques	175
Figure 65 : Carte de la donnée de référence "indice de battance" sur la parcelle "La Jaillière"	176
Figure 66 : Zonages correspondants à la simulation {résistivité électrique + rendement + image satellitaire} pour la référence "indice de battance" pour un petit et un grand nombre de zones sur la parcelle "La Jaillière"	177
Figure 67 : Cartes de zonage à petit nombre de zones obtenues avec l'image satellitaire	184

Figure 68 : Cartes de zonage à grand nombre de zones obtenues avec l'image satellitaire _ 184

Figure 69 : Cartes de zonage à petit nombre de zones obtenues avec les cartes de rendement
_____ 185

Figure 70 : Cartes de zonage à grand nombre de zones obtenues avec les cartes de rendement
_____ 185

Figure 71 : Cartes de zonage à petit nombre de zones obtenues avec les cartes de résistivité
électrique _____ 186

Figure 72 : Cartes de zonage à grand nombre de zones obtenues avec les cartes de résistivité
électrique _____ 186

Figure 73 : Cartes de zonage à grand nombre de zones obtenues avec l'ensemble des données
_____ 187

Figure 74 : Cartes de zonage à grand nombre de zones obtenues avec l'ensemble des données
_____ 187

Table des tableaux

Tableau 1 : Mise en relation des principales caractéristiques des problèmes décisionnels à référence spatiale avec certains avantages du recours à l'A.M.C.D.	84
Tableau 2 : Valeurs seuils des catégories et poids attribués aux critères.	92
Tableau 3 : Données retenues lors de l'étude initiale de pré-diagnostic	95
Tableau 4 : Données retenues lors de l'approche par A.M.C.D.	96
Tableau 5 : Poids et sens des préférences des critères de l'approche par A.M.C.D.	97
Tableau 6 : Catégories définies pour l'approche par A.M.C.D.	97
Tableau 7 : Coefficients de variation des onze données de référence pour chaque parcelle étudiée.	149
Tableau 8 : Indices d'autocorrélation de MORAN I des données tests sur les parcelles étudiées.	150
Tableau 9 : Corrélations entre les données de référence et les données tests sur la parcelle "La Jaillière".	153
Tableau 10 : Corrélations entre les données de référence et les données tests sur la parcelle "Imbault".	154
Tableau 11 : Corrélations entre les données de référence et les données tests sur la parcelle "Magneraud".	155
Tableau 12 : Corrélations entre les données de référence et les données tests sur la parcelle "Ouzouer".	156
Tableau 13 : Simulations retenues pour la référence "pourcentage d'argile"	158
Tableau 14 : Configurations de données recommandées pour la référence "Pourcentage d'argile" pour un petit nombre de zones	159
Tableau 15 : Configurations de données recommandées pour la référence "pourcentage d'argile" pour un grand nombre de zones	160
Tableau 16 : Simulations retenues pour la référence "teneur en P2O5"	162
Tableau 17 : Configurations de données recommandées pour la référence "teneur en P2O5" pour un petit nombre de zones	163
Tableau 18 : Configurations de données recommandées pour la référence "teneur en P2O5" pour un grand nombre de zones	164
Tableau 19 : Simulations retenues pour la référence "pH H2O"	165
Tableau 20 : Configurations de données recommandées pour la référence "pH H2O" pour un petit nombre de zones	166
Tableau 21 : Configurations de données recommandées pour la référence "pH H2O" pour un grand nombre de zones	167
Tableau 22 : Configurations de données recommandées pour la référence "Pourcentage de limon fin" pour un petit nombre de zones	168
Tableau 23 : Configurations de données recommandées pour la référence "pourcentage de limon fin" pour un grand nombre de zones	169
Tableau 24 : Configurations de données recommandées pour la référence "Pourcentage de limon grossier" pour un petit nombre de zones	169
Tableau 25 : Configurations de données recommandées pour la référence "pourcentage de limon grossier" pour un grand nombre de zones	169
Tableau 26 : Configurations de données recommandées pour la référence "Pourcentage de sable fin" pour un petit nombre de zones	170
Tableau 27 : Configurations de données recommandées pour la référence "pourcentage de sable fin" pour un grand nombre de zones	170

Tableau 28 : Configurations de données recommandées pour la référence "Pourcentage de sable grossier" pour un petit nombre de zones	171
Tableau 29 : Configurations de données recommandées pour la référence "pourcentage de sable grossier" pour un grand nombre de zones	171
Tableau 30 : Configurations de données recommandées pour la référence "teneur en CaO" pour un petit nombre de zones	172
Tableau 31 : Configurations de données recommandées pour la référence "teneur en CaO" pour un grand nombre de zones	172
Tableau 32 : Configurations de données recommandées pour la référence "teneur en K ₂ O" pour un petit nombre de zones	172
Tableau 33 : Configurations de données recommandées pour la référence "teneur en K ₂ O" pour un grand nombre de zones	173
Tableau 34 : Configurations de données recommandées pour la référence "pH KCl" pour un petit nombre de zones	173
Tableau 35 : Configurations de données recommandées pour la référence "pH KCl" pour un grand nombre de zones	173
Tableau 36 : Configurations de données recommandées pour la référence "profondeur de sol" pour un petit nombre de zones	174
Tableau 37 : Configurations de données recommandées pour la référence "profondeur de sol" pour un grand nombre de zones	174
Tableau 38 : Résultats des quelques simulations effectuées pour la référence "indice de battante" sur les parcelles "La Jaillière" et " Imbault"	177
Tableau 39 : Types de caractéristiques sur lesquelles fonder des entités homogènes en agriculture de précision.	193
Tableau 40 : Opérations culturales ayant fréquemment recours à la modulation et informations couramment utilisés pour définir les zones homogènes.	194
Tableau 41 : Partie du tableau de résultats des simulations pour la référence "pourcentage d'argile" sur la parcelle "La Jaillière"	195
Tableau 42 : Tableau de synthèse des résultats des parcelles pour la référence "pourcentage d'argile"	197
Tableau 43 : Tableau de synthèse des résultats des parcelles pour la référence "teneur en K ₂ O"	198
Tableau 44 : Tableau de synthèse des résultats des parcelles pour la référence "pH H ₂ O"	199

Depuis quelques années sont apparus en agriculture des outils techniques, tels que les capteurs de rendement ou l'imagerie satellitaire, qui ont permis de mettre en évidence des différences sensibles du comportement de certaines informations à l'échelle d'une parcelle. Qu'il s'agisse du rendement observé, de la profondeur de sol, de la teneur en nitrates ou d'autres informations, cette variabilité intraparcellaire a amené les acteurs de l'agriculture et de l'agronomie à percevoir différemment l'espace et à se poser de nouvelles questions.

L'espace, en passant du simple support de la culture à un ensemble de continuités et de discontinuités pouvant l'influencer, se voit pris en considération, et ce à une échelle très grande : la parcelle.

L'agriculture de précision s'appuie donc sur la gestion de la variabilité intraparcellaire mise en évidence afin d'améliorer l'efficacité des différentes opérations culturales. Ceci doit permettre à la fois de limiter les coûts d'exploitation et de réduire l'impact environnemental des pratiques agricoles.

Répondant aux exigences sociétales modernes en termes de souci de protection de l'environnement et de traçabilité, l'agriculture de précision se base sur certains des principes de l'agriculture durable et en constitue donc une des nombreuses approches.

Mais la pratique de l'agriculture de précision pose plusieurs problèmes complexes dont certains relèvent directement de problématiques spatiales. C'est spécialement le cas lors de la phase de décision définissant la stratégie de gestion de la variabilité intraparcellaire.

L'un de ces problèmes, en particulier, a été soumis à notre attention : la définition d'entités homogènes. Cette étape, visant à déceler au sein de la parcelle des zones aux caractéristiques jugées suffisamment proches pour être considérées comme semblables, nécessite une restitution précise de l'information géographique disponible mais aussi le respect de certaines contraintes techniques. Ces deux conditions doivent permettre de s'assurer de l'opérationnalité de la carte de préconisation finale. Celle-ci définit en effet les doses d'intrants à apporter en différents lieux de la parcelle et constitue donc la clé de voûte de la pratique de l'agriculture de précision.

Nous présenterons dans la première partie les grandes étapes de l'évolution de l'agriculture et les événements et principes qui ont amené à l'émergence de l'agriculture de précision. Cette dernière sera alors définie et présentée en détails, en particulier les problèmes qui se posent lors de sa pratique et les verrous limitant sa diffusion chez les exploitants agricoles.

Dans la seconde partie, nous verrons quelles relations l'agriculture de précision entretient avec les outils géographiques nécessaires à sa pratique : l'information géographique et les S.I.G. (Systèmes d'Information Géographique). Nous constaterons également quels peuvent être les apports de l'analyse spatiale et les limites des fonctionnalités proposées par les S.I.G. utilisés en agriculture de précision. Enfin, nous relèverons les contraintes de notre travail et feront le point sur les méthodes de définition d'entités homogènes existantes.

La troisième partie exposera l'une des deux approches que nous avons retenues pour mettre au point une méthode de définition d'entités homogènes : l'Aide MultiCritère à la Décision. Après avoir présenté les principales caractéristiques de cette dernière, nous en viendrons aux nombreux avantages que l'A.M.C.D. et les S.I.G. ont à être intégrés avant de donner deux exemples d'applications.

La quatrième partie présente en détail la méthode de définitions d'entités homogènes que nous avons mis au point en nous appuyant sur la statistique multivariée. La méthode de validation des cartes de zonage est également décrite.

La cinquième et dernière partie exposera une étude que nous avons menée en fin de doctorat et qui constitue un exemple d'application de notre méthode de zonage. Cette étude visait à tenter de reproduire, à partir d'informations géographiques peu coûteuses, la structure spatiale de données bien plus chères. Après avoir présenté le contexte, les objectifs et le déroulement de l'étude, nous révélerons une partie des résultats obtenus.

Première partie :

L'agriculture de précision

Signes distinctifs les plus visibles de l'agriculture de précision, certaines nouvelles technologies (en agriculture, s'entend) tels que les capteurs de rendement, le G.P.S. (Global Positioning System), l'imagerie satellitaire et les S.I.G. (Systèmes d'Information Géographique) en sont presque devenu les symboles. Pourtant, le perfectionnement et la diffusion croissante de ces technologies ne suffisent pas à expliquer l'émergence de l'agriculture de précision.

Plusieurs autres facteurs sont intervenus pour favoriser cette émergence. Et d'autres encore devront intervenir pour que l'agriculture de précision se diffuse plus largement.

Dans cette première partie, nous allons voir comment l'agriculture et la perception qu'en avait la société se sont modifiées jusqu'à aboutir à la situation actuelle : l'intégration croissante des notions de développement durable aux pratiques agricoles mais aussi à l'ensemble de la chaîne de production agro-alimentaire.

Nous présenterons alors ce mode de conduite des cultures fondé sur les principes de l'agriculture durable qu'est l'agriculture de précision.

1. De l'agriculture à l'agriculture de précision

1.1. Introduction

1.1.1. Une activité très ancienne...

Pendant des centaines de milliers d'années, durant le paléolithique, nos lointains ancêtres ont fondé leur subsistance sur la chasse, la pêche et la cueillette. Ce mode de subsistance n'était ni plus ni moins qu'un mode de vie, faisant d'eux des nomades. Au Néolithique, il y a près de dix mille ans, nos aïeux homo sapiens se sont mis à cultiver certaines plantes et à élever des animaux pour consommer leur chair : une véritable révolution. L'agriculture permit aux nomades de se libérer de la contrainte des migrations des troupeaux et donc de se sédentariser.

Aujourd'hui encore, notre mode de vie découle directement de certains choix qu'ont faits nos ancêtres à cette époque. L'agriculture, une des toutes premières activités à laquelle la race humaine se soit livrée, en fait partie.

Activité ancienne nécessitant de larges surfaces spécialement aménagées pour être mise en oeuvre, l'agriculture a exercé et exerce toujours une influence majeure sur l'espace de vie et les cultures des sociétés humaines.

1.1.2. Qui influence les espaces...

La principale phase d'expansion spatiale de l'agriculture remonte au XIII^{ème} siècle avec la mise en culture des plateaux et des versants pentus boisés. Par la suite, l'évolution des systèmes de culture a consisté en la réduction des surfaces en herbe au profit des cultures céréalières et industrielles que nous connaissons encore aujourd'hui [Gaillard 2004].

Néanmoins, la part relative de l'agriculture dans les activités humaines diminue depuis le XIX^{ème} siècle, de même que la part de la production agricole dans le P.I.B. (Produit Intérieur Brut) [Laurent 2005].

L'espace rural européen reste toutefois très majoritairement occupé par des activités agricoles qui contribuent à la production et à l'évolution des paysages. L'influence de l'agriculture sur le développement territorial est particulière, cette activité ne pouvant pas être déplacée du territoire dans lequel elle prend place [Laurent 2005]. De plus, l'agriculture, en substituant aux écosystèmes initiaux des "agrosystèmes", produit de la nature [Chaléard, Charvet 2004].

1.1.3. ...Et les sociétés

Etant une activité très ancienne et ayant rapidement concerné de vastes territoires, et donc des populations diverses, l'agriculture s'est développée dans des contextes extrêmement variés. Dans tous les cas, les hommes ont tenté d'adapter au mieux leurs techniques pour tirer ressources des conditions spécifiques des différents terroirs. Cette perpétuelle recherche d'adaptation des techniques à des conditions de production locale spécifiques a amené au développement de nombreuses formes d'activités agricoles. Les différences peuvent porter sur les modalités d'accès aux moyens de production, l'organisation du travail, les méthodes et la nature des productions et la finalité de l'activité [Laurent Thinon 2005].

Cette multitude de formes d'activités agricoles contribue à son tour à caractériser les paysages et les sociétés. Car l'impact culturel de l'agriculture, bien que moins visible, n'en est pas moins réel. Au-delà des différences dans les méthodes de travail et dans le rapport émotionnel des agriculteurs à leur terre, il est peu de personnes que l'évocation de l'agriculture laisse indifférentes. D'aucuns l'associeront à des souvenirs de vacances passées chez des grands-parents travaillant selon les pratiques ancestrales. D'autres penseront d'emblée à de vastes champs de céréales en openfield moissonnés par d'imposantes machines. D'autres encore l'associeront plutôt à la pollution... Les sensibilités et les expériences amènent à envisager individuellement ou partiellement les différents traits de ce qui reste pourtant une seule et même activité.

Des chefs de clan aux seigneurs locaux puis aux élus internationaux, les influences multiples de l'agriculture sur notre vie (mode de subsistance, ancienneté, impact sur les espaces de vie, les cultures et les sociétés) ont amené cette activité à être de plus en plus prise en compte dans les décisions des dirigeants. De nos jours, les marchés des produits agricoles sont très instables et les politiques agricoles jouent donc un rôle décisif dans la gestion de l'offre. En Union Européenne, l'activité agricole est l'une des composantes essentielles de la politique. La Politique Agricole Commune (P.A.C.) représente 43% du budget européen et l'élaboration de cette politique suscite bien des débats.

Selon Laurent, les fonctions sociales actuelles de l'agriculture sont telles qu'elle est l'objet de mesures politiques répondant à deux logiques distinctes. La première concerne le développement sectoriel par accumulation et reproduction de l'agriculture à partir des revenus de l'activité agricole. La seconde concerne le développement territorial dans le cadre duquel l'agriculture, en tant que système social localisé, doit permettre de garantir un certain niveau d'accumulation sur son territoire, à y maintenir des activités économiques et à protéger ses ressources naturelles [Laurent 2005]. Au final, l'agriculture se trouve chargée de multiples fonctions : productions de denrées, sécurité alimentaire, entretien du territoire, protection de l'environnement, maintien d'un tissu économique et social rural.

1.2. L'émergence et les limites de l'agriculture productiviste

1.2.1. La première révolution agricole

C'est en Angleterre au XVII^{ème} siècle qu'a débuté la première révolution agricole. Elle est caractérisée par la disparition des jachères (qui concernaient entre un tiers et la moitié des terres cultivées), la création de prairies artificielles et le développement des cultures fourragères. Ces dernières vont permettre un développement conséquent de l'élevage, qui, à son tour, va fournir le fumier et la force de travail nécessaires à un accroissement sensible des rendements des cultures céréalières [Chaléard, Charvet 2004].

Cette complémentarité entre les cultures et les élevages va faire la force de l'agriculture européenne dès la fin du XVIIIème siècle.

1.2.2. La seconde révolution agricole

La fin de la seconde guerre mondiale n'a pas résolu le problème de l'alimentation de la population qui sévissait depuis des décennies dans de nombreux pays. Peu d'entre nous n'ont pas entendu parler des tickets de rationnement par leurs parents ou leurs grands-parents. Face à cette pénurie, plusieurs nations, dont la France, ont eu pour objectif de produire des aliments en quantité suffisante et à des prix abordables pour toute la population.

Interviennent alors la motomécanisation et la généralisation de l'utilisation de produits chimiques tels les engrais azotés et les produits phytosanitaires. Ces nouveaux moyens de production agricole vont contribuer à augmenter dans d'importantes proportions les rendements à l'hectare. La motorisation du machinisme agricole va favoriser la réelle mise en place du remembrement de 1941. L'élevage, pour sa part, bénéficie d'aliments élaborés par de grands groupes industriels. C'est la seconde révolution agricole [Chaléard, Charvet 2004].

Le Plan Marshall, les progrès du machinisme agricole, la création de l'Institut National de Recherche Agronomique en 1946 vont permettre aux agriculteurs, dans un premier temps, d'intensifier leurs cultures et élevages afin d'augmenter sensiblement leurs rendements. En 1957, la Politique Agricole Commune (P.A.C.) est mise en place. Ses objectifs sont clairement l'augmentation de la productivité de l'agriculture et l'assurance d'un niveau de vie équitable aux agriculteurs. Elle contribuera grandement à la diffusion des bouleversements de l'agriculture à l'échelle européenne.

C'est ainsi que se généralise l'agriculture productiviste (aussi appelée agriculture intensive). Elle se caractérise par la recherche de la maximisation de la production par rapport aux facteurs de production, qu'il s'agisse de la main d'œuvre, du sol ou des autres moyens. Il est possible de chercher à maximiser la productivité du sol, en augmentant les facteurs humains et financiers (cas des élevages hors-sol ou des cultures en serre). A l'inverse, il est possible de chercher à réduire la main d'œuvre par le recours à la mécanisation sur de grandes surfaces (cas de la céréaliculture des pays développés). Ce mode de production implique, en particulier, le recours massif aux intrants.

En ce qui concerne les grandes cultures, la mécanisation croissante et cette recherche de la productivité maximale vont conduire les agriculteurs à traiter au même moment et de la même manière des parcelles culturales¹ de plus en plus grandes. Pourtant, la taille croissante des parcelles a pour conséquence une augmentation sensible de l'hétérogénéité de leurs caractéristiques [Jullien, Huet 2005].

Alors que l'agriculture se mécanise, le développement de la recherche agronomique, en particulier en génétique à travers les biotechnologies, permet d'effectuer des croisements pour sélectionner des espèces plus compétitives. Toutefois, il ne sera réellement question d'Organismes Génétiquement Modifiés (O.G.M.) que dans les années 1970, lorsque des plantes disposant d'attributs spécifiques inédits (résistance à un herbicide, un insecte, une maladie...) feront leur apparition.

L'agriculture productiviste a permis de diminuer les coûts de production, principalement par le biais d'une augmentation sensible des rendements. Ainsi, à la fin des années 1990, les rendements de blé et de maïs pourront atteindre, dans certaines régions, les 100 quintaux à l'hectare contre moins de 15 au début du siècle.

¹ Mathieu, Thion et Havet définissent une parcelle culturale "comme une portion de territoire, d'un seul tenant, qui porte la même culture, et sur laquelle est appliqué un itinéraire technique homogène." Ils précisent "qu'en production végétale, la parcelle culturale constitue un découpage essentiel de l'espace" [Mathieu, Thion, Havet 2005].

Dans le même temps, l'industrialisation, l'amélioration des techniques d'élevage et d'agriculture, les innovations dans les procédés de conditionnement et de conservation des aliments vont contribuer à accentuer cette optique productiviste et à l'étendre du système de production agricole à l'ensemble du domaine agro-alimentaire.

Peu à peu, au fil des siècles, la nature et son influence sur les conditions de production auront donc été amoindries, voire asservies, jusqu'à être totalement ignorées dans les années 1960 [Chaléard, Charvet 2004].

1.2.3. Les limites de l'agriculture productiviste

La pérennité de ces progrès et de ce mode de production intensif a eu, et a encore, de fâcheuses conséquences dans des domaines ayant souvent trait à la santé publique. La qualité de l'environnement, la qualité des produits, les conditions de vie des animaux d'élevage sont autant d'aspects qui furent longtemps négligés. De nombreux milieux naturels sont pollués. La Bretagne est particulièrement touchée avec les "marées vertes" (voir Figure 1). La prolifération de l'ulve, algue opportuniste à la croissance très rapide, sur les plages est liée à trois facteurs : l'ensoleillement, la présence de nutriments et le faible déplacement de la masse d'eau. 95% des nutriments azotés nécessaires à l'épanouissement de cette algue sont apportés par le lessivage des sols gorgés de l'azote des fertilisants agricoles. La Bretagne est également soumise à la pollution de l'eau par les nitrates issus des effluents de l'élevage porcin. Par ailleurs, certains sols réputés fertiles sont épuisés. Plusieurs affaires concernant le secteur de la production alimentaire ont alerté l'opinion publique : bœuf aux hormones, maladie de la "vache folle", poulets à la dioxine... En dehors de la santé publique, l'évènement le plus marquant a été la diminution importante de la population agricole suite aux gains de productivité, avec toutes les conséquences territoriales qu'on peut constater.



Figure 1 : Exemple de "marée verte" près de la Pointe de Guettes, en Bretagne

Pendant de nombreuses années, l'évolution du marché alimentaire national et mondial n'a pas été prise en compte. L'agriculture européenne est devenue une activité parmi d'autres, là aussi non sans conséquences. Ainsi, la mise en place des quotas laitiers en 1984 est consécutive à une surproduction dont le taux atteint, pour le lait en poudre, 400%. De même, des terres sont laissées en jachères pour limiter les volumes de production. La France est ainsi devenue le second exportateur mondial sur le plan agro-alimentaire.

Si l'agriculture productiviste fut indubitablement nécessaire dans l'immédiat après-guerre pour assurer la subsistance des populations et si elle a indéniablement contribué à l'essor économique de l'Europe, elle n'en est pas moins, pour les raisons précédemment évoquées, de moins en moins bien perçue par la population. D'autres modes de production existent, parfois depuis plusieurs décennies, et connaissent un réel essor depuis quelques années, comme l'agriculture biologique et l'agriculture raisonnée.

1.3. L'agriculture durable

1.3.1. Définition et objectifs

L'agriculture durable est la mise en application dans le domaine agricole de la notion de développement durable définie en 1992 à Rio [Thévenet 2005]¹. Le développement durable vise à *"répondre aux besoins du présent sans compromettre la capacité des générations futures à répondre à leurs propres besoins"*. Il s'agit donc d'opérer les différentes productions agricoles tout en respectant les limites écologique, économique et sociale afin d'assurer la durabilité dans le temps de ces productions.

Les principes de l'agriculture durable se fondent sur le fait que les ressources sont limitées et que celles-ci doivent donc être utilisées raisonnablement pour favoriser le bien-être économique et social. L'agriculture durable doit viser à améliorer la soutenabilité du système de production, en permettant un développement économique équitable et ce de façon respectueuse de l'environnement [Wikipedia 2005]².

Comme le souligne Chaléard et Charvet, c'est paradoxalement à l'heure où la science permet aux exploitants agricoles de se libérer des contraintes des milieux naturels comme jamais auparavant qu'émerge la prise en considération de ces milieux [Chaléard, Charvet 2004].

Selon Thévenet, cette notion recouvre les réalités suivantes à l'échelle d'une exploitation [Thévenet 2005] :

- Le système de production doit permettre de sécuriser les revenus des exploitants face aux aléas climatiques et aux variations des marchés.
- La dimension écologique doit s'entendre au sens large en incluant la préservation des ressources naturelles (énergie, sol...), de la qualité de l'eau ou de l'air, de la biodiversité et des paysages.
- Quant à la dimension sociale, elle est certainement la plus difficile à appréhender. Elle se réfère à des valeurs telles que la qualité des produits, la solidarité, la citoyenneté ou la qualité de vie.

L'agriculture durable ne fait pas l'objet d'une législation précise, en tout cas en France. Néanmoins, cette façon d'aborder l'agriculture se retrouve dans plusieurs modes de production dont certains disposent, eux, de cadres réglementaires bien formalisés.

¹ [Thévenet 2005] : <http://www.arvalisinstitutduvegetal.fr>

² [Wikipedia 2005] : <http://fr.wikipedia.org>

1.3.2. Deux exemples de modes de production relevant de l'agriculture durable

Il est possible d'aborder l'agriculture durable de nombreuses façons différentes. Parmi celles-ci, deux approches sont à ce jour relativement bien connu du grand public et font l'objet de réglementations précises en France : l'agriculture biologique et l'agriculture raisonnée.

1.3.2.1. *L'agriculture biologique*

L'agriculture biologique est un mode de production recourant à des pratiques culturales et d'élevage soucieuses du respect des équilibres naturels. L'utilisation des produits chimiques de synthèse et des O.G.M. est donc proscrite, tandis que celle des intrants est limitée. De même, les modes de transformation privilégient la mise en valeur des caractéristiques naturelles des produits.

Ce mode de production permet d'expérimenter en situation réelle des pratiques innovantes respectueuses de l'environnement et susceptibles d'être développées plus largement dans les autres modes de production agricoles.

La charte éthique de l'agriculture biologique, publiée en 1972 par l'I.F.O.A.M. (International Federation of Organic Agriculture Movements), définit treize objectifs écologiques, sept objectifs sociaux et humanistes, et cinq objectifs économiques. La simple dénomination des groupes d'objectifs et leur ordre dans la charte en disent long sur la philosophie du document et son positionnement par rapport à l'agriculture productiviste.

La première officialisation des principes de l'agriculture biologique en France date de la réglementation de 1980 concernant la Loi d'Orientation Agricole (L.O.A.). Mais le terme "agriculture biologique" n'apparaît qu'en 1991 dans le règlement européen CEE 2092/91. Depuis, la Surface Agricole Utilisée (S.A.U.) dédiée à ce mode de production n'a cessé de croître en France (voir Figure 2).

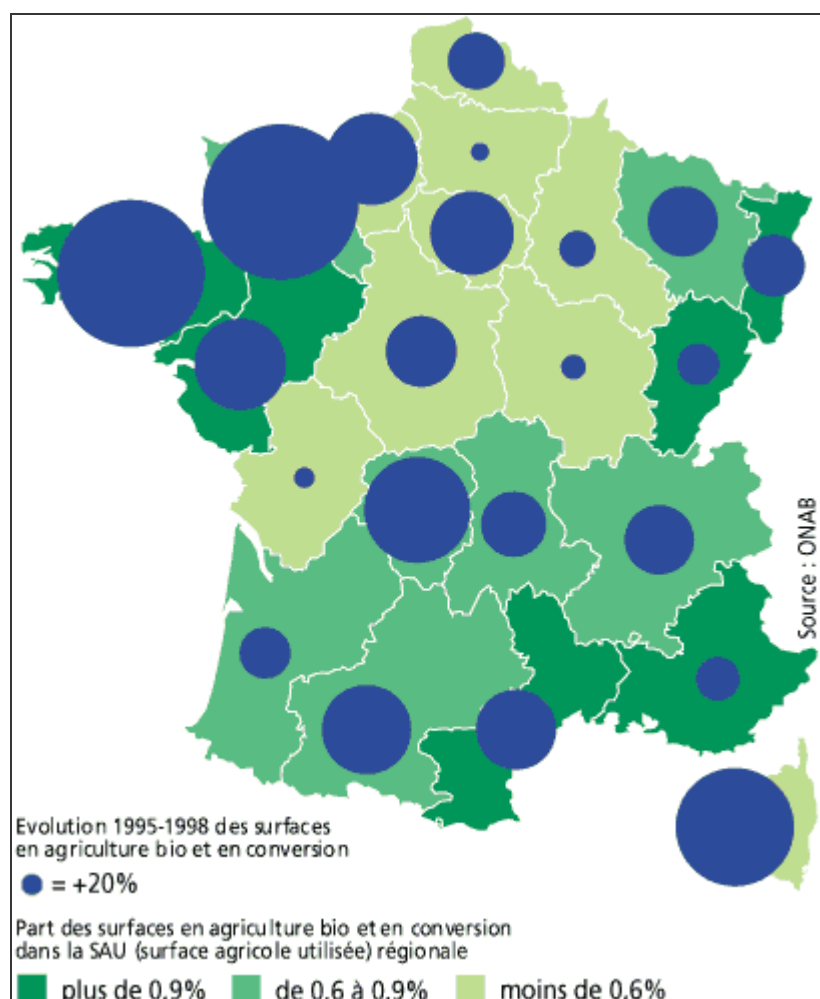


Figure 2 : Evolution des surfaces en agriculture biologique entre 1995 et 1998

De nos jours, l'agriculture biologique fait l'objet de règlements spécifiques à la fois aux niveaux européen et français. Les règles de production biologiques sont consignées dans des cahiers des charges établis par les professionnels et homologués par l'État français et/ou par l'Union Européenne :

- Le règlement CEE 2092/91, comportant des règles européennes pour les productions végétales et animales biologiques.
- Le cahier des charges français complémentaire (CC REPAB F) aux règles européennes pour les productions animales.
- Les guides de lectures officiels français qui complètent et explicitent ces textes.

Tous ces documents ont été mis à jour ou complétés par divers textes au cours de l'année 2004.

Le respect de ces règles permet d'obtenir la certification des produits. Des contrôles sont donc effectués par des organismes certificateurs indépendants qui contrôlent chaque unité de production et de transformation de produits biologiques. Ces contrôles, payés par l'opérateur, sont effectués au minimum une fois par an et autant que nécessaire en rapport avec la complexité du processus de fabrication, tant pour la production que pour la transformation.

Dans la mesure où les divers règlements ont pour objectif d'attribuer aux produits la certification "agriculture biologique", de nombreuses et complexes dispositions concernent les produits d'importation.

Les contrôles et la certification sont pour le consommateur la garantie que les produits qu'il achète auront été élaborés en respectant des cahiers de charges rigoureux.

1.3.2.2. L'agriculture raisonnée

Le concept est né d'une démarche conjointe de la F.N.S.E.A. (Fédération Nationale des Syndicats d'Exploitants Agricoles), de l'industrie agro-alimentaire, de l'industrie agro-pharmaceutique et de la grande distribution. En France, c'est le groupe F.A.R.R.E. (Forum de l'Agriculture Raisonnée Respectueuse de l'Environnement) qui se charge de promouvoir l'agriculture raisonnée en insistant sur sa compétitivité face à l'agriculture productiviste tout en soulignant les bénéfices environnementaux qui l'accompagnent. Des associations comparables à F.A.R.R.E. existent dans six autres pays européens. Elles se sont regroupées au sein de l'E.I.S.A. (European Initiative for Sustainable development in Agriculture).

L'agriculture raisonnée est un mode de production agricole dont l'objectif premier est d'optimiser le résultat économique tout en gérant au mieux les quantités d'intrants utilisées, notamment les substances chimiques, et ce dans le but de limiter leur impact sur l'environnement. L'agriculture raisonnée s'appuie sur la notion de plafond d'efficacité. Celle-ci considère qu'au delà d'un certain seuil, les moyens mis en œuvre pour augmenter le rendement coûtent trop cher économiquement et écologiquement et ne sont donc plus rentables.

Bien que prudente en matière d'utilisation d'intrants, l'agriculture raisonnée n'en admet pas moins l'utilisation, de même pour la plupart des techniques de l'agriculture intensive : engrais chimiques, aliments énergétiques pour les animaux, sélection des plantes et des animaux, produits phytosanitaires de synthèse pour la protection des plantes, antibiotiques pour les animaux. Cependant, ces techniques sont ici mises à contribution dans un cadre de gestion globale de l'exploitation agricole¹ prenant en compte l'ensemble de l'environnement physique, social, économique, ainsi que les spécificités biologiques du territoire (voir Figure 3).

¹ La définition d'une exploitation agricole émane du décret 2000-60 et de l'arrêté du 24 janvier 2000 prescrivant le recensement : "*unité économique qui répond simultanément aux trois conditions suivantes* :

- Elle produit des produits agricoles (précisés sur une liste positive pour les cas litigieux).
- Elle atteint ou dépasse une certaine dimension (minimum un hectare de superficie agricole utile (S.A.U.) ou moins pour les cultures spécialisées, les productions hors-sol, etc...).
- Elle est soumise à une gestion courante indépendante de toute autre unité".

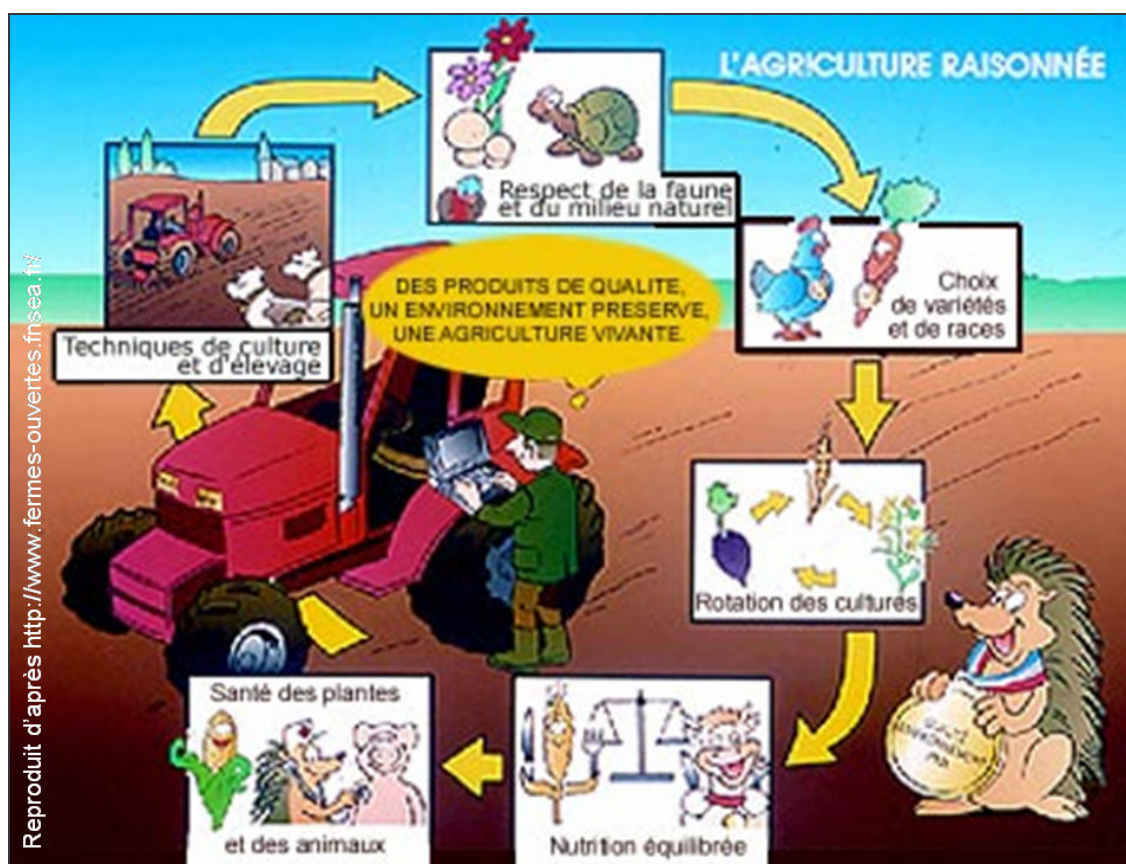


Figure 3 : Principe de l'agriculture raisonnée

En France, le décret n° 2002-631 du 25 avril 2002 et l'arrêté du 30 avril 2002 offrent un cadre officiel et réglementé à la pratique de l'agriculture raisonnée. A l'échelle européenne, il n'existe pas encore de législation commune sur ce sujet.

Le décret fixe un référentiel permettant à une exploitation agricole d'obtenir une qualification au titre de l'agriculture raisonnée. Ce référentiel porte sur le respect de l'environnement, la maîtrise des risques sanitaires, la santé et la sécurité au travail et le bien-être des animaux. Les exigences concernent plus précisément :

- L'accès de l'exploitant et de ses salariés à l'information et à la formation nécessaires à la conduite de l'exploitation agricole.
- La mise en œuvre d'un système d'enregistrement et de suivi des opérations effectuées et des produits utilisés pour les besoins des cultures et des animaux.
- La maîtrise des intrants agricoles ainsi que des effluents et des déchets produits par l'exploitation.
- L'usage justifié de moyens appropriés de protection des cultures et de la santé des animaux de l'exploitation.
- L'équilibre de la fertilisation des cultures.
- La mise en œuvre de pratiques culturales permettant la préservation des sols et limitant les risques de pollutions.
- La participation à une gestion économe et équilibrée des ressources en eau.
- La prise en compte de règles dans les domaines de la sécurité sanitaire et de l'hygiène.

- La prise en compte des besoins des animaux en matière d'alimentation et de bien-être.
- La contribution de l'exploitation à la protection des paysages et de la diversité biologique.

L'arrêté définit précisément quatre-vingt dix-huit points dans quatorze catégories se rapportant à tous les aspects d'une gestion globale : connaissance de l'exploitation, traçabilité des pratiques, santé et sécurité au travail, gestion des sols, fertilisation minérale et organique, protection des cultures, irrigation, identification des animaux, santé des animaux, alimentation des animaux, bien-être des animaux, hygiène, gestion des déchets de l'exploitation, paysage et biodiversité.

Les textes laissent une certaine liberté d'appréciation à l'agriculteur, dans une philosophie de responsabilité et de gestion raisonnable qui caractérise ce type d'agriculture. De ce fait, celle-ci attire de nombreux agriculteurs qui peuvent ainsi satisfaire la demande de certification de leurs clients sans devoir modifier trop profondément leurs méthodes de travail.

En ce sens, l'agriculture raisonnée constitue une sorte de compromis entre l'agriculture productiviste qui prévalait jusqu'alors et l'agriculture biologique, aux normes particulièrement sévères.

Même si elle n'intègre pas ses principes sociaux, l'agriculture raisonnée demeure l'un des principaux moyens pour les exploitants agricoles de mettre en œuvre l'agriculture durable [Thévenet 2005]¹.

2. L'agriculture de précision, une approche de l'agriculture durable

2.1. L'émergence de l'agriculture de précision

Selon Jullien et Huet, l'agriculture de précision est apparue en 1983 lorsque deux chercheurs de l'Université du Minnesota ont utilisé une carte d'acidité pour moduler les apports d'amendement calciques [Jullien, Huet 2005].

Au début des années 1990, plusieurs facteurs ont contribué à l'émergence de l'agriculture de précision aux Etats-Unis [Robert 2000] [Jullien, Huet 2005]. Les principaux étaient :

- Le besoin d'améliorer la gestion des fermes.
- La diffusion et l'utilisation de capteurs de rendement.
- La disponibilité d'ordinateurs relativement bon marché et fiables.
- Le développement d'outils complémentaires tels les Systèmes d'Information Géographique (S.I.G.) et le Global Positionning System (G.P.S.).

A cette époque, les technologies ont permis l'acquisition, la gestion et l'utilisation de vastes bases de données géographiques sur l'état des sols et des cultures. Ces bases de données ont permis la mise en évidence du fait qu'une parcelle n'a pas les mêmes caractéristiques en tous points et est donc soumise à variabilité intraparcellaire.

L'ensemble de ces facteurs a amené l'idée de moduler les pratiques culturales en fonction de cette variabilité pour améliorer les revenus et limiter les risques de pollution.

¹ [Thévenet 2005] : <http://www.arvalisinstitutduvegetal.fr>

Ce principe, fondement de l'agriculture de précision, est de nos jours mis à l'essai ou effectivement pratiqué pour la majorité des cultures importantes dans le monde. Son adoption reste cependant très variable en fonction des pays et des cultures.

L'agriculture de précision a atteint l'Europe au début des années 1990. Elle y est surtout abordée par son aspect environnemental.

Jusqu'à aujourd'hui, ce sont 31 000 capteurs de rendement qui ont été vendus aux Etats-Unis pour une surface totale de 10 000 000 d'hectares cartographiés. En France, les premières cartes de rendement datent de 1997 [Jullien, Huet 2005]. En fonction des estimations, on peut considérer que la France compte entre 100 et 150 systèmes de cartographie de rendement mis en place chez les agriculteurs [Boisgontier 2000] [Jullien, Huet 2005]. Boisgontier constatait, en 2000, que le principe de l'agriculture de précision avait du mal à s'imposer en France [Boisgontier 2000].

On notera comme l'affirment Jullien et Huet, que le principe d'adaptation des pratiques culturales à la variabilité intraparcellaire a toujours été d'usage pour certaines cultures peu mécanisées comme les cultures fruitières [Jullien Huet 2005]. En effet, dans ce cas, les traitements sont appliqués individuellement à chaque plant, ce qui induit l'adaptation implicite de ces traitements par l'exploitant agricole.

2.2. Définition de l'agriculture de précision

Il existe de nombreuses définitions de l'agriculture de précision. On peut toutefois les distinguer selon leur orientation technologique ou agronomique. Faire la distinction entre ces deux orientations nous semble important afin de ne pas confondre les outils avec les finalités de la démarche.

L'orientation technologique de certaines définitions est due au fait que l'agriculture de précision est apparue suite à un progrès technique, comme nous l'avons vu en section 2. Elles considèrent donc que la simple utilisation de capteurs de rendement, de G.P.S. ou de systèmes de cartographie suffit pour "pratiquer" l'agriculture de précision. La grande majorité de ces définitions émanent de constructeurs ou prestataires fournissant ces technologies [R.A.C.A.L. 1999] ou de documents largement diffusés (revues, sites Web...) [Johanssen 1998] [Lefaucheur 1999].

Les définitions à orientation agronomique émanent en majorité de la communauté scientifique agricole et/ou agronomique [Doerge 2001], [Boisgontier 2000], [Thévenet 1999], [Fridgen, Fraisse, Kitchen, Sudduth 2000] [Hoskinson, Hess, Hempstead 1998] [Jullien, Huet 2005] [Laurent, Thion 2005]. Certains auteurs de cette communauté regrettent d'ailleurs l'orientation technologique de certaines définitions, la jugeant trop réductrice [Gay 1997] [Boisgontier, Ariaux 1998]. Cependant, leur propos n'est pas de nier l'implication de la technologie dans l'agriculture de précision mais de considérer son utilisation pour ce qu'elle est, un moyen, et non comme une fin en soi.

La définition de Jullien et Huet est représentative de cette approche [Jullien, Huet 2005]. Pour ces auteurs, *"l'agriculture de précision vise à adapter les pratiques culturales à l'hétérogénéité intraparcellaire du sol et des cultures"*. Et ils affirment plus loin que *"le développement de l'agriculture de précision s'appuie sur des progrès technologiques récents, notamment dans le domaine de l'informatique et de l'analyse spatiale"*. Bien que leur définition de l'agriculture de précision ait une orientation agronomique, ces auteurs reconnaissent que l'orientation technologique reste encore très présente dans l'acception générale actuelle du terme. Selon eux, *"le terme générique "agriculture de précision" désigne l'ensemble des techniques culturales fondées sur l'utilisation des nouvelles technologies de mesure et de traitement de l'information spatialisée"*.

Dans un article de Delpech et Biaggini, Boissongontier définit l'agriculture de précision comme "la prise en compte des variations intraparcélaires dans l'itinéraire [technique¹, N.D.A.], une démarche agronomique basée sur des nouvelles technologies". Il résume le principe en ces termes : "Apporter la bonne dose, au bon endroit et au bon moment" [Delpech, Biaggini 2000].

Gay affirme que "la gestion de la variabilité intraparcéliaire est au cœur de l'agriculture de précision. Elle en constitue le fondement et la justification. L'agriculture de précision regroupe l'ensemble des outils et des connaissances qui permettent :

- De collecter les données à une échelle intraparcéliaire.
- De les analyser et de les comprendre.
- D'orienter une action dans l'itinéraire technique, dont les effets se situent au même niveau d'échelle." [Gay 1997]

Laurent et Thinon considèrent que le principe de l'agriculture de précision est de "*renouer avec une tradition ancienne d'ajustement des pratiques culturales à l'hétérogénéité intraparcéliaire*" [Laurent Thinon 2005]. Ces auteurs considèrent, eux aussi, la nécessité de la mobilisation des nouvelles technologies dans cette tâche. Ils insistent enfin sur l'importance des compétences que les agriculteurs devront mettre en œuvre dans le cadre de la pratique de l'agriculture de précision.

Considérant que la technologie reste un outil, notre approche de l'agriculture de précision sera agronomique, même si notre travail vise, à sa façon, à améliorer certaines technologies.

2.3. Les enjeux de l'agriculture de précision

De nombreux auteurs s'accordent sur les deux enjeux majeurs de l'agriculture de précision, à savoir un gain économique et la préservation de l'environnement. [Malzer 1996] [Chung, Sudduth, Drummond, Kitchen 2001] [Doerge 2001] [Gay 1997] [Van Alphen, Stoorvogel 1999] [Boissongontier 1997b] [Jullien, Huet 2005].

Ainsi, Thévenet [Thévenet 1997] [Thévenet 1999] considère que l'agriculture de précision doit être pratiquée lorsqu'elle est agronomiquement nécessaire, techniquement possible et économiquement intéressante. Il espère également qu'elle permettra de passer du "toujours plus" au "toujours mieux".

2.3.1. Un gain économique

Le recours à une gestion améliorée de l'itinéraire technique, en particulier pour les semis, le désherbage et les intrants, doit permettre à l'exploitant agricole d'accroître la qualité des produits et le rendement de la parcelle et donc constituer un gain économique. En particulier, l'ajustement des intrants aux potentialités locales de production doit permettre de diminuer les coûts de production.

Cependant, cet impact économique reste encore à démontrer aujourd'hui. Selon Gay [Gay 1997], "la démonstration de l'intérêt économique de l'agriculture de précision la fera évoluer d'un développement poussé par une technique à une demande effective des utilisateurs".

Certains résultats récents tendent à prouver qu'il y a bel et bien un gain économique [Masson 2000] [Baratte 2002] (voir aussi en section 2.5).

¹ Jouanon, Souchère et Tichit définissent l'itinéraire technique comme "l'enchaînement logique et ordonné d'opérations culturales appliquées à une culture en vue d'en obtenir une production" [Jouanon, Souchère, Tichit 2005].

2.3.2. La préservation de l'environnement

La gestion raisonnée des intrants, en fonction des besoins réels pour les cultures, permet de limiter l'impact des pratiques culturales sur l'environnement, en particulier pour l'irrigation et la fertilisation.

Cet objectif est à l'heure actuelle celui dont les résultats sont les plus tangibles et les plus évidents. En effet, la réduction des intrants entraîne de facto une réduction des résidus non consommés par la culture dans le sol, et donc une réduction de l'assimilation de ces résidus par l'environnement.

2.3.3. D'autres objectifs connexes

Certains auteurs considèrent que l'agriculture peut servir d'autres objectifs en plus des deux précédemment cités.

Ainsi, Boigontier dans [Delpech, Biaggini 2000], Masson [Masson 2000], Bichat [Bichat 1999] et Mérour [Mérour 2002] estiment que la traçabilité est un autre objectif, ou une autre conséquence selon le cas, de l'agriculture de précision. L'obligation induite par la gestion de la variabilité intraparcellaire de garder une base de données à jour spatialement et temporellement se prête également à un suivi précis des itinéraires techniques réalisés. Dans l'article de Masson, la remarque concernant l'intérêt de l'agriculture de précision pour la traçabilité émane directement d'un exploitant agricole. Celui-ci considérerait qu'au delà des avantages économiques et environnementaux qu'il avait effectivement relevés, c'était surtout la traçabilité qui constituait "*le meilleur retour sur investissement*" [Masson 2000].

Thévenet parle, quant à lui, d'améliorer la compétitivité de l'exploitant agricole, cet objectif étant très lié à l'objectif de gain économique [Thévenet 1997]. Cette augmentation de la productivité peut passer entre autres, par une augmentation des rendements [Mérour 2002].

Jullien et Huet considèrent que la mise en œuvre des méthodes de l'agriculture de précision peut s'avérer profitable sans devoir nécessairement aller jusqu'à la modulation. Selon eux, beaucoup d'agriculteurs utilisent l'information spatialisée sans faire de modulation. Cette information spatialisée peut en effet servir à effectuer un diagnostic agronomique ou environnemental sur une parcelle, permettant d'identifier des zones de verse ou de maladie. Elle peut également permettre de comparer plusieurs techniques sur une parcelle. L'utilisation de l'information géographique à ces fins permet déjà de répondre aux exigences sociétales en termes de respect de l'environnement, de qualité des produits et de traçabilité [Jullien, Huet 2005].

2.4. Une approche de l'agriculture durable

Nous avons vu en section 1.3.1, en section 2.2 et surtout en section 2.3, que les objectifs de l'agriculture durable et de l'agriculture de précision sont très proches, voire se recoupent au moins en partie. De plus, ne tenant pas compte des aspects sociaux de l'agriculture durable, l'agriculture de précision se rapproche sur ce point de l'agriculture raisonnée.

Pourtant, il existe une différence capitale entre ces deux approches de l'agriculture durable. En effet, depuis que l'agriculture raisonnée a fait l'objet d'une législation spécifique en 2002 (voir section 1.3.2.2), les liens qui unissent les deux approches sont moins pratiques et légaux qu'idéologiques. Si l'on peut s'attendre à ce que les exploitations agricoles ayant reçu la qualification "agriculture raisonnée" ne pratiquent pas l'agriculture de précision, les agriculteurs pratiquant l'agriculture de précision ne sont pas qualifiés "agriculture raisonnée" non plus. Et ce d'autant plus que certains exploitants agricoles pratiquant l'agriculture de précision ne le font que sur certaines parcelles culturales, alors que la qualification "agriculture raisonnée" doit concerner toute l'exploitation.

Par conséquent, l'agriculture de précision n'est pas, comme le précisent Jullien et Huet un mode de conduite des cultures (comme le sont l'agriculture biologique et l'agriculture raisonnée) mais *"un outil supplémentaire de raisonnement des itinéraires techniques qui s'appuie sur les outils classiques de l'agronomie"* [Jullien, Huet 2005]. En s'efforçant de répondre aux exigences actuelles en termes de gestion économique, de protection de l'environnement et de traçabilité, l'agriculture de précision respecte scrupuleusement certains principes de l'agriculture durable. On peut donc la considérer, sinon dans la lettre, du moins dans l'esprit, comme une des approches de l'agriculture durable.

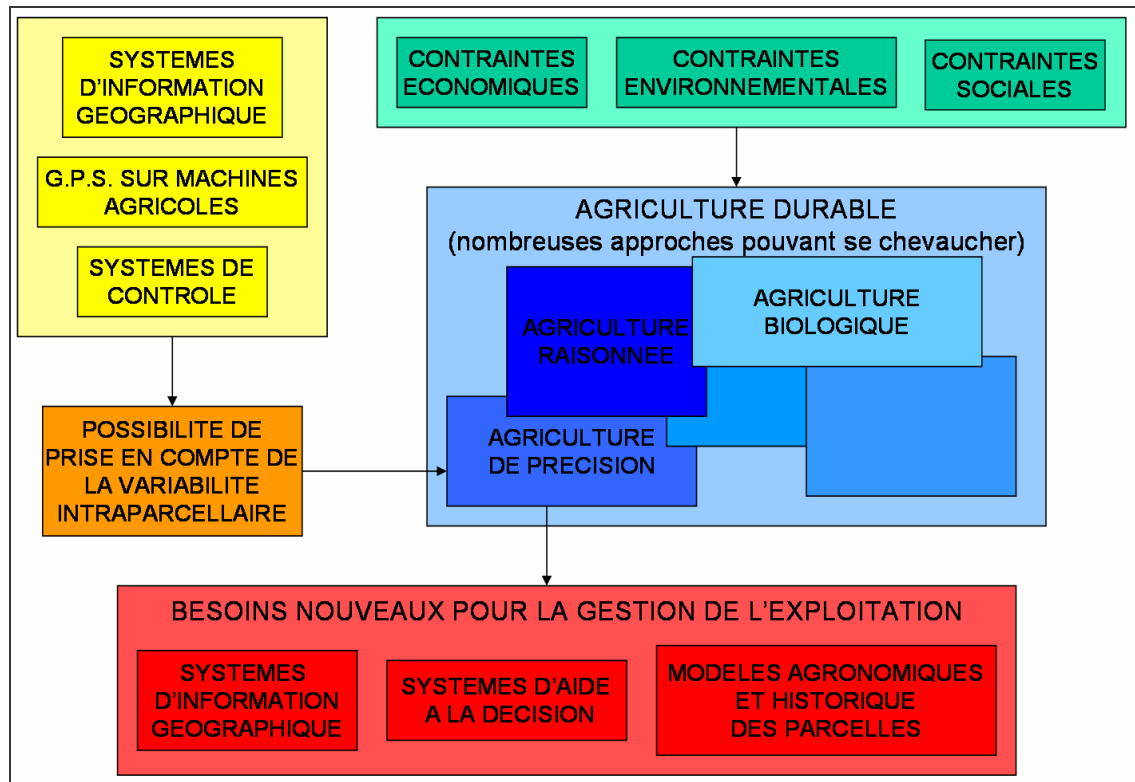


Figure 4 : Contexte de l'agriculture de précision

2.5. Un exemple de mise en œuvre de l'agriculture de précision

Dans un article de 2002, Baratte fait état des conclusions concernant l'agriculture de précision de M. Benoît Raffin, agriculteur dans la région de Béthune (Pas-de-Calais) [Baratte 2002].

Tout d'abord, cet agriculteur affirme que l'acquisition d'un G.P.S. et d'un système de cartographie du rendement lui a permis de confirmer ses observations empiriques. Il a pu constater, en particulier, que le rendement pouvait varier de 40 à 140 quintaux par hectare d'un bout à l'autre d'un même champ.

Mais surtout, en cinq ans, la mise en œuvre du mode opératoire propre à l'agriculture de précision a permis à Benoît Raffin d'augmenter son rendement moyen de 8 quintaux par hectare sur l'ensemble de son exploitation de 250 hectares tout en réduisant des deux tiers les apports d'engrais phospho-potassiques.

Cet agriculteur est parvenu à améliorer l'efficacité des intrants tout en diminuant les coûts de production. Mais selon ses propres termes, sa *"satisfaction première est d'avoir réussi, à partir d'impératifs économiques, à faire de l'environnemental"*. Dans ce cas précis, la quasi-totalité des objectifs de l'agriculture de précision tels que nous les avons définis en section 2.3 sont atteints.

3. Types de modulation et étapes de la pratique de l'agriculture de précision

L'agriculture de précision vise à moduler les pratiques culturales en fonction de la variabilité intraparcellaire. Cette modulation peut prendre deux formes dites directe et indirecte.

On notera cependant que, dans la très grande majorité des articles et ouvrages, l'agriculture de précision est implicitement considérée comme s'appuyant sur une modulation indirecte. Les définitions de l'agriculture de précision (voir section 2.2) et les descriptions de son mode opératoire en trois étapes sont très claires sur ce point. C'est également sur ce type de modulation que s'appuie notre travail.

3.1. La modulation directe

La modulation directe ou modulation en temps réel vise à effectuer les mesures sur la parcelle, leurs analyses et le calcul des doses à apporter au cours d'un seul et unique passage (par opération culturale). Dans ce cas, un système adapté à l'opération culturale récolte l'information par le biais d'un capteur, l'analyse et se charge enfin d'agir directement sur la machine agricole pour moduler l'application.

C'est le cas du N-sensor, commercialisé par Yara¹. Cet appareil mesure les besoins en azote de la culture et module les apports en conséquence. Il permet ainsi d'éviter de laisser des endroits dans la parcelle sous-fertilisés (conduisant à des pertes de rendement) et d'autres sur-fertilisés (synonymes de gaspillage d'azote néfaste pour l'environnement et augmentant le risque de verse).

Le N-Sensor réalise en temps réel trois opérations (voir Figure 5 et Figure 6) :

- Mesure de la réflectance du couvert végétal : elle permet d'évaluer les quantités de chlorophylle et la biomasse du couvert végétal et donc d'estimer l'état de nutrition azotée de la culture.
- Traitement du signal optique et traduction en une dose d'épandage.
- Dialogue et envoi des instructions de dose à l'épandeur d'engrais.

Les variations de l'intensité lumineuse sont intégrées en permanence et permettent de travailler en toutes conditions (temps couvert, temps clair, alternance de nuages...). Depuis le terminal embarqué à bord du tracteur, l'opérateur contrôle et maîtrise avec facilité toutes les étapes de l'utilisation de l'équipement.

¹ <http://fert.yara.fr/fr/index.html>

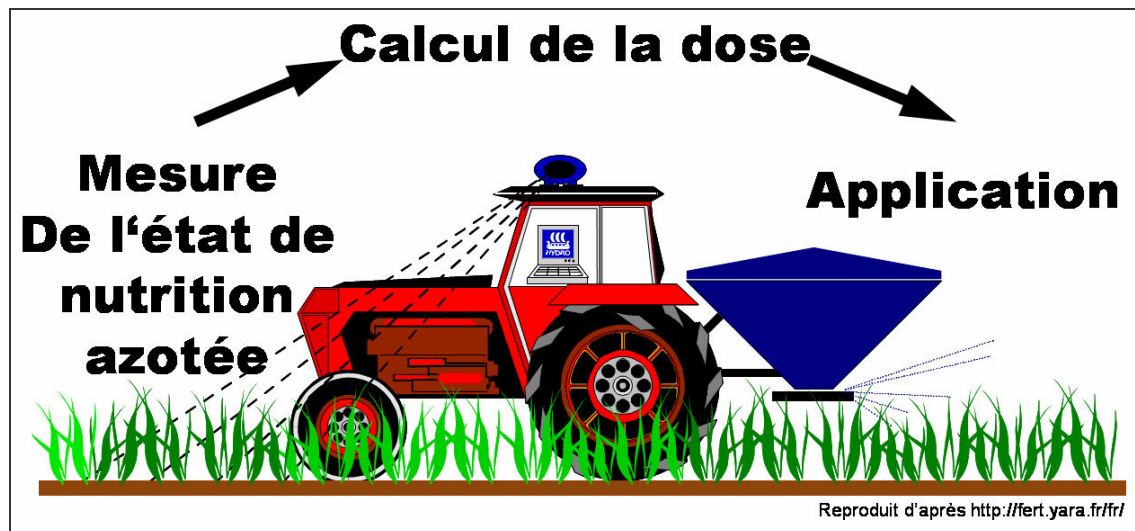


Figure 5 : Fonctionnement du système N-Sensor

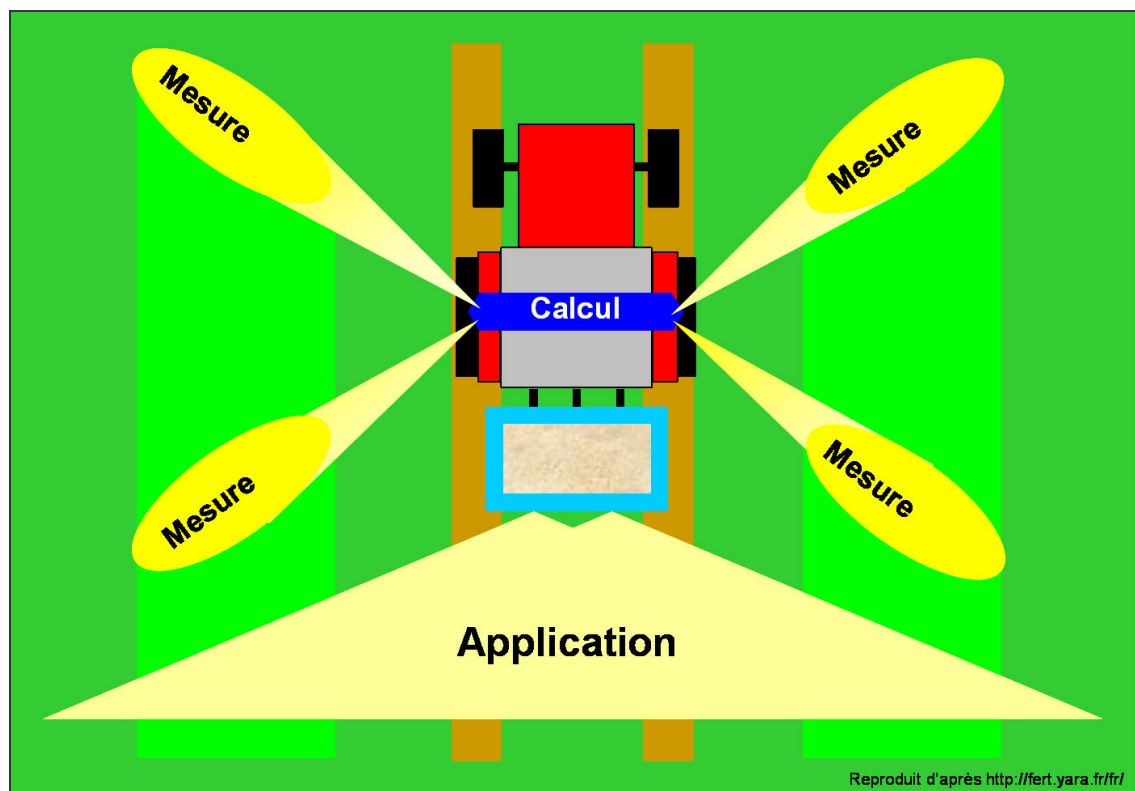


Figure 6 : Localisation des mesures de réflectance avec le système N-Sensor

Le N-Sensor, comme les autres systèmes permettant la modulation directe, permet de gagner du temps car mesure et épandage sont réalisés en une seule opération, sans élaboration préalable de cartes de préconisation.

La modulation directe se heurte toutefois à une limite opérationnelle, comme le notent Jullien et Huet. En effet, l'hétérogénéité spatiale de la culture peut être, au moins en partie, liée à une différence de développement entre les plantes, cette différence de développement étant elle-même liée en partie à l'âge des plantes. La logique de l'agriculture de précision voudrait alors que les plantes dont les stades de développement différent soient traitées à des moments différents. Toutefois, la planification de passages multiples sur une même parcelle pour une même opération culturale constitue une importante difficulté au vu des calendriers déjà chargés [Jullien, Huet 2005].

3.2. La modulation indirecte

Plusieurs définitions de l'agriculture de précision distinguent trois étapes dans sa pratique (voir Figure 7). Bichat les aborde sous la forme : voir, juger, agir [Bichat 1999]. Grenier, pour sa part, les ramène au triptyque classique : mesure, décision, action [Grenier 1997].

Il existe quelques variantes comme celle de Pilesjö, Thylén et Persson qui distingue quatre étapes, la seconde étape "classique" de jugement ou de décision étant divisée en deux [Pilesjö, Thylén, Persson 2000].

Pour notre part, nous nous en tiendrons aux trois étapes sur lesquelles la grande majorité des auteurs s'accordent.

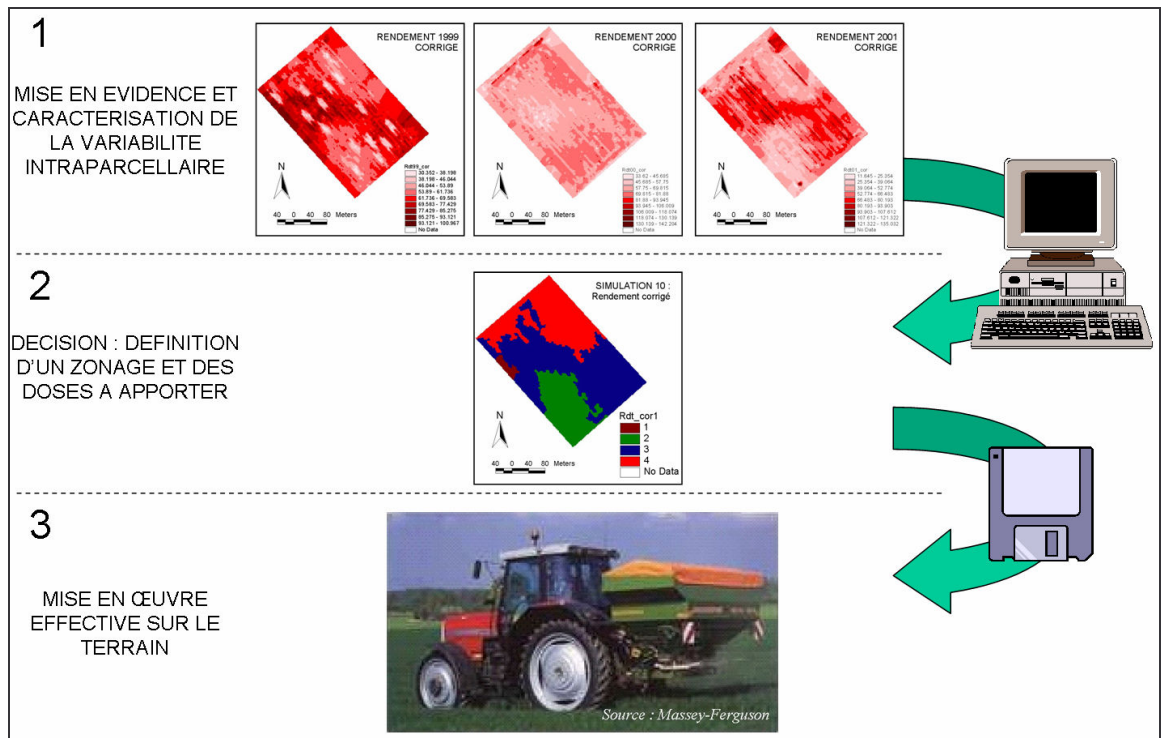


Figure 7 : Les trois étapes de l'agriculture de précision

3.2.1. La mise en évidence et la caractérisation de la variabilité intraparcellaire

3.2.1.1. Ses implications pour l'agronome

La première étape doit permettre la mise en évidence de la variabilité intraparcellaire. Pour ce faire, l'acquisition de données sur la parcelle et leur organisation en une base de données est un passage obligé.

Cette étape préliminaire se caractérise donc par l'accumulation d'information géographique sur la parcelle. A ce stade, il n'est pas encore question de la traiter pour en extraire d'autres informations.

Les données acquises sur la parcelle pour identifier la variabilité intraparcellaire vont être utilisées de différentes manières lors de la deuxième étape. Ces données :

- Concernent le sol (origine de la variabilité) ou la plante (qui exprime la variabilité selon les conditions climatiques et les pratiques de l'exploitant agricole).
- Proviennent de différentes sources (analyses de sol, images satellitaires...).

- Sont de nature qualitative ou quantitative, analytique ou globale.
- Sont géoréférencées au G.P.S. ou pas.
- Sont collectées à des échelles spatiales différentes selon les paramètres mesurés, les zones de la parcelle ou les différentes parcelles de l'exploitation.
- Sont collectées à des échelles de temps différentes, tous les ans ou presque pour certains paramètres (rendement par exemple) et certaines années pour d'autres.
- Ont un caractère pérenne ou non, c'est à dire que certaines des informations sont invariables dans le temps (profondeur et texture du sol par exemple) et d'autres pas.

Les technologies permettant d'acquérir ces données existent déjà puisqu'elles sont à l'origine de la prise en compte de la variabilité intraparcellaire et donc de l'agriculture de précision (voir section 2). Mais y avoir accès est un autre problème comme nous le verrons en section 4.1.

McKinion, Olson, Sequeira, Whisler, Varco et Williford introduisent une nuance intéressante entre l'agriculture et l'agriculture de précision en affirmant que c'est la masse d'informations requise pour prendre des décisions qui les distingue [McKinion, Olson, Sequeira, Whisler, Varco, Williford 1997].

La constitution et le stockage des bases de données sur la variabilité intraparcellaire s'appuient logiquement sur les Systèmes d'Information Géographique (S.I.G.). Ces logiciels capables de stocker des bases de données et de gérer des informations localisées sont adaptés à ce besoin de l'agriculture de précision [Ariaux 2000]. Les relations entre l'information géographique, les S.I.G. et l'agriculture de précision seront abordées plus précisément dans les sections 5.1 et 5.2.

Au final, la base de données constituée permettra d'identifier la variabilité intraparcellaire et servira de point de départ aux deux étapes suivantes.

3.2.1.2. Ses implications pour le géographe

Par rapport à l'agriculture, ce n'est pas tant la mise en évidence de la variabilité intraparcellaire que sa caractérisation qui importe pour le géographe.

Les exploitants agricoles savent que les caractéristiques physico-chimiques des sols de leurs parcelles ne sont pas uniformes. L'information intéressante est l'estimation des valeurs de ces différentes caractéristiques.

Les technologies utilisées pour mettre en évidence la variabilité intraparcellaire sont souvent capables de produire une information géographique précise à grande échelle. C'est cette précision à grande échelle qui présente un intérêt car elle va permettre de caractériser cette variabilité.

C'est sur cette caractérisation de la variabilité intraparcellaire que le géographe va pouvoir ensuite appuyer son travail.

3.2.2. La prise de décision

3.2.2.1. Ses implications pour l'agronome

Il s'agit, lors de cette étape, de décider de la manière dont la variabilité intraparcellaire mise en évidence va être gérée lors d'une opération culturale. Sous cette formulation simple se cachent en fait plusieurs problèmes complexes que de nombreux auteurs ont identifié [Lark, Stafford 1998] [Boisgontier 1997a] [Boisgontier 1997b] [Boisgontier 1999] [Boisgontier 2000] [Thévenet 1999] [Kachanoski, O'halloran, Aspinall, Von Bertoldi 1996] [Van Alphen, Stoorvogel 1999] [Gay 1997].

On peut considérer que le problème vient du fait que "l'information fournie par les nouvelles technologies est abondante mais insuffisante à l'état brut pour moduler les techniques culturales" [Jullien, Huet 2005]. Dès lors, "la détermination des modalités de la modulation nécessite de mobiliser les outils de l'agronomie pour interpréter les données fournies par les capteurs et identifier correctement les facteurs à l'origine des hétérogénéités de rendement observées".

Robert cite les propos d'un agriculteur : "Tous les fermiers ont les mêmes tracteurs, les mêmes machines à récolter, les mêmes terrains, et les mêmes pluies. Aujourd'hui, la seule chose qui vous différencie de votre concurrent, c'est la connaissance". Dans le contexte de l'article, on comprend que l'agriculteur signifiait par "connaissance" à la fois la connaissance de ses parcelles, de leur potentiel, de leur variabilité intraparcellaire, mais aussi la connaissance de méthodes à mettre en œuvre pour gérer au mieux cette variabilité [Robert 2000]. Ici, la collecte et la valorisation de l'information sont considérées comme des éléments déterminants de la compétitivité de l'exploitant agricole.

Peu d'auteurs ont ordonné les différents problèmes identifiés en une démarche logique. Selon Boisgontier, Hébrard [Boisgontier, Hébrard 2001] et Thévenet [Thévenet 1997] [Thévenet 1999], elle devrait être la suivante :

- Décider s'il faut moduler ou non. La modulation, selon Thévenet, est "l'adaptation des pratiques culturales (travail du sol, dose de semis, quantité d'engrais, dose de phytosanitaire, réglage de la moissonneuse-batteuse,...) aux hétérogénéités décelées [sur la parcelle, N.D.A.]" [Thévenet 1997]. Rappelons que Thévenet parle d'appliquer l'agriculture de précision lorsqu'elle est "agronomiquement nécessaire" [Thévenet 1999].
- Si l'on considère que la variabilité intraparcellaire justifie la modulation des opérations culturales, il faut alors décider si l'on recourt à des entités homogènes¹ ou pas. Si oui, il faut alors décider d'une méthode de zonage pour identifier des entités homogènes. Une entité homogène est *"une sous région d'un champ exprimant une combinaison relativement homogène de facteurs limitant le rendement et pour laquelle un taux unique d'intrant pour une culture spécifique est approprié"* [Doerge 2001]². Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers et Payton sont plus précis : *"les zones de gestion, dans le contexte de l'agriculture de précision, se réfèrent à des secteurs géographiques possédant des attributs de condition de sol et de terrain homogènes. Quand ils sont homogènes dans un secteur spécifique, ces attributs devraient fournir des résultats similaires en termes de potentiel de rendement, d'efficacité des intrants et d'impact environnemental"* [Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers, Payton 2000]³.
- Une fois les entités homogènes définies, il s'agit alors de décider ou de calculer la valeur (d'engrais ou de semis ou autre selon l'opération culturale) à appliquer à chaque zone.

Au final, la prise de décision doit déboucher sur la réalisation d'une carte de préconisation. Cette carte reprendra, pour chaque entité homogène définie, la valeur à appliquer.

¹ Chez certains auteurs, c'est le terme "zone de gestion" ("management zones" en anglais) qui est utilisé pour désigner une entité homogène. Chez de nombreux auteurs, les deux termes sont parfois usités indistinctement. Ce sera le cas dans ce document, selon que nous insisterons sur l'aspect géographique (zone) ou agronomique (entité homogène).

² Traduction libre de [Doerge 2001] : "A management zone is a sub-region of a field that expresses a relatively homogeneous combination of yield-limiting factors for which a single rate of a specific crop input is appropriate."

³ Traduction libre de [Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers, Payton 2000] : "Management zones, in the context of precision agriculture, refer to geographic areas that possess homogeneous attributes in terrain and soil condition. When homogeneous in a specific area, these attributes should lead to the same results in crop yield potential, input-use efficiency, and environmental impact."

3.2.2.2. Ses implications pour le géographe

Définir des entités homogènes implique de traiter l'information géographique, acquise sur la parcelle durant la première étape, pour y déceler des zones dont les caractéristiques sont suffisamment semblables pour être considérées uniformes. Mais cela reviendrait à opérer simplement une classification, quelle qu'elle soit, et ne justifierait pas l'intervention du géographe.

Car établir une classification ne suffit pas. Cette classification doit pouvoir être appliquée concrètement sur la parcelle. C'est à dire que les caractéristiques des matériels agricoles qui se chargeront de mettre en œuvre la modulation sur le terrain durant la troisième phase doivent être prises en compte. Jullien et Huet sont très clairs sur ce point : *"le nombre et la taille des unités de gestion doivent tout d'abord être en adéquation avec les caractéristiques techniques des engins agricoles"* [Jullien, Huet 2005]. De même, Mérour considère que les limites des techniques actuelles des matériels agricoles rendent inutile le travail à une résolution trop fine [Mérour 2002].

Ainsi, la largeur d'outils tels que les rampes de fertilisation ou les semoirs peut varier entre 3 et 24 mètres. L'inertie longitudinale, elle, varie entre 1 et 10 mètres. Enfin, la majorité des systèmes de modulation actuels permettent de modifier les paramètres sur une largeur ou une demie largeur de l'outil. A cela s'ajoutent l'erreur de positionnement G.P.S. et le temps de réponse de la machine agricole qui induisent un décalage rendant très incertain le point d'épandage de la dose préconisée.

Cela signifie que les différentes entités homogènes résultant du zonage ne doivent pas être trop petites, être relativement compactes et ne pas être trop nombreuses. Il est donc nécessaire d'introduire d'autres contraintes à la classification : contiguïté, nombre de zones... Et ces contraintes seront le plus souvent liées à la position ou à la forme d'objets dans l'espace.

Malgré ces contraintes, plusieurs S.I.G. ou logiciels de cartographie disponibles actuellement élaborent des cartes de préconisation en traitant les données en mode image (voir section 5.2.2). Chaque couche d'informations est une grille de points et chaque point a une valeur. La carte de préconisation résultante est donc une grille de points et là aussi, chaque point a une valeur. La carte préconise donc une modulation en continu car il est rarissime que deux points voisins aient la même valeur. Ce mode de réalisation de la carte de préconisation n'a que peu de sens dans un cadre opérationnel, comme nous venons de le montrer.

Jullien et Huet concluent sur ce point qu'il faut trouver un optimum entre précision et faisabilité [Jullien, Huet 2005]. Boigontier considère la recherche de cet optimum comme l'acceptation d'une part d'imprécision pour rendre opérationnelle l'agriculture de précision [Boigontier 2002]¹.

La définition d'entités homogènes doit donc s'appuyer sur le traitement conjoint des caractéristiques attributaires et spatiales des objets, et relève donc de l'analyse spatiale, comme nous le verrons en section 5.3.

Visant à la reconnaissance d'aires homogènes spatialement continues, la définition d'entités homogènes en agriculture de précision relève d'une démarche chorologique pour le géographe [Dumolard 1981]. Un algorithme chorologique aurait ici pour but de construire la "meilleure" typologie en blocs spatiaux. Si la configuration de p variables (cartes de rendement, teneur en CaO...) combinées sur n unités spatiales se rapproche d'une répartition en zones, cela signifie que l'algorithme chorologique n'a qu'à réduire certaines irrégularités dues à des erreurs (de mesure ou autre...). Dans ce cas, son utilisation est pertinente.

¹ [Boigontier 2002] : Communication privée.

On retrouve alors le point de vue de Jullien, Huet [Jullien, Huet 2005] et Boisgontier [Boisgontier 2002]¹ sur la notion d'équilibre entre applicabilité et précision de la carte de préconisation. En effet, la répartition d'un contenu spatial en zones est une forme d'organisation, sinon pauvre, du moins simplificatrice d'une réalité souvent plus complexe, combinant des effets de processus différents [Dumolard 1981]. Il s'agit ici, comme le dit Camille Vallaux dans [Gay 1995], de "*chercher à passer suffisamment de diversité pour faire croire au contraste*". C'est bien de cette simplification qu'ont besoin les données initiales afin d'aboutir à la carte de préconisation. Car si les centres des zones sont souvent bien identifiés, leurs marges sont beaucoup plus incertaines. Dès lors, chercher des types spatialement continus (c'est-à-dire des zones) séparés par des frontières fixes ou floues, revient à chercher des discontinuités spatiales avec l'idée que celles-ci traduisent un changement de nature de la combinaison des données. La principale difficulté est alors de savoir ou de choisir où placer la limite même si "*le géographe doit savoir que le flou n'est pas synonyme d'échec et qu'il doit se décider, dans bien des cas, pour ce qui est indécis*" [Gay 1995].

Comme le dit Gay, "*l'espace géographique est hétérogène*" et "*l'organisation spatiale du monde est faite de ruptures qui peuvent être nettes ou floues*" [Gay 1995]. Les contenus spatiaux étant très rarement organisés en aires homogènes et continues, la diversité paraît plutôt être la règle. Pour autant, elle ne signifie pas anarchie ou absence de logique, ni inorganisation [Dumolard 1981]. Car l'espace, et donc les objets spatiaux, sont très majoritairement soumis à l'autocorrélation spatiale. Cette propriété veut que plus des objets spatiaux sont proches, plus leurs caractéristiques le sont également. Les données à référence spatiale sont aussi soumises à l'impact spatial [Openshaw 1991] : par le biais tout d'abord de la dépendance spatiale qui est la propension d'entités voisines à s'influencer, à posséder des attributs similaires ou des potentialités complémentaires ; ensuite par le biais de l'hétérogénéité spatiale qui intervient quand la dépendance spatiale manque d'uniformité. Il faut noter que les effets de l'impact spatial concernent plusieurs attributs et doivent donc être visibles sur plusieurs couches d'information. L'impact spatial a une conséquence directe sur l'analyse statistique des données puisque leur nature géographique veut qu'elles soient en général bien corrélées.

Il est donc possible d'analyser les données afin d'y déceler les discontinuités spatiales qui permettront à leur tour de délimiter des sous-ensembles spatialement continus et isoschèmes : des zones. En agriculture de précision, l'autocorrélation spatiale et l'impact spatial sont moins des propriétés que les conditions sine qua non de la définition d'entités homogènes.

3.2.3. La modulation effective

La dernière étape consiste logiquement à la mise en application, sur le terrain, de la carte de préconisation réalisée durant l'étape précédente.

Pratiquement, cela signifie qu'il faut exporter la carte du S.I.G. vers un système de modulation de machine agricole relié à un G.P.S. Ceci fait, au fur et à mesure de la progression dans la parcelle, le système de modulation de la machine agricole va se repérer dans l'espace, rechercher sur la carte de préconisation la dose à appliquer et commander l'application.

Tout comme pour la première étape, l'application effective de la modulation fait appel à plusieurs technologies, en particulier en matière de positionnement G.P.S. et de machinisme agricole. Mais, là encore, la disponibilité de ces technologies n'est pas un problème. C'est leur coût, encore une fois, qui constitue un frein évident à la diffusion de ces méthodes et techniques (voir section 4.1).

4. Les verrous au développement de l'agriculture de précision

Dans un article datant de 2000, Robert identifie plusieurs verrous au développement de l'agriculture de précision [Robert 2000].

¹ [Boisgontier 2002] : Communication privée.

4.1. Les verrous socio-économiques

Le principal obstacle au développement de l'agriculture de précision reste l'investissement nécessaire à sa pratique. Celui-ci n'est souvent perçu que comme un coût supplémentaire par les exploitants agricoles.

Parmi les dépenses à engager, on trouve l'indispensable capteur de rendement, le G.P.S. et l'abonnement à sa correction, et enfin l'ordinateur avec logiciel de type S.I.G. pour mettre en place la cartographie. A l'heure actuelle, on peut estimer le prix d'acquisition de ces différents outils à 15000€. Il reste ensuite à développer la base de données mettant en évidence la variabilité intraparcellaire : analyses de sol, zones d'adventices... Or, les coûts de certaines données restent très élevés (voir section 5.1). Par conséquent, l'investissement nécessaire est souvent trop important pour un exploitant agricole, d'autant plus que le retour sur investissement reste, nous l'avons vu, peu ou pas démontré. Le coût d'acquisition des données reste donc un facteur critique à la mise en évidence de la variabilité intraparcellaire et donc, par extension, à la pratique de l'agriculture de précision [Stafford, Lark 1997] [Stafford, Lark, Bolam ?] Boisgontier dans [Delpech, Biagini 2000].

On notera que le fait que l'acquisition des données induise des coûts élevés n'est pas spécifique à l'agriculture de précision. Dans toutes les études S.I.G. nécessitant l'acquisition de données, cette étape demande des investissements conséquents [Mainguenaud 2002].

Pawlak a identifié plusieurs des facteurs contribuant à faire diminuer les coûts d'exploitation au niveau du champ [Pawlak 2003]. Cette diminution intervient lorsque :

- Le prix des équipements pour l'agriculture de précision diminue.
- La taille des champs augmente.
- L'utilisation durant l'année des machines équipées pour l'agriculture de précision augmente.
- Les économies d'intrants augmentent.
- Le rendement augmente.
- La valeur économique accordée à l'environnement augmente.

Outre le fait que le prix des équipements devrait logiquement diminuer au cours des années à venir, on remarque que plusieurs des facteurs identifiés ici correspondent aux constats positifs effectués par Benoît Raffin sur sa propre expérience de l'agriculture de précision (voir section 2.5) [Baratte 2002]. On peut donc considérer que le verrou économique au développement de l'agriculture de précision devrait se faire moins prégnant dans le futur.

Le principal verrou sociologique est l'attitude de l'exploitant agricole. Cette attitude est souvent associée à son niveau de formation : les exploitants agricoles ayant une formation limitée s'intéressent rarement à l'agriculture de précision.

4.2. Les verrous technologiques

Nous avons vu plus haut que les première et troisième étape de l'agriculture de précision étaient déjà relativement opérationnelles (voir sections 3.2.1 et 3.2.3). Cependant, il existe encore de nombreux facteurs limitant l'utilisation de ces technologies au-delà de leur simple coût.

Parmi ceux-ci, Robert identifie, concernant les machines agricoles :

- L'inexistence de certains équipements destinés à des opérations culturales ou à des cultures précises. C'est le cas, en particulier, des machines de travaux aratoires variables ou de celles destinées aux légumineuses ou arbres fruitiers.
- L'absence de standard pour la transmission de l'information entre les différents éléments du système.
- Le manque de formation à l'utilisation et à la maintenance des nouveaux systèmes chez les exploitants agricoles et les détaillants. L'exemple le plus flagrant concerne le nécessaire calibrage des capteurs de rendement.
- Le manque de précision pour certaines machines. Dans un article de Pavard, un technicien déclare : *"Avant de pouvoir moduler l'intérieur du champ, il faut déjà s'assurer de mettre la bonne dose uniformément !"* [Pavard 1996]. Nous avons également vu en section 3.2.2.2 les limites de la précision des appareils destinés à l'agriculture de précision.

Robert identifie, concernant les capteurs :

- L'inexistence de certains équipements destinés à des tâches telles que la mesure de la fertilité des sols ou la détection des mauvaises herbes ou à des cultures autres que les céréales.
- L'imprécision de certains capteurs. Si des progrès ont été accomplis, il reste encore des sources d'erreurs, notamment au niveau de la conception des capteurs, du signal G.P.S. et des manœuvres de l'opérateur. Pour le positionnement G.P.S., les corrections différentielles ne sont pas disponibles partout et demandent souvent un abonnement annuel.
- La nécessité de calibrer et de paramétrer les capteurs à chaque utilisation [Jullien, Huet 2005].

D'importantes lacunes concernent également les logiciels :

- La compatibilité hasardeuse des formats de données. L'échange de données entre différents logiciels peut entraîner leur dégradation.
- Les performances des logiciels. Les fonctionnalités proposées par les différents logiciels dédiés à l'agriculture de précision sont très variables. De plus, sur un même jeu de données, les résultats donnés par les mêmes fonctions de deux logiciels différents peuvent parfois être très dissemblables.
- La prise en main difficile. Les différents logiciels, et à plus forte raison ceux proposant des fonctionnalités avancées, sont souvent difficiles à prendre en main pour le profane.

Enfin, Robert identifie les facteurs limitants suivants concernant la télédétection :

- Le coût élevé des images satellitaires, en particulier pour les hautes résolutions nécessaires à la caractérisation de la variabilité intraparcellaire.
- L'absence de coopératives compétentes dans l'interprétation et l'utilisation de ces images.
- Les services de photographie aérienne sporadiques et de qualités très variables.

4.3. Les verrous liés à la gestion agronomique

Les cartes de sol ont souvent été réalisées à des échelles trop petites pour pouvoir être efficacement utilisées en agriculture de précision. Il manque donc, à l'heure actuelle, des données élémentaires essentielles à sa pratique.

De même, en cours de campagne, les techniques visuelles utilisées par les consultants pour le suivi des besoins en éléments nutritifs et le contrôle des adventices et nuisibles manquent de précision.

L'agriculture de précision introduit une nouvelle catégorie d'activités pour lesquelles la plupart des exploitants agricoles n'ont pas de qualification. Ces activités concernent la gestion de l'information géographique à l'aide d'ordinateurs et son utilisation à des fins agronomiques. Les exploitants dépendent donc des services offerts par les distributeurs et les consultants agronomes.

Devant le manque de recommandations agronomiques, il est nécessaire d'expérimenter au niveau des fermes et cette expérimentation doit être faite par les exploitants agricoles. Ce sont eux qui pourront définir la meilleure pratique pour leur exploitation. Cette démarche demande un effort de vulgarisation considérable.

Ce manque de recommandations entraîne également des utilisations inadéquates de l'information. L'exemple le plus fréquent est l'apport massif d'engrais supplémentaire dans les zones peu productives pour en accroître à tout prix le rendement. L'utilisation abusive des cartes de rendement, sans tenir compte, par exemple, de l'influence du climat, constitue un second exemple.

Jullien et Huet insistent sur la nécessité d'interpréter les informations fournies par les différents capteurs mis à disposition de l'exploitant agricole afin de définir les facteurs limitant le rendement [Jullien, Huet 2005]. Mais la plupart des capteurs mesurent plusieurs facteurs du milieu. Par conséquent, le problème de l'identification des facteurs limitants reste le même qu'en agronomie classique non modulée. Les nouvelles technologies utilisées en agriculture de précision ne font qu'améliorer la précision des mesures et agrandir l'échelle de travail.

L'agriculture de précision marquera-t-elle la troisième révolution agricole ? Si cela n'est pas impossible sur le long terme, la situation actuelle rend néanmoins cette éventualité peu probable et ce pour plusieurs raisons. Tout d'abord, l'agriculture de précision s'appuie sur des techniques qui, si elles sont novatrices dans le domaine agricole, sont utilisées depuis bien plus longtemps dans d'autres domaines. Ensuite, d'autres progrès, en particulier dans le domaine des biotechnologies, ont des impacts bien plus flagrants. Au final, la nouveauté introduite par l'agriculture de précision est moins technique que méthodologique et conceptuelle : elle concerne principalement la prise en compte de la préservation de l'environnement.

De fait, si cette nouveauté devait constituer une troisième révolution agricole, cette dernière incomberait alors à l'agriculture durable, dont l'agriculture de précision constitue une approche. Et pour l'heure, si troisième révolution agricole il y a, elle s'appuie plus sur les biotechnologies et la génétique.

Pourtant, cela ne signifie en rien que l'agriculture de précision soit une approche stérile ou peu innovante. Elle a ses atouts et ses particularités : ses besoins et son intérêt pour les outils et les méthodes relevant des S.I.G. et de l'analyse spatiale n'étant pas la moindre. A travers certaines problématiques spatiales de l'agriculture de précision, le géographe se voit donc offrir une opportunité de contribuer, à son échelle, au développement d'une approche de l'agriculture durable.

Seconde partie :

Géomatique et agriculture de précision

Pour être mise en œuvre, l'agriculture de précision a besoin, comme nous l'avons dit dans la partie précédente, d'information géographique à grande échelle et des S.I.G. pour gérer cette information géographique. Dans un second temps, l'analyse spatiale intervient à son tour, en particulier lors de la prise de décision concernant la carte de préconisation.

Dans cette partie, nous allons examiner quels sont les différents types de données utilisés en agriculture de précision. Nous aborderons plus en détail les points forts qui font des S.I.G. et de l'analyse spatiale des outils et méthodes utiles à cette approche de l'agriculture. Nous verrons également qu'elles peuvent être les limites d'utilisation de l'analyse spatiale des S.I.G.

Nous en viendrons ensuite à présenter les contraintes de notre travail avant d'établir un état des lieux des méthodes de définition d'entités homogènes existantes.

5. Information géographique, S.I.G., analyse spatiale et agriculture de précision

Nous l'avons vu en section 3, la mise en œuvre et la pratique de l'agriculture de précision s'appuie très largement sur des données géoréférencées. Il est ensuite possible, une fois celles-ci correctement agencées au sein d'un S.I.G., d'en extraire de l'information par le biais de différentes méthodes dont certaines relèvent de l'analyse spatiale. Il existe donc de nombreuses relations entre l'agriculture de précision et certaines thématiques et outils géographiques. Dans ce chapitre, nous allons expliciter ces relations.

5.1. Information géographique et agriculture de précision

L'information géographique et les données géoréférencées sont à la base de la pratique de l'agriculture de précision. Leur collecte en constitue la première étape (voir section 3.2.1). Sans elles, il est impossible de mettre en évidence la variabilité intraparcellaire, et donc, au final, de pratiquer la modulation. Par conséquent, comme pour toutes les analyses de données, connaître le mode d'acquisition de ces données et être conscient de leurs limites est primordial pour juger et nuancer la validité des analyses.

Laurent et Thinon s'interrogent sur le degré d'autonomie des agriculteurs devant l'importance de l'échange d'informations [Laurent Thinon 2005]. En accordant une importance stratégique à l'acquisition et à l'échange de l'information géographique, l'agriculture de précision a modifié les relations entre les acteurs du monde agricole [Jullien, Huet 2005]. Ainsi, un exploitant agricole peut adopter une attitude de consommateur d'information géographique ou s'investir dans l'élaboration de cette information :

- Dans le premier cas, l'information est fournie par des sociétés privées par le biais de différents services. L'exploitant agricole n'a pas à se soucier de l'acquisition physique de la mesure et de l'interprétation des données. Il bénéficie d'une information directement utilisable. Cependant, il ne peut exercer aucun contrôle sur l'interprétation des données et n'est pas propriétaire de l'information.
- Dans le second cas, un investissement temporel et intellectuel non négligeable intervient en plus de l'inévitable investissement financier. L'exploitant, indépendant, peut intégrer son point de vue sur l'information dans l'interprétation des résultats. Il a toutefois besoin d'une aide technique pour exploiter les données. Là encore, des prestations existent.

On retrouve donc ici l'un des verrous liés à la gestion agronomique de l'information identifié par Robert (voir section 4.3) [Robert 2000].

Les données utilisées en agriculture de précision sont nombreuses mais on peut globalement les classer dans quatre grandes catégories : les analyses de sol, les données de rendement, les mesures de conductivité/résistivité électrique et l'imagerie satellitaire. Au sein de chaque catégorie, on retrouvera ainsi des informations ayant le même libellé ou appartenant à un même type de mesure. Pourtant, ces données pourront, et devront, être différenciées en fonction de plusieurs critères décrits en section 3.2.1.1.

Jullien et Huet distinguent, parmi les outils de mesure, ceux qui relèvent des méthodes de mesure directes et ceux qui relèvent des méthodes de mesure indirectes [Jullien, Huet 2005] :

- Les méthodes de mesures directes sont fastidieuses et destructrices (prélèvements et analyse d'échantillons de plantes et de sol), et donc réalisées en petit nombre sur la parcelle. Elles fournissent un semis de points trop peu dense pour en établir une cartographie mais restent utilisées pour les caractéristiques permanentes des sols ou pour calibrer les capteurs embarqués sur les machines agricoles. Les informations fournies par les méthodes de mesure directes sont souvent des variables utilisées dans les modèles agronomiques.
- Les méthodes de mesures indirectes s'appuient sur des nouvelles technologies qui vont permettre de récupérer une information avec un semis de points beaucoup plus dense sur la parcelle. La mesure étant indirecte, aucun échantillon n'est prélevé. Toutefois, une interprétation est nécessaire, l'information fournie ne proposant qu'une estimation d'une ou plusieurs variables utilisées dans les modèles agronomiques.

5.1.1. Les analyses de sol

Les analyses de sol sont généralement réalisées à la tarière, bien que certains équipements plus perfectionnés existent actuellement. Elles consistent à prélever un échantillon sur les vingt ou trente premiers centimètres du sol puis à l'analyser en laboratoire pour en définir plusieurs caractéristiques granulométriques et physico-chimiques. Les points d'analyse sont géolocalisés par G.P.S.

Dans la mesure où il y a prélèvement, les analyses de sol sont des méthodes de mesures directes.

5.1.1.1. Une information capitale

Les informations fournies par les analyses de sol sont capitales en agriculture. En effet, ces informations mesurent les paramètres granulométriques (teneur en argile, en sable...) et physico-chimiques (teneur en potassium, azote, valeur du pH...) du sol. Et ces paramètres contribuent en partie à conditionner la croissance de la plante (avec le climat, les maladies ...), et par conséquent le potentiel de rendement de la culture [Jullien, Huet 2005].

Les pratiques culturales sont donc définies en fonction des valeurs de ces paramètres. Ceux-ci sont fréquemment utilisés par les modèles agronomiques pour définir les doses à apporter, que l'on travaille à l'échelle du parcellaire ou de l'intraparcellaire.

5.1.1.2. Mais une information coûteuse

Le prix moyen d'une analyse de sol est de 23€. Cette mesure est proposée par de nombreux prestataires, comme Satplan, sur l'ensemble du territoire français.

En raisonnant à l'échelle de la parcelle, le coût des analyses de sol reste abordable. Mais il en va autrement lorsque l'on souhaite disposer d'une information suffisamment fine pour pouvoir gérer la variabilité intraparcellaire. En effet, la grande majorité des campagnes d'analyses de sol destinées à l'agriculture de précision se font selon une grille de points régulière. Or, aux Etats-Unis, Franzen, Halvorson et Hofman considèrent que la grille doit avoir un pas de 66 mètres [Franzen, Halvorson, Hofman 2000]. En France, les spécialistes d'Arvalis Institut du Végétal estiment qu'un minimum de cinq à dix analyses de sol par hectare doit être effectué pour obtenir un semis de points d'une densité adaptée à la pratique de l'agriculture de précision [Piriaux 2003]¹. La Figure 8 illustre une localisation quasi régulière des analyses de sol sur la parcelle expérimentale "Imbault" du réseau variabilité intraparcellaire d'Arvalis Institut du Végétal.

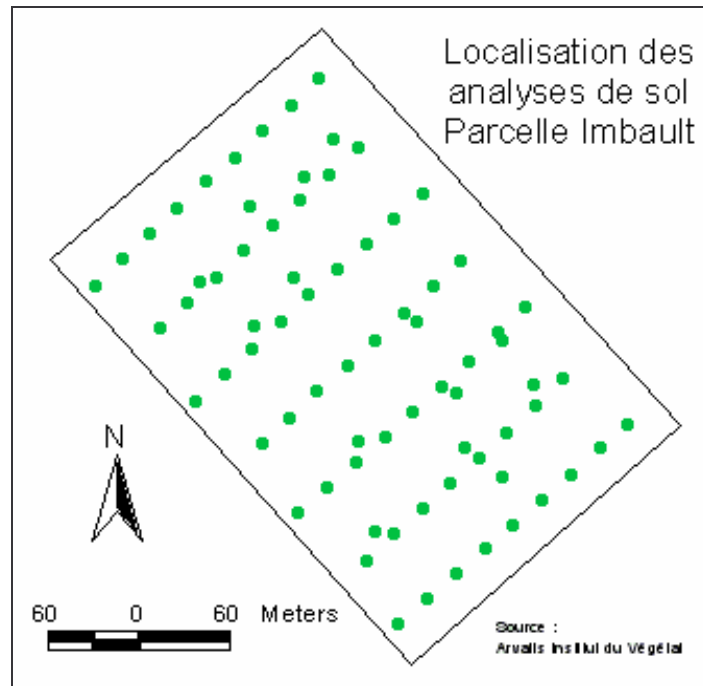


Figure 8 : Exemple de localisation d'analyses de sol sur la parcelle Imbault

Kravchenko a étudié l'influence de la densité de points d'analyse de sol sur les résultats d'interpolation. Il a montré que la précision dans la cartographie des propriétés du sol dépend fortement de leur structure spatiale. Les propriétés avec des structures spatiales nettes peuvent être cartographiées facilement à partir d'un nombre réduit d'analyses de sol. A l'inverse, les propriétés ayant des structures mal définies ne peuvent être cartographiées qu'avec de nombreuses analyses de sol [Kravchenko 2003].

Chang, Clay, Carlson et Reese considèrent que l'échantillonnage par grille fournit une information suffisante dès lors que les analyses sont suffisamment proches pour que leur dépendance spatiale (voir section 3.2.2.2) soit restituée [Chang, Clay, Carlson, Clay, Reese 2000]. Si ce principe est parfaitement fondé, sa mise en application est loin d'être évidente. En effet, elle impliquerait d'effectuer une première campagne d'analyses avec une grille de points très lâche. Il faudrait ensuite attendre les résultats des analyses (quelques jours) puis, le cas échéant, effectuer une nouvelle campagne pour compléter la grille de points... Et ainsi de suite jusqu'à ce que la dépendance spatiale soit restituée. Au-delà du fait que chaque nouvelle campagne va augmenter significativement les coûts, c'est la multiplication des campagnes d'analyses de sol qui rend cette méthode très peu pratique dans un cadre opérationnel.

¹ [Piriaux 2003] : Communication privée.

Effectuer les analyses de sol par le biais d'une grille régulière de points sur la parcelle, et ce à une échelle propice à la gestion de la variabilité intraparcellaire, engendre donc des coûts très importants. Ce problème est relevé par de nombreux auteurs [Fleming, Westfall, Bausch 2000] [Franzen, Halvorson, Hofman 2000] [Chang, Clay, Carlson, Clay, Reese 2000] [Zimmermann, Plöchl, Luckhaus, Domsch 2003] [Frogbook, Oliver, Derricourt 2003]. Il constitue l'un des verrous économiques identifiés en section 4.1.

Pour obtenir les informations issues des analyses de sol en restituant correctement leur structure spatiale tout en diminuant les coûts d'obtention de ces informations, plusieurs auteurs ont recours à l'échantillonnage par zone [Franzen, Halvorson, Hofman 2000] [Chang, Clay, Carlson, Clay, Reese 2000]. Cette méthode consiste à établir un zonage en utilisant une donnée peu onéreuse, puis à affecter à chaque zone une analyse de sol (ou quelques unes si l'on souhaite augmenter la précision).

Franzen, Halvorson et Hofman partent du principe qu'il y a une raison logique pour que les caractéristiques des nutriments soient visibles dans le champ et que ces caractéristiques doivent être stables dans le temps [Franzen, Halvorson, Hofman 2000]. Ces auteurs ont effectué des essais sur les apports en azote et en potassium sur trois parcelles. Ils ont effectué un zonage d'après la topographie et ont ainsi localisé une quinzaine d'échantillons. Ils ont montré que ces quinze points restituaient la même information qu'une grille régulière de 51 points. Le problème est que la méthode de zonage utilisée nécessitait de fixer le nombre de zones arbitrairement. Rien ne prouve que le nombre de zones réel de la parcelle était le même que celui choisi par les auteurs.

Zimmermann, Plöchl, Luckhaus et Domsch ont recours à la conductivité électrique, par le biais de la méthode E.M. 38 (voir section 5.1.4), pour obtenir une information sur l'hétérogénéité de la structure du sol de la parcelle [Zimmermann, Plöchl, Luckhaus, Domsch 2003]. Ils appliquent ensuite plusieurs traitements aux données collectées (krigeage, Classification Ascendante Hiérarchique, Analyse en Composantes Principales) avant d'utiliser leur propre méthode pour positionner les points d'échantillonnage. Là encore, le nombre de zones est défini arbitrairement par les auteurs sans assurance qu'il soit le même que le nombre de zones réel de la parcelle. Enfin, cette démarche présente les mêmes avantages et inconvénients que la conductivité électrique sur laquelle elle s'appuie : bonne précision des mesures avec un semis de points très dense, mais un coût d'acquisition qui, tout en restant inférieur à l'échantillonnage par grille, n'en demeure pas moins conséquent. D'autres auteurs utilisent la conductivité électrique pour pratiquer l'échantillonnage par zone [Frogbook, Oliver, Derricourt 2003].

5.1.2. Les cartes de rendement

Lors de la récolte, un capteur placé dans la machine va mesurer la quantité de grain moissonnée. Ce capteur, relié à un G.P.S., permet d'associer la quantité de grain à un point géolocalisé. Les semis de points obtenus de cette manière sont donc très denses. Les cartes de rendement sont ensuite obtenues par interpolation des valeurs des semis de points géolocalisés.

La mesure du rendement s'effectuant en même temps que la moisson de la plante, cette méthode de mesure est directe.

5.1.2.1. Données d'entrée ou information synthétique

Les données de rendement relevées présentent plusieurs défauts qui font que les cartes de rendement obtenues sont souvent de piètre qualité et, par conséquent, difficilement exploitables, en particulier pour définir un zonage :

- Les mesures relevées par les capteurs de rendement sont souvent entachées d'erreurs, au sens statistique et spatial du terme, liées aux conditions d'utilisation et au mode de fonctionnement même de ces capteurs [Mérour 2002]. Ainsi, les rangs de collecte et les tournières, en particulier, sont nettement visibles sur les cartes. Et le temps de montée en charge de la machine induit un décalage des valeurs de rendement entre deux rangs voisins moissonnés dans des directions opposées.

- Les cartes de rendement sont instables d'une année sur l'autre [Dabas, Tabbagh, Boisgontier 2001]. C'est-à-dire que les zones de bons et de mauvais rendements ne sont pas toujours les mêmes et que leurs limites sont rarement nettes. Parmi les facteurs expliquant cette instabilité, on trouve bien évidemment les conditions météorologiques au cours de la saison, mais aussi les adventices et les maladies, ainsi que d'éventuelles erreurs humaines.

Il existe donc à l'heure actuelle plusieurs attitudes de la part des spécialistes de l'agriculture de précision vis à vis de l'utilisation des cartes de rendement pour la définition d'entités homogènes.

Un premier groupe considère que cette information doit être prise en compte dans la mesure où elle synthétise la variabilité intraparcellaire globale au moment de la récolte [Boydell, McBratney 1999] [Cupitt, Whelan 2001] [McKinion, Olson, Sequeira, Whisler, Varco, Williford 1997]. De plus, les cartes de rendement sont souvent les seules cartes dont dispose l'exploitant agricole qui commence à pratiquer l'agriculture de précision. Elles sont par conséquent les seules informations sur lesquelles il peut appuyer sa réflexion.

Le second groupe considère, en revanche, que cette donnée ne doit pas être prise en compte car elle est, justement, synthétique. Ce groupe argue que cette information fondamentalement "finale" ne saurait être considérée comme une information "initiale" pour la définition d'entités homogènes [Gay 1997] [Thévenet 1999] [Rouiller, Dabas, Tabbagh, Perrard 2001] [Dabas, Tabbagh, Boisgontier 2001] [Doerge 2001]. Gay précise que *"connaître la variabilité d'un résultat ne permet que des hypothèses tant que l'on n'a pas autant de précisions sur la variabilité des facteurs qui ont produit ce résultat"* [Gay 1997].

Enfin, un troisième groupe, plus nuancé, se propose d'utiliser les cartes de rendement avec prudence. Ainsi, Lark et Stafford utilisent des cartes de rendement relevées sur plusieurs années successives pour définir des zones qui ont une réponse uniforme aux saisons. Ils insistent cependant sur les limites d'une telle méthode [Stafford, Lark 1997] [Lark, Stafford 1998]. Domsch, Kaiser, Witzke et Zauer s'appuient sur le concept de potentiel de rendement qui permet de moyenniser les conditions du site en quelques zones. Le potentiel de rendement reprend des conditions naturelles (relief, réserve utile, climat...), les espèces cultivées et la stratégie de l'exploitant agricole [Domsch, Kaiser, Witzke, Zauer 2003]. Le concept de potentiel de rendement est repris par d'autres auteurs [Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers Payton 2000]. Hornung, Koshla, Reich et Westfall adoptent une position assez proche. Ils considèrent le rendement comme une des mesures les plus précises de la variabilité intraparcellaire. Mais, en pratique, l'utilisation du rendement pour définir des zones est délicate de par la variation spatiale et temporelle de celui-ci. Les auteurs se limitent donc à l'identification des régions de forts et faibles rendements [Hornung, Koshla, Reich, Westfall 2003].

5.1.2.2. La correction des cartes de rendement

L'utilisation nuancée des cartes de rendement décrite plus haut induit que certaines corrections aient été apportées aux données initiales. Ces corrections se font généralement par application(s) de techniques diverses telles que les filtres issus du traitement d'images ou des méthodes statistiques... Pour les méthodes les plus complexes, l'objectif est d'éliminer toutes les erreurs de mesure. Cependant, la grande majorité des méthodes de correction des données de rendement visent moins à supprimer les données aberrantes qu'à limiter leur influence sur la carte. La correction des cartes de rendement constitue une réponse à plusieurs verrous technologiques identifiés en section 4.2, en particulier ceux ayant trait à l'imprécision des capteurs et aux performances des logiciels.

Ainsi, Dobermann, Ping, Simbahan et Adamchuk considèrent que les données de rendement contiennent trois types d'informations : la variabilité stable du rendement due au climat et au sol ; la variabilité variable du rendement due à la gestion ; et enfin les erreurs de mesure. Afin d'éliminer ces dernières, ils procèdent à une correction succincte des données de rendement par le biais d'une standardisation avant de les exploiter [Dobermann, Ping, Simbahan, Adamchuk 2003].

Gerwig, Sadler et Evans emploient une méthode de lissage consistant en une interpolation en fonction du rendement pondéré par la distance. La pondération est de 5 pour les valeurs les plus proches et de 1 pour les valeurs les plus éloignées. Les tests ont été réalisés pour des rayons de six et douze mètres [Gerwig, Sadler, Evans 2000]. Dobermann, Ping, Simbahan et Adamchuk procèdent à une correction succincte des données de rendement par le biais d'une normalisation [Dobermann, Ping, Simbahan, Adamchuk 2003].

Pour Mérour, la valeur d'une carte de rendement réside dans son interprétation, et l'interprétation doit se fonder sur des données et des cartes fiables [Mérour 2002]. Cet auteur a donc mis au point une méthode complexe de correction des cartes de rendement pour Arvalis Institut du Végétal [Mérour 2002]. Cette méthode porte sur les points suivants :

- Elimination des rangs avec trop peu de points.
- Recalage des données de rendement (basé sur le temps de transit).
- Elimination des points trop proches (basé sur la vitesse de bourrage).
- Elimination des points invraisemblables (basé sur la vitesse maximale).
- Elimination des relevés en début de rang (basé sur le temps de montée en charge).
- Elimination des rendements extrêmes.

Cette méthode fait partie de celles qui visent à éliminer l'ensemble des erreurs de mesure. Un exemple des résultats fournis est présenté en Figure 9 pour la parcelle expérimentale "Imbault" du réseau variabilité intraparcellaire d'Arvalis Institut du Végétal.

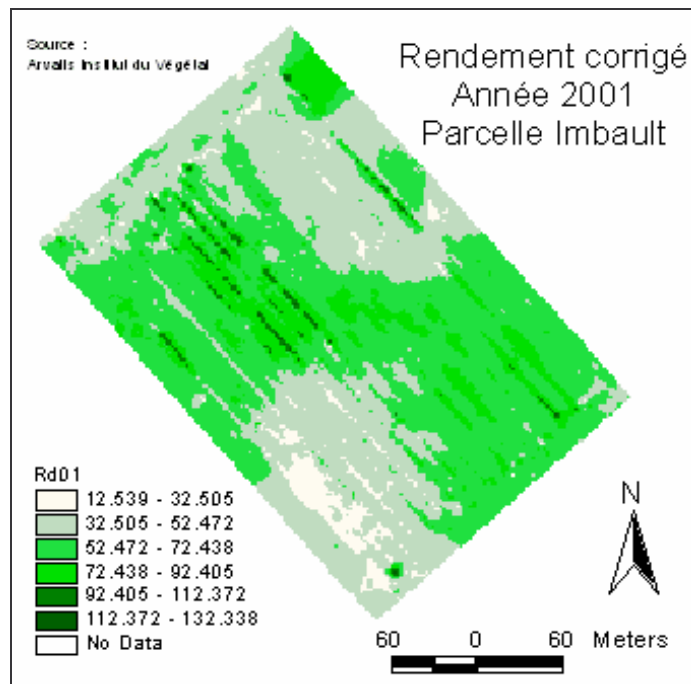


Figure 9 : Exemple de carte de rendement corrigée sur la parcelle Imbault

5.1.3. Les images satellitaires et aériennes issues de la télédétection

Pour Robin, "l'objet de la télédétection est de fournir des informations sur les paysages sous la forme de données images, en utilisant le rayonnement électromagnétique comme véhicule de ces informations"[Robin 1995]. On distingue deux types de télédétection [Minvielle, Souiah 2003] :

- La télédétection active où les capteurs mesurent la réponse de l'objet par rapport à un rayonnement qu'ils émettent. C'est le cas des radars.
- La télédétection passive où les capteurs mesurent la réponse de l'objet par rapport à un rayonnement naturel fourni par la lumière ou la chaleur (du soleil, en général).

En ce qui concerne l'agriculture de précision, le signal émis par un couvert végétal est corrélé aux variables biophysiques qui le caractérisent [Jullien, Huet 2005].

La Figure 10 montre une image satellitaire Landsat obtenue dans le rayonnement proche infra-rouge en août 2000 (résolution de trente mètres) sur la parcelle expérimentale "Imbault" du réseau variabilité intraparcellaire d'Arvalis Institut du Végétal.

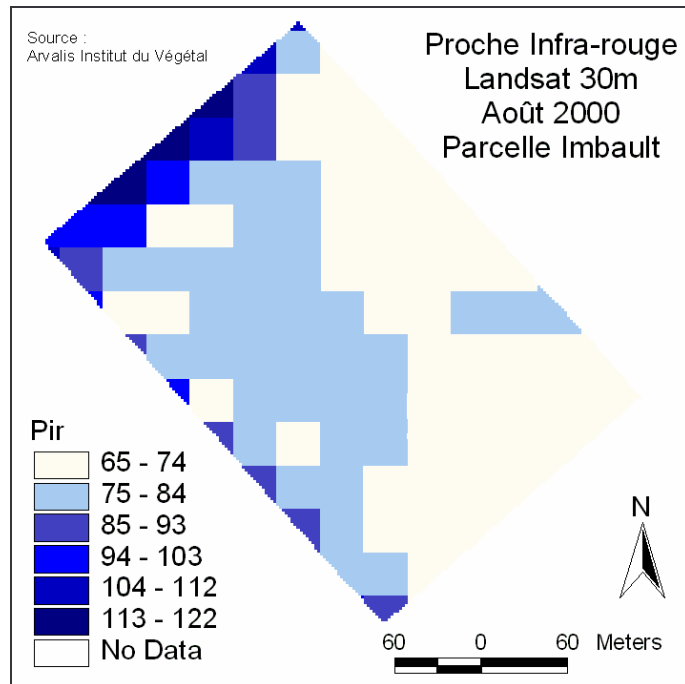


Figure 10 : Exemple d'image satellitaire sur la parcelle Imbault

L'un des avantages de l'imagerie satellitaire et aérienne est que l'ensemble de la parcelle est couvert par la mesure. Par conséquent, le problème habituel de densité du semis de points ne se pose pas ici car aucune interpolation des données n'est nécessaire. Se pose en revanche le problème de la résolution des images. Au final, ces deux problèmes, densité du semis de points de mesure et résolution des images, relèvent du même besoin de disposer de données dont la précision finale permettra d'identifier et de gérer la variabilité intraparcellaire.

Jullien et Huet considèrent que des images d'une bonne résolution spatiale et d'une fréquence d'acquisition hebdomadaire sont nécessaires pour assurer le suivi des cultures [Jullien, Huet 2005]. Mais l'acquisition des images dépend des conditions météorologiques, en particulier de la nébulosité. Il faudrait donc un satellite ayant une fréquence de passage de deux ou trois jours pour disposer d'une image exploitable par semaine. Actuellement, ces conditions ne sont pas remplies en France. De plus, le coût des images satellitaires est resté très élevé dans notre pays.

Ce sont donc à la fois les verrous économiques et technologiques (voir sections 4.1 et 4.2) qui sont concernés avec l'imagerie satellitaire.

Afin de se libérer, dans une certaine mesure, des contraintes météorologiques, techniques et financières, certaines entreprises et instituts de recherche se sont tournés vers les drones.

Ainsi, en France, le drone Pixy a été développé par le centre I.R.D. (Institut de Recherche pour le Développement) de Montpellier, l'E.N.S.I.C.A. (Ecole Nationale Supérieure d'Ingénieurs de Constructions Aéronautiques), A.B.S. Aérolight (société spécialisée dans la conception, la réalisation et la formation en matière d'ultra léger motorisé) et GEOSYS (société spécialisée dans l'information géographique, l'imagerie aérienne pour l'agriculture et l'environnement) qui a développé des logiciels et différentes configurations d'aide à la navigation adaptée au drone¹.

Pixy est un petit paramoteur radiocommandé conçu spécialement pour permettre des prises de vues aériennes à très basse altitude (voir Figure 11). Son objectif est de faciliter l'accès aux photos, vidéos et images numériques ou radiométriques aériennes en proposant un nouveau support de coût réduit, de grande sécurité, de simplicité d'utilisation et dépendant de réglementations peu contraignantes.



Figure 11 : Le drone Pixy au décollage

Dans la limite de quatre kilogrammes, l'utilisateur peut embarquer les capteurs de son choix, photo, vidéo, infrarouge, radiométrie, etc... L'acquisition de vues à très basse altitude et à faible vitesse, saisies sur divers supports (photographie argentique, numérique ou vidéo), dans les spectres visible, invisible, thermique ou infrarouge ouvre l'accès à de nombreuses applications :

- Suivi de parcelles agronomiques, de leur hétérogénéité (voir Figure 12).
- Appréciation de la variabilité de la surface du sol et du couvert végétal.
- Détection d'anomalies des systèmes d'irrigation, semis, traitements....
- Acquisition numérique sous nuages et faible visibilité.
- Relevés cadastraux, surveillance d'ouvrages d'arts.
- Cartographie magnétique en géophysique.

¹ <http://www.drone-pixy.com/FTECH.html#>

- Repérage de sites en archéologie, cinéma et télévision.
- Contrôles aériens ponctuels, travaux de spatialisation.
- Comptage de populations concentrées ou déplacées.

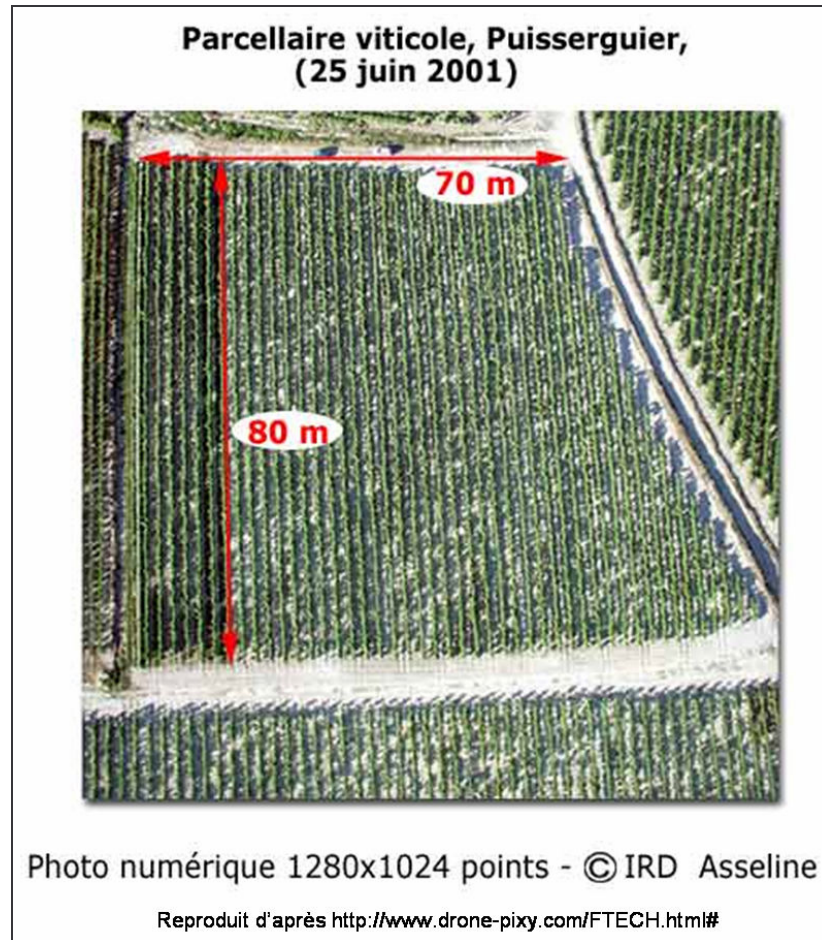


Figure 12 : Exemple de photo réalisée avec le Pixy

5.1.4. La conductivité et la résistivité électrique du sol

La conductivité électrique d'un sol représente sa capacité à laisser passer un courant électrique. Elle permet d'estimer des propriétés du sol importantes agronomiquement telles que la salinité, la teneur en eau et le pourcentage d'argile. Elle permet aussi d'estimer les niveaux de présence de nutriments dans le sol. Si cette méthode est très facile à employer, il faut prendre garde à l'interprétation des résultats. En effet, la conductivité électrique est sensible au type de sol et aux conditions de mesure [Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers Payton 2000].

La résistivité électrique d'un sol représente sa capacité à limiter le passage d'un courant électrique. Cette faculté est étroitement liée aux caractéristiques intrinsèques et pérennes des sols tels que la texture, la teneur en argile, la pierrosité, la profondeur et le type de substrat géologique [Geocarta 2004]¹. La résistivité électrique est aussi corrélée avec l'humidité et la fertilité chimique du sol [Rouiller, Dabas, Tabbagh, Perrard 2001]. Les différents facteurs définissant la fertilité chimique du sol ont une influence directe sur le rendement et sont souvent des paramètres d'entrée des modèles agronomiques. La résistivité électrique est donc communément considérée comme une information synthétique fiable agronomiquement.

¹[Geocarta 2004] : <http://www.geocarta.net/>

Que l'on mesure la conductivité ou la résistivité électrique du sol, les méthodes de mesure n'induisent pas de prises d'échantillons et sont réalisées avec des engins limitant les tassements du sol (quads, traineaux, remorques spécialement conçues). Ces méthodes de mesure sont donc indirectes.

On utilise l'induction électromagnétique pour mesurer la conductivité électrique. Chaque mesure est géolocalisée par G.P.S. L'interpolation des valeurs du semis de points permet d'obtenir une carte.

Le principe de la mesure de la résistivité électrique consiste à injecter du courant dans le sol via deux électrodes et de mesurer le potentiel résultant grâce à deux autres électrodes (voir Figure 4). C'est la géométrie de la disposition des électrodes qui détermine la profondeur d'investigation ainsi que le volume de terrain prospecté. La résistivité électrique est mesurée sur trois voies, ces trois voies correspondant respectivement à des mesures prises à cinquante centimètres, un mètre et deux mètres de profondeur. Chaque mesure est géolocalisée par G.P.S.

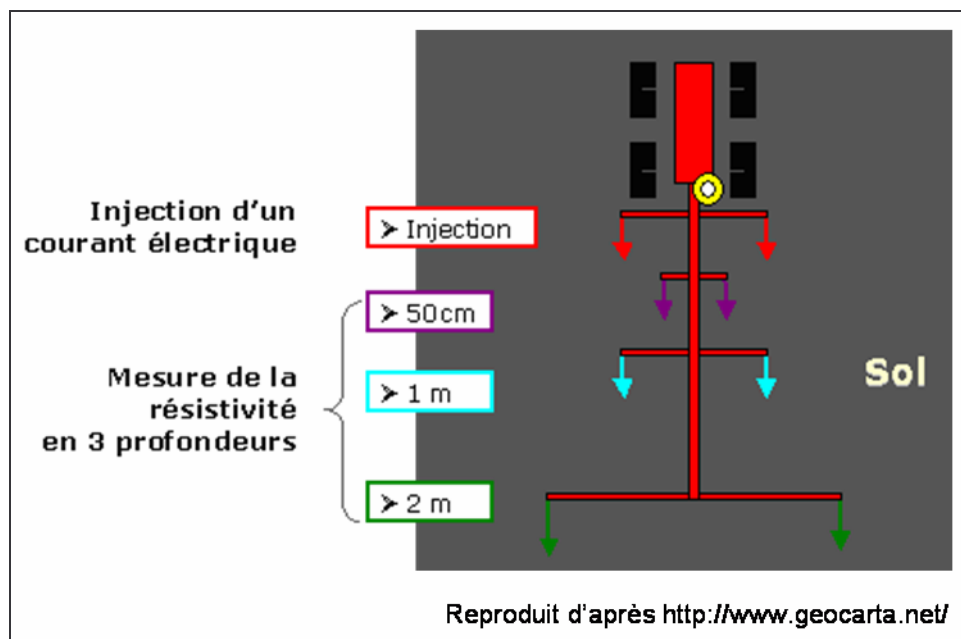


Figure 13 : Méthode de mesure de la résistivité électrique

Les cartes de résistivité électrique sont obtenues par interpolation des semis de points fournis par la méthode MuCEP [Dabas, Rouiller, Tabbagh 2000] [Dabas, Tabbagh, Boigontier 2001].

La mesure de la conductivité et de la résistivité électrique présente de nombreux avantages :

- Le paramètre mesuré est corrélé avec de nombreuses caractéristiques granulométriques et physico-chimiques pertinentes agronomiquement.
- Les mesures effectuées sont géolocalisées très précisément sur la parcelle.
- Le semis de points est très dense et assure une bonne restitution de la structure spatiale de la parcelle.
- Le coût est de l'ordre d'une centaine d'euros par hectare.
- Le paramètre mesuré varie peu dans le temps et, par conséquent, l'information obtenue reste valable très longtemps.
- Le paramètre est mesuré dans le sol même, contrairement aux cartes de rendement ou à de nombreuses images satellitaires. Cela explique les bonnes corrélations avec les informations fournies par des analyses de sol.

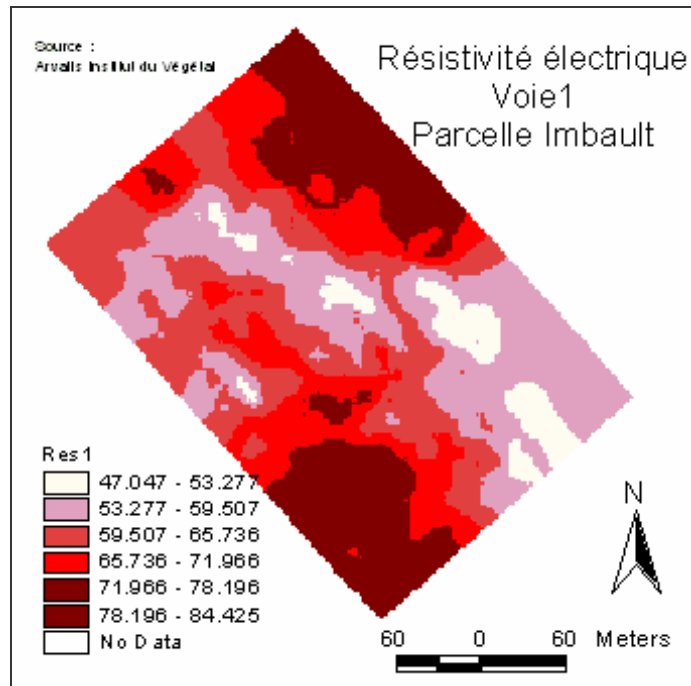


Figure 14 : Exemple de carte de résistivité électrique sur la parcelle Imbault

5.1.5. Les croisements d'informations

Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers et Payton [Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers Payton 2000] ont montré que l'altitude, la couleur du sol, l'aspect, la pente et la conductivité électrique expliquaient près de 65% de la variation du rendement pour cinq années de culture de maïs. Ils considèrent donc ces données comme adaptées à la définition d'entités homogènes.

5.2. S.I.G. et agriculture de précision

Comme la grande majorité des démarches s'appuyant sur l'information géographique, l'agriculture de précision tire partie des S.I.G., Systèmes d'Information Géographique.

5.2.1. Définition des S.I.G.

Les toutes premières utilisations de S.I.G. remontent aux années 1960. La plus grande partie d'entre elles sont originaires d'Amérique du Nord. On peut distinguer plusieurs périodes dans le développement et la diffusion des S.I.G. [Coppock, Rhind 1991]. La première va des années 1960 aux alentours de 1975, époque de découverte où quelques personnalités et organismes ont joué un rôle crucial. La seconde va de 1975 au début des années 1980 et a vu les S.I.G. faire l'objet d'utilisations et d'expérimentations plus régulières, en particulier au sein des organismes d'état. La troisième période, plus courte, a duré d'environ 1982 à la fin de la même décennie et a été marquée par la diffusion et l'utilisation commerciale des S.I.G. La quatrième période, qui se poursuit encore, s'oriente vers les besoins et les attentes de l'utilisateur et est marquée par une compétition commerciale. Parallèlement à cette compétition interviennent des débuts de standardisation des formats d'échange de données.

Goodchild considère que c'est la technologie qui est à l'origine de l'émergence des S.I.G. L'augmentation de la puissance de calcul des ordinateurs et de leurs capacités graphiques a permis de développer ces logiciels sur la base des Systèmes de Gestion de Base de Données (S.G.B.D.) [Goodchild 1991].

Pour Fischer et Nijkamp, "un S.I.G. peut être défini comme un S.G.B.D. conçu pour saisir, stocker, manipuler, analyser et afficher des données à référence spatiale en vue de résoudre des problèmes complexes de gestion et de planification" [Fischer, Nijkamp 1993].

Denègre et Salgé considèrent qu'un S.I.G., comme tout S.G.B.D., doit pouvoir "*acquérir, archiver, accéder, analyser et afficher*" les données [Denègre, Salgé 1996]. Cela signifie que le S.I.G. doit permettre :

- L'acquisition des données d'entrée.
- Le stockage, la récupération et la gestion de bases de données.
- La manipulation et l'analyse des données.
- L'affichage et la génération de produits.
- Et doit offrir une interface à l'utilisateur.

Enfin, Girard et Girard définissent un S.I.G. comme un "système informatique permettant, à partir de diverses sources, de rassembler et organiser, gérer, analyser et combiner, élaborer et présenter des informations localisées géographiquement, contribuant notamment à la gestion de l'espace" [Girard, Girard 1999].

5.2.2. Modes de représentation de l'information géographique des S.I.G.

5.2.2.1. Le modèle matriciel

Une matrice, dans son acception géomatique, est normalement formée d'éléments de forme et de taille uniformes appelés pixels (pour picture element) ou points. Avec la modélisation cartographique matriciel (couramment appelée "mode image" ou "mode raster"), ces pixels sont localisés par des coordonnées mais, contrairement à celles-ci, possèdent une superficie.

Par rapport au mode vectoriel, les points, polygones et polygones sont donc toujours représentés par des surfaces (voir Figure 15). Une polygones sera par exemple un long groupe de pixels très étroit. Ici, la notion d'entité géographique, de même que la notion d'objet graphique, n'existe pas vraiment. On retrouve simplement des groupes de pixels contigus ayant les mêmes valeurs pour une même variable.

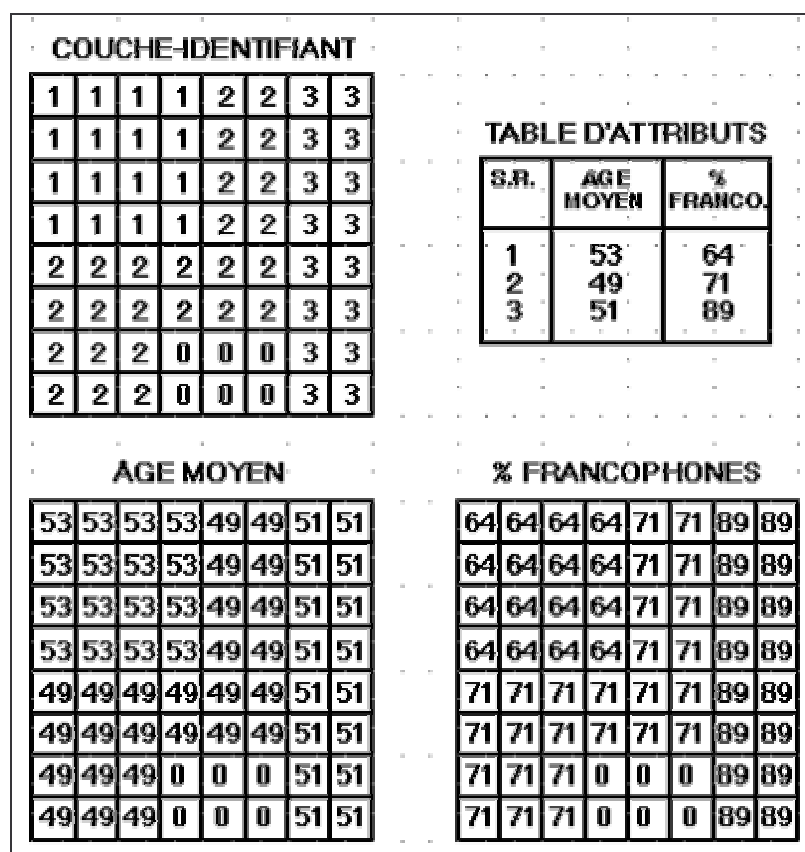


Figure 15 : Exemples de représentation de données en mode image

Les S.I.G. utilisant ce mode de représentation des données offrent souvent des méthodes pour développer assez facilement des opérations analytiques. Toutefois, l'orientation du mode image tend à s'intéresser davantage au contenu des aires (autrement dit les valeurs des points) qu'aux limites entre ces aires. Toutes les opérations ayant trait aux limites ou à des figurés linéaires, comme c'est souvent le cas pour les problématiques socio-économiques, ne sont donc pas facilitées par ce mode de représentation. En revanche, il s'accorde généralement bien des problématiques environnementales [Maguire, Dangermond 1991].

5.2.2.2. Le modèle vectoriel

La modélisation cartographique vectorielle (couramment appelée "mode vectoriel" ou "mode vecteur") demande que l'on associe un objet géographique à une entité géographique.

Les objets sont donc tous décrits dans l'espace soit par des lignes, des points ou des polygones (voir Figure 16). Un bâtiment peut ainsi devenir un point, une rivière peut devenir une polyligne, un lac un polygone, etc...

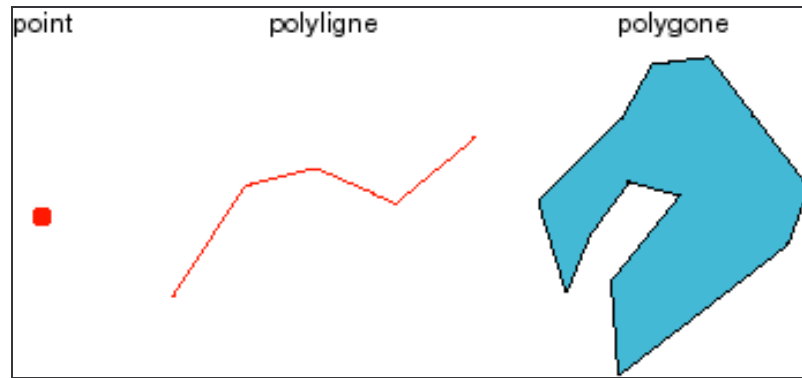


Figure 16 : Figurés élémentaires permettant la représentation vectorielle des objets géographiques en mode vecteur

Ce mode de représentation est plus complexe que le mode image. La structuration des données en une base de données est souvent plus longue et plus difficile à réaliser. La complexité de l'organisation des données induit que la plupart des opérations analytiques (voir section 5.3) sont conceptuellement complexes et lourdes à mettre en œuvre [Maguire, Dangermond 1991]. Le mode de représentation vectoriel est bien adapté aux problématiques socio-économiques, en particulier celles concernant des phénomènes de densité variables et des réseaux.

5.2.3. L'utilité des S.I.G. en agriculture de précision

Plusieurs caractéristiques des S.I.G. justifient leur utilisation en agriculture de précision.

5.2.3.1. L'intégration, la gestion et l'affichage des données

Le principal atout des S.I.G. est leur rôle intégrateur de données [Maguire 1991] [Maguire, Dangermond 1991]. En effet, ils acceptent de travailler avec de nombreuses données, quelque que soit leur origine (carte numérisée, imagerie satellitaire, données G.P.S. ou station totale...), leur structure (vectorielle ou en tessellation) et leur type (quantitative ou qualitative). Ils permettent donc de réunir sous un même format les différentes formes que peuvent prendre les données à référence spatiale.

Une fois les données intégrées, les S.I.G. offrent également de nombreuses fonctionnalités nécessaires à la gestion et à la mise à jour des bases de données constituées.

Enfin, il est possible, avec un minimum de manipulation, de mettre en forme les données et de les appréhender aisément sous leur forme cartographique.

Devant la diversité, et parfois la quantité, de données géoréférencées utilisées en agriculture de précision (voir section 3.2.1), c'est ce rôle intégrateur de données qui a amené les acteurs de l'agriculture de précision à s'intéresser aux S.I.G. [Reetz 1998] [Ariaux 2000]. Cette caractéristique fait qu'ils sont mis à contribution dès la première étape d'identification et de caractérisation de la variabilité intraparcellaire (voir section 3.2.1).

5.2.3.2. L'aspect spatial des données gérées

Par rapport à un S.G.B.D. classique, le S.I.G. a la particularité de gérer les données à référence spatiale. Les données sont donc associées à des objets ayant une position, et souvent aussi, une forme dans l'espace. Les données qui concernent donc la position et la forme de l'objet sont appelées "données géométriques".

Pour Maguire, la principale différence entre les S.I.G. et les S.G.B.D. dont ils sont issus réside dans le fait que les éléments spatiaux des objets sont plus importants que leurs attributs [Maguire 1991].

De fait, par rapport à un S.G.B.D. classique, la considération spatiale des seules données géométriques peut mettre en valeur de nouveaux aspects de la réalité peu ou pas abordés avec les attributs descriptifs classiques. En allant encore un peu plus loin, la topologie, accessible sur certains S.I.G. comme Arc-Info, permet la gestion des liens directement issus des formes géométriques des objets.

Pour sa part, Grenier remarque que "le S.I.G. est l'outil qui s'impose lorsqu'on intègre dans la démarche agronomique la dimension spatiale des phénomènes observés". Il ajoute : "Il s'agit d'un outil qui permet d'aller beaucoup plus loin que la simple représentation cartographique de ces phénomènes" [Grenier 2000].

5.2.4. Les logiciels dédiés à l'agriculture de précision

Ariaux considère "qu'une bonne maîtrise et une exploitation valorisée des relations spatiales entre objets repose sur une bonne connaissance des S.I.G." [Ariaux 2000]. Elle est l'un des rares auteurs à avoir établi une typologie des S.I.G. utilisés en agriculture de précision. Parmi ces logiciels, elle distingue :

- Les logiciels de cartographie : Ces logiciels sont principalement destinés au stockage et à la visualisation de l'information collectée sur les parcelles. Exemples : Instant Yield Map (Case IH), Micosat (Micotron), GPX (Agrotronix).
- Les S.I.G. sans modèle agronomique : Le terme S.I.G. est ici adéquat puisque ces logiciels ajoutent aux fonctionnalités cartographiques quelques possibilités d'analyse spatiale. Ils permettent également de créer manuellement des cartes de préconisation et de les transférer vers un système de modulation. Exemples : Instant Application Map (Case IH), Field Star (Massey-Ferguson), JD Map (John Deere), Plot Plan (RDS), ProFas (Kverneland), AgroMap Basic (Claas), Agri 4d (Agri 4d).
- Les S.I.G. avec modèle agronomique : Ce sont les outils les moins nombreux et aussi les plus puissants. Ils permettent de générer automatiquement des cartes de préconisation à partir de modèles agronomiques travaillant sur la base de données. Exemples : Agriprecis (CDER Informatique), Loris (Kemira), Agri+ (ESRI-Arvalis Institut du Végétal-GEOSYS).

5.3. Analyse spatiale et agriculture de précision

Lors de l'étape de prise de décision de l'agriculture de précision (voir section 3.2.2), les données collectées sur la parcelle et structurées en une base de données au sein d'un S.I.G. sont sensées être analysées afin d'en extraire des informations supplémentaires. Cette étape s'appuie donc largement sur l'analyse spatiale pour établir la carte de préconisation.

5.3.1. Définition de l'analyse spatiale

L'analyse spatiale trouve ses origines dans l'émergence de la géographie quantitative et statistique des années 1950 [Openshaw 1991]. Elle s'appuyait initialement sur l'application de méthodes statistiques à des données géographiques. Les préceptes de l'analyse spatiale sont donc antérieurs à l'apparition des S.I.G. et les deux domaines ont évolué indépendamment l'un de l'autre pendant un certain temps [Joërin 1997].

Cependant, l'intégration de certaines méthodes d'analyse spatiale dans les S.I.G. était inévitable. Les S.I.G. offraient l'outil informatique adéquat pour effectuer d'importants calculs tandis que l'analyse spatiale amenait un ensemble de méthodes susceptible de mettre en valeur les bases de données géographiques [Openshaw 1991]. Ainsi, la croissance des S.I.G. a contribué à l'augmentation de la recherche en analyse spatiale [O'Kelly 1994].

Les modèles d'analyse spatiale et les S.I.G. ont des similarités, comme le besoin en données à référence spatiale et les méthodes d'analyse pour les traiter, mais des différences inhérentes aux approches suivies demeurent. Ainsi, un modèle d'analyse spatiale insiste sur les traitements des phénomènes et des dynamiques spatiales alors qu'un S.I.G. s'attache plus aux éléments et aux objets spatiaux.

Les relations qui unissent les S.I.G. et les outils de la géomatique en générale d'une part et l'analyse spatiale d'autre part font que les définitions de l'analyse spatiale peuvent grandement varier selon les origines disciplinaires de leurs auteurs, comme d'autres auteurs l'ont déjà constaté [Joërin 1997] [Laaribi 2000].

Ainsi, Bavoux, Chapuis, Delmer, Mannone, Passegué et Volpoët définissent l'analyse spatiale comme suit : "l'analyse spatiale est science nomothétique. Elle essaie d'identifier des configurations spatiales qui se répètent, des structures que l'on retrouve en plusieurs lieux ou à plusieurs niveaux. A partir de là, elle s'efforce, à travers l'élaboration de modèles, de formuler des lois générales, ou au moins, des principes d'organisation de l'espace. Appuyée sur une réflexion théorique et bien que ne méconnaissant pas la riche diversité de l'espace, elle se différencie de l'approche idiographique qui vise à identifier des particularités spatiales." [Bavoux, Chapuis, Delmer, Mannone, Passegué, Volpoët 1998]. Ici, c'est l'aspect modélisation de phénomènes spatiaux qui est mis en avant. Aucun lien n'est établi avec quelque type de logiciels ou d'outils géomatiques que ce soit.

Pour Girard et Girard, l'analyse spatiale "consiste à étudier les relations spatiales qui existent entre divers objets répartis dans un plan (ou un volume). Les objets peuvent être étudiés en fonction de leur abondance, évaluée par des aires ou des pourcentages de superficie. Il est également très intéressant d'évaluer la nature des voisinages entre les divers types d'objets" [Girard, Girard 1999]. Cette définition n'aborde pas les questions de phénomènes spatiaux. Elle limite l'analyse spatiale à l'étude des relations spatiales entre objets sans recherche de modélisation. Elle est en cela assez proche de certaines notions d'analyse d'images.

Pour Haining [Haining 1994], "par contraste avec d'autres formes d'analyse, l'analyse spatiale requiert l'information à la fois sur les valeurs attributaires et sur la localisation géographique des objets auxquels les attributs sont liés" [Haining 1994]¹. Pour cet auteur, l'un des objectifs de l'analyse spatiale est "l'exploration systématique des "patterns" d'événements et l'association entre ces événements dans l'espace afin d'accéder à une meilleure compréhension des processus qui pourraient être responsables de la distribution observée de ces événements"². Cette définition fait le lien entre l'analyse spatiale en tant que recherche de modèles concernant des phénomènes spatiaux (telle que Bavoux, Chapuis, Delmer, Mannone, Passegué et Volpoët la définissent) et le mode de représentation des données géoréférencées dans les S.I.G. (liens entre les informations attributaires et les informations de localisation) (voir section 5.2.2).

Openshaw nous incite à établir clairement une distinction entre les opérations relevant de l'analyse spatiale et celles relevant de la manipulation des données [Openshaw 1991]. Pour ces dernières, telles la création de zones tampon, les superpositions de couches et autres, on ne peut parler de méthodes d'analyse spatiale qu'au sens de description des données ou au sens cartographique.

Parmi les définitions existantes de l'analyse spatiale, plusieurs caractéristiques ressortent [Laaribi 2000] : "Ces analyses touchent aux attributs descriptifs des objets, à leur localisation et leur forme (attributs géométriques) ainsi qu'à leur topologie (connectivité). Toutes ces opérations vont tirer parti de la distribution spatiale des objets ou des phénomènes sur le territoire, soit pour des simples analyses géométriques et arithmétiques, soit pour des opérations relativement plus sophistiquées relatives à des calculs d'optimisation ou de simulation."

¹ Traduction libre de [Haining 1994] : "In contrast with other forms of analysis, therefore, spatial analysis requires information both on attribute values and the geographical locations of the objects to which the collection of attributes are attached."

² Traduction libre de [Haining 1994] : "systematic exploration of the pattern of events and the association between events in space in order to gain a better understanding of the processes that might be responsible for the observed distribution of events."

Ainsi, l'analyse spatiale doit permettre de développer des aperçus ou de la connaissance à partir des caractéristiques et des associations trouvées dans les données. Ces types de conclusion peuvent alors être utiles et éventuellement servir à orienter des analyses plus fines. Selon Openshaw [Openshaw 1991], ce rôle qui consiste à indiquer quelle est la direction à suivre est un objectif important de l'analyse spatiale.

De nombreux auteurs s'accordent pour considérer qu'un S.I.G. se distingue d'un logiciel de cartographie numérique par ses capacités d'analyse spatiale [Maguire 1991] [Maguire, Dangermond 1991] [Sheperd 1991]. En effet, celles-ci offrent de puissants outils pour valoriser les données et sont donc l'un des aspects fondamentaux des S.I.G.

5.3.2. Les principales méthodes d'analyse spatiale des S.I.G.

Laaribi avance un classement des capacités d'analyse spatiale [Laaribi 2000] :

- Explorations topologiques et métriques relatives à la recherche d'objets par la création de zones tampon (buffer zones en anglais) et aux calculs de distance ou de surface.
- Explorations topologiques et métriques relatives à des analyses d'impact.
- Superposition de "couches" de données par des opérations booléennes qui permettent d'identifier et d'afficher des entités géographiques ayant des caractéristiques communes.
- Calculs statistiques sur certaines variables.
- Calculs d'optimisation relatifs à l'analyse de réseaux.

Pornon considère que "pour certains organismes (agriculture, environnement...), elles [les fonctions d'analyse spatiales, N.D.A.] sont indispensables" [Pornon 1992].

5.4. Limites et lacunes des capacités d'analyse spatiale des S.I.G.

Openshaw [Openshaw 1991] affirme que les outils d'analyse spatiale des S.I.G. doivent maintenant se mettre à portée des utilisateurs en adoptant des méthodes plus souples et flexibles, à l'inverse des méthodes principalement statistiques actuelles. Si celles-ci fournissent d'excellents moyens d'exploration des données lorsque aucune hypothèse n'a pu être formulée, elles demeurent relativement complexes à utiliser et requièrent certaines connaissances préalables pour interpréter les résultats.

Openshaw distingue huit facteurs limitant les possibilités de l'analyse spatiale [Openshaw 1991] :

- L'absence d'une théorie ou d'hypothèses préalables, ce qui force l'utilisateur à commencer par rechercher des caractéristiques ou des relations sans savoir ce qu'il cherche, ou seulement sans savoir s'il y a quelque chose à trouver.
- La difficulté d'opérer dans un contexte exploratoire pour lequel la connaissance des données complique la mise à l'essai d'hypothèses et de modèles post hoc.
- Les données géographiques disponibles sont souvent des substituts aux données dont on a réellement besoin, ces dernières étant souvent indisponibles voire inexistantes.
- Les S.I.G. sont souvent riches en termes de données mais, pour la plupart des applications, les méthodes qui seraient vraiment utiles sont rarement disponibles.

- La plupart des analyses étant écologiques ou environnementales, l'agrégation de données peut être pénalisante dans la mesure où elle peut changer la nature de relations à très grande échelle ou en créer d'autres totalement fausses.
- Les données spatiales sont caractérisées par une très grande complexité.
- Le problème classique de la validité et de la pertinence des données.
- Les difficultés à prendre en compte le temps.

Parmi ces huit facteurs limitants évoqués par Openshaw, il nous semble nécessaire d'opérer une distinction. En effet, certains de ces facteurs sont plus ou moins directement liés à l'analyse spatiale et constituent en ce sens des approfondissements ou des compléments de ces modes d'analyse. Ces facteurs constituent donc des limites à l'analyse spatiale des S.I.G. D'autres facteurs correspondent plus à des manques ou à des lacunes des S.I.G. qu'à des besoins d'approfondissement ou de compléments.

5.4.1. Les limites

5.4.1.1. *Le manque de fonctionnalités analytiques*

Comme nous l'avons vu, les deux principaux atouts des S.I.G. résident dans leurs capacités d'intégration et de manipulation de données et dans leurs capacités d'analyse spatiale. Malgré cela, ces dernières restent relativement élémentaires et limitées par rapport aux premières.

Openshaw dénonce le manque de distinction entre l'analyse spatiale et la simple manipulation des données [Openshaw 1991]. Selon lui, des fonctions telles que les zones tampon ou les superpositions de couches ne sont des analyses qu'au sens descriptif ou cartographique. Cela s'explique par le fait que les S.I.G. sont originellement basés sur la carte. Ils s'attachent donc plus à la manipulation cartographique qu'à une véritable analyse des données [Laaribi 2000]. Et ils sont par conséquent plus efficaces lorsqu'il s'agit de gérer et manipuler les données que lorsqu'il s'agit de contribuer à l'analyse et à la planification.

La technique de superposition de couches (overlay en anglais) illustre bien cette dualité entre la capacité de combiner de l'information et les capacités analytiques :

- Bien que les S.I.G. soient capables de superposer une grande quantité de couches d'information, le résultat devient inextricable et donc difficile à comprendre au delà de quatre ou cinq couches [Openshaw 1991].
- Chaque couche d'information est considérée d'égale importance avec les autres. Il n'est donc pas possible de pondérer les informations entre elles [Densham 1991].
- Le recours à des valeurs de seuil et à la définition de classes engendre une diminution de la qualité de l'information.

5.4.1.2. *La difficile intégration de l'analyse spatiale aux S.I.G.*

L'analyse spatiale et les S.I.G. sont deux domaines dont l'émergence et l'évolution sont distinctes (voir section 5.4). Ainsi, les contraintes d'application de certains modèles d'analyse spatiale peuvent ne pas correspondre aux modèles de données ou aux modes d'analyse des S.I.G. (superposition de couches non pondérées, polygones exclusifs entre eux...).

Triboulet, Lardon et Langlet [Triboulet, Lardon, Langlet 1996] insistent sur le fait qu'il est important de disposer de techniques travaillant à la fois sur les aspects statistiques et spatiaux de l'information géographique. Mais ils reconnaissent qu'il existe peu de méthodes fiables permettant de prendre les deux aspects en compte simultanément. Pour eux, *"la prise en compte du spatial dans le statistique ne permet pas de mesurer le poids respectif de l'information liée au lieu et de celle liée à son voisinage. Les résultats sont donc difficilement validables et interprétables car les règles d'agrégation sont peu explicites"*. Ils finissent par conclure que *"séparer le statistique du spatial permet donc de mieux contrôler les processus mis en jeu et donc d'asseoir leur validité respective."*

5.4.1.3. L'absence de méthodologies

Confrontés à un même problème spatial, différents utilisateurs vont recourir à des fonctionnalités différentes du S.I.G. ou, s'il utilisent les mêmes, ils ne le feront pas de la même manière. Au final, dans la grande majorité des cas, les résultats obtenus seront différents. Si cette démarche intuitive peut être jugée acceptable pour des opérations simples, elle s'avère vite peu efficace pour des analyses plus complexes s'intégrant dans un mode opératoire lui-même bien formalisé [Openshaw 1991].

Ainsi, Zheng, Kainz et Groten considèrent que les S.I.G. restent des outils de spécialistes en dépit de leur diffusion croissante [Zheng, Kainz, Groten 2000].

Les techniques d'analyse spatiale des S.I.G., souvent développées dans des domaines différents, souffrent donc de l'absence d'un cadre théorique et méthodologique global. Un utilisateur peut appliquer une méthode d'analyse spatiale sans même en maîtriser les principes.

5.4.2. Les lacunes

5.4.2.1. Les classifications avec contrainte de contiguïté

A l'heure actuelle, quelques S.I.G. proposent des méthodes de classification. *k*-moyennes, classification ascendante hiérarchique, ces méthodes d'analyse statistiques offrent l'avantage de pouvoir extraire de l'information a posteriori du jeu de données initial.

Ces méthodes sont toutefois souvent appliquées telles quelles et ne prennent pas en compte d'éventuelles relations liées à l'impact spatial que l'utilisateur pourrait souhaiter conserver dans le résultat final. Autrement dit, ces méthodes de classification sont rarement subordonnées ou appliquées conjointement à une contrainte de contiguïté. Cette lacune a pourtant été rapidement identifiée par certains auteurs [Openshaw 1991].

Ainsi, à l'heure actuelle, la grande majorité des S.I.G. ne propose pas de classification avec contrainte de contiguïté. En agriculture de précision, cela pose un réel problème car la modulation directe, comme nous l'avons vu en section 3.1, est peu opérationnelle.

On notera toutefois que plusieurs méthodes de zonage spécifiquement destinées à l'agriculture de précision existent. Mais ces méthodes n'intègrent pas toutes des contraintes spatiales, sont rarement implémentées dans un S.I.G. et ne sont pas exemptes de défauts. Plusieurs d'entre elles seront présentées en section 6.2.

5.4.2.2. Le problème des modèles

Un S.I.G. sert à gérer l'information géographique afin de manipuler un type de modèle de la réalité de l'espace. Mais manipuler un modèle est différent de le créer [Laaribi 2000]. Ainsi, si un S.I.G. est bel et bien capable de manipuler différents types de modèles, il n'est pas pour autant capable d'en créer. Il est donc nécessaire de concevoir de nouvelles méthodes d'exploration des données pour créer de nouveaux outils analytiques qui seront à plus ou moins long terme intégrés aux S.I.G.

Toutefois, une certaine prudence s'impose. La relative facilité qu'ont les S.I.G. à manipuler les résultats de plusieurs modèles et le fait que certains logiciels commerciaux commencent à s'ouvrir à des langages de programmation puissants et répandus (tel Arc-GIS avec le C++) peuvent avoir des conséquences néfastes. Ainsi, il devient possible d'implémenter rapidement dans un S.I.G. une méthode bien étudiée dans un autre contexte sans pour autant s'assurer de la pertinence de sa nouvelle application.

5.4.2.3. L'absence d'explications sur la causalité

Openshaw considère que les S.I.G. ne fournissent aucune information sur la causalité des phénomènes spatiaux analysés [Openshaw 1991]. S'il est possible de mettre en évidence des formes ou des relations entre des objets géographiques, on ne dispose en revanche d'aucune explication sur les raisons qui ont pu générer ces formes ou ces relations.

5.4.2.4. L'absence d'aspects temporels

Openshaw remarque également que les S.I.G. ne sont pas capables de réaliser des analyses dynamiques dans le temps sur des objets spatiaux [Openshaw 1991]. Pourtant, ce type de problématique, évolution dans le temps et dans l'espace d'un phénomène donné, est très souvent traitée, en particulier par les décideurs sur le territoire.

S'il reste possible d'entreprendre une étude comparative sur des cartes à des dates différentes, les S.I.G. n'offrent cependant pas d'outils d'analyses spécifiques pour ces études.

5.4.2.5. Un outil d'aide à l'information plus qu'un outil d'aide à la décision

Les décideurs mettent souvent l'accent sur leurs besoins de capacités analytiques de prédiction ou de simulation afin de pouvoir évaluer les effets possibles de leur décision. Ce type de besoin est rarement pris en compte dans les S.I.G., induisant plusieurs manques qui concernent :

- La prise en compte des solutions proposées par le décideur.
- La prise en compte des préférences du décideur.
- La prise en compte des aspects conflictuels.
- La comparaison des solutions potentielles.

De nombreux auteurs, qu'ils appartiennent à la communauté géomatique ou agriculture de précision, vont dans ce sens en affirmant que les S.I.G. ne permettent pas de tenir compte des attentes et des besoins des décideurs [McRae, Cleaves 1986] [Maguire 1991]. Sheperd [Sheperd 1991] et Densham [Densham 1991] rejoignent ce point de vue en déclarant que les S.I.G. s'attachent plus aux besoins techniques qu'aux besoins humains.

Ainsi, les S.I.G. s'appuient bien plus sur les données disponibles que sur la connaissance et les préférences du décideur. Même s'ils répondent à plusieurs caractéristiques des problèmes à référence spatiale, leur rôle est souvent limité à l'aide à l'information et non à l'aide à la décision.

Cette distinction entre aide à l'information et aide à la décision est extrêmement importante, "*une véritable aide à la décision suppose le recours à une "modélisation sémantique" et à la simulation*" [Dumolard, David 1992]. Trop souvent, des utilisateurs de S.I.G. considèrent leur outil comme un outil d'aide à la décision alors que l'utilisation qu'ils en font relève de l'aide à l'information. Cette erreur est fréquente dans le milieu de l'agriculture de précision où les utilisateurs des S.I.G. sont rarement des spécialistes de ces outils.

Il ne faut cependant pas nier le rôle que peuvent jouer les S.I.G. dans la décision. Comme le souligne Didier et Bouveyron, "*l'information améliore la décision*" [Didier, Bouveyron 1993]. Si le S.I.G. ne peut pas à proprement parler être abordé comme un outil d'aide à la décision, il contribue néanmoins à valoriser l'information disponible et donc à améliorer indirectement la décision [Dumolard, David 1992].

6. Le développement d'une méthode de zonage

Notre travail va consister en la conception et le développement d'une méthode de zonage propice à la définition d'entités spatiales homogènes en agriculture de précision. La définition d'entités homogènes pourra à son tour servir les deux besoins identifiés dans les sections 3.2.2.2 et 5.1.1, à savoir la réalisation d'une carte de préconisation et l'aide à la localisation des analyses de sol.

6.1. Les contraintes de travail

Notre but premier concerne donc la mise au point d'une méthode de zonage. Autrement dit, il s'agit de déceler des discontinuités spatiales sur la parcelle afin de représenter son hétérogénéité de façon lisible, comme nous l'avons vu en section 3.2.2.2.

A ce but premier s'ajoutent plusieurs contraintes spécifiques concernant la méthode de zonage. Ces contraintes forment autant d'objectifs secondaires :

- Les différents exploitants agricoles ne disposeront pas tous des mêmes jeux de données (résistivité électrique, rendement, image satellitaire, analyses de sol...). Et une même donnée peut prendre plusieurs formes selon son mode de mesure, la précision de cette mesure, l'éventuelle méthode d'interpolation utilisée... De plus, l'exploitant agricole doit avoir la possibilité d'intégrer des cartes issues de ses propres observations sur le terrain. Notre méthode doit être en mesure de traiter des données d'entrée nombreuses et diverses.
- La méthode doit permettre d'éviter tout a priori sur les données traitées. Cet objectif a été formulé initialement avec l'espoir que la méthode de zonage pourrait servir à mettre en évidence des relations entre des variables qui auraient semblé sans cela indépendantes. Il nous faudra également envisager l'automatisation d'un certain nombre de choix relevant de l'utilisateur, en particulier celui du nombre de zones à conserver sur la carte finale.
- A l'opposé, si l'exploitant agricole est en mesure d'établir une échelle d'importance entre les différentes données traitées, la méthode de zonage doit pouvoir l'intégrer. En effet, à l'heure actuelle, de nombreux auteurs considèrent que la connaissance et l'avis de l'exploitant agricole ne sont pas assez mis en valeur, même si ceux-ci ont leurs limites [Boisgontier 1997a] [Boisgontier 2000] [Thévenet 1999] [Fleming, Westfall 1999] [Fleming, Westfall, Bausch 2000] [Dillon, Mueller, Shearer 2003].
- La méthode de zonage doit être aisée d'utilisation. Domsch, Kaiser, Witzke et Zauer affirment que les exploitants agricoles qui souhaitent réaliser quelques tests sur leurs parcelles veulent pouvoir utiliser des méthodes simples et des logiciels courants [Domsch, Kaiser, Witzke, Zauer 2003]. De plus, nous avons vu en section 4.2 que Robert considère que les logiciels difficiles à prendre en main constituent des freins au développement de l'agriculture de précision [Robert 2000]. Le nombre de paramètres à renseigner doit être donc réduit à son minimum.
- La carte de zonage finale doit être aisément applicable sur le terrain. Elle doit proposer un nombre restreint de zones et éviter les points isolés. Qui plus est, les limites et contraintes techniques des appareils agricoles ne permettent pas de travailler à une échelle trop grande [Mérour 2002].

- Quelques informations permettant d'appréhender la qualité du zonage doivent être fournies.

La méthode doit offrir la possibilité d'intégrer de nouvelles données au fur et à mesure des acquisitions de l'exploitant agricole.

Prenant en compte ces objectifs, nous avons envisagé différentes approches pour mettre au point notre propre méthode de zonage. Les approches qui n'ont pas été retenues et les raisons justifiant ce choix sont présentées en section 6.3.

6.2. Les méthodes de zonage existantes

A l'heure actuelle, plusieurs méthodes de zonage et/ou de classification sont utilisées en agriculture de précision. Nous présentons ici certaines de ces méthodes décrites dans la littérature. Nous nous intéresserons en particulier à celles qui disposent de contraintes spatiales et/ou qui se rapprochent de notre propre problématique.

6.2.1. Méthodes sans contraintes spatiales

6.2.1.1. Méthodes fondées sur les *k*-moyennes

La plupart des méthodes font appel à la classification par *k*-moyennes (voir section 10.2.4). Cette méthode dérive de la méthode des centres mobiles [Minvielle, Souiah 2003] [Lebart, Morineau, Piron 2000].

La classification par *k*-moyennes est notamment utilisée par Cupitt et Whelan [Cupitt, Whelan 2001] et Berry [Berry 1998]. L'utilisation des *k*-moyennes dans ces méthodes consiste en l'établissement d'une seule classification sans vérification particulière de la qualité du résultat ni de la justesse du nombre de classes retenu. Dans le cas du Cupitt et Whelan, les données traitées sont des cartes de rendement.

6.2.1.2. La méthode FuzME et les classifications floues

Le programme FuzME vise à pallier la trop grande rigidité des méthodes de classification habituelles et permet donc d'établir des classifications floues [Boydell, McBratney 1999] [Fridgen, Fraisse, Kitchen, Sudduth 2000]. Il a été développé par l'Australian Center for Precision Agriculture (A.C.P.A.)¹.

En classification floue, chaque point de la carte (en mode image) est susceptible d'appartenir à toutes les classes à la fois. Ce sont donc ses degrés d'appartenance aux différentes classes qui caractérise un point, et non son appartenance à une classe unique. La somme des degrés d'appartenance d'un point est égale à 1, chaque degré d'appartenance étant donc compris entre 0 et 1. Plus le degré d'appartenance d'un point à une classe est élevé, plus ce point a des valeurs semblables aux valeurs moyennes des autres points de la classe.

Des indices mesurent les performances de l'agrégation et permettent donc d'estimer le meilleur nombre de zones à retenir sur la parcelle.

Mais l'utilisation du flou présente toutefois un inconvénient majeur : elle établit une carte de zonage par classe. Ainsi, si l'on traite trois cartes initiales mais que l'on désire effectuer un zonage à cinq zones, le programme restitue cinq cartes finales, une par classe. Et chaque point se voit attribuer un degré d'appartenance à chacune des cinq classes. L'information initiale est donc résumée mais il reste ensuite à synthétiser les cartes de zonage finales, ou au moins à en choisir une, ce pour quoi aucune démarche n'est proposé.

Cette approche, pour intéressante qu'elle soit, nous semble donc peu opérationnelle dans le cadre d'une problématique visant à résumer l'information initialement disponible.

¹ <http://www.usyd.edu.au/agric/acpa/welcome.html>

De plus, Dobermann, Ping, Simbahan et Adamchuk ont montré que la méthode "fuzzy k -means" est très sensible aux données d'entrée et au nombre de classes retenu [Dobermann, Ping, Simbahan, Adamchuk 2003].

Lark et Stafford utilisent une méthode similaire de classification floue [Stafford, Lark 1997] [Lark, Stafford 1998]. Le fonctionnement de la méthode et ses inconvénients sont les mêmes.

Mérour [Mérour 2002] a proposé une méthode simple pour concaténer les différentes cartes floues obtenues à l'issue de la classification. Sa méthode repose toutefois sur l'arbitraire de l'utilisateur et est donc difficile à mettre en œuvre dans le cadre d'une méthode automatisée.

6.2.2. Méthodes avec contraintes spatiales

Comme nous l'avons vu en section 6.2.1, de nombreuses méthodes utilisées en agriculture de précision ne s'appuient que sur des classifications sans aucune contrainte spatiale. Ce point est abordé par de nombreux auteurs et souvent identifié comme une lacune [Stafford, Lark 1997] [Stafford, Lark, Bolam, ?] [Chung, Sudduth, Drummond, Kitchen 2001] [Van Alphen, Stoorvogel 1999] [Fridgen, Fraisse, Kitchen, Sudduth 2000] [Boydell, McBratney 1999] [Cupitt, Whelan 2001] [Triboulet, Lardon, Langlet 1996] [Berry 1998] [Grenier 1997] [Openshaw 1991] [Shatar, McBratney 2001]. Plusieurs auteurs ont donc tenté d'intégrer aux méthodes de classification des contraintes spatiales. Quelques méthodes sont présentées ici.

6.2.2.1. La méthode k -zones de Shatar et McBratney

Shatar et McBratney [Shatar, McBratney 2001] ont conçu une méthode de classification avec contrainte de contiguïté qui suit le mode opératoire suivant :

- k centres de zones sont définis par une classification par k -moyennes appliquée aux coordonnées géographiques et aux attributs des points. Une triangulation de Delaunay permet d'attribuer les points de la parcelle aux k centres de zones.
- L'algorithme du "convex hull"¹ est alors appliqué à chaque classe et un traitement permet d'éventuellement réaffecter certains points en fonction de la distance moyenne entre leurs attributs.

La méthode initiale, à savoir la classification par k -moyennes appliquée aux coordonnées géographiques, est loin d'être pertinente car elle ne considère, justement, que les coordonnées géographiques, d'où :

- Si les zones obtenues sont spatialement continues, les auteurs reconnaissent que les résultats sont fortement influencés par la prise en compte des coordonnées géographiques. Or, celles-ci sont prises en compte au même titre qu'une information attributaire alors qu'elles sont de nature radicalement différente.
- La classification obtenue, si elle est cartographiquement et visuellement très lisible, peut toutefois n'avoir aucun rapport avec les attributs des points. Il est en effet rarissime que les valeurs d'un attribut soient directement corrélées aux coordonnées géographiques.
- A partir du moment où l'on classe les coordonnées géographiques par k -moyennes, des expériences personnelles montrent qu'il n'est pas nécessaire d'appliquer ensuite les techniques de triangulation de Delaunay et de "convex hull" pour obtenir une carte lisible.

¹ La traduction française littérale de "convex hull" est "enveloppe convexe". Cette traduction respecte l'acception première de l'expression et de la technique. Toutefois, le terme français n'est quasiment pas usité dans la littérature francophone que nous avons consulté sur cette technique sauf, justement, à titre d'exemple de traduction. Nous conserverons donc dans ce document le terme de "convex hull".

Au final, il apparaît que le traitement de la position dans l'espace des points et celui de leurs attributs sont distincts et fournissent donc des résultats fort différents. Les cartes ont donc des structures spatiales très différentes à l'issue des deux étapes. La méthode accorde à la position dans l'espace des points un rôle au moins aussi important que celui de leurs attributs. Cela contribue à limiter l'adéquation du zonage final à une application nécessitant une restitution précise de l'information attributaire, comme c'est le cas en agriculture de précision.

6.2.2.2. Les méthodes filtrées de Stewart et McBratney

Stewart et McBratney ont mis au point deux méthodes de classification avec contrainte de taille s'appuyant sur des classifications [Stewart, McBratney 2001].

La première s'appuie sur une méthode de classification supervisée (maximum de vraisemblance), la seconde sur une méthode de classification non supervisée (k -moyennes, voir aussi en section 10.2.4). La contrainte de taille intervient après la classification et repose sur un filtre majoritaire appliqué à une fenêtre 5x5. C'est à dire qu'on considère une fenêtre de 5 pixels sur 5 et on classe les 25 pixels en question dans la classe qui avait la majorité avant l'application du filtre. Cette méthode équivaut donc bel et bien à l'introduction d'une contrainte de contiguïté. Aussi élémentaire qu'elle soit, elle n'en fournit pas moins des résultats satisfaisants.

6.2.2.3. La méthode filtrée de Dobermann, Ping, Simbahan et Adamchuk

Dobermann, Ping, Simbahan et Adamchuk ont mis au point une méthode de zonage afin de déterminer les tendances fortes du rendement sur une parcelle à partir de plusieurs cartes de rendement [Dobermann, Ping, Simbahan, Adamchuk 2003].

On distingue quatre étapes dans la méthode de zonage :

- La standardisation qui va permettre de corriger les erreurs présentes sur les cartes de rendement.
- Une interpolation des données de rendement par la méthode Inverse Distance Weight (I.D.W.).
- La classification : plusieurs techniques de classification ont été testées et validées par une analyse de variance en fonction du nombre de zones retenu (entre trois et huit). Les auteurs recommandent l'utilisation d'une méthode de classification hiérarchisée et notent que les classes sont spatialement fragmentées. Ce dernier point les a amené à introduire des filtres spatiaux dans leur démarche.
- L'application de contraintes spatiales par le biais des fonctions "focal analysis", "clump" et "eliminate" du logiciel Erdas Imagine.

La méthode proposée par Dobermann, Ping, Simbahan et Adamchuk présente l'avantage d'être issue de tests ayant confronté différentes techniques de classification et d'être donc parfaitement adaptée à son objectif. De plus, de toutes les méthodes décrites ici, c'est celle qui combine les traitements attributaires et spatiaux de la façon la plus performante. L'inconvénient de cette méthode est qu'elle repose manifestement sur un ou plusieurs logiciels coûteux.

6.2.2.4. Le logiciel eCognition

eCognition est un logiciel développé et distribué par Definiens Imaging. Il permet de faire de l'analyse d'image orientée objet.

Dans le cadre de notre travail, nous nous sommes intéressés aux fonctionnalités de eCognition susceptibles d'être mises en relation avec notre propre méthode de zonage. Deux points ont attiré notre attention : le premier concerne la segmentation ; le second concerne la classification orientée objet. Ces deux fonctionnalités ont fait l'objet d'une analyse réalisée par Antunes pour Arvalis Institut du Végétal [Antunes 2004].

L'opération de segmentation va extraire des objets (non classifiés à ce stade) d'images de différentes résolutions. L'algorithme va réduire au minimum l'hétérogénéité locale des objets en fonction :

- Du poids affecté à chaque carte initiale.
- Du paramètre d'échelle qui précise le niveau auquel l'objet va être construit (du point à l'image entière).
- Du réglage des paramètres de la balance. La balance est une limite abstraite qui détermine l'hétérogénéité maximum permise pour les objets extraits des images initiales. Par conséquent, elle influe sur la taille moyenne des objets en fonction de l'hétérogénéité des données initiales. La balance est définie par deux paramètres qui sont le critère d'homogénéité et le critère de forme. Ce dernier est lui-même scindé en un critère de finesse et un critère de compacité.

Itération après itération, l'algorithme va assurer une croissance égale et simultanée des segments (au départ des points simples) pour aboutir à des objets adjacents de taille semblable.

Deux méthodes de classification sont proposées par le logiciel eCognition. On notera que, quelle que soit la méthode de classification proposée, il est nécessaire de définir au préalable les classes par le biais de propriétés et/ou de conditions spécifiques. Ces conditions peuvent être spectrales, de texture, contextuelles...

La première classification repose sur le principe du voisin le plus proche ("nearest neighbour") où l'appartenance d'un objet à une classe est décrite de manière booléenne (0 ou 1). La seconde repose sur la fonction d'adhésion ("membership function") où l'appartenance d'un objet à une classe est décrite de manière floue (entre 0 et 1).

eCognition permet donc de définir des segmentations et des classifications de façon extrêmement précise et ce en accord avec les préférences de l'utilisateur final. Toutefois, cette précision requiert de cet utilisateur une grande maîtrise des différents concepts permettant le paramétrage des nombreux critères. Avant d'obtenir des résultats satisfaisants, plusieurs tests sur les paramètres sont souvent nécessaires. Enfin, l'appréciation du meilleur nombre de classes à retenir lors de la classification se fait de façon empirique en consultant quelques informations. Aucune méthode automatisée n'est ici proposée.

6.2.2.5. Autres méthodes

D'autres méthodes que celles présentées ci-dessus existent, mais il ne nous a pas toujours été possible de nous renseigner de manière approfondie sur leur mode de fonctionnement.

Parmi celles-ci, on notera celle proposée par Gangloff, Reich, Westfall et Koshla qui s'appuie sur une analyse par classification multivariée spatiale sur la base d'images satellitaires [Gangloff, Reich, Westfall, Koshla 2001]. Ces auteurs considèrent que les zonages définis sur la base de semis d'analyses de sol réguliers sont coûteux et laborieux à réaliser. La méthode semble, sur la base d'une simple analyse visuelle, fournir des résultats convenables. Toutefois, aucune information concernant la validité des zonages fournis n'est disponible.

6.3. Approches non retenues

Il existe bien entendu plusieurs techniques ou combinaisons de techniques permettant de définir des zonages. Nous allons présenter ici rapidement trois techniques auxquelles nous nous sommes intéressés avant, finalement, de les écarter.

6.3.1. Les classifications floues

La présentation des méthodes fondées sur les classifications floues, et en particulier du programme FuzME est explicite (voir section 6.2.1.2). Si le principe de la classification floue est intéressant lorsqu'on veut étudier le degré d'appartenance d'une aire d'une parcelle à une ou plusieurs classes, se retrouver avec plus de cartes de classification que de cartes initiales pose un sérieux problème d'applicabilité. De plus, aucune indication sur le nombre de classes "naturel" présent dans les données initiales n'est fournie.

Mérour propose une méthode de concaténation des différentes cartes de classification floue mais cette méthode reste très frustrée et repose beaucoup sur l'arbitraire de l'utilisateur [Mérour 2002].

Nous n'avons donc pas retenu les classifications floues pour notre méthode de zonage.

6.3.2. Les réseaux neuronaux

Les réseaux de neurones sont une des nombreuses ramifications du vaste domaine qu'est l'intelligence artificielle. Néanmoins, les méthodes d'apprentissage auxquelles ils ont recours sont généralement de type statistique, ce qui leur vaut d'être également classés dans cette discipline [Wikipédia 2005]¹. Les réseaux de neurones disposent d'une capacité de généralisation qui les rend adaptés au traitement de problèmes de nature statistiques et perceptives, telles que la classification et l'évaluation. L'organisation et le principe de fonctionnement des réseaux neuronaux leur confèrent des propriétés particulières qui leur permettent de traiter certains problèmes, en particulier non linéaires, de manière plus efficace que d'autres types de modèles [Bourret, Reggia, Samuelides 1991] [Fourty 1996].

S'inspirant des réseaux de neurones naturels, un réseau de neurones artificiel s'articule autour d'un ensemble de cellules ou neurones dont le nombre et les détails de fonctionnement peuvent varier d'une méthode à l'autre (voir Figure 17). Le principe général consiste à calibrer le réseau pour qu'il associe des données initiales (appelées "entrées") aux résultats finaux correspondants (appelés "sorties"). La mémoire du calibrage du réseau consiste en une matrice de poids et de biais. Une fois le réseau calibré, on le valide en le faisant travailler sur d'autres données d'entrée (ayant évidemment le même format que les données d'entrée du calibrage) dont on connaît les sorties correspondantes, et en comparant les sorties obtenues par le réseau à celles connues. Le réseau peut ensuite estimer des données de sorties à partir de données d'entrée similaires à celles utilisées lors du calibrage.

¹ [Wikipédia 2005] : http://fr.wikipedia.org/wiki/R%C3%A9seau_de_neurones

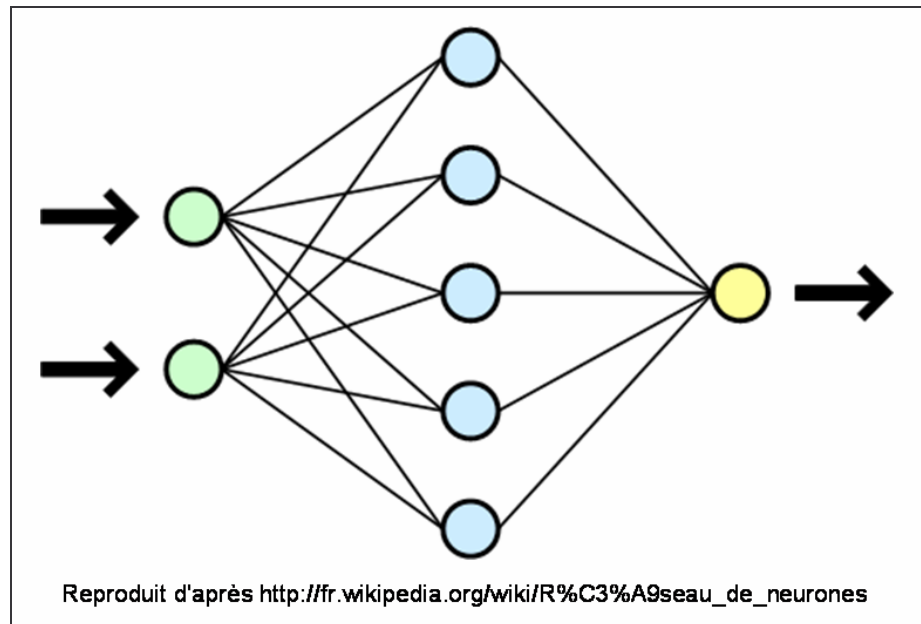


Figure 17 : Vue simplifiée d'un réseau de neurones artificiel

Les réseaux neuronaux ont été utilisés par l'Association Inter-régionale Européenne pour la Promotion de la Cartographie Tridimensionnelle des Sols Agricoles (A.I.R.E./C.T.S.) pour mettre en place des outils et des méthodes peu coûteux et non destructeurs pour caractériser la variabilité intraparcellaire du sol. A l'issue de l'étude, il est apparu que, si un modèle de prédiction universel n'était pas envisageable, un modèle parcellaire donnait des résultats très satisfaisants pour la plupart des variables pédologiques traitées [A.I.R.E./C.T.S. 2001].

Les réseaux neuronaux ont également été utilisés avec succès dans un cadre assez proche par Sudduth, Drummond, Birrell et Kitchen pour modéliser les relations entre les caractéristiques du sol et le rendement [Sudduth, Drummond, Birrell, Kitchen 1996]. Godard a utilisé la résistivité électrique pour produire des modèles prédictifs de paramètres pédologiques [Godard 2001].

Les réseaux neuronaux ont également été utilisés par Joërin, Thériault et Musy pour la détection d'entités homogènes dans un cadre toutefois différent de l'agriculture de précision [Joërin, Thériault, Musy 2001].

Nous n'avons toutefois pas retenu la technique des réseaux neuronaux pour notre méthode de zonage pour les raisons suivantes :

- Les réseaux de neurones ont "*une sensibilité aux conditions initiales qui peut changer leur performance d'estimation*" [A.I.R.E./C.T.S. 2001]. Autrement dit, il est indispensable, pour utiliser efficacement un réseau de neurones, d'utiliser des données d'entrée ayant rigoureusement le même format que celles utilisées pour la calibration et la validation du modèle. Dans notre cas, l'un des objectifs était de mettre au point une méthode pouvant travailler avec des données d'entrée de qualité et de contenu fort différents (voir section 6.1). Les réseaux neuronaux n'étaient donc pas adaptés à nos besoins sur ce point.
- Il existe un risque de "sur-ajustement" [Godard 2001]. Cela signifie que le réseau a été tellement bien calibré pour le jeu de données de calibration qu'il ne parvient pas à prédire le jeu de données de validation. Ce risque constitue la conséquence la plus extrême de la sensibilité aux conditions initiales décrite ci-dessus.
- Le mode d'apprentissage et de validation d'un réseau de neurones induit de disposer d'un jeu de données conséquent (pouvant fournir un jeu de données de calibration et un jeu de données de validation) si l'on veut s'assurer de la portabilité du modèle obtenu [Godard 2001]. Or, notre jeu de données n'était pas assez vaste pour nous garantir la portabilité du modèle.

Ainsi, si les réseaux neuronaux sont adaptés à la modélisation d'une variable à partir d'une ou de plusieurs autres variables, leur mode de fonctionnement induit certaines contraintes qui les rendent inadaptés à nos besoins.

6.3.3. La segmentation d'images

La segmentation d'images, comme son nom l'indique, a pour objectif de diviser une même image en régions réputées homogènes ou semblables. En général, le but est de déceler des objets (bâtiments, routes ou autres...) sur une image satellitaire ou aérienne.

Young, Gerbrands et Van Vliet précisent que la segmentation d'images est un terme regroupant un ensemble de techniques d'analyse d'images dont le principe consiste à conserver ce qui est intéressant (appelé le "foreground") en le séparant du reste, moins intéressant (le "background") [Young, Gerbrands, Van Vliet 1998].

Ces auteurs affirment également que le problème de la segmentation d'images est qu'il n'existe pas de méthodes pouvant s'appliquer universellement sur toutes les images et qu'aucune technique n'est parfaite. Par exemple, la méthode du seuillage ("thresholding") consiste à choisir un seuil de brillance ("brightness") et à l'appliquer à l'image pour distinguer les objets en fonction de ce seuil. L'inconvénient est que le seuil est fixé arbitrairement ou par une étude automatique ou visuelle de l'histogramme des valeurs des pixels de l'image. De même, la méthode de détection des contours ("edge finding") nécessite l'application de lissages, dont le "zero crossing", car elle est peu efficace sur les images bruitées.

Ruefenacht, Vanderzanden, Morrison et Golden définissent plutôt la segmentation d'images comme le partitionnement d'une image en régions ne se chevauchant pas [Ruefenacht, Vanderzanden, Morrison, Golden 2002]. Ils font un parallèle entre la segmentation d'images et des méthodes de classification auxquelles auraient été ajoutées des contraintes spatiales.

Au final, nous n'avons pas retenu la segmentation d'images pour les raisons suivantes :

- Les différentes méthodes sont conçues pour être appliquées, en premier lieu, à des images satellitaires ou aériennes. Or, nous devons mettre au point une méthode de zonage susceptible de traiter des données d'origines bien plus variées.
- Comme le disent Young, Gerbrands et Van Vliet, l'objectif de la segmentation d'images est, entre autres, de distinguer plus particulièrement certains objets par rapport à d'autres. Nous ne pouvons pas considérer que, sur une parcelle agricole, certaines zones méritent plus d'attention que les autres et faire en sorte que notre méthode traite les données de façon différenciée, eu égard à nos objectifs (voir section 6.1).
- Il existe peu de logiciels et de méthodes permettant de faire de la segmentation sur plusieurs images à la fois alors que nous sommes sensés traiter plusieurs sources d'informations en même temps. eCognition, le logiciel dont nous avons exploré les possibilités en 6.2.2.4, fait figure d'exception dans ce domaine.

Comme de nombreux domaines ayant recours à l'information géographique, l'agriculture de précision a des besoins spécifiques en analyse spatiale. Dès lors, le géographe, et plus particulièrement le géomaticien et l'analyste spatial, peuvent apporter leur concours au développement de cette approche de l'agriculture durable.

L'un des besoins de l'agriculture de précision concerne la définition d'entités homogènes. Non seulement les S.I.G. ne proposent pas ce genre de fonctions, mais nous avons vu qu'il existe peu de méthodes réellement satisfaisantes pour ce faire à l'heure actuelle.

Notre travail va donc consister en la mise au point d'une méthode de zonage répondant à plusieurs contraintes d'applicabilité et d'accessibilité. Pour ce faire, nous avons adopté deux approches, l'une par Aide MultiCritère à la Décision et l'autre par statistique multivariée, qui seront exposées dans les parties suivantes.

Troisième partie :

Approche de la mise au point d'une méthode de zonage par Aide MultiCritère à la Décision

Le problème qui nous occupe ici, la conception de la carte de préconisation et plus particulièrement la mise au point d'une méthode de définitions d'entités homogènes, concerne la valorisation de l'information géographique lors de la phase de décision de la pratique de l'agriculture de précision. Nous nous sommes donc intéressés aux méthodes dites "d'aide à la décision" dans une acception large du terme.

Parmi celles-ci, l'Aide MultiCritère à la Décision (A.M.C.D.) nous a semblé prometteuse. Cette approche prend en compte de manière approfondie les préférences du décideur et restitue de façon réaliste les relations entre les différents facteurs en jeu, et donc entre les solutions potentielles.

Qui plus est, les méthodes issues de cette approche présentent plusieurs complémentarités avec les fonctionnalités des S.I.G., outils indispensables à la pratique de l'agriculture de précision.

Dans cette partie, nous présentons dans un premier temps l'Aide MultiCritère à la Décision avant d'analyser les avantages que les S.I.G. obtiendraient à intégrer certaines de ses méthodes. Enfin, nous développerons deux exemples d'applications où des méthodes issues des S.I.G. et de l'A.M.C.D. sont utilisées conjointement.

7. L'Aide MultiCritère à la Décision

Vers la fin des années 1960 et le début des années 1970, des études ont été menées pour tenter de repousser les limites des méthodes d'aide à la décision principalement utilisées jusque là. L'Aide MultiCritère à la Décision (A.M.C.D.) fait partie des méthodes de travail qui ont émergé suite à ces études.

7.1. Vocabulaire

Comme de nombreuses disciplines, l'A.M.C.D. possède un vocabulaire très précis. Ce dernier possède néanmoins une spécificité : les termes utilisés ont parfois plusieurs acceptions dans d'autres disciplines, voire dans le langage commun.

Il nous semble donc nécessaire de préciser ici le sens de quelques termes auxquels nous aurons recours par la suite. Nous tentons de rester le plus fidèle possible à l'esprit des définitions de Bernard Roy [ROY 1985], par ailleurs exposées en section 19 en annexe :

- Acteur : individu ou groupe d'individus influençant directement ou indirectement la décision.
- Action : solution potentielle au problème décisionnel posé.
- Aide à la décision : Activité permettant d'obtenir des éléments de réponse éclairant la décision.
- Conséquence : Tout effet d'une action susceptible d'interférer avec les objectifs ou le système de valeurs d'un acteur.
- Critère : paramètre permettant de distinguer deux actions entre elles en fonction des objectifs et du système de valeur d'un acteur.

- Echelle de préférence : Echelle représentant les différents jugements qu'un acteur porte sur les différentes valeurs prises par un critère.
- Tableau de performances : Tableau représentant en ligne les actions et en colonnes les critères sur lesquels on porte les jugements. Au croisement d'une ligne et d'une colonne se trouve la performance d'une action pour un critère.

7.2. La critique de la recherche opérationnelle

Cette partie est inspirée de l'ouvrage pédagogique d'Alain Schärlig : "*décider sur plusieurs critères, panorama de l'aide à la décision multicritère*" [Schärlig 1985].

7.2.1. La recherche opérationnelle : réussites et échecs

Pendant longtemps, l'aide à la décision s'est faite par le biais de la recherche opérationnelle. Cette méthode a connu de nombreux succès, mais aussi de nombreux échecs.

Il est apparu que la recherche opérationnelle était efficace lorsqu'il était possible d'isoler le problème du processus de gestion de l'organisme. En revanche, lorsqu'on ne pouvait pas isoler le processus de gestion de son contexte, la recherche opérationnelle ne donnait pas satisfaction. A cela une explication simple : on avait fixé à la recherche opérationnelle un objectif trop ambitieux : la recherche d'un optimum en toutes circonstances, même lorsque cela n'avait aucun sens.

7.2.2. Le postulat de la recherche d'un optimum

Selon Bernard Roy, cité par Alain Schärlig, la recherche d'un optimum repose sur le postulat suivant : "Dans toute situation devant entraîner une décision, il existe au moins une décision qui, avec suffisamment de temps et de moyens, puisse être objectivement démontrée comme étant optimale, et ceci restant neutre par rapport au processus de décision."

Ce postulat implique à son tour trois hypothèses implicites.

7.2.2.1. L'hypothèse de globalité

Chercher une solution optimale suppose qu'une action unique pourra être considérée comme la meilleure. Cela implique que chaque action potentielle englobe tous les aspects de la question et que les actions sont donc mutuellement exclusives.

Dans la réalité, les mesures susceptibles de résoudre un problème sont souvent complémentaires et partielles, et donc rarement globales.

7.2.2.2. L'hypothèse de stabilité

La seconde hypothèse suppose que l'ensemble des actions potentielles ne peut pas être remis en cause en cours d'étude, que ce soit sous l'effet de constatations internes ou du contexte extérieur.

Mais l'étude d'un problème, dans un organisme tel une entreprise ou une administration, suffit à elle seule à faire jaillir de nouvelles idées qui sont autant d'actions potentielles. Et les délais d'étude étant relativement longs, l'environnement est lui aussi susceptible de générer de nouvelles actions en cours d'étude, ou d'en faire disparaître d'autres.

7.2.2.3. L'hypothèse de complète comparabilité transitive

C'est cette troisième hypothèse qui est la plus discutable.

En mathématiques, une relation R est dite transitive si aRb et bRc implique que aRc . C'est le cas des relations "égal à", "inférieur à", "est le frère de"...

Une relation peut aussi ne pas vérifier cette propriété. C'est le cas des relations "vaut le double de" ou "est le père de". Ces relations sont donc intransitives.

Dans le cadre de la recherche d'une solution optimale, cela signifie qu'entre deux actions potentielles, seules deux situations sont permises : la préférence stricte de l'une par rapport à l'autre, ou l'indifférence entre les deux. Ces deux situations sont des relations obligatoirement transitives.

Cette hypothèse de complète comparabilité transitive est critiquable sur trois points qui sont représentatifs de la réalité humaine :

- Elle ne tient pas compte de la situation d'incomparabilité : deux actions peuvent être différentes au point que l'acteur ne trouve aucun point de comparaison entre elles.
- Elle ne considère pas que l'indifférence peut être intransitive : on peut être indifférent entre a et b d'une part, entre b et c d'autre part tout en n'étant pas indifférent entre a et c . Si la relation d'indifférence évoque l'égalité, elles ne doivent cependant pas être confondues. Ce qui distingue l'indifférence de l'égalité mathématique, c'est qu'elle recouvre la notion de préférence faible, où la préférence n'est pas assez sensible pour être humainement ressentie et exprimée bien qu'elle existe réellement.
- Elle ne considère pas que la préférence elle-même n'est pas nécessairement transitive : a peut être préférée à b et b préférée à c alors que c est préférée à a . Ici, encore, le surclassement d'une action par une autre action ne doit pas être confondu avec une relation mathématique de supériorité.

Ainsi, pour la troisième hypothèse, la recherche opérationnelle, à travers l'optimisation, considère que toutes les actions sont comparables entre elles. Pour cela, il faut que ces actions soient ramenées à une même échelle de valeurs. Par conséquent, on aborde les différentes actions sous un angle monocritère. Cette critique étant la plus grave adressée à la recherche opérationnelle, il n'est pas rare qu'il lui soit fait référence en parlant de "paradigme monocritère".

La recherche opérationnelle construit donc les raisonnements sur deux relations :

- P Préférence transitive
- I Indifférence transitive

Au vu des limites humaines de l'hypothèse de complète comparabilité transitive, il nous faut revoir ces deux relations "classiques" en leur en ajoutant deux autres pour obtenir :

- P Préférence intransitive
- I Indifférence intransitive
- Q Préférence faible intransitive
- R Incomparabilité intransitive

7.2.3. Conclusion sur les domaines d'échecs de la recherche opérationnelle

Au terme de cette partie, nous sommes en mesure de définir quels sont les domaines et les raisons des échecs de la recherche opérationnelle.

Il s'agit des domaines où les problèmes posés n'ont pas de sens à l'égard de l'objectif d'optimisation de la recherche opérationnelle, les hypothèses implicites n'étant pas vérifiées. Les causes de ces échecs se recoupent avec le fait que les problèmes posés ne pouvaient pas être isolés de leur contexte.

7.3. L'intérêt de l'approche par Aide MultiCritère à la Décision

7.3.1. La prise en compte de la multiplicité des critères

Dans les décisions de tous les jours, nous sommes obligés de tenir compte de plusieurs critères pour faire un choix.

Par exemple, si le seul critère d'achat d'une voiture était son prix, nous roulerions tous dans le même véhicule. Ou tout au moins, à un moment donné, tous les gens achèteraient la même. Or, ça n'est pas le cas. C'est donc que d'autres critères entrent en ligne de compte, comme le confort, la sécurité, l'habitabilité, l'impression produite sur autrui, le plaisir de conduite...

Mis à part pour quelques problèmes particuliers... qu'il est possible d'isoler de leur contexte, la réalité d'une décision est donc très souvent multicritère, que ce soit au niveau personnel ou pour une grande entreprise ou une administration.

7.3.2. La prise en compte de la multiplicité des acteurs

Lorsqu'une décision doit être prise dans une grande entreprise ou une administration, le processus met souvent en action plusieurs personnes, chacune parlant en son nom ou se faisant le dépositaire des préférences d'autres individus. Plusieurs systèmes de valeurs interviennent donc, chacun jugeant les actions possibles d'une manière différente.

La multiplicité des acteurs et la multiplicité des critères ne se superposent pas, elles se multiplient car le nombre de points de vue devient d'autant plus grand que le nombre d'intervenants s'accroît.

7.3.3. La non commensurabilité des critères

A cela s'ajoute le fait que certains critères ne peuvent pas être directement mesurés. Il est donc nécessaire, pour décider, de ne pas recourir à une méthode qui se contente de tout ramener à un seul critère généralement pécuniaire. Ceci afin d'éviter la perte d'informations non négligeable qui en résulte.

Dans l'exemple de la voiture, comment mesurer l'impression faite sur autrui ? Ou le plaisir de conduite ?

7.3.4. Les oppositions entre critères

Il arrive aussi que certains critères soient en conflit.

Ainsi, vouloir s'acheter une voiture sportive, spacieuse, économique à l'usage et confortable tout en étant peu chère à l'achat est utopique. On remarque d'emblée que plusieurs critères sont en conflit dans des proportions variables. Ce qui n'empêche pas chaque individu de les prendre en considération lorsqu'il choisit un nouveau véhicule.

7.3.5. Conclusion sur l'intérêt de l'A.M.C.D.

L'objectif de l'aide multicritère à la décision n'est donc pas de trouver la solution optimale à un problème donné. Se basant sur une approche plus humaine et réaliste que mathématique, elle vise à proposer une ou des actions adéquates au vu des critères retenus et des systèmes de valeur des acteurs concernés.

7.4. Les étapes d'une étude par A.M.C.D.

7.4.1. La définition des actions potentielles

La première étape consiste à établir une sorte d'inventaire des actions potentielles. Bien que le principe soit simple, sa mise en application dans un cas concret est loin d'être évidente.

La seule chose qui soit certaine, c'est que l'ensemble A des actions ne sera pas complet, même si l'on doit s'efforcer d'être aussi exhaustif que possible.

7.4.2. La définition de la famille des critères

En demandant aux acteurs de s'exprimer sur les conséquences des actions définies précédemment, on les amène à désigner les critères qui devront être pris en compte pour l'étude.

Au contraire de l'étape précédente, il s'agit ici plus d'un choix raisonné que d'un inventaire. L'exhaustivité est là aussi recherchée, même si elle est encore une fois pratiquement inatteignable. Quitte à regrouper plusieurs critères mineurs en un seul plus général, il faut éviter de mettre trop de critères de côté. Car dans le cas contraire, on retomberait dans le travers du monocritère (de l'optimisation) : ne voir qu'une partie du problème.

Concernant le type de variables représentant les critères et l'acceptation de ces derniers par les décideurs, Maystre, Pictet et Simos considèrent qu'*"il est illusoire, dans les problèmes complexes comme ceux de la gestion environnementale, de chercher à s'appuyer uniquement sur des critères quantitatifs. Dans l'optique constructiviste, ce qui caractérise la pertinence d'un critère, c'est son acceptation par les acteurs ; cette acceptation dépend, à notre avis, de la capacité de ce critère à transcrire la réalité ou tout au moins sa perception : si un critère quantitatif existe, il faut l'utiliser ; dans le cas contraire, il ne faut pas craindre de recourir à un critère qualitatif, pour autant que sa définition soit claire et son codage sous forme numérique transparent."* [Maystre, Pictet, Simos 1994].

7.4.3. La restitution des échelles de préférence

On considère ensuite chaque critère individuellement et on établit une échelle "technique" qui représente les "états" ou les "valeurs" que peut prendre ce critère.

On transforme alors chaque échelle "technique" en échelle de "préférences" qui traduit les évaluations de la première échelle dans le système de valeurs de l'acteur. Une fois encore, passer de la théorie à la pratique peut être très délicat.

Dans l'exemple de la voiture, l'échelle technique du critère "habitabilité" peut reprendre le volume intérieur du véhicule. De même, l'échelle technique du critère "puissance" reprendra la puissance des différents véhicules.

L'échelle des préférences va reprendre l'échelle technique, c'est à dire que plus le véhicule est habitable, plus il intéresse l'acteur. Mais en ce qui concerne la puissance, on pourrait considérer qu'au-delà d'une certaine valeur (par exemple 200 chevaux), le véhicule devient difficile à conduire et est donc moins intéressant (voir Figure 18).

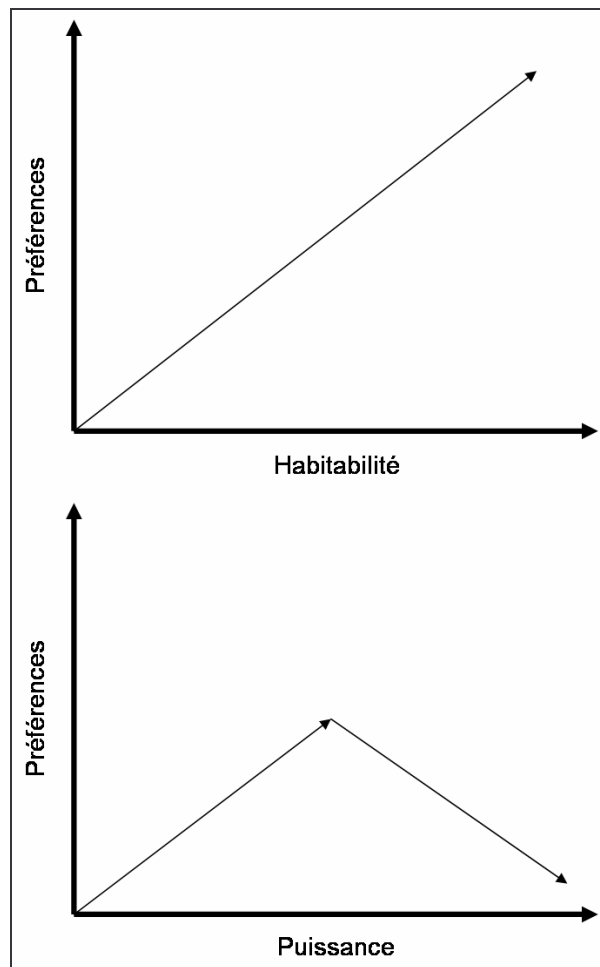


Figure 18 : Restitution des sens de préférence pour les critères "habitabilité" et "puissance"

Dans ce cas, l'échelle des préférences va progresser avec l'échelle technique jusqu'à la limite des 200 chevaux. Au delà, l'échelle des préférences prendra des valeurs inverses à celles de l'échelle technique (plus on ira au-delà des 200 chevaux, moins le véhicule intéressera l'acteur).

7.4.4. Le choix des pondérations

Il reste à présent à envisager dans quelle mesure l'importance des critères varie aux yeux du ou des acteurs, ce qui constitue le choix des pondérations de critères.

Une fois de plus, bien que cette démarche paraisse intuitivement simple, l'utilisateur doit se montrer prudent. Il arrive que des acteurs donnent des poids différents à un même critère à différents stades d'avancée de l'étude.

A cela s'ajoute la difficulté de fixer les poids lorsque plusieurs acteurs, donc plusieurs systèmes de valeurs interviennent dans le processus.

Enfin, les poids n'auront pas la même signification ou la même utilisation en fonction de la méthode d'agrégation retenue, c'est à dire en fonction de la méthode de comparaison des actions (voir section 7.4.6).

Il existe plusieurs méthodes plus ou moins complexes qui permettent de fixer les poids. Il existe également des méthodes qui postulent l'impossibilité de fixer des poids correctement.

7.4.5. La saisie de la matrice des jugements

A ce stade intervient la construction d'un tableau à double entrée dans lequel figurent en ligne les actions et en colonne les critères. A l'intersection d'un critère et d'une action figure la performance de l'action considérée pour le critère considéré. C'est ce tableau qui est appelé matrice des jugements.

Chaque ligne de la matrice est composée d'une suite de valeurs (ou d'appréciations pour des critères non commensurables) qui constitue un vecteur. Ce vecteur représente donc l'ensemble des conséquences de l'action envisagée.

7.4.6. L'agrégation

Une fois que toutes les étapes précédentes ont été effectuées, il reste alors à agréger les différents jugements des différentes actions. En effet, l'information contenue par la matrice des jugements suffit rarement pour prendre une décision. Dès que l'on considère plus de quatre ou cinq critères sur plus de quatre ou cinq actions, la lecture directe de la matrice de jugement devient très complexe. De Montgolfier et Bertier considèrent que l'agrégation est une manière d'apporter de nouvelles informations [De Montgolfier, Bertier 1978]. Ces nouvelles informations doivent faire ressortir la solution la "plus adéquate" aux yeux du problème posé et des systèmes de valeurs en présence.

Mais le choix d'une procédure d'agrégation multicritère est lui-même un problème multicritère ! Cette étape est donc souvent très complexe, pour des raisons humaines, comme cela a été le cas pour la majorité des problèmes déjà évoqués liés à l'utilisation de l'A.M.C.D., mais aussi pour des raisons techniques et méthodologiques.

Ainsi, à l'heure actuelle, aucune procédure d'agrégation multicritère n'est en mesure de répondre à l'ensemble des exigences qu'un acteur serait en droit d'avoir à son égard. Par conséquent, lors du choix de la procédure, on est fatalement contraint de céder sur telle ou telle exigence.

Il est cependant possible de classer les différentes procédures de deux façons complémentaires. En effet, chaque méthode nous obligeant à céder sur tel ou tel point, le fait de bien cerner une problématique permet d'éviter d'utiliser une méthode qui ne lui soit pas adaptée.

7.4.6.1. Classement des procédures par problématique

La première façon de classer les procédures d'agrégation consiste à prendre en compte le type de problématique à l'éclaircissement desquelles elles peuvent contribuer. Les différentes problématiques, relevées par Bernard Roy, sont les suivantes :

- La problématique de choix : elle consiste à rechercher un sous-ensemble de A , aussi restreint que possible, qui contienne les actions les plus satisfaisantes. On peut bien sûr essayer de réduire ce sous-ensemble jusqu'à ce qu'il ne contienne plus qu'une seule action. Cela revient alors à pratiquer une sorte d'optimisation mais la méthode utilisée pour cela est nettement moins réductrice et beaucoup plus "humaine".
- La problématique de tri : elle cherche à affecter chacune des actions potentielles à une des catégories définies au préalable.
- La problématique de rangement : elle consiste à regrouper les actions en classes d'équivalence, ces classes étant ordonnées conformément aux préférences de l'acteur.

7.4.6.2. Classement des procédures par attitude d'agrégation

C'est également à Bernard Roy que l'on doit ce classement des différentes procédures d'agrégation en attitudes. Il distingue ainsi l'attitude d'agrégation complète, celle d'agrégation partielle et celle d'agrégation locale et itérative. Chaque attitude a bien sûr ses avantages et ses inconvénients.

Attitude dite de "l'école américaine" ou de "l'école anglophone" (à cause de la nationalité de ses concepteurs et premiers utilisateurs), l'agrégation complète vise à mesurer l'utilité totale d'une action. Pour cela, elle considère dans un premier temps les utilités élémentaires de chaque action aux yeux de chaque critère avant de les combiner pour en déduire l'utilité totale. L'action ayant l'utilité totale la plus importante est alors considérée comme la meilleure.

Le fait d'agréger des utilités induit deux hypothèses :

- D'abord, les jugements des différents critères doivent être commensurables. Nous avons vu en section 7.3.3 que l'une des raisons d'être du multicritère est la non commensurabilité de tous les critères. Conformément au fait que chaque attitude cède sur un point, l'agrégation complète cède donc sur la non commensurabilité des critères.
- Ensuite, on est censé "combiner" toutes les utilités, ce qui revient à considérer tous les critères comme réductibles à leur combinaison et cela implique que tous les jugements soient transitifs.

Au final, l'attitude d'agrégation complète part d'une approche multicritère pour revenir ensuite à du monocritère qui peut s'apparenter à de l'optimisation. Il y a cependant une nuance : l'homme d'étude est cette fois conscient du côté réducteur de l'attitude adoptée. Cette attitude peut parfois s'avérer très utile ou être tout simplement la seule utilisable.

Il faut noter ici que la distinction de la problématique telle qu'elle a été définie plus haut (choix, tri, rangement) n'a pas d'importance lorsqu'on a recours à une agrégation complète puisque les résultats sont toujours présentés de la même manière. Ainsi, en fournissant un classement unique, on peut à la fois désigner la meilleure action (problématique du choix), les bonnes actions (problématique du tri) et les meilleures actions (problématique du rangement).

Attitude dite de "l'école française" ou de "l'école francophone" (encore une fois à cause de l'origine de ces concepteurs et premiers utilisateurs), l'agrégation partielle consiste à adopter les deux nouveautés portées par le multicritère : l'intransitivité et l'incomparabilité.

L'agrégation partielle procède en deux temps :

- On commence par prendre deux actions et à les comparer pour voir laquelle surclasse l'autre. On répète cette opération pour tous les couples d'actions possibles. On considère qu'une action en surclasse une autre si elle est au moins aussi bonne pour une majorité de critères (concordance) sans être trop nettement plus mauvaise sur les autres critères (discordance).
- On procède ensuite à une synthèse de ce que l'on a ainsi constaté.

Avec cette attitude, on décide de ne conserver que les éléments sûrs, même s'ils ne sont pas assez nombreux ou assez probants pour trancher de manière nette et précise. C'est à dire que les actions ne seront pas classées de la meilleure à la pire mais sous la forme d'un graphe ou certaines actions pourront être placées au même niveau si elles sont "ex-aequo" ou incomparables. On cède ainsi sur la lisibilité des résultats au profit de leur qualité.

Lorsque l'ensemble A des actions devient trop grand ou s'il est continu, ou encore si l'acteur a du mal à exprimer ses préférences, l'attitude d'agrégation locale et itérative va permettre, plus encore que les autres attitudes, de trouver un compromis.

Un ensemble A continu suppose que chaque action est définie par un même nombre de variables, chaque variable ayant sa propre échelle, et qu'elle prend des valeurs en continu entre les deux bornes de chaque échelle.

L'homme d'étude va prendre une première action dans A qui devra être la moins mauvaise possible. Il la soumet à l'acteur qui doit alors exprimer en quoi elle lui convient ou non. L'homme d'étude tient alors compte des préférences émises pour rechercher, dans l'environnement de l'action initiale, une nouvelle action qui soit encore moins mauvaise. Il soumet alors cette nouvelle action à l'acteur qui donne son avis et on procède ainsi jusqu'à ce que l'acteur soit satisfait de l'action proposée.

Le fait que l'on explore l'environnement d'une action pour en trouver une moins mauvaise fait que cette attitude d'agrégation est dite "locale". Les discussions successives qui s'établissent entre l'homme d'étude et l'acteur font qu'elle est "itérative". La limitation évidente d'une telle attitude est qu'elle est bien évidemment cantonnée à la problématique de choix. En effet, rien ne permet de considérer que l'avant-dernière action découverte appartient à la catégorie des meilleures (problématique de tri) ni qu'elle est la seconde meilleure (problématique de rangement).

8. L'intérêt d'une intégration S.I.G. - A.M.C.D.

Cette quatrième partie est inspirée du livre d'Amor Laaribi "S.I.G. et analyse multicritère" [Laaribi 2000].

Il s'agit dans cette section d'exposer en quoi l'intégration de l'aide multicritère à la décision aux systèmes d'information géographique est susceptible d'améliorer l'implication de ces derniers en matière d'aide à la décision à référence spatiale. Cela nécessite toutefois, dans un premier temps, d'identifier les caractéristiques des problèmes décisionnels à référence spatiale.

8.1. Les caractéristiques des problèmes décisionnels à référence spatiale

8.1.1. Le besoin en données à référence spatiale

Au delà du fait que les données ont souvent des sources, des projections, des échelles, des précisions géométriques et sémantiques différentes (voir section 3.2.1.1 dans notre cas), les problèmes décisionnels à référence spatiale ont plusieurs exigences concernant les données :

- L'information, même si elle est présente en grande quantité, est souvent insuffisante par rapport à la complexité du problème.
- L'information nécessaire à la résolution d'un problème à référence spatiale n'est pas seulement quantitative mais aussi qualitative. Or cette dernière est en général très difficile à appréhender, non seulement parce que son mode d'acquisition peut être peu explicite (étude ? enquête ? sondage ?), mais aussi parce qu'elle est souvent plus imprécise qu'une information quantitative.

8.1.2. Des décisions qui concernent le territoire

Les impacts spatiaux des décisions touchent toujours plusieurs domaines et/ou activités. Autrement dit, les problèmes décisionnels concernent toujours un ou plusieurs territoires. Ils nécessitent donc une grande quantité d'informations.

Les territoires sont des systèmes emboîtés complexes et hétérogènes. Y interagissent des aspects environnementaux et paysagers, biophysiques au sens large, socio-économiques et surtout humains, les sociétés humaines étant elles-mêmes hétérogènes et changeantes. A cela s'ajoute le fait que ces aspects sont parfois contradictoires et toujours en évolution.

C'est cette complexité qui explique le fait que l'information, même disponible en grande quantité, est souvent insuffisante pour englober tous les aspects du problème.

8.1.3. La multiplicité des intervenants

Les problèmes à référence spatiale, ayant des conséquences sur plusieurs aspects du territoire, impliquent toujours plusieurs partis. Ceux-ci peuvent être des représentants de l'Etat, des institutions, des entrepreneurs, des associations... Bien entendu, les points de vue des intervenants sur le problème sont toujours différents, et souvent conflictuels.

8.1.4. La nécessité de tenir compte de l'existant

Il y a toujours une situation de départ sur un territoire concerné par une décision. On ne part jamais d'un environnement neutre. Il est donc indispensable de tenir compte de l'existant.

Cette nécessité s'ajoute au problème de la multiplicité des facteurs à prendre en compte. En effet, il faut non seulement disposer d'une information suffisamment complète pour pouvoir prendre en compte l'existant, mais il faut aussi que cette information sur l'existant embrasse tous les aspects spatiaux susceptibles de subir les conséquences de la décision.

8.2. Les problèmes décisionnels à référence spatiale sont souvent des problèmes multicritères

Le Tableau 1 met en relation certaines des caractéristiques des problèmes décisionnels à référence spatiale exposées en section 8.1 et certains des avantages d'une approche par A.M.C.D. exposés en section 7.3.

Caractéristiques des problèmes décisionnels à référence spatiale	Avantages d'une approche par Aide MultiCritère à la Décision
Multiplicité des facteurs à prendre en compte	Prise en compte de la multiplicité des critères
Multiplicité des intervenants	Prise en compte de la multiplicité des acteurs
Nécessité de prendre en compte l'existant	Adaptée aux problèmes difficiles à isoler de leur contexte
Grande complexité	Adaptée aux problèmes peu structurés

Tableau 1 : Mise en relation des principales caractéristiques des problèmes décisionnels à référence spatiale avec certains avantages du recours à l'A.M.C.D.

Il apparaît clairement que les grandes caractéristiques des problèmes décisionnels à référence spatiale sont de nature multicritère et peuvent être efficacement abordées par une approche de type A.M.C.D.

A ces conditions fondamentales de conjonction entre les problèmes décisionnels à référence spatiale et les atouts de l'A.M.C.D. viennent s'ajouter d'autres facteurs ou conditions qui montrent l'intérêt d'une intégration S.I.G. - A.M.C.D.

8.3. La complémentarité S.I.G. - A.M.C.D.

8.3.1. Une puissante interface graphique

L'aide multicritère à la décision entraînant inévitablement un dialogue avec les acteurs, l'interface graphique des S.I.G. révèle alors son potentiel (voir sections 5.2.1 et 5.2.3.1). En effet, les résultats, affichés sous forme de cartes, sont interprétables rapidement et simplement. Les acteurs sont alors en mesure d'appréhender les conséquences de leurs choix et il leur est possible d'effectuer plusieurs séries de tests.

Pourtant, les S.I.G. manquent d'interactivité. Ils ne sont conçus que pour un utilisateur unique et ne permettent pas de prendre en compte les points de vue de plusieurs acteurs. L'interface graphique des S.I.G. n'est donc souvent mise à contribution que pour la communication à des fins d'acceptation des projets et de limitation des conflits [Molines 2003]

8.3.2. Un système de gestion de base de données

Les S.I.G., nous l'avons vu en section 5.2.3.1, sont de puissants outils d'intégration et de manipulation des données. En cela, ils complètent parfaitement l'analyse multicritère où ces éléments font cruellement défaut.

De plus, les tableaux de données dont disposent les S.I.G. sont une excellente base pour la constitution, le cas échéant, de la matrice des jugements.

Suhaedi, Metternicht et Lodwick considèrent précisément que ce sont la manipulation aisée des données, l'affichage aisée des scénarios et la flexibilité dans l'intégration des informations sur les actions qui sont les atouts majeurs de la combinaison des S.I.G. et de l'A.M.C.D. [Suhaedi, Metternicht, Lodwick 2001].

8.3.3. La technique de superposition de couches, les contraintes et les critères

Nous avons parlé dans la section 5.4.1.1 de la technique de superposition de couches des S.I.G. et nous avons remarqué qu'elle ne permettait pas, en mode vecteur, de pondération entre les couches concernées.

On ne peut donc pas parler de "critère" pour la superposition de couches dans la mesure où un critère, au sens de l'A.M.C.D., est justement susceptible d'être pondéré. En cela, les conditions de superposition de couches des S.I.G. s'apparentent plus à des contraintes qu'à des critères. Si cette différence terminologique n'a que peu d'importance en géomatique, elle est primordiale en aide multicritère à la décision.

Par conséquent, en combinant la pondération de l'A.M.C.D. à la technique de superposition de couches des S.I.G. vectoriels, on peut transformer des contraintes en critères. On dépasse alors le caractère booléen de cette technique pour en arriver à une appréciation plus souple et plus humaine de l'information disponible.

Il s'agit ici d'un point capital de la complémentarité entre l'analyse spatiale des S.I.G. et l'A.M.C.D. Et de ce point découlent plusieurs conséquences très importantes, elles-mêmes étroitement liées.

8.3.3.1. De la définition de l'ensemble des actions $A...$

Grâce à la technique de superposition de couches des S.I.G., il est possible de sélectionner, sur l'ensemble de l'espace, des entités spatiales répondant à une ou plusieurs contraintes. L'ensemble des entités spatiales résultant pourra donc constituer une base solide pour la constitution de l'ensemble des actions potentielles A .

Dans la majorité des études entreprises par A.M.C.D., cette définition des actions potentielles reste étrangère à tout avis du ou des acteurs. Il s'agit souvent d'une étude technique complètement extérieure au processus de décision. Dès lors, à ce stade, le S.I.G. s'est contenté de faire de l'aide à l'information, c'est à dire qu'il a géré un vaste ensemble de données pour en faire émerger des informations nécessaires à l'avancée du processus de décision.

Le simple fait de faire intervenir le ou les acteurs dès ce stade d'étude préliminaire contribuerait déjà à faire du S.I.G. un outil d'aide à la décision, certes encore modeste. Les acteurs fixeraient certains critères pour les modalités desquels aucune concession n'est envisageable. Ces critères deviendraient alors des contraintes d'admissibilité qui leur permettraient d'effectuer une première approche du problème et pourraient éventuellement les amener à modifier sa formulation.

8.3.3.2. ... Au choix d'une solution

Une fois l'ensemble A constitué, il est alors possible de l'explorer plus finement en prenant en compte les critères plus complexes pour lesquels les préférences des acteurs sont moins formalisées. On pondère alors les couches d'informations par le biais de l'A.M.C.D. et on évalue les actions pour rechercher le meilleur compromis.

Ici, la méthode prend en compte de manière approfondie les systèmes de valeurs du ou des acteurs et relève alors moins de l'aide à l'information que de l'aide à la décision.

8.3.4. La comparaison des solutions

Si un S.I.G. est capable, par le biais d'une superposition de couches, de considérer plusieurs "critères-contraintes" et de dégager plusieurs solutions à un problème donné, il n'offre pas de méthodes de comparaison des résultats. Et cela est d'autant plus vrai que les "critères-contraintes" sont conflictuels.

Ainsi, il est nécessaire, pour pouvoir choisir parmi les solutions envisageables, d'avoir recours à d'autres méthodes, ce qui signifie concrètement à d'autres logiciels. L'intégration de méthodes issues de l'A.M.C.D. aux S.I.G. permet de pallier ce problème et les pertes de temps inhérentes [Molines 2003].

8.3.5. La prise en compte de l'espace

8.3.5.1. Des actions qui restent des entités spatiales

Dans les études menées par A.M.C.D. concernant des problèmes à référence spatiale, les S.I.G. ne sont pas toujours mis à contribution. Et lorsque c'est le cas, c'est souvent, comme nous l'avons dit, dans le cadre d'une étude préliminaire qui n'intervient pas à proprement parler dans le processus de prise de décision.

Nous avons vu que l'intervention des S.I.G. et des acteurs dès la définition des actions va permettre, en prenant en compte leurs préférences, de faire de l'aide à la décision. A cela va s'ajouter une prise en compte de l'espace en tant que tel. En effet, en travaillant sur des cartes, les acteurs vont vraiment appréhender les actions comme des entités spatiales. Dans des études sans S.I.G., les actions définies dans l'étude préliminaire sont ensuite traitées sans rapport à l'espace, si ce n'est qu'elles ont des coordonnées géographiques.

8.3.5.2. L'exploration des relations spatiales

Contrairement aux autres études menées par A.M.C.D., les données traitées pour les problèmes à référence spatiale sont bien entendu à référence spatiale. Il est alors possible, l'espace étant réellement pris en compte, de faire émerger des actions nouvelles ou d'éclairer le problème sous un jour nouveau par la simple étude des relations entre entités (voir section 5.2.3.2).

8.4. L'adaptation de la procédure d'agrégation multicritère au problème décisionnel à référence spatiale

Nous avons constaté en section 7.4.6.2 qu'il existe plusieurs attitudes d'agrégation en A.M.C.D. Et pour chaque attitude, il existe de nombreuses procédures d'agrégation. Dès lors, devant la complexité et la variété des problèmes décisionnels à référence spatiale, il est difficile de choisir une procédure d'agrégation adaptée.

En général, une procédure d'agrégation est sélectionnée dès le début de l'étude et ce de façon arbitraire. Les raisons les plus courantes sont que l'homme d'étude est familier avec cette procédure, ou que celle-ci a été développée ad hoc ou encore qu'elle était disponible aisément.

Ici, il s'agit plutôt d'analyser précisément le problème décisionnel à référence spatiale pour en formaliser les différentes caractéristiques afin de les mettre en relation avec les caractéristiques des procédures d'agrégation multicritère.

- Type de problématique : il s'agit d'identifier quelle est la problématique du problème décisionnel par rapport aux différentes problématiques multicritères définies dans la section 7.4.6.1.
- Nature de l'ensemble des actions $\mathcal{A}$: il faut essayer d'estimer la taille de l'ensemble $\mathcal{A}$ et de savoir en particulier s'il va être continu ou discret.
- Nature de l'information à obtenir sur les critères : il faut définir si l'information est de type déterministe ou non ; si elle est de type nominale, ordinale ou si elle prend la forme d'intervalles ou de ratios.

A ce stade, il faut comparer les caractéristiques du problème mises en valeur avec les différentes procédures d'agrégation. Cela reste délicat. En effet, il existe plusieurs dizaines de procédures d'agrégation multicritère et la littérature les concernant ne fournit pas toujours les informations nécessaires pour juger de leur adéquation aux yeux des caractéristiques d'un problème décisionnel.

A cela s'ajoutent d'autres facteurs beaucoup moins formels. La procédure retenue doit être compréhensible par les acteurs et leur inspirer confiance. Et elle doit être, si possible, facile à programmer.

Il n'est donc pas possible d'être complètement objectif dans le choix de la procédure d'agrégation multicritère. Néanmoins, se poser un minimum de questions sur les caractéristiques du problème décisionnel permettra sans doute une meilleure adéquation qu'un choix arbitraire effectué en tout début d'étude.

8.5. Vers des Systèmes d'Aide à la Décision à Référence Spatiale

Un S.I.G. constitue une excellente base pour la mise au point de Systèmes d'Aide à la Décision à Référence Spatiale (S.A.D.R.S.). Ce dernier doit avoir, en plus des fonctionnalités d'un S.I.G. telles que nous les avons définies en section 5.2, les caractéristiques suivantes :

- Etre capable de résoudre des problèmes peu ou pas structurés : c'est bien entendu le cas des problèmes décisionnels à référence spatiale et l'A.M.C.D. est alors très utile.
- Disposer d'une interface à la fois puissante et simple d'utilisation.
- Permettre de manière flexible la combinaison des modèles et des données.

- Permettre l'exploration de l'ensemble des solutions : ici, c'est par la combinaison de l'interface des S.I.G. et de la prise en compte des préférences des acteurs par l'A.M.C.D. que va se faire cette exploration.
- S'adapter à l'évolution des besoins de l'utilisateur.

Ainsi, l'intégration des systèmes d'information géographique et de l'aide multicritère à la décision pourrait déboucher sur des systèmes d'aide à la décision à référence spatiale.

Les S.I.G. ont tout à gagner d'une telle intégration car ils passeraient du statut d'outil d'aide à l'information en décision opérationnelle au statut d'outil d'aide à la décision en décision tactique voire stratégique.

8.6. Pourquoi l'intégration S.I.G. - A.M.C.D. ne s'est pas faite avant ?

Plusieurs facteurs expliquent le peu d'intégration de l'A.M.C.D. aux S.I.G. :

- L'aide multicritère à la décision s'est développée dans les sciences de gestion de la décision et n'a encore que peu touché d'autres domaines.
- De leur côté, les S.I.G. sont issus d'un progrès technologique (voir section 5.2.1). Ils n'ont pas été développés pour répondre aux besoins spécifiques en matière d'aide à la décision des géographes et des aménageurs du territoire. Il a également fallu un certain temps avant que ceux-ci ne les utilisent.
- L'analyse spatiale, qui est la principale composante de l'aide à l'information et, dans quelques cas, de l'aide à la décision des S.I.G., est elle-même peu intégrée à ces outils, comme nous l'avons vu dans la section 5.4.
- Certains auteurs considèrent que les S.I.G. commerciaux répondent aux besoins immédiats du marché, qui sont généralement peu sophistiqués et plus axés sur la gestion des données et la représentation cartographique que sur leur analyse.
- Les domaines de l'aide multicritère à la décision et de la géomatique sont relativement jeunes et ont encore besoin de reconnaissance dans leur propre discipline. Les spécialistes de ces domaines travaillent donc encore dans un milieu relativement cloisonné et se connaissent peu. En France, depuis quelques années, les colloques GeoAD, organisés par le L.A.M.S.A.D.E. (Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision), contribuent à faire se rencontrer les membres des deux domaines.
- Enfin, les difficultés méthodologiques et techniques induites par une telle intégration ne favorisent pas la recherche multidisciplinaire dans cette voie.

9. Exemples d'application

9.1. La définition d'entités homogènes

9.1.1. Objectifs

Dans cet exemple d'application, l'A.M.C.D. a été appliquée seule à la définition d'entités homogènes afin de vérifier son degré d'adéquation à cette problématique et de pouvoir envisager d'éventuelles méthodes devant lui être adjointes pour améliorer ses résultats. Les objectifs étaient de vérifier l'adéquation technique et la faisabilité d'un zonage avec une méthode émanant de l'A.M.C.D. Aucune préconisation ou opération culturale précise n'a donc été définie.

La parcelle retenue pour cet exemple est la parcelle expérimentale "Baziège" du réseau variabilité intraparcellaire d'Arvalis Institut du Végétal. Les cartes retenues pour tenter d'obtenir un zonage sont des cartes de K₂O, de CaO et de pourcentage d'argile (voir Figure 19). Ces cartes ont été obtenues par interpolation I.D.W. (Inverse Distance Weight) sous Arcview des valeurs fournies localement par les analyses de sol (voir section 5.1.1).

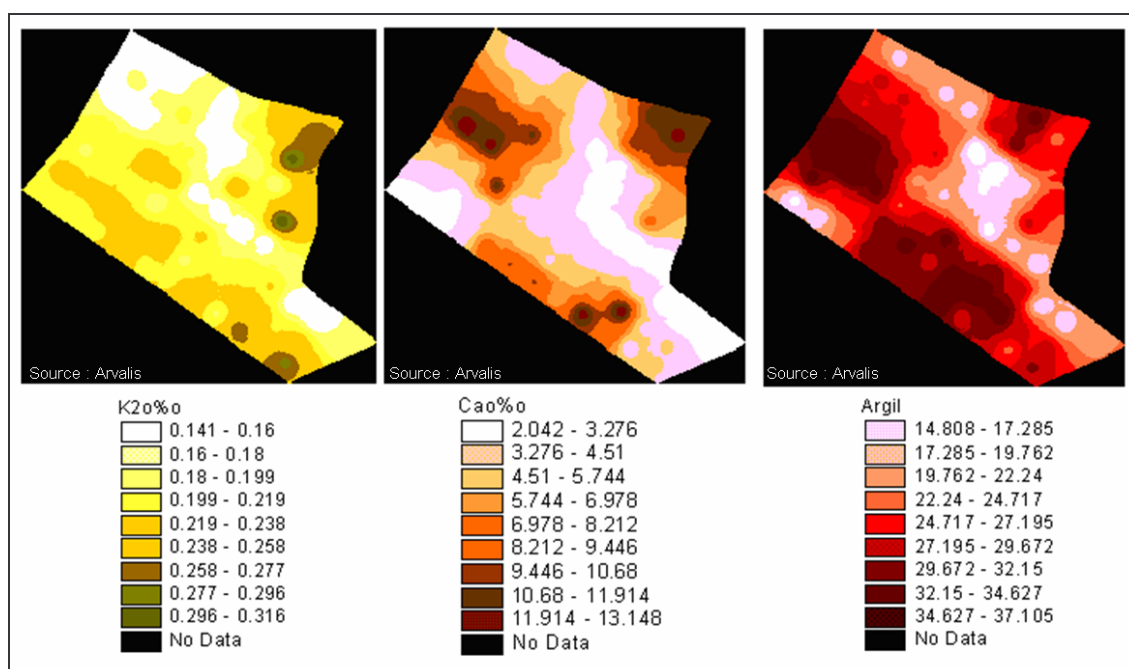


Figure 19 : Cartes de K₂O, de CaO et de pourcentage d'argile sur la parcelle "Baziège"

9.1.2. Méthode d'A.M.C.D. retenue

La problématique de tri permet, en tenant compte de plusieurs critères, de trier des aires géographiques dans différentes catégories définies au préalable. Pour un S.I.G. en mode image, ces aires peuvent être des petits amas de pixels déjà regroupés par une autre méthode ou les pixels eux-mêmes. Pour un S.I.G. en mode vecteur, ces aires sont des polygones de surface variable.

Ainsi, si on considère un nombre de catégories suffisamment restreint, le principe de dépendance spatiale (voir section 3.2.2.2) contribuera à ce que des aires triées dans une même catégorie soient contiguës, et donc à ce que des zones plus ou moins vastes soient définies sur la parcelle.

Afin de préserver au mieux l'information disponible et le système de valeurs du décideur, nous avons décidé d'opter pour une méthode relevant de l'agrégation partielle. Concernant la problématique, nous avons choisi d'utiliser une méthode relevant de la problématique de tri.

En plus d'être adaptée, comme les autres problématiques et les méthodes correspondantes, au problème de l'agriculture de précision, la problématique de tri dispose de quelques caractéristiques intéressantes supplémentaires :

- Elle permet de juger une solution pour elle-même, indépendamment des autres solutions. Là où les autres problématiques comparent les valeurs des solutions entre elles, la problématique de tri les juge par rapport aux valeurs seuils des catégories prédéfinies. Il est donc possible de déterminer les catégories en fonction de normes légales ou de résultats obligatoires (comme une teneur minimale en azote résiduel).
- Les solutions étant jugées pour elles-mêmes, le nombre de comparaisons nécessaires pour la répartition des actions est nettement diminué par rapport aux deux autres problématiques. Il est donc possible de considérer un nombre d'actions bien plus important qu'avec les autres problématiques. Or, travailler sur des aires géographiques, qu'il s'agisse de points en mode "image" ou de polygones en mode "vecteur", implique de travailler sur un très grand nombre d'actions, une aire étant une action.

Au final, c'est la méthode E.L.E.C.T.R.E. TRI (ELimination Et Choix Traduisant la REalité) qui a été retenue.

9.1.3. Renseignement des paramètres d'E.L.E.C.T.R.E. TRI

Cette étape permet la prise en compte du système de valeurs du décideur par le biais de nombreux paramètres. L'utilisation d'E.L.E.C.T.R.E. TRI pour notre travail implique donc qu'un interlocuteur soit en mesure d'apporter une réelle expertise en agronomie pour pouvoir tirer parti de la méthode.

Les paramètres à renseigner sous E.L.E.C.T.R.E. TRI sont :

- La liste des critères. Dans ce cas précis, les critères sont les cartes présentées en section 9.1.1 dans la Figure 19.
- Les poids à appliquer aux cartes.
- Les seuils de préférence, et éventuellement d'indifférence et de veto, à appliquer aux cartes. Les valeurs seuils de référence peuvent s'appuyer sur des zones types ou être définies a priori.
- Le nombre de catégories dans lesquelles ventiler les aires.
- Les valeurs seuils des catégories.
- Le seuil de coupe.

Le choix des poids à attribuer aux différentes cartes repose directement sur des appréciations subjectives et arbitraires de la part de l'utilisateur, exploitant agricole, ingénieur agronome ou autre.

Une démarche permettant d'estimer la pondération a donc été mise au point avec Denis Boisgontier, d'Arvalis Institut du Végétal (voir Figure 20). Parmi les cartes initiales, il faut tout d'abord distinguer celles dont la donnée ou variable est utilisée par le modèle agronomique correspondant à l'opération culturale et celles dont la donnée ou variable n'est pas utilisée par ce modèle agronomique. Les poids des cartes doivent ensuite être attribués en fonction de leur pertinence, de leur fiabilité, et de leur pérennité :

- S'interroger sur la pertinence revient à se poser la question "L'utilisation de cette donnée a-t-elle un sens pour la préconisation que nous nous proposons de faire ?".
- S'interroger sur la fiabilité¹ revient à se poser la question "Les valeurs prises par la carte peuvent-elles être considérées comme fiables par rapport à la valeur réelle sur le terrain ?".
- S'interroger sur la pérennité revient à se poser la question "Combien de temps l'information fournie par cette carte va-t-elle être valable ?".

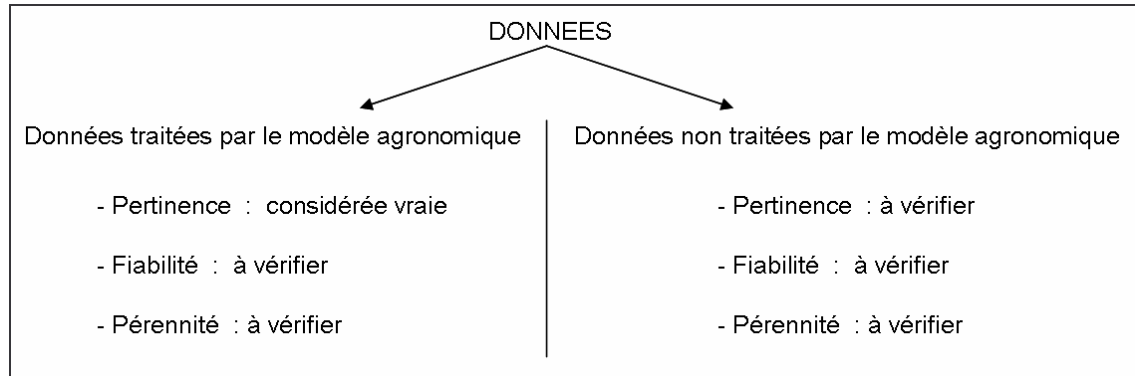


Figure 20 : Les trois interrogations permettant d'estimer la pondération des cartes initiales

En fonction des considérations sur la pertinence, la fiabilité et la pérennité de chaque carte, il est possible d'établir un ordre d'importance des différentes cartes, et donc une échelle de pondération.

On notera que la question de la pertinence des données révèle une utilisation un peu singulière des critères. En effet, la constitution de la famille des critères se fait généralement par évocation des conséquences possibles d'une action (voir section 7.4.2). Dans le cas de la définition d'entités homogènes pour une opération culturale donnée, les critères retenus sont à l'inverse ceux qui vont influencer le paramètre faisant l'objet de l'opération culturale. Ce paramètre va à son tour influencer certaines variables de la parcelle, le rendement n'étant pas la moindre. Ce sont donc ces variables "conséquences" du paramètre de l'opération culturale qui devraient normalement être prises en compte et non les variables "causes" du paramètre. Notre utilisation de la famille de critères est donc quelque peu détournée de son optique initiale.

C'est ce détournement qui justifie l'existence de la question sur la pertinence de la donnée. En effet, dans le cas d'une famille de critères constituée par la voie normale de réflexion, chaque critère est sensé être influencé par les actions. Par conséquent, la question de la pertinence, au sens où nous l'avons définie plus haut, ne se pose pas. Dans le cas d'une famille de critères constituée pour la définition d'entités homogènes, il est possible de prendre en compte des critères influençant peu, voire pas du tout, certaines aires de la carte (points ou polygones). La question de la pertinence de la donnée permet alors de prendre en compte l'impact du critère sur le paramètre de l'opération culturale et de nuancer son poids en fonction de cet impact.

¹ On notera que c'est à l'utilisateur qu'incombe la responsabilité de juger de la fiabilité des données. Cette fiabilité dépend en particulier de facteurs tels que la qualité de la correction G.P.S. au moment de la campagne de mesure, la densité du semis de points et la méthode d'interpolation utilisée ainsi que son paramétrage, la résolution d'une image satellitaire... L'application d'E.L.E.C.T.R.E. TRI ne vise donc en aucun cas à jauger la qualité des données utilisées. Elle se sert de ces données pour définir des entités homogènes qui se veulent, toutes considérations sur l'adéquation des données aux objectifs fixés mises à part, les plus satisfaisantes possibles.

Doerge a mis au point une méthode permettant de sélectionner les informations utilisables pour définir des entités homogènes en fonction de l'opération culturale et de considérations pratiques [Doerge 2001]. Bien que son jugement sur l'information s'attache moins à son origine qu'à son adéquation à l'opération culturale concernée, son approche n'en conserve pas moins certaines similitudes avec la notre. Ainsi, la densité de l'information disponible peut être mise en relation avec notre prise en compte de la fiabilité de la donnée. De même, la distinction entre informations stables ou dynamiques peut être mise en relation avec notre prise en compte de la pérennité de la donnée (voir aussi la section 20 en annexe).

9.1.4. Application de la méthode

Nous avons choisi de fonder le zonage sur trois catégories. E.L.E.C.T.R.E. TRI étant appliqué sans contrainte spatiale, il est plus que probable que les points d'une même catégorie ne seront pas continus spatialement. Une même catégorie sera donc représentée sur la carte par plusieurs zones. Par conséquent, si l'on souhaite conserver une certaine lisibilité à la carte de zonage finale, il importe de limiter le nombre de catégories afin de limiter le nombre de zones. Le Tableau 2 résume les seuils retenus pour les trois catégories sur lesquelles nous avons fondé notre zonage ainsi que le poids attribué à chaque critère.

Critère	Poids	Catégorie 1	Catégorie 2	Catégorie 3
Profondeur de sol	4	< 20 cm	20-50 cm	> 50 cm
Reliquats azotés	6	< 10 kg/ha	10-25 kg/ha	> 25 kg/ha
Réserve utile	7	< 50 mm	50-100 mm	> 100 mm

Tableau 2 : Valeurs seuils des catégories et poids attribués aux critères.

Après renseignement des autres paramètres, nous avons obtenu la carte présentée en Figure 21.

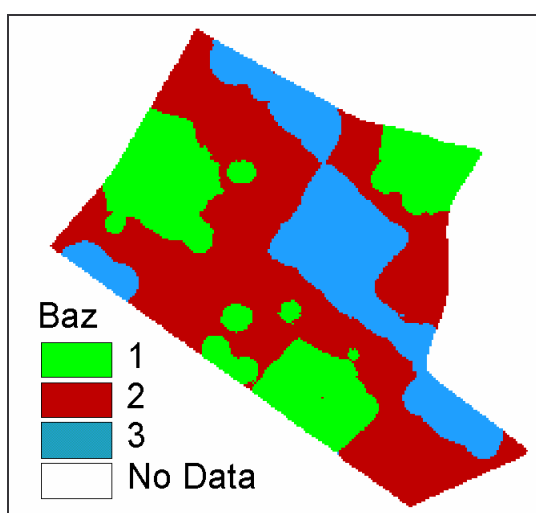


Figure 21 : Carte de zonage obtenue sur la parcelle "Baziège"

Comme on peut le voir sur la Figure 21, le nombre d'individus par catégorie est suffisant pour que les trois catégories soient bien représentées sur la carte. Néanmoins, plusieurs zones différentes peuvent appartenir à une même catégorie. Parmi ces zones, plusieurs sont trop petites pour être prises en compte dans un cadre opérationnel et mériteraient d'être supprimées ou, le cas échéant, regroupées avec des zones voisines très proches de la même catégorie.

9.1.5. Conclusion

E.L.E.C.T.R.E. TRI est donc en mesure d'établir une carte de zonage à partir des renseignements saisis par l'utilisateur. Mais le recours à la méthode E.L.E.C.T.R.E. TRI s'adresse plus à des experts agronomes étant en mesure d'établir clairement des contraintes de succès pour la carte de préconisation. Le nombre de paramètres à renseigner et leur impact sur le résultat final exigent une bonne maîtrise des concepts de l'A.M.C.D. et une part importante de réflexion en amont à laquelle va encore s'ajouter celle concernant les questions de pertinence, de fiabilité et de pérennité. Par rapport à nos contraintes de travail (voir section 6.1), ce point constitue un défaut majeur.

S'il reste possible de procéder à plusieurs simulations pour affiner les résultats, il faut pour cela s'attacher à améliorer la prise en compte de l'expertise en modifiant les paramètres. Il pourrait être tentant, de cette façon, de modifier les paramètres plus ou moins au hasard jusqu'à obtenir la carte "que l'on avait à l'esprit". Pour éviter cet écueil, il est donc préférable de s'assurer que les paramètres ont été mûrement réfléchis avant d'appliquer la méthode de zonage. La carte obtenue reste ainsi la plus proche possible de l'expertise de l'agronome et non d'un éventuel a priori.

De plus, dans le cadre d'une utilisation combinant strictement un S.I.G. et une méthode d'A.M.C.D. (Arcview et E.L.E.C.T.R.E. TRI ici) sans aucun apport extérieur, deux aspects constituent des défauts importants par rapport à nos contraintes de travail (voir section 6.1) :

- Le nombre de zones, bien que n'étant pas fixé par l'utilisateur, n'en demeure pas moins directement lié au nombre de catégories définies. Le problème est donc le même que celui relevé en section 6.2 pour de nombreuses méthodes de zonage existantes : le choix d'un nombre de zones, ou dans ce cas de catégories, pouvant être fort différent du nombre de zones réel des données initiales.
- Les actions d'une même catégorie ne sont pas forcément contiguës spatialement.
- Pratiquement, dans le cadre de notre travail, l'A.M.C.D. en général et la méthode E.L.E.C.T.R.E. TRI en particulier présentent un inconvénient important : elles nécessitent de bonnes connaissances théoriques et pratiques pour pouvoir renseigner correctement les paramètres et s'assurer que le point de vue du décideur a été correctement retranscrit en poids, seuils et autres... Cela suppose donc qu'un spécialiste de la méthode travaille avec l'utilisateur ou que celui-ci s'est vu inculquer les notions élémentaires.

9.2. Le pré-diagnostic d'un risque de pollution

Bien que cet exemple d'application de l'A.M.C.D. ne relève pas de l'agriculture de précision, il reste néanmoins appliqué au domaine de l'agriculture et son mode opératoire pourrait être reproduit à d'autres échelles. L'intérêt principal de cet exemple réside dans le fait qu'il a été bien formalisé grâce au concours de François, chef de projet chez GEOSYS S.A.

L'utilisation conjointe de méthodes relevant de l'A.M.C.D. et des S.I.G. est de plus en plus fréquente dans le domaine de la gestion environnementale, et en particulier dans l'évaluation des risques. On trouve à présent de nombreux articles, émanant de différentes disciplines, traitant de ce sujet avec ce type d'approche [Rousseau, Martel 1994] [Joërin 1997] [Culligan Dunsmore, Salt 1998] [Varma, Ferguson, Wild 1998] [Laaribi 2000] [Joërin, Thériault, Musy 2001] [Suhaedi, Metternicht, Lodwick 2001] [Adinarayana, Maitra, Venkataraman 2001] [Molines, Chevallier 2001] [Molines 2003] [Graillot, Londiche, Dechomets, Batton-Hubert ?].

9.2.1. Objectif

L'objectif de cette étude, menée par François (GEOSYS S.A.) et Réal (Arvalis Institut du Végétal), était d'effectuer un pré-diagnostic des risques de pollution diffuse par des produits phytosanitaires lié au fonctionnement hydraulique d'un bassin versant. Le site de l'étude est le bassin versant de la Nonette, dans l'Oise.

Le pré-diagnostic doit prendre la forme d'une carte identifiant plusieurs types de risques de pollution diffuse. Les zones à haut risque identifiées serviront de base à des diagnostics d'exploitation plus précis.

La mise en place du pré-diagnostic en amont du diagnostic d'exploitation présente de nombreux avantages liés à la complémentarité des investigations :

- Gain de temps pour réaliser les diagnostics d'exploitation.
- Etablissement plus aisé des priorités d'action au niveau du bassin versant.
- Possibilité d'extrapoler les solutions à l'intérieur d'un type d'exploitation caractérisé par une nature de risque.
- Support de communication pédagogique, montrant la logique d'une évaluation partant du global vers le particulier.

L'étude s'appuie sur un certain nombre de données jugées pertinentes pour ce type d'étude :

- Nature de la nappe : il s'agit ici de considérer si la nappe est captive ou libre.
- Occupation des sols : l'objectif du pré-diagnostic concernant directement des activités agricoles (épandage de produits phytosanitaires), la définition de zones d'occupation agricole des sols permet de limiter la zone d'étude.
- Eloignement au réseau hydrographique.
- Hydromorphie
- Texture des sols
- Niveau de réserve utile
- Pente
- Perméabilité du substrat

9.2.2. Etude initiale GEOSYS S.A. – Arvalis Institut du Végétal

L'étude conduite initialement reposait sur une partie seulement des données jugées pertinentes (voir Tableau 3).

Donnée	Retenue
Nature de la nappe	Non
Occupation des sols	Oui
Eloignement au réseau hydrographique	Oui
Hydromorphie	Oui
Texture des sols	Oui
Niveau de réserve utile	Oui
Pente	Oui
Perméabilité du substrat	Non

Tableau 3 : Données retenues lors de l'étude initiale de pré-diagnostique

Ce sont des contraintes méthodologiques qui ont amené à ne pas retenir certaines des données initiales.

Bien que cette étude n'ait pas été conçue selon une approche par A.M.C.D., on notera que la simple définition des données retenues nous en éloigne de toute façon pour au moins deux raisons :

- Les contraintes méthodologiques ayant amené à mettre de côté deux données pourtant pertinentes font que l'on ne peut prétendre à l'exhaustivité devant normalement caractériser la famille de critères telle que nous l'avons définie en 7.4.2.
- La notion même de critère, dans son acception propre à l'A.M.C.D., est ici inapplicable puisque certaines données ne sont abordées que de façon booléenne. C'est le cas de l'occupation des sols, qui doit uniquement servir à délimiter les zones à activité agricole dans la région d'étude.

A partir des données retenues, GEOSYS S.A. et Arvalis Institut du Végétal ont procédé à une classification des risques des zones du bassin versant par un arbre dichotomique. Au final, le bassin versant de la Nonette a été divisé en plusieurs régions¹.

Au final, on peut faire les remarques suivantes sur l'approche adoptée :

- Discretisation importante de l'information par l'approche dichotomique : si ce point n'est pas gênant pour des informations telles que l'occupation du sol, il en va autrement pour l'éloignement au réseau hydrographique ou le niveau de réserve utile pour lesquels une prise en compte plus fine permettrait de nuancer les résultats.
- Pas de considération formelle de l'importance relative des informations : l'approche dichotomique écarte d'emblée toute possibilité de pondération des données retenues.
- Bonne caractérisation du risque mais faible appréciation du niveau de risque : si les différents types de risques sont bien identifiés, il n'est pas possible de conclure précisément sur leur éventuelle gradation.

¹ L'arbre dichotomique et la carte finale ne sont pas présentés ici pour des raisons de confidentialité.

9.2.3. Approche par A.M.C.D.

L'approche par A.M.C.D. a eu pour objectif de pallier à la faible appréciation du niveau de risque due à l'approche dichotomique. L'objectif était donc de cartographier différents niveaux de risque de pollution diffuse sur le bassin versant de la Nonette. Pour cela, il était nécessaire dans un premier temps de définir ces niveaux de risque en dehors du contexte de l'étude. En effet, l'approche dichotomique peut être appliquée à différentes zones d'étude. Pour les niveaux de risque, procéder à un simple classement sur le bassin versant de la Nonette n'aurait pas suffi à assurer l'adaptabilité de la méthode à d'autres sites.

Dès lors que notre problème concernait la catégorisation des niveaux de risque sur un site quelconque et que cette catégorisation allait s'appuyer sur des catégories prédéfinies, nous avons choisi d'utiliser la méthode E.L.E.C.T.R.E. TRI.

Afin de définir les catégories correspondantes aux niveaux de risque, nous nous sommes appuyés sur les données utilisées lors de l'étude par approche dichotomique afin de définir les critères, avec toutefois quelques différences (voir Tableau 4).

Donnée	Retenue
Nature de la nappe	Oui
Occupation des sols	Oui comme contrainte
Eloignement au réseau hydrographique	Oui
Hydromorphie	Oui
Texture des sols	Oui
Niveau de réserve utile	Oui
Pente	Non
Perméabilité du substrat	Oui

Tableau 4 : Données retenues lors de l'approche par A.M.C.D.

On notera en particulier que la nature de la nappe et la perméabilité du substrat ont ici été prises en considération. En revanche, la pente ne l'a pas été, la donnée n'étant pas disponible. L'occupation des sols n'a pas été considérée comme un vrai critère (au sens A.M.C.D. du terme) mais a été traitée comme une contrainte afin de limiter l'application d'E.L.E.C.T.R.E. TRI aux aires du bassin versant effectivement occupées par l'activité agricole.

Une fois les critères définis, ceux-ci ont reçu les poids et les sens de préférence présentés dans le Tableau 5. Cette pondération présente l'avantage, par rapport à l'approche dichotomique, de nuancer l'influence des différentes données disponibles afin de se rapprocher de leur influence réelle sur le terrain, ou de pouvoir modifier leur influence par rapport à des facteurs externes dont les données ne sont pas disponibles (saison, événement météorologique et/ou hydrologique exceptionnel...).

Critères	Poids	Sens des préférences
Nature de la nappe	7	Décroissant
Eloignement au réseau hydrographique	12	Croissant
Hydromorphie	10	Décroissant
Texture des sols	10	Décroissant
Niveau de réserve utile	10	Croissant
Perméabilité du substrat	7	Croissant

Tableau 5 : Poids et sens des préférences des critères de l'approche par A.M.C.D.

Cinq catégories ont ensuite été définies avec leurs valeurs seuils pour chaque critère, comme présenté dans le Tableau 6. L'approche dichotomique entraîne une importante perte de nuance du résultat par rapport à la richesse des données initiales. Avec l'approche par A.M.C.D., la possibilité de définir autant de catégories qu'on le souhaite (même si au-delà d'un certain nombre, le procédé perd de son intérêt) permet de tirer parti de cette richesse.

Critères	Risque Très Fort	Risque Fort	Risque Moyen	Risque Assez Faible	Risque Faible
Nature de la nappe	Alluviale	Libre	Libre	Libre	Perchée
Eloignement au réseau hydrographique	$x < 50m$	$50m < x < 100m$	$100m < x < 200m$	$x > 200m$	$x > 200m$
Hydromorphie	Très Forte	Forte	Assez Forte	Modérée	Faible / Aucune
Texture des sols	L	E	AIL	A / I / U	A / I / U
Niveau de réserve utile	$x < 50mm$	$50mm < x < 80mm$	$80mm < x < 110mm$	$110mm < x < 140mm$	$x > 140mm$
Perméabilité du substrat	Interstices / Fissures	Interstices / Fissures	Interstices / Fissures	Imperméable	Imperméable

Tableau 6 : Catégories définies pour l'approche par A.M.C.D.

La contrainte "occupation des sols" et les six critères définis ont été cartographiés sous Arcview afin de répondre aux exigences de cette étude (voir Figure 22). La contrainte a été dans un premier temps utilisée comme un masque, seules les aires étant soumises à une activité agricole nous intéressant. Ensuite, sur l'aire du bassin versant restante après application du masque, toutes les cartes ont été superposées deux à deux afin d'obtenir une carte unique de synthèse reprenant l'ensemble des polygones identifiables sur toutes les cartes de critères. Cette carte unique de synthèse dispose donc d'une table attributaire à six attributs, un par critère, et chaque polygone possède une valeur pour chaque attribut. La table attributaire ainsi créée constitue donc un tableau de performances au sens A.M.C.D. du terme.

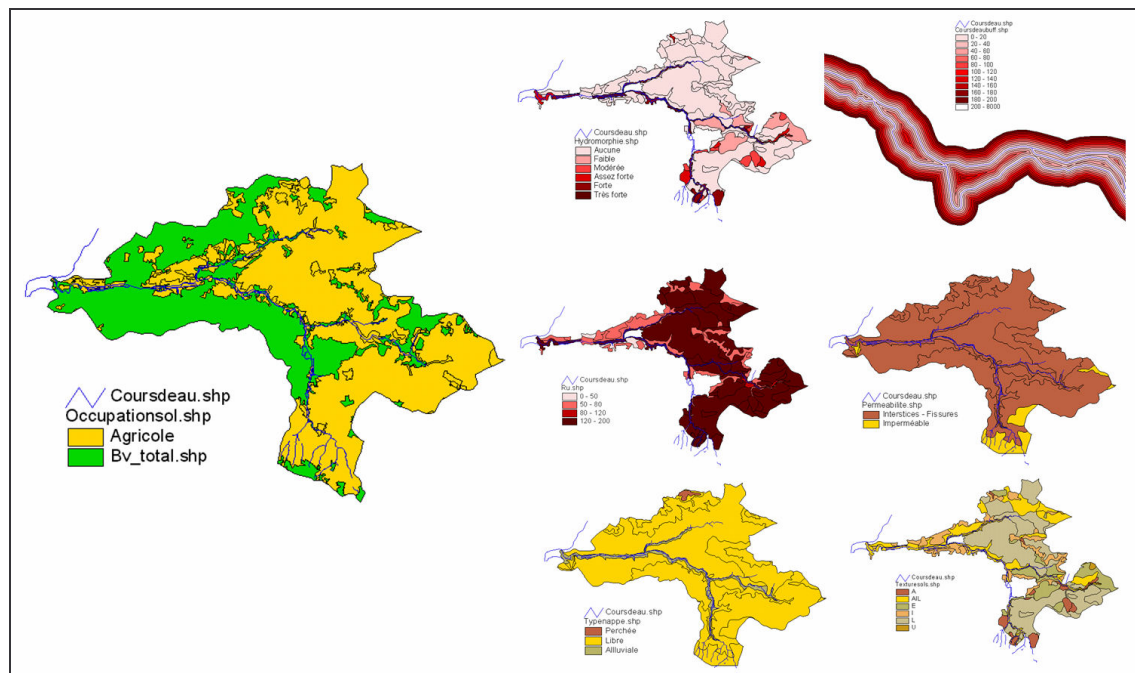


Figure 22 : Cartes correspondantes à la contrainte et aux critères retenus pour l'approche par A.M.C.D.

Après application de la méthode E.L.E.C.T.R.E. TRI selon les paramètres définis plus haut, on obtient la carte de niveaux de risque présentée en Figure 23.

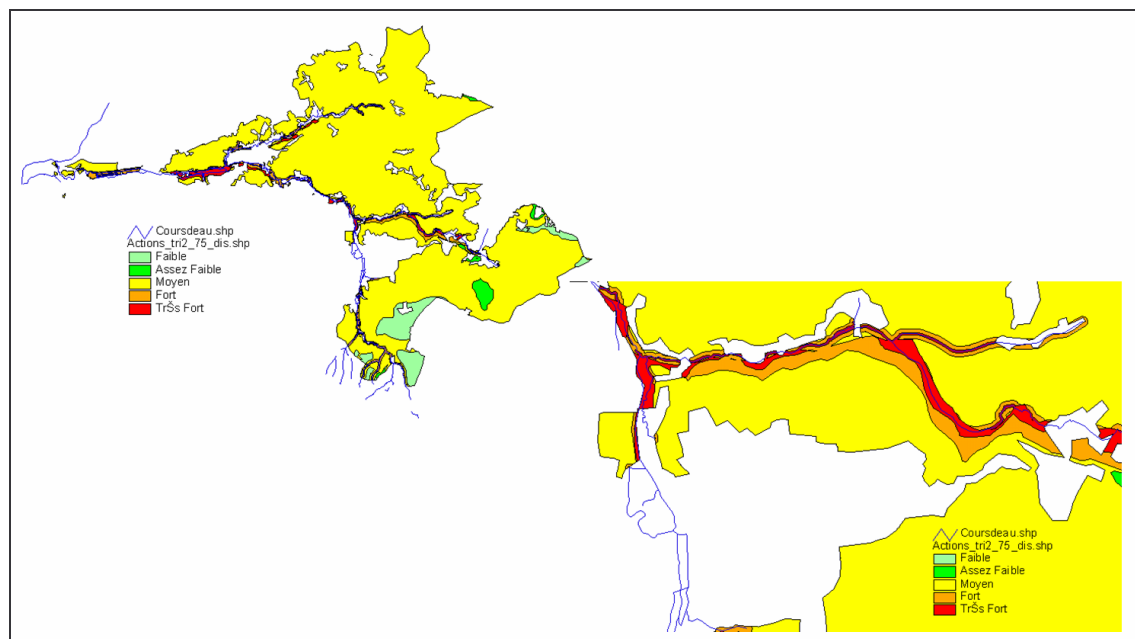


Figure 23 : Cartographie des niveaux de risque de pollution diffuse sur le bassin versant de la Nonette

On obtient une carte au rendu très différent de la carte des types de risque présentée en . Ici, les différentes zones identifiées s'échelonnent à partir du cours d'eau, restituant plus nettement le souci d'identifier rapidement des aires sur lesquelles procéder à des diagnostics plus précis. On note toutefois que la majeure partie de la carte est classée en risque moyen. Il est nécessaire de considérer la carte à plus grande échelle à proximité des cours d'eau, très localement, pour saisir l'intérêt véritable de l'approche par A.M.C.D. Dans ce cas, il est possible d'identifier des zones de risque fort ou très fort plus ou moins larges et donc d'orienter plus finement les diagnostics complets à venir.

9.2.4. Conclusion

Pour le pré-diagnostic des risques de pollution diffuse sur la bassin versant de la Nonette, l'A.M.C.D. permet de compléter l'approche initiale par arbre dichotomique. En effet, celle-ci avait permis d'identifier clairement les différents types de risque susceptibles d'entraîner une pollution. L'utilisation que nous avons faite de l'A.M.C.D. permet de compléter cette typologie en l'associant à un niveau de risque. La consultation des deux documents simultanément facilite donc l'association des types de risques aux niveaux de risque.

Pour intéressante qu'elle soit, cette double approche d'un même problème présente toutefois un inconvénient. Les deux méthodes employées reposent sur la connaissance d'un ou de plusieurs experts et nécessitent un effort réel de formalisation de la part de ces derniers. La méthode à double approche a néanmoins été mise au point sans trop prendre en considération les conditions locales, mais plutôt en s'efforçant d'ignorer tout contexte. Elle devrait donc pouvoir être appliquée à d'autres bassins versants au prix d'un minimum d'adaptation. L'investissement en temps nécessaire à la formalisation de l'expertise peut donc s'avérer rentable de par la portabilité que l'on peut attendre de la méthode.

L'Aide MultiCritère à la Décision s'avère être une approche très intéressante de l'aide à la décision, tant par les principes sur lesquels elle repose que par la finesse de la restitution des systèmes de valeurs qu'elle permet. De plus, l'intégration de certaines de ses méthodes aux S.I.G. est un atout indéniable pour la résolution de problèmes décisionnels à référence spatiale.

Notre travail consiste à valoriser l'information géographique dans le cadre d'une phase de décision. Pourtant, nous n'avons pas pu retenir les méthodes issues de l'A.M.C.D. pour notre méthode de zonage, et ce pour deux raisons.

Tout d'abord, la méthode que nous avons choisi d'employer, E.L.E.C.T.R.E. TRI, est une méthode complexe à utiliser, en particulier dans notre cas. Nous l'avons mis en évidence dans les exemples d'application de la section 9. Globalement bien adaptée à notre problème, cette méthode ne répond en revanche pas du tout à l'un de nos objectifs : la simplicité d'utilisation de la méthode de zonage (voir section 6.1). C'est d'ailleurs une caractéristique commune à plusieurs méthodes relevant de l'A.M.C.D. de l'école francophone. Devant la nécessité exprimée par les acteurs de manipuler des outils intégrés et très simples d'utilisation, cette caractéristique devient une grave faiblesse et un important facteur d'inapplicabilité [Joërin, Thériault, Musy 2001]. Dès lors, on peut s'interroger sur la possibilité de simplifier la méthode ou d'en automatiser certains paramètres. Mais une simplification ou une automatisation rendrait la méthode moins adaptée "*à la résolution d'un problème spatial par essence complexe*" [Joërin, Thériault, Musy 2001]. Simplifier une méthode dont la qualité repose sur sa complexité, c'est prendre le risque de dégrader fortement l'intérêt et la pertinence de l'utilisation de cette méthode.

La seconde raison pour laquelle nous n'avons pas retenu l'A.M.C.D. est l'acceptation de la subjectivité inhérente à l'utilisation des méthodes relevant de cette approche. Toute décision contient une part de subjectivité. Celle-ci peut être liée au décideur lui-même mais aussi au lieu et au moment où il prend sa décision [Joërin 1997]. Et comme le souligne Joërin, *"cette subjectivité est souvent perçue à tort comme un défaut de la décision"*. Nous avons été confrontés à ce problème de l'appréhension de la subjectivité et ce point a également contribué à nous faire abandonner l'approche par A.M.C.D. pour la mise au point de notre méthode de zonage.

Quatrième partie :

Approche de la mise au point d'une méthode de zonage par la statistique multivariée

Devant les difficultés à mettre en œuvre de façon opérationnelle une approche s'appuyant sur l'Aide MultiCritère à la Décision, nous avons décidé d'explorer d'autres approches. Nous en avons sélectionné quelques unes qui étaient à la fois adaptées à nos contraintes de travail tout en étant moins complexes d'utilisation (et/ou paramétrables sans perte de qualité) que l'A.M.C.D.

Au final, nous avons décidé de fonder notre méthode de zonage sur un ensemble de méthodes relevant de la statistique multivariée. Ces méthodes, modifiées plus ou moins profondément et combinées à un ensemble de règles spécifiques à la définition d'entités homogènes, nous permettent de respecter l'ensemble de nos contraintes de travail.

Dans cette partie, nous exposerons en détail le fonctionnement de notre méthode de zonage. Nous présenterons ensuite plusieurs techniques venant en amont et en aval du zonage afin respectivement d'analyser les données initiales et de valider le zonage obtenu. Enfin, un exemple d'application de notre méthode est décrit.

10. Notre méthode de zonage

Cette partie contient les descriptions et justifications méthodologiques de notre méthode de zonage.

10.1. Fonctionnement global de notre méthode de zonage

10.1.1. Une méthode fondée sur une classification mixte...

Nous avons décidé d'appuyer notre méthode de zonage sur le principe de la classification mixte. Ce principe est recommandé par Lebart, Morineau et Piron [Lebart, Morineau, Piron 2000] et par Triboulet, Lardon et Langlet [Triboulet, Lardon, Langlet 1996].

La classification mixte relève du bon sens en ce qu'elle consiste à combiner deux algorithmes de classification afin de tirer le meilleur parti de leurs avantages respectifs :

- Dans un premier temps, une méthode de classification telle que la méthode des "centres mobiles" [Lebart, Morineau, Piron 2000] va réduire la dimension de l'ensemble des individus à classer en opérant des regroupements préalables.
- Dans un second temps, une classification hiérarchique "déterministe" [Lebart, Morineau, Piron 2000] va achever la classification en regroupant de nouveau les regroupements issus de la première étape.

Ce principe de la classification mixte est particulièrement adapté à notre propos pour deux raisons :

- Le recours à une méthode telle que les "centres mobiles" ou variantes permet de traiter de vastes recueils de données. Or les cartes que nous traitons comptent souvent plusieurs dizaines de milliers de points. Toutefois, une telle méthode impose de fixer à priori le nombre de classes et la classification obtenue dépend beaucoup des premiers centres de classes définis.

- Plusieurs auteurs considèrent que le recours à une classification hiérarchique permet d'obtenir certaines informations sur le nombre de classes à retenir par le biais du dendrogramme [Minvielle, Souiah 2003] [Lebart, Morineau, Piron 2000].

Bien que la classification mixte soit adaptée à notre problématique, les deux raisons évoquées ci-dessus n'en présentent pas moins toutes deux des caractéristiques qui se heurtent à l'une de nos contraintes de travail (voir section 6.1). En effet, le choix de paramètres tels que le nombre de classes des centres mobiles ou le nombre de classes de la classification hiérarchique entrent en conflit avec notre objectif de simplicité d'utilisation de la méthode de zonage.

Mais l'adéquation de la classification mixte sur les autres caractéristiques de notre problème nous a amené à l'adopter pour notre méthode de zonage. Les quelques limites évoquées ici ont trouvé des solutions qui seront décrites en section 10.2.

10.1.2. Appuyée par une analyse factorielle préalable...

A ce principe de la classification mixte, nous avons ajouté celui de la complémentarité entre analyse factorielle et classification [Lebart, Morineau, Piron 2000].

Effectuée en préalable à une classification mixte, une analyse factorielle présente les avantages suivants :

- Elle est adaptée à l'exploration de vastes recueils de données et est donc adaptée à notre besoin (voir section 10.1.1).
- Elle va permettre d'effectuer la classification mixte sur des coordonnées factorielles moins nombreuses que les variables de départ en ne conservant que les axes significatifs.
- Le fait de ne conserver que les axes factoriels significatifs va permettre d'éliminer les fluctuations mineures qui caractérisent la variance des derniers axes factoriels.
- Ces axes factoriels sont indépendants.

10.1.3. Et soumise à des contraintes spatiales...

Le problème des méthodes purement statistiques que sont la classification mixte et l'analyse factorielle est que les classes qu'elles définissent sont spatialement fragmentées. Or, nous avons vu en section 3.2.2.2 que les cartes de zonage finales doivent justement être vierges de tout point isolé ou d'amas de points de trop petite taille.

Nous allons donc devoir adjoindre aux différentes méthodes statistiques retenues des contraintes à composante spatiale afin de nous assurer de l'opérationnalité des cartes issues de la classification mixte.

10.1.4. Synthèse et fonctionnement global

Notre méthode de zonage va donc s'appuyer en grande partie sur des méthodes statistiques multidimensionnelles. Bien que les techniques multivariées ne soient pas initialement orientées pour des données dépendantes spatialement, elles sont toutefois très utiles comme outils de réduction des données et pour identifier des combinaisons de variables [Bailey 1994]. L'adaptation de ces techniques aux données de nature géographique vient du fait, comme l'affirment Lebart, Morineau et Piron, "*qu'il est fréquent que le tableau de données actives soit déjà structuré : c'est le cas des données géographiques ou temporelles où la structure intervient au niveau des observations (individus voisins ou consécutifs)*" [Lebart, Morineau, Piron 2000]. Autrement dit, notre approche du zonage s'appuie sur la notion d'autocorrélation spatiale qui suppose que "*deux lieux proches tendent à se ressembler davantage que deux lieux éloignés*" (voir section 3.2.2.2) [Bavoux, Chapuis, Delmer, Mannone, Passequé, Volpoët 1998].

Les choix que nous avons opérés pour développer notre méthode de zonage se sont voulus les plus près possibles de nos contraintes de travail (voir section 6.1). Dans son ensemble, notre méthode de zonage suit donc le mode opératoire présenté en Figure 24.

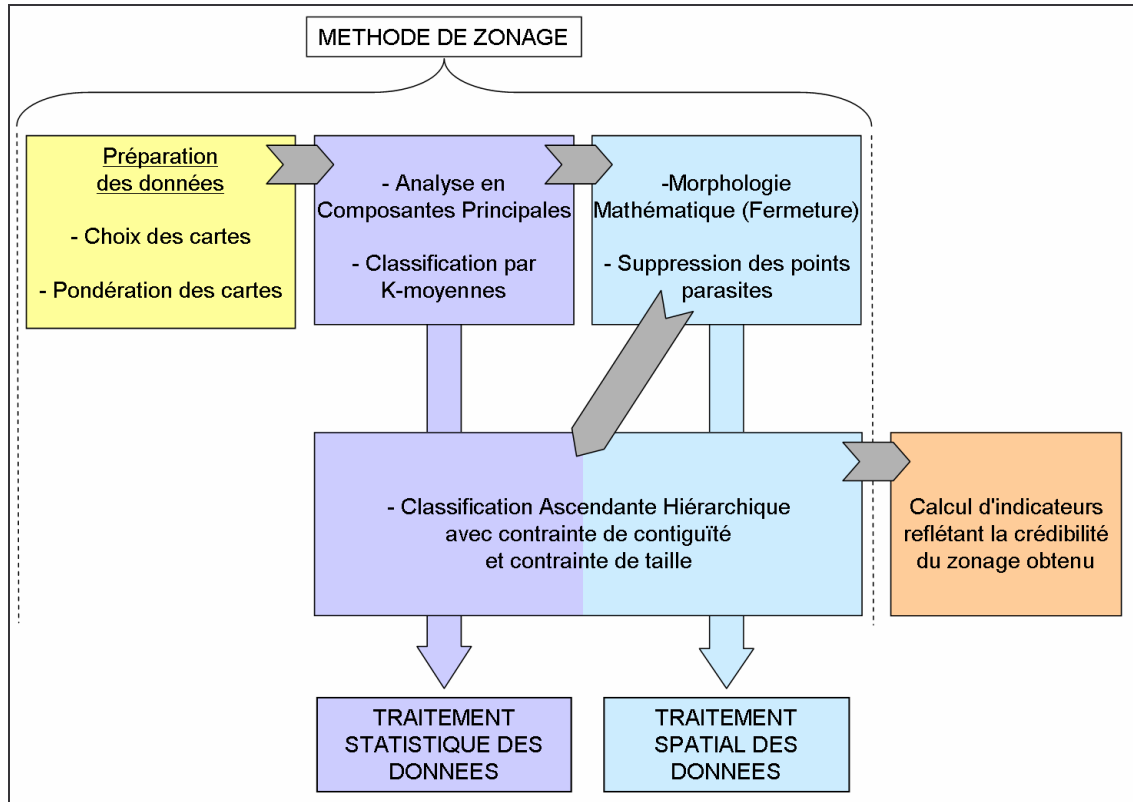


Figure 24 : Fonctionnement global de la méthode de zonage

La méthode de zonage telle que présentée en Figure 24 traite conjointement des caractéristiques attributaires (par le biais des traitements statistiques) et spatiales des objets spatiaux afin de répondre à nos contraintes d'applicabilité. Ce traitement conjoint des caractéristiques attributaires et spatiales induit que la mise au point de cette méthode de zonage relève directement de la mise au point d'une méthode d'analyse spatiale telle qu'elle est définie par Haining [Haining 1994].

Dans ce cadre, notre méthode répond à un besoin identifié par Triboulet, Lardon et Langlet qui insistent sur le nécessaire traitement à la fois statistique et spatial de l'information géographique et expliquent *"qu'il existe peu de méthodes éprouvées travaillant simultanément sur les deux"* [Triboulet, Lardon, Langlet 1996]. Ces auteurs notent également que *"la prise en compte du spatial dans le statistique ne permet pas de mesurer le poids respectif de l'information liée au lieu et de celle liée à son voisinage"*. Ils en concluent que *"les résultats sont donc difficilement validables et interprétables car les règles d'agrégation sont peu explicites"*. Et ils considèrent par conséquent que *"séparer le statistique du spatial permet donc de mieux contrôler les processus mis en jeu et donc d'asseoir leur validité respective."*

Nous avons pu tenir compte de ce principe de précaution en ce qui concerne la première phase de traitement statistique et la première série de contraintes spatiales. En effet, la première phase de traitements statistiques vise à réduire le nombre d'individus et le nombre de variables et constitue donc une approche statistique classique. De même, la première série de contraintes spatiales repose sur des filtres, certes élaborés, mais ne prenant pas directement en compte l'information attributaire puisqu'ils considèrent uniquement les classes statistiques auxquelles ont été affectés précédemment les individus. En revanche, nous avons dû intégrer une contrainte de contiguïté à la C.A.H., comme nous le verrons en section 10.2.8. Car si le principe de précaution de Triboulet, Lardon et Langlet permet de mieux valider les méthodes, il préconise de se livrer à des recherches ayant directement trait à l'analyse spatiale... sans utiliser de méthodes d'analyse spatiale ! C'est donc moins sur l'existence de méthodes d'analyse spatiale éprouvées que sur l'existence de méthodes permettant de les valider qu'il faut s'interroger.

Afin que l'A.C.P., la classification mixte et les différentes contraintes spatiales opèrent correctement, il est indispensable que les cartes aient le même nombre de points. Les cartes doivent donc être structurées en mode image et sont rééchantillonnées au besoin pour avoir les mêmes dimensions (largeur et hauteur). Notre méthode de zonage travaille donc sur des points auxquels sont associés des vecteurs d'informations. Les surfaces de points disposant de vecteurs d'informations et celles sans (dites "no data") doivent également être parfaitement superposées.

10.2. Fonctionnement détaillé de notre méthode de zonage

10.2.1. Mise en forme des cartes

Le programme de zonage a été conçu pour travailler sur des fichiers exportés à partir d'Arcview 3.2. L'ensemble des cartes initiales doit être représenté en mode image. Si certaines cartes initiales sont en mode vecteur et représentent des semis de points (cas des cartes d'analyses de sol, de rendement et de résistivité électrique sous leur forme originale), les valeurs doivent être interpolées et les cartes ensuite converties en mode image. Les cartes ainsi obtenues doivent avoir le même nombre de lignes et de colonnes et la même taille de point. Au final, on dispose, sous Arcview 3.2, d'un ensemble de cartes en mode image appelées "grids".

Pratiquement, le programme de zonage développé a été conçu pour traiter des fichiers "A.S.C.I.I. Raster" (extension ".asc") exportés depuis les "grids" d'Arcview 3.2 (voir Figure 25).

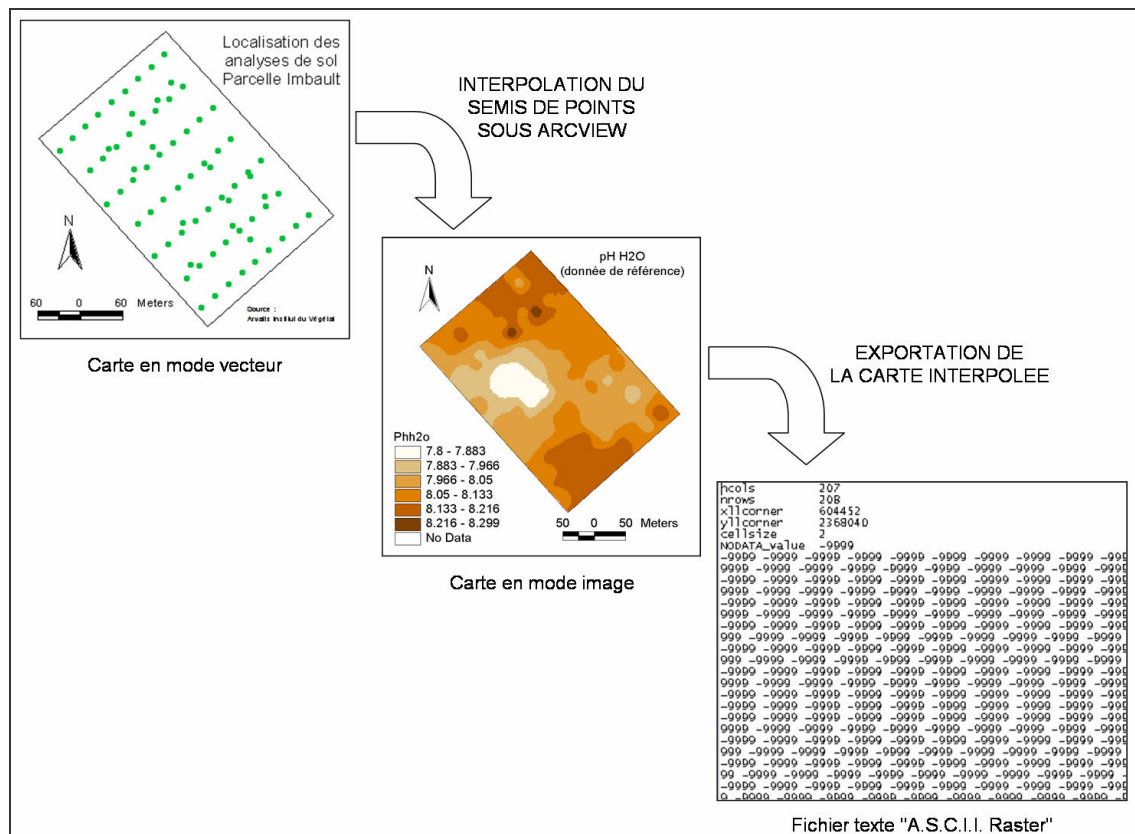


Figure 25 : Création d'un fichier "A.S.C.I.I. Raster" à partir d'une carte de semis de points

A l'issue du zonage, la ou les cartes de zonage finales sont enregistrées dans ce même format "A.S.C.I.I. Raster" et peuvent être à leur tour converties en "grids" sous Arcview. Le cas échéant, il est possible de reconverter les "grids" des zonages en mode vecteur (voir Figure 26).

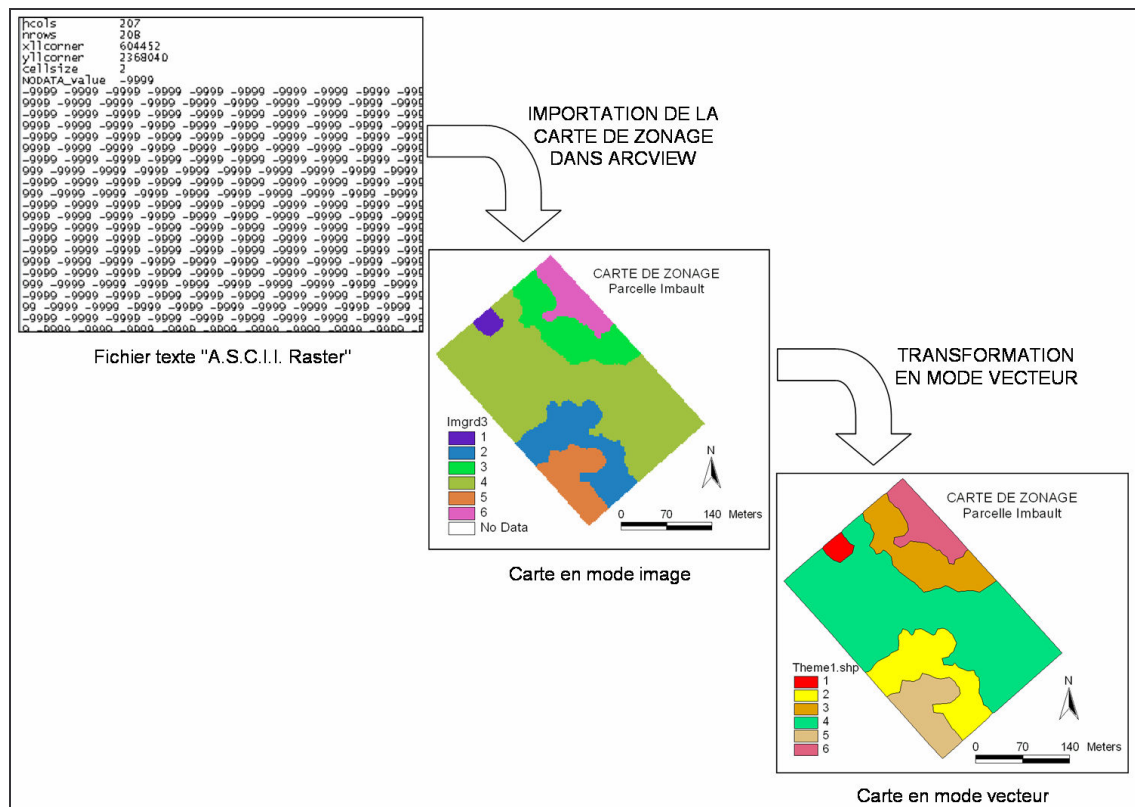


Figure 26 : Création d'une carte de zonage en mode vecteur à partir d'un fichier "A.S.C.I.I. Raster"

Les "grids" initiaux et finaux étant donc en format image, il est rarissime que le rectangle les représentant soit parfaitement superposé à la région sur laquelle nous souhaitons établir le zonage (voir). Les grids contiennent donc :

- La région que nous qualifions d'"utile" de la carte, c'est-à-dire la partie de la carte dans laquelle se situent les points que nous devons traiter afin d'établir le zonage.
- Une ou plusieurs régions de "no data" dans les angles de la carte, représentées souvent en noir ou sans couleur.

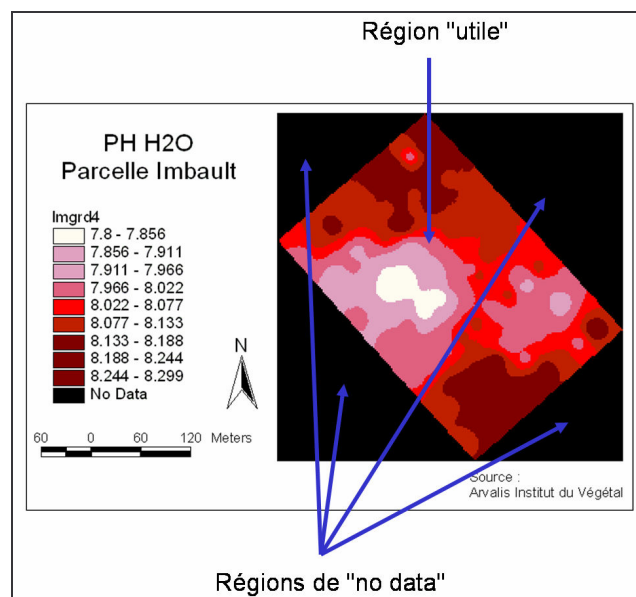


Figure 27 : Région "utile" et régions de "no data" d'une carte en mode image

10.2.2. Le choix des cartes initiales et leur pondération

Même si la méthode de zonage elle-même est différente de celle utilisée en section 9.1 (à savoir l'A.M.C.D. par le biais de la méthode E.L.E.C.T.R.E. TRI), la démarche adoptée pour estimer l'importance des cartes initiales afin de leur affecter un poids reste la même.

Cette étape est d'une importance capitale car elle va influencer directement la carte de zonage obtenue. Elle permet en particulier d'intégrer la connaissance et les préférences de l'exploitant agricole en lui offrant la possibilité d'opérer certains choix.

L'utilisateur doit tout d'abord choisir quelles sont les cartes, et donc les informations, qu'il va considérer. S'il en a plusieurs à disposition, les utiliser toutes pour définir la carte de zonage n'est pas obligatoirement un gage de qualité.

L'utilisateur doit alors choisir le poids qu'il va attribuer aux différentes cartes. Cette étape repose donc directement sur des appréciations subjectives et arbitraires de la part de l'exploitant agricole.

Une démarche permettant d'estimer la pondération a donc été mise au point avec Denis Boissongotier, d'Arvalis Institut du Végétal (voir Figure 28). Le poids des cartes doit être attribué en fonction de leur pertinence, de leur fiabilité, mais aussi de leur pérennité :

- S'interroger sur la pertinence revient à se poser la question : "L'utilisation de cette donnée a-t-elle un sens pour la préconisation que nous nous proposons de faire ?". Bien entendu, dans le cas d'une donnée directement prise en compte par le modèle agronomique, la question de la pertinence ne se pose pas.
- S'interroger sur la fiabilité revient à se poser la question "Les valeurs prises par la carte peuvent-elles être considérées comme fiables par rapport à la valeur réelle sur le terrain ?".
- S'interroger sur la pérennité revient à se poser la question "Combien de temps l'information fournie par cette carte sera-t-elle être valable ?".

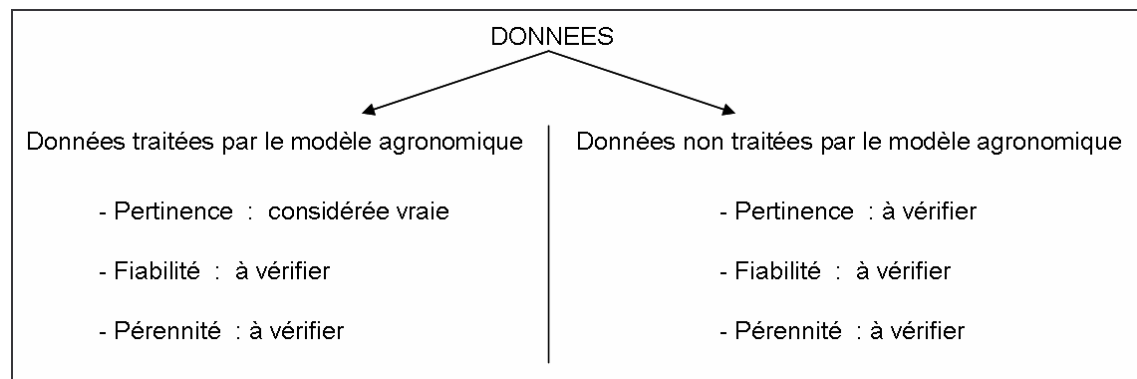


Figure 28 : Les trois interrogations permettant d'estimer la pondération des cartes initiales

En fonction des considérations sur la pertinence, la fiabilité et la pérennité croissante de chaque carte, il est possible d'établir un ordre d'importance des différentes cartes, et donc une échelle de pondération¹.

¹ On notera que c'est à l'utilisateur qu'incombe la responsabilité de juger de la fiabilité des données. Cette fiabilité dépend en particulier de facteurs tels que la qualité de la correction G.P.S. au moment de la campagne de mesure, la densité du semis de points et la méthode d'interpolation utilisée ainsi que son paramétrage, la résolution d'une image satellitaire... Notre méthode de zonage ne vise donc en aucun cas à jauger la qualité des données utilisées. Elle se sert de ces données pour définir des entités homogènes qui se veulent, toutes considérations sur l'adéquation des données aux objectifs fixés mises à part, les plus satisfaisantes possibles.

10.2.3. L'Analyse en Composantes Principales

Méthode statistique d'analyse multivariée, l'Analyse en Composantes Principales (A.C.P.) remplace les variables qui sont corrélées par de nouvelles variables : les composantes principales. Ainsi, elle permet de décorrélérer les données et de réduire le nombre de dimensions à traiter tout en préservant au maximum l'information initiale [Bouroche, Saporta 1989] [Minvielle, Souiah 2003] [Lebart, Morineau, Piron 2000] [Groupe Chadule 1987]. De plus, les composantes principales sont indépendantes.

A l'issue de la transformée, les différents axes ont la propriété suivante : le premier axe est celui sur lequel les informations sont le plus étalées, et dont la valeur propre est la plus grande (voir Figure 29). Les axes suivants, orthogonaux aux précédents, comportent tous une part d'information décroissante [Girard, Girard 1999]. Par conséquent, les premières composantes regroupent une grande partie de la variance des cartes initiales. En conservant un nombre de composantes inférieur au nombre de cartes initiales, on réduit le nombre de dimensions tout en conservant une proportion significative de l'information initiale et en laissant de côté une bonne proportion des fluctuations mineures.

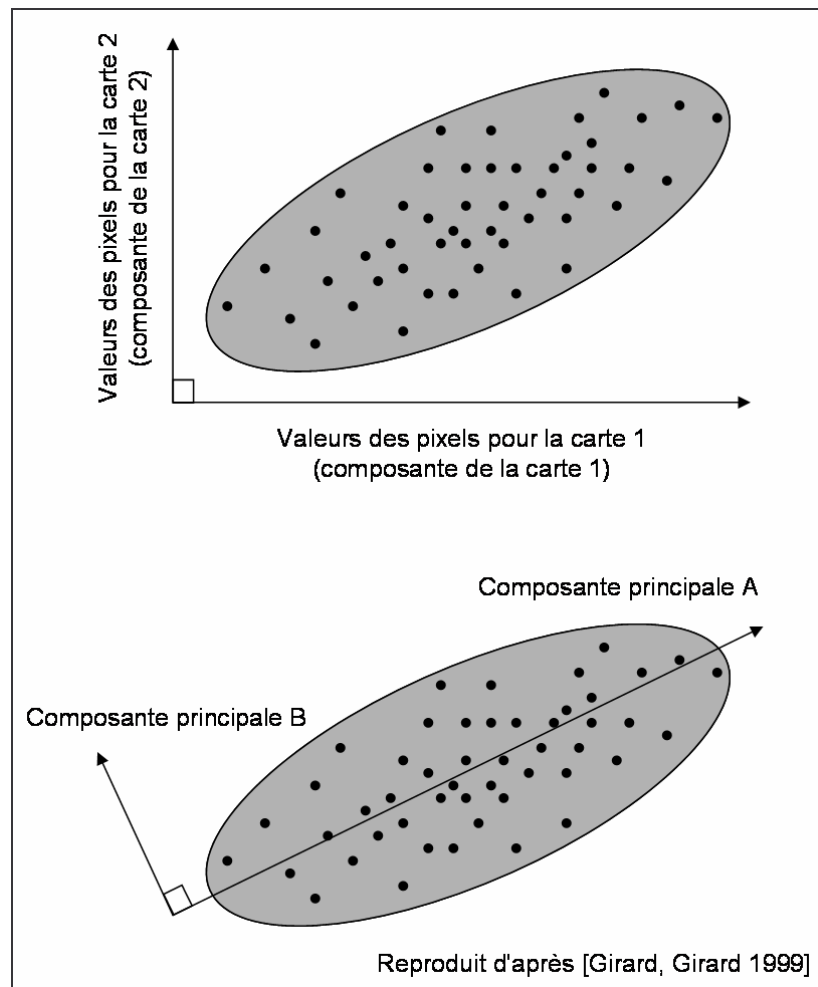


Figure 29 : Principe du changement de système d'axes de l'A.C.P.

Sur la Figure 29, la composante principale A contient le maximum de variance des données des cartes initiales. La composante principale B possède la variance la plus grande après la composante A. S'il y avait encore d'autres composantes, la variance diminuerait ainsi pour chacune d'entre elles.

L'A.C.P. est par exemple utilisée par Chung, Sudduth, Drummond et Kitchen dans la recherche de structures spatiales à partir de données d'analyses de sol, de topographie et de rendement [Chung, Sudduth, Drummond, Kitchen 2001].

Pour notre méthode de zonage, l'intérêt de l'A.C.P. est de diminuer le nombre de dimensions et donc de créer des supervariables pour faciliter la classification par k -moyennes [Dumolard 1981]. Cette réduction du nombre de dimensions est rendue possible, et qui plus est efficace, par le fait que les tableaux de données géographiques sont souvent déjà structurés, cette structure intervenant au niveau des individus voisins ou consécutifs [Lebart, Morineau, Piron 2000].

C'est au cours de l'A.C.P. que la pondération définie par l'utilisateur (telle que définie en section 10.2.1) va être mise en valeur. Afin de conserver la maîtrise des échelles de valeurs des cartes, conformément à leur pondération, l'A.C.P. pratiquée est une A.C.P. centrée et non centrée réduite.

Toutes les valeurs des différentes cartes initiales subissent d'abord une translation pour que toutes les valeurs soient positives (afin d'être exploitables lors du futur calcul d'indicateurs de crédibilité, voir section 12.1). Elles sont ensuite normées entre 1 et 2 avant d'être multipliées par leur poids respectif. Le fait que des poids différents multiplient les valeurs des différentes cartes initiales va contribuer à augmenter la variance des données de ces cartes. A leur tour, ces augmentations de la variance différenciées par la pondération vont contribuer à étirer le nuage de points des données initiales et à influencer le positionnement des composantes principales sur la matrice de variance-covariance. Dès lors, c'est l'ensemble de la méthode de zonage qui subit l'influence de la pondération choisie, ce qui correspond à l'un de nos objectifs présenté en section 6.1.

Le nombre de composantes principales retenues à l'issue de l'A.C.P. est défini automatiquement de sorte à conserver 90% de l'information initiale.

10.2.4. La classification par k -moyennes

Première étape de notre classification mixte (voir section 10.1.1), la classification par k -moyennes est une technique d'agrégation autour de centres mobiles [Lebart, Morineau, Piron 2000] [Bouroche, Saporta 1989] [Minvielle, Souiah 2003] [Groupe Chadule 1987]. Elle permet le traitement statistique, dans un but d'intelligibilité, d'un ensemble d'observations dont la structure profonde n'est pas immédiatement décelable [Groupe Chadule 1987]. Cette méthode est actuellement la mieux adaptée pour le partitionnement de vastes recueils de données [Lebart, Morineau, Piron 2000]. Par rapport aux méthodes fondées sur les nuées dynamiques ou les centres mobiles, la méthode des k -moyennes a la particularité de recalculer à chaque réaffectation d'individu la position du centre de classe correspondant. En quelques itérations, cette procédure peut donc donner une partition de qualité [Lebart, Morineau, Piron 2000].

Cette classification permet de partitionner en k classes n individus caractérisés par p variables. D'une manière générale, un algorithme de partitionnement de ce genre (k -moyennes, nuées dynamiques, centres mobiles) procède de la manière suivante (voir Figure 30) [Lebart, Morineau, Piron 2000] :

- 1) Il détermine k centres provisoires de classes.
- 2) Il calcule la distance (souvent la distance euclidienne) entre chaque individu n du nuage de points et chaque centre de classe.
- 3) Il affecte chaque individu n à la classe c dont le centre de classe est le plus proche.
- 4) Il calcule le centre de gravité de chaque classe et déplace le centre de classe sur ce centre de gravité.
- On itère jusqu'à ce qu'une condition d'arrêt prédéfinie soit remplie.

On procède ainsi jusqu'à ce que la classification soit stable. Cette stabilité peut être définie de différentes façons : deux itérations successives conduisent à la même partition, la courbe des rapports des variances inter et intra-classes atteint une asymptote, un nombre maximal d'itérations a été fixé...

Pour un individu i et un centre de classe c , la distance euclidienne $d(i, c)$ sur la base de p variables vaut :

$$d(i, c) = \sqrt{\sum_{j=1}^p (x_{ij} - c_j)^2} \quad (\text{Équation 1})$$

Où x_{ij} est la valeur de l'individu i et c_j la valeur du centre de classe c pour chacune des p variables.

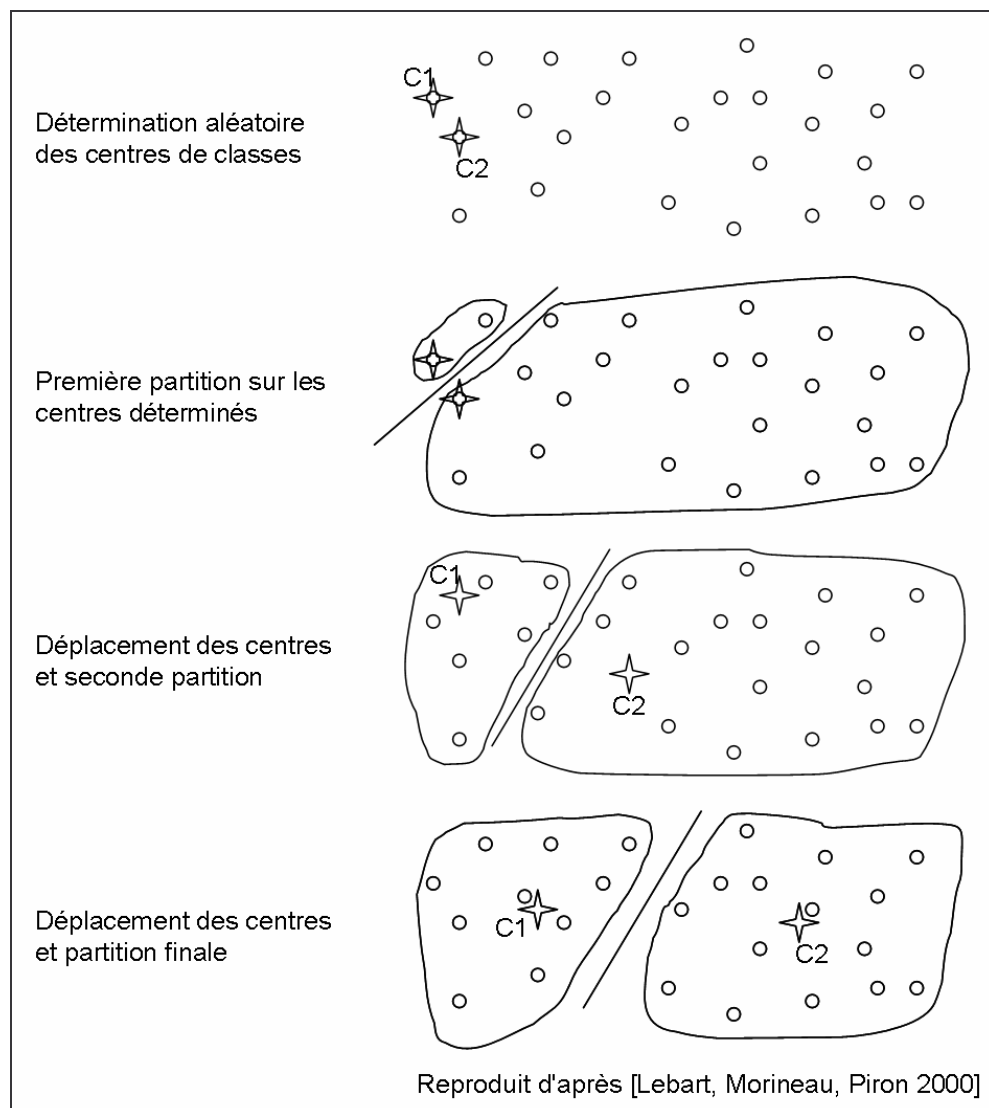


Figure 30 : Exemple de classification par k -moyennes sur deux classes dans un espace à deux variables en trois itérations

Pour notre méthode de zonage, les p variables sont les axes de l'A.C.P. permettant de regrouper 90% de l'information initiale. Les n individus sont les points des cartes et leurs valeurs dans l'espace à p dimensions seront celles calculées lors du changement de système d'axes de l'A.C.P.

La classification par k -moyennes étant initialisée aléatoirement, deux classifications peuvent donner deux cartes différentes. Cela constitue, à notre sens, une faiblesse. Dans notre méthode, le procédé de classification par k -moyennes est donc répété une dizaine de fois avec des initialisations chaque fois différentes. A l'issue de chaque classification, la variance intra-classes et la variance inter-classes sont calculées afin de juger de la robustesse de la classification. Au final, la classification retenue correspond au plus petit rapport variance intra-classes / variance inter-classes obtenu (soit l'inverse du F de Fischer Snedecor).

Pour notre méthode de zonage, le nombre de classes a été fixé à quinze. Boisgontier et Piraux estiment qu'une parcelle peut être divisée en trois à huit zones [Boisgontier, Piraux 2003]¹. En fixant le nombre de classes à quinze, une valeur relativement élevée par rapport au seuil de huit zones et à la taille des parcelles françaises, nous nous assurons que les classes auront une taille modeste. Mais une même classe statistique est toujours fragmentée spatialement, se présentant sous forme d'amas de points discontinus sur la carte, donc se concrétisant par plusieurs zones. Par conséquent, on obtient à l'issue de la classification une carte comportant un grand nombre de zones de tailles réduites. La réduction de l'information initiale par les k -moyennes est donc modeste, conformément au principe de la classification mixte (voir section 10.1.1).

10.2.5. La morphologie mathématique

A l'issue de la classification par k -moyennes, la carte contient souvent des points isolés ou des petits amas de points qui n'apportent pas d'information et nuisent à la lisibilité de la carte. Un traitement spatial s'appuyant sur la morphologie mathématique est donc appliqué pour pallier à ce problème.

La morphologie mathématique considère une image comme un ensemble de points ayant des coordonnées discrètes ou continues. Un objet est formé de points contigus ayant les mêmes caractéristiques. On distingue l'objet de son arrière-plan et il est alors possible d'étudier la façon dont les deux sont connectés. Les opérations peuvent donc être effectuées uniquement sur les points des limites entre les objets et leur arrière-plan [Young, Gerbrands, Van Vliet 1998] [Minvielle, Souiah 2003].

A partir des opérations mathématiques élémentaires que sont l'intersection, l'union, le complément et la translation, la morphologie mathématique définit deux nouvelles opérations : la dilatation et l'érosion. Elles s'opèrent sur la base d'un élément structurant qui peut prendre deux formes (voir Figure 31) et qui s'applique à chaque point de l'objet. La dilatation et l'érosion font respectivement se grossir ou s'amincir l'objet auquel elles sont appliquées (voir Figure 32). Ainsi, si l'on considère un objet A et son arrière-plan sur une image binaire, la dilatation de l'un donne un résultat équivalent à l'érosion de l'autre.

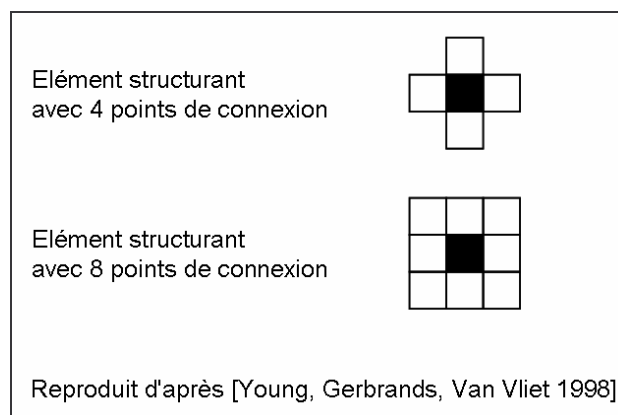


Figure 31 : Eléments structuraux à quatre et huit connecteurs permettant de considérer respectivement le 4-voisinage et le 8-voisinage du point central

¹ [Boisgontier, Piraux 2003] : Communication privée.

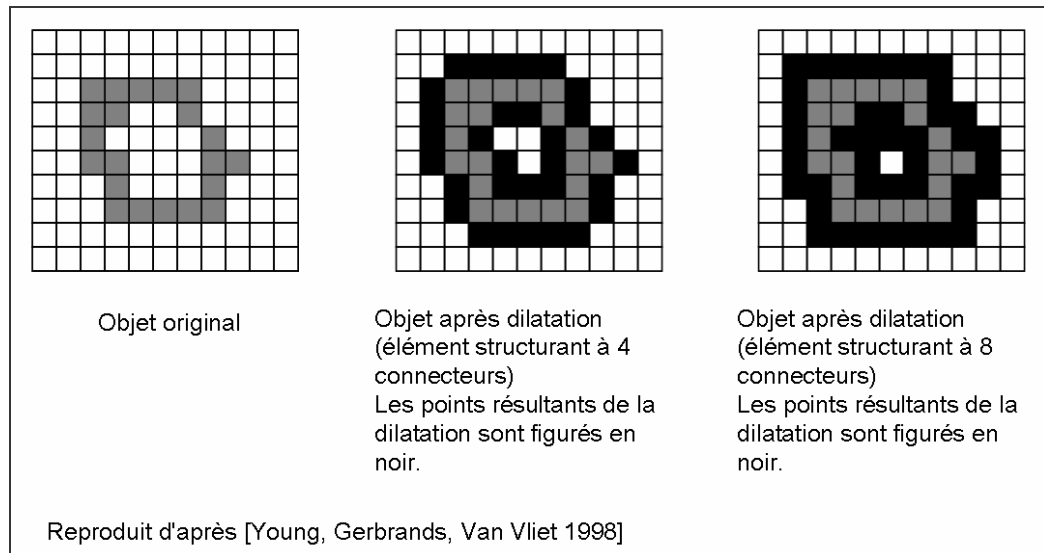


Figure 32 : Exemple de dilatation d'un objet avec les éléments structurants à 4 et 8 connecteurs

A partir de ces deux opérations de dilatation et d'érosion, la morphologie mathématique propose deux autres opérations plus complexes : l'ouverture et la fermeture. L'ouverture est en fait une érosion suivie d'une dilatation, les deux étant effectuées avec le même élément structurant. La fermeture est une dilatation suivie d'une érosion. Là encore, le même élément structurant est appliqué. L'ouverture permet de séparer des objets qui sont connectés alors que la fermeture permet de combler des petits "trous" dans l'image. Les deux opérations consistent donc en lissage utilisant un élément structurant.

Nous allons donc introduire une première contrainte de contiguïté dans notre méthode de zonage par le biais de la morphologie mathématique. Bien que la morphologie mathématique ne tienne pas compte des coordonnées géographiques, elle considère, comme nous l'avons vu, la forme des jonctions entre deux objets. Elle tient donc compte des positions relatives et des valeurs des points situés autour du point concerné. Autrement dit, les valeurs des points contigus vont influencer sur la valeur finale que prendra le point central. La morphologie mathématique intervient donc à la fois comme une contrainte morphologique et comme une contrainte de contiguïté, même si cette dernière intervient à très grande échelle et prend la forme d'un filtre.

Pour notre méthode de zonage, c'est l'opération de fermeture qui est utilisée. Elle permet de faire disparaître une grande partie des points isolés ou les petits amas de points qui font transition entre deux zones plus importantes. Les objets que nous cherchons à éliminer étant de très petite taille, la taille de la fenêtre de fermeture appliquée est la taille minimale de trois points sur trois points (3x3 ou trois pixels sur trois pixels). L'élément structurant est à huit connecteurs.

La taille de la fenêtre de fermeture retenue, trois points sur trois points, ne permet toutefois pas d'éliminer totalement les points isolés. Ils restent souvent quelques points relevant de structure de points "en ligne" ou "en damier". Il suffirait d'augmenter la taille de la fenêtre de fermeture pour résoudre ce problème. Mais nous n'avons pas retenu cette solution pour deux raisons :

- L'opération de fermeture s'applique à tous les points et ne prend pas en compte l'information attributaire si ce n'est par le biais de la classification. Cela n'occasionne pas de réelle perte d'information attributaire lorsque la taille de fenêtre est modeste puisque peu de points sont concernés. En revanche, en augmentant la taille de la fenêtre de fermeture, le traitement purement spatial de chaque point par rapport à un large voisinage peut induire des changements de classe intempestifs. Ceux-ci pourraient à leur tour entraîner une perte d'information attributaire pour les traitements suivants.

- L'augmentation de la taille de la fenêtre de fermeture contribue à ce que les limites des zones prennent des formes "carrées". En effet, les limites de zones s'appuient alors sur les coordonnées lignes-colonnes de la carte qui sont rarement en adéquation avec la structure spatiale ou l'orientation de la parcelle. La carte résultante a alors un aspect peu naturel, voire choquant, ce qui pourrait nuire à son acceptation par l'exploitant agricole.

Nous avons donc préféré développer un filtre supplémentaire pour supprimer les quelques points isolés restants à l'issue de la fermeture.

10.2.6. La suppression des points isolés restants

A l'issue de la fermeture, la carte de zonage présente souvent encore quelques points isolés, voire quelques amas de points "en ligne" ou "en damier". Nous avons donc mis au point un filtre qui permet d'introduire un second traitement spatial.

Ce filtre repose sur un filtre majoritaire ne s'appliquant qu'à une condition précise. Il considère le 8-voisinage de chaque point de la carte par le biais d'un élément à huit connecteurs (voir Figure 31 en section 10.2.5). Si sept ou huit de ces huit points voisins n'ont pas la même valeur que le point concerné, ce dernier prend alors la valeur modale du voisinage. En appliquant ce filtre à plusieurs reprises, on élimine non seulement les points isolés, mais aussi les petites lignes et les structures "en damier" laissés par la fermeture, et ce sans changer l'allure générale de la carte (voir Figure 33).

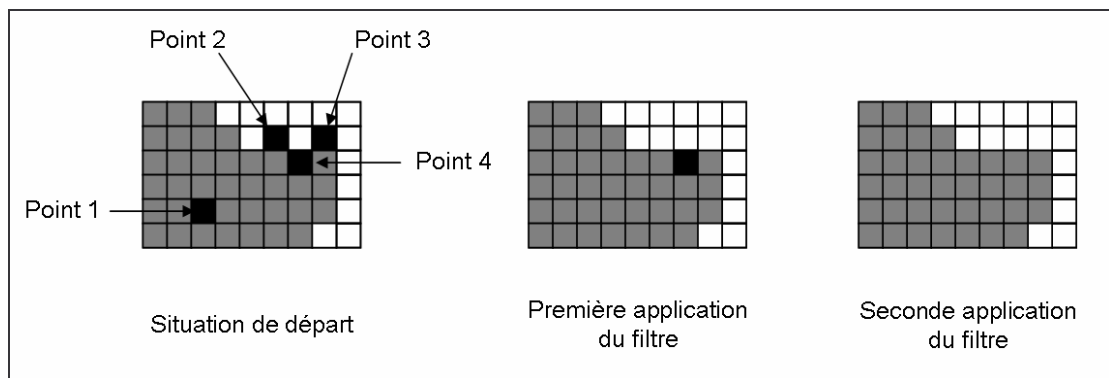


Figure 33 : Exemple d'applications successives du filtre

Sur la Figure 33, les points que nous voulons supprimer sont figurés en noir. Au premier lissage, les points 1, 2 et 3 seront retenus car sept ou huit des points voisins ont des couleurs différentes. Le point 1, entouré de points gris, devient gris. Les points 2 et 3, entourés de 7 voisins de couleurs différentes, prennent la couleur la plus représentée chez leurs voisins, soit ici le blanc. Au premier lissage, le point 4 n'a que six voisins de couleurs différentes, cinq gris et un blanc, et n'est donc pas modifié. En revanche, au second lissage, il ne compte plus que des voisins de couleurs différentes et il prend donc la couleur la plus représentée, soit ici le gris.

La Figure 34 représente la carte de zonage après les trois dernières étapes présentées : *k*-moyennes, fermeture par morphologie mathématique et suppression des points isolés restants.



Figure 34 : Exemple d'application de la fermeture et du lissage à la carte de zonage issue des k -moyennes

10.2.7. La différenciation des entités statistiques en entités spatiales

A ce stade de la méthode de zonage, une même classe statistique est fragmentée en plusieurs entités spatiales différentes (voir Figure 34). Sur la carte, cela signifie que plusieurs zones ont la même couleur. La distinction des entités statistiques en entités spatiales va permettre d'attribuer un identifiant différent à chaque zone homogène continue spatialement.

Simple modification de la mise en forme de la carte, cette étape est indispensable au bon fonctionnement de l'étape suivante, la Classification Ascendante Hiérarchique avec contrainte de contiguïté.

10.2.8. La Classification Ascendante Hiérarchique avec contrainte de contiguïté

Seconde étape de la classification mixte (voir section 10.1.1), la Classification Ascendante Hiérarchique (C.A.H.), méthode statistique d'analyse multivariée, permet d'opérer une hiérarchie de partitions en regroupant deux à deux des individus ou groupes d'individus sur la base d'un critère de ressemblance [Bouroche, Saporta 1989] [Minvielle, Souiah 2003] [Lebart, Morineau, Piron 2000] [Groupe Chadule 1987].

Dans notre cas, la C.A.H. va permettre de regrouper les amas de points qui caractérisent la carte de zonage à ce stade en entités homogènes ayant une taille suffisante pour une mise en œuvre opérationnelle, conformément à nos objectifs exposés en section 6.1.

La C.A.H. va calculer la distance euclidienne entre les différentes zones de la carte et opérer la fusion des deux zones contiguës minimisant cette distance (voir Figure 35). Le calcul de la distance euclidienne se base sur les valeurs d'axes de l'A.C.P. On tient ainsi compte, implicitement, de la pondération des cartes initiales. On réitère le processus de fusion jusqu'à n'avoir qu'une seule zone.

Entre une zone i et une zone k , la distance euclidienne $d(i, k)$ sur la base de p variables vaut :

$$d(i, k) = \sqrt{\sum_{j=1}^p (\bar{x}_{ij} - \bar{x}_{kj})^2} \quad (\text{Équation 2})$$

Où $\bar{x}_{ij}$ et $\bar{x}_{kj}$ sont les moyennes des zones i et k pour chacune des p variables.

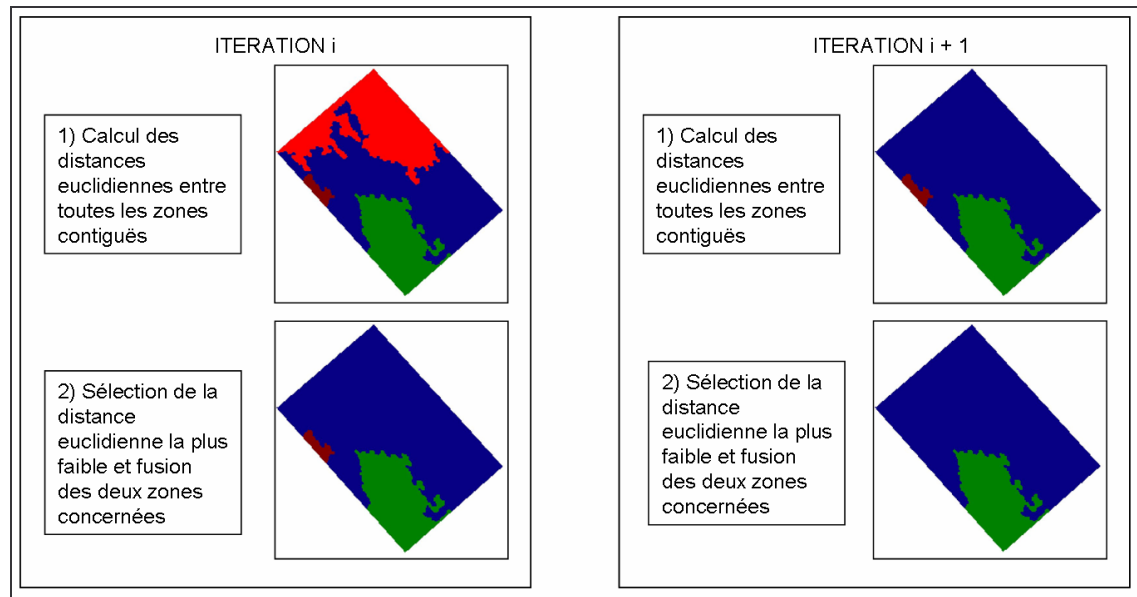


Figure 35 : Représentation de deux itérations successives lors de la C.A.H.

A ce stade, de nombreux auteurs considèrent qu'il est aisé de définir le bon nombre de classes (ou de zones, dans notre cas) à conserver sur la base du dendrogramme [Minvielle, Souiah 2003] [Lebart, Morineau, Piron 2000]. Mais les démarches pour ce faire sont toutes subjectives, faisant appel au jugement de l'utilisateur. De plus, aucune d'entre elles ne s'appuie sur une démarche automatisable et/ou reproductible.

Nous avons donc mis au point une technique afin de définir automatiquement le nombre de zones à conserver pour la carte finale. L'historique des distances de fusion (voir Figure 36) est consulté. L'objectif étant de constituer des entités homogènes, on souhaite conserver des cartes pour lesquelles : les fusions précédentes concernaient des zones se ressemblant ; les fusions suivantes auraient concerné des zones ne se ressemblant pas. Par conséquent, lorsqu'une fusion concerne une carte définissant des entités homogènes crédibles (par rapport aux données utilisées), les distances euclidiennes des fusions précédentes décroissent et les distances euclidiennes des fusions suivantes croissent. On retient donc la carte correspondant à un minimum local dans l'historique des distances de fusion de la C.A.H.

Plus le nombre d'itérations réalisées est important, moins la carte de zonage compte de zones. Or nous cherchons à conserver un nombre restreint de zones. L'évolution des distances de fusion est donc consultée en commençant par la dernière itération et en remontant l'historique jusqu'à un minimum local.

A travers cette méthode de recherche de minimum local, c'est donc la C.A.H. qui va déterminer automatiquement le nombre de zones de la carte finale en recherchant le meilleur compromis entre la restitution de l'information attributaire et le respect des contraintes spatiales.

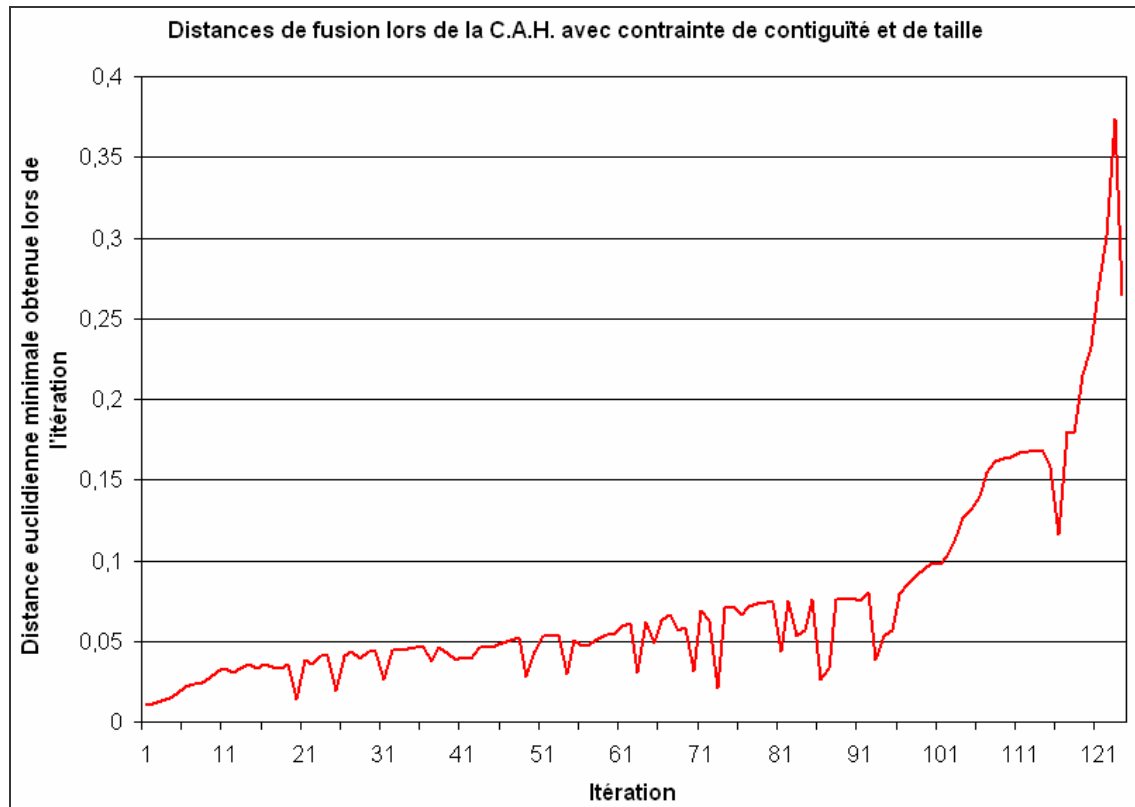


Figure 36 : Historique des distances de fusion de la C.A.H. avec contrainte de contiguïté et de taille

Comme on peut le constater sur la Figure 36, il est possible, et même fréquent, que la courbe des distances de regroupement présente plusieurs minima locaux. Cela signifie que plusieurs zonages peuvent être fournis lors d'une même application de la méthode de zonage, chacun d'entre eux correspondant à un zonage satisfaisant pour un nombre de zones précis.

Toutes les cartes correspondant à un minimum local subissent alors un dernier traitement spatial. Celui-ci s'assure que les cartes de zonage ne contiennent pas de zones trop petites nuisant à leur applicabilité. Le seuil retenu pour la suppression des zones trop petites est de 1% de la surface totale de la parcelle.

10.2.9. Le choix des cartes de zonage

À l'issue de la C.A.H. avec contrainte de contiguïté et de taille, plusieurs zonages sont proposés à l'utilisateur. Ceux-ci comportent plus ou moins de zones et c'est à l'utilisateur de choisir le nombre de zones qu'il souhaite voir apparaître sur la ou les cartes de zonage sélectionnées. Ce choix n'est pas anodin. Plus le nombre de zones est important, plus on affine la restitution des transitions entre les tendances fortes de la structure spatiale [Gerwig, Sadler, Evans 2000].

En fonction de son implication dans l'agriculture de précision, c'est à l'exploitant agricole de définir le compromis qu'il souhaite entre : d'une part la qualité de la restitution de l'information attributaire et spatiale initiale ; d'autre part la lisibilité et l'exploitabilité de la carte de zonage finale.

L'exploitant agricole est ensuite libre de ne pas tenir compte ou d'affiner certaines recommandations de la carte de zonage : fusion de zones, découpage supplémentaire sur la base de sa propre expertise, adoucissement des limites entre les zones... L'exploitant agricole doit garder l'initiative et pouvoir modifier le zonage à son gré. Cela peut se faire simplement sous Arcview en utilisant la carte de zonage comme un modèle sur la base duquel il fondera son propre zonage.

Il est important de noter que la numérotation des différentes zones des cartes de zonage obtenues avec notre méthode n'a aucune signification qualitative. Cette numérotation est établie par les classifications et ne permet en rien de juger, par exemple, du potentiel de rendement de telle ou telle zone.

10.2.10. Calculs d'indicateurs de crédibilité

Une fois la ou les cartes de zonage finales sélectionnées par l'exploitant agricole, une dernière étape (voir section 10.1.4), consiste à calculer différents indicateurs statistiques afin de refléter, dans la mesure du possible, la crédibilité du ou des zonages retenus. Ces indicateurs vont donc nous permettre de procéder à une validation spatiale pour chaque carte de zonage.

Bien que ces calculs soient techniquement intégrés à notre méthode, ils en sont exclus en ce sens qu'ils n'influent pas sur la carte de zonage finale. Ces calculs seront donc présentés en section 12, avec les autres méthodes de validation.

11. L'analyse des données d'entrée

L'analyse des données concernant les cartes initiales et d'éventuelles cartes de référence a un intérêt double :

- Elle permet d'étudier, avant même d'appliquer la méthode de zonage, les correspondances et/ou les corrélations entre les différentes données. Il est ainsi possible de relever des erreurs lors de certains zonages et de corriger ces erreurs.
- Elle permet de disposer d'un outil sur lequel appuyer l'interprétation finale des résultats afin de nuancer la validité de telle ou telle conclusion.

11.1. L'Analyse en Composantes Principales

11.1.1. Principes

Le fonctionnement général de l'A.C.P. a été décrit en section 10.2.3. En réalisant le plan factoriel des variables à l'issue de l'A.C.P., il est possible d'étudier les attractions et répulsions entre variables. A leur tour, ces attractions et répulsions permettent d'établir des hypothèses de redondance entre différentes variables [Mallarino, Hinz, Oyarzabal 1996].

L'interprétation se fait, en particulier, par le biais d'un plan factoriel. Les valeurs des différentes modalités des cartes sont représentées dans l'espace factoriel en fonction de leurs coordonnées de projection sur deux ou trois composantes. Ce diagramme présente l'avantage d'offrir une grande lisibilité par le biais du nuage des modalités.

Chaque composante explique une certaine proportion de la variance totale exprimée en pourcentage.

11.1.2. Exemple d'application

La Figure 37 reprend le diagramme à colonnes de l'A.C.P. réalisée sur des données de la parcelle expérimentale "Imbault" du réseau variabilité intraparcellaire d'Arvalis Institut du Végétal. L'objectif de cette A.C.P. était de déterminer d'éventuelles attractions entre plusieurs types de données dont nous disposons : deux canaux d'une image satellitaire (rouge et proche infra-rouge), des cartes de résistivité électrique (une carte par voie) et des cartes de rendement (sur trois années, corrigées et corrigées moyennées par zonage).

Les trois premières composantes regroupent plus de 72% de l'information totale. On note que les trois cartes de résistivité électrique forment un ensemble bien distinct. De même, les cartes de rendement corrigées des différentes années sont toutes proches de leur homologue corrigée et moyennée par zonage. Enfin, les deux canaux de l'image satellitaire n'ont aucune relation entre elles, ni avec aucune autre donnée.

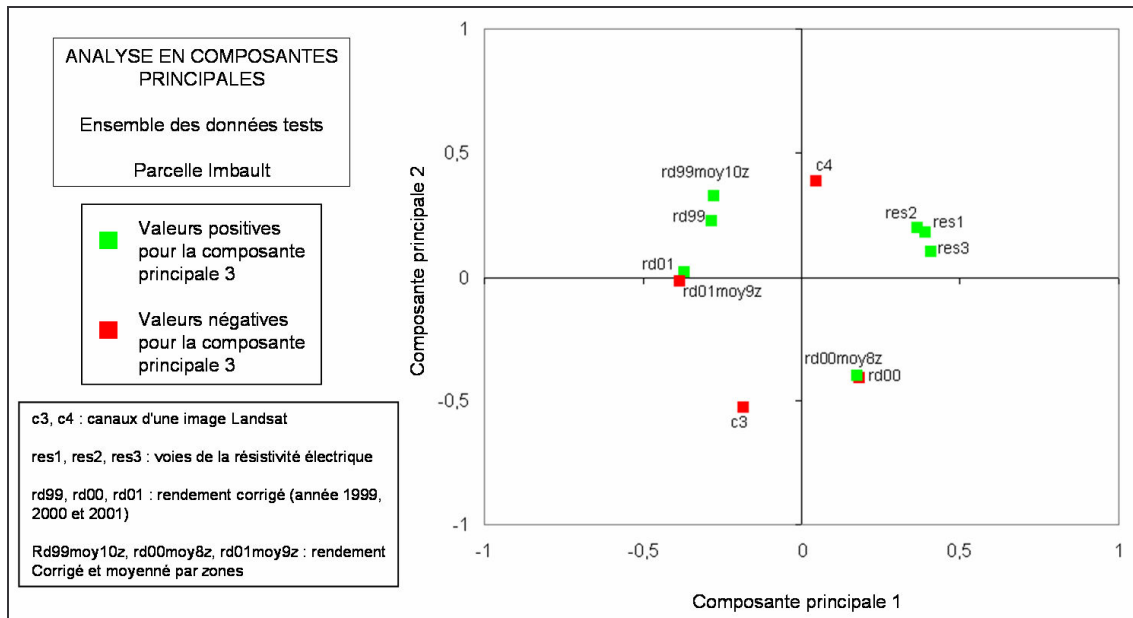


Figure 37 : Analyse en Composantes Principales de données tests sur la parcelle Imbault

Au final, il apparaît qu'aucune relation ne lie les différents types de données entre eux, même si des relations existent entre les différentes cartes de rendement d'une part et les cartes de résistivité électrique d'autre part. Toutefois, malgré le fait que le plan factoriel des variables regroupe près des trois quarts de l'information initiale, les différentes variables sont proches du centre du diagramme.

11.2. L'Analyse des Correspondances Multiples

11.2.1. Principes

L'analyse des correspondances multiples (A.C.M.) est une méthode statistique exploratoire et descriptive, comme l'A.C.P. [Minvielle, Souiah 2003] [Lebart, Morineau, Piron 2000] [Groupe Chadule 1987].

D'un fonctionnement assez proche de l'A.C.P. (voir section 10.2.3 et section 11.1), l'A.C.M. permet d'étudier les attractions et répulsions entre les différentes modalités des variables initiales. A leur tour, ces attractions et répulsions permettent d'établir des hypothèses de redondance entre différentes variables [Mallarino, Hinz, Oyarzabal 1996].

L'interprétation se fait, là aussi, par le biais d'un plan factoriel des variables. Les valeurs des différentes modalités des cartes sont représentées dans un plan en fonction de leurs coordonnées de projection sur les deux premières composantes. Ce plan factoriel présente l'avantage d'offrir une grande lisibilité par le biais du nuage des modalités.

Chaque composante explique une certaine proportion de la variance totale qui est exprimée en pourcentage.

L'inconvénient majeur de ce type d'étude est que les deux premiers axes résument rarement plus de 30% de la variance. De plus, les attractions et/ou répulsions visibles dépendent directement des composantes représentées sur le diagramme des colonnes. Il est donc nécessaire d'interpréter un nombre conséquent de diagrammes.

Pour pouvoir prendre en compte à la fois les données initiales (variable quantitative) et les cartes de zonage (variable qualitative), l'ensemble des variables est codé en tableau disjonctif complet. Cette opération entraîne toutefois une perte d'informations non négligeables qui induit des pourcentages d'inertie expliquée et des R^2 faibles. La représentation des modalités de variables retenue est la représentation par quartiles. Le nombre de classes induit permet, une fois ces quartiles représentés sur une carte, d'obtenir des zones de compacité et de surface suffisantes pour pouvoir interpréter leur influence sur la carte de zonage.

Le recours à l'A.C.M. avait pour objectif d'étudier l'influence des différentes cartes initiales sur chaque zone afin de valider l'influence de la pondération sur la carte de zonage (voir section 12).

11.2.2. Exemple d'application

Cet exemple d'application est issu d'une A.C.M. réalisée sur des données concernant la parcelle "Imbault" du réseau "variabilité intraparcellaire" d'Arvalis Institut du Végétal. L'exercice visait à étudier les relations entre une carte de profondeur de sol et trois cartes de résistivité électrique, aucune carte de zonage n'est donc prise en compte. Les cartes considérées ici sont issues de semis de points qui ont été interpolés, comme décrit en sections 5.1.1 et 5.1.4

Les deux premières composantes issues de l'analyse des correspondances multiples représentent 48% de la variance totale (28% pour la première, 20% pour la seconde) (voir Figure 38). Ce pourcentage, très élevé pour une analyse basée sur un tableau disjonctif complet, traduit, à lui seul, les relations assez étroites qui caractérisent la profondeur de sol et la résistivité électrique.

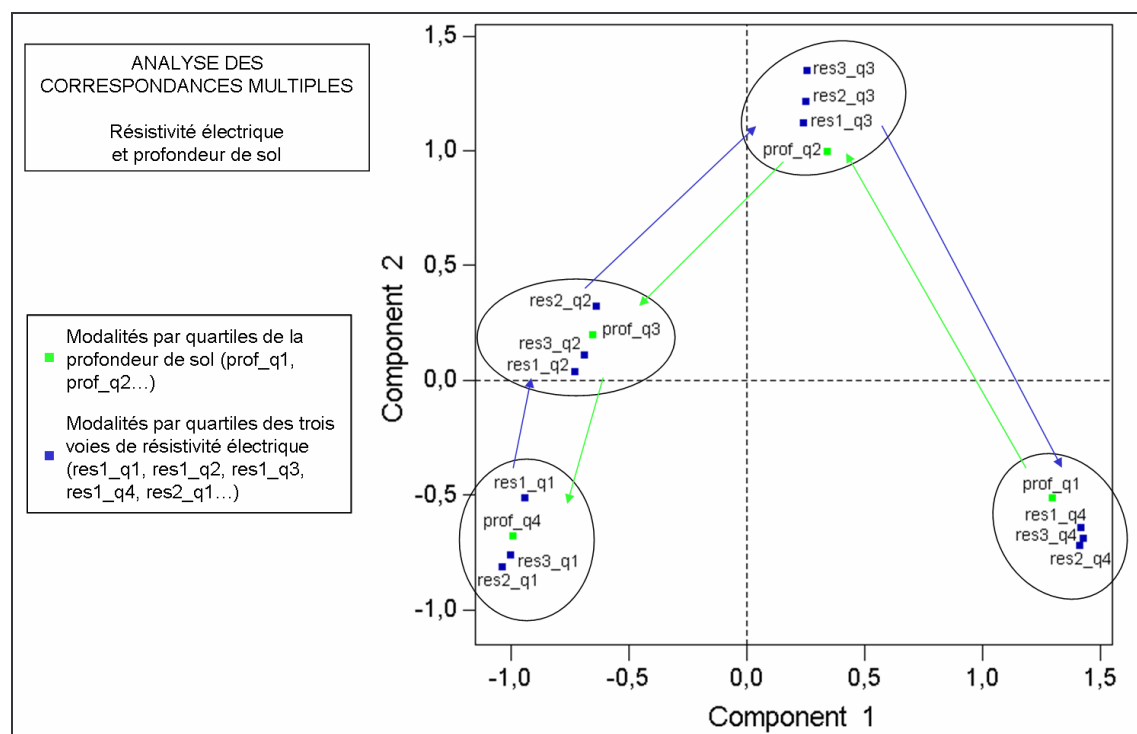


Figure 38 : Diagramme des colonnes (composantes 1 et 2) issu de l'A.C.M. sur les données de profondeur de sol et les trois voies de résistivité électrique

En consultant les contributions des variables à l'inertie des composantes, on constate qu'il n'est pas possible d'attribuer directement l'inertie des deux premières composantes à la profondeur de sol ou à la résistivité électrique ni de l'attribuer à une voie de résistivité bien particulière.

Les relations entre profondeur de sol et résistivité électrique sont évidentes dans ce cas. Les ensembles de modalités et leurs oppositions par rapport aux axes apparaissent de façon très nette sur le diagramme des colonnes : forte profondeur de sol avec faible résistivité électrique et inversement.

L'A.C.M. permet donc, comme une Analyse en Composantes Principales, de mettre en évidence les relations entre variables. Mais elle permet aussi de mettre en évidence dans l'A.C.M. les relations entre modalités de variables, en précisant les variations de grandeurs entre, dans notre cas, les quartiles. Toutefois, pour une information quantitative continue, la mise en évidence de ces relations et variations de grandeurs se traduit par une perte d'informations lors de la réalisation du tableau disjonctif complet.

11.3. L'autocorrélation spatiale

11.3.1. Principes

La mesure de l'autocorrélation spatiale (voir section 3.2.2.2) va nous permettre de juger de la ressemblance ou de la dissemblance d'une même variable sur un intervalle spatial. Dans le cadre de notre travail, cette mesure présente deux avantages :

- Plus une variable est spatialement autocorrélée, plus la carte correspondante est susceptible de fournir facilement un bon zonage. Toutefois, dès que l'on prend en compte deux cartes ou plus, il faut nuancer ce propos. En effet, même si ces cartes ont toutes un bon indice d'autocorrélation spatiale, elles ne représenteront pas forcément la même structure spatiale, ne serait-ce que dans les zones de transition entre les tendances fortes de la parcelle. Par conséquent, si l'autocorrélation spatiale fournit un bon indicateur sur la possibilité d'obtenir aisément un zonage, cela est valable lorsque peu de cartes, voire une seule, sont prises en compte.
- De nombreuses méthodes statistiques s'appuient sur le principe d'indépendance. Or, appliquer ces méthodes à des données géoréférencées fortement susceptibles, en particulier dans le cas d'informations sur une parcelle agricole, d'être soumises à l'autocorrélation spatiale peut amener à des résultats fortement biaisés. En particulier, pour les modèles tels que les modèles de régression, les résidus ne doivent pas être spatialement autocorrélés. Ici, l'autocorrélation spatiale peut servir à indiquer une faiblesse dans l'application de certaines méthodes statistiques.

Le calcul de l'indice d'autocorrélation spatiale d'une carte impose que cette dernière soit représentée en mode image (voir section 5.2.2). Il est possible d'appliquer le calcul avec le 4-voisinage ou le 8-voisinage (voir section 10.2.5).

La formule du calcul de l'autocorrélation spatiale de MORAN I est la suivante :

$$\text{MORAN I} = n \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \mu)(y_j - \mu)}{\left(\sum_{i=1}^n (y_i - \mu)^2 \right) \left(\sum_{i \neq j} \sum w_{ij} \right)} \quad (\text{Équation 3})$$

Où n est le nombre de points de la carte, w le poids spatial des points (1 pour les points haut, bas, droite et gauche ; 0.70711, la racine carrée de 2 pour les points voisins en diagonale dans le cas d'un 8-voisinage), $y_{i/j}$ est la valeur du point et μ la moyenne des valeurs de y .

Dans la plupart des cas, l'indice d'autocorrélation spatiale de MORAN I varie en -1 et +1. +1 indique une autocorrélation spatiale totale entre les valeurs de la carte alors que 0 indique qu'aucune autocorrélation spatiale n'existe. Une valeur négative peut indiquer qu'il existe une sorte de motif régulier sur la carte.

11.3.2. Exemple d'application

Cet exemple d'application est un peu particulier dans la mesure où il présente deux cas extrêmes¹.

La première carte concerne un échantillon du Modèle Numérique de Terrain (M.N.T.) de la forêt de Harvard à Petersham dans le Massachusetts aux États-Unis (voir Figure 39). L'altitude d'un point étant directement liée à l'altitude des points voisins (sauf dans des cas très précis comme les failles ou d'autres accidents de terrain de ce genre), les M.N.T. constitue un type de donnée particulièrement représentatif de ce que peut être l'autocorrélation spatiale.

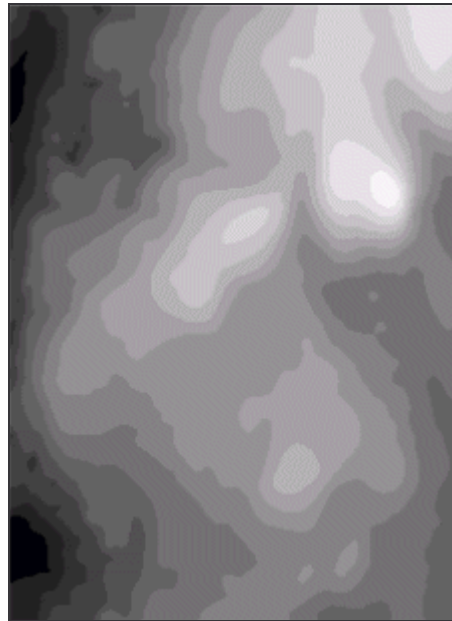


Figure 39 : Echantillon du modèle numérique de terrain de la forêt de Harvard

Ainsi, cette carte obtient un indice d'autocorrélation spatiale de MORAN I de 0.9955, soit une très forte valeur.

Si l'on applique le même calcul sur une carte générée avec une fonction aléatoire (voir Figure 40), l'indice d'autocorrélation spatiale de MORAN I est de -0.0029.

¹ Cette exemple d'application est extrait de la page web
http://geography.laurentian.ca/resources/idrisi/s_tools7.htm



Figure 40 : Carte générée avec une fonction aléatoire

12. La validation de la méthode de zonage

Cette étape de notre travail est aussi nécessaire que délicate à mener. En effet, comme nous l'avons vu en section 10.1.4, Triboulet, Lardon et Langlet expliquent que *"la prise en compte du spatial dans le statistique ne permet pas de mesurer le poids respectif de l'information liée au lieu et de celle liée à son voisinage. Les résultats sont donc difficilement validables et interprétables car les règles d'agrégation sont peu explicites"* [Triboulet, Lardon, Langlet 1996].

Notre méthode propose pourtant les deux approches. Certaines méthodes statistiques et spatiales sont clairement séparées (A.C.P. et *k*-moyennes distinctes de la fermeture et du filtre), respectant ainsi le principe de précaution énoncé ci-dessus. En revanche, d'autres sont intimement liées et c'est ce qui fait leur intérêt (C.A.H. à contrainte de contiguïté). Toutefois, pour ce dernier point, les règles d'agrégation sont particulièrement explicites.

La validation de notre méthode de zonage repose sur le calcul de quelques indicateurs qui permettent de juger de la qualité du zonage obtenu (voir section 12.1). Ces calculs sont intégrés dans la méthode de zonage et effectués pour chaque carte réalisée.

Enfin, une seconde série de tests a consisté à comparer les zonages obtenus avec notre méthode à certaines classifications et segmentations obtenues avec le logiciel eCognition (voir section 6.2.2.4). Toutefois, ne disposant que de la version d'essai de ce logiciel, aucun indicateur statistique sur les classifications obtenues n'était disponible. De même, aucune exportation de cartes depuis eCognition n'était possible. Par conséquent, les comparaisons ont été uniquement opérées de visu. Dans la mesure où la possibilité d'effectuer des comparaisons n'est pas directement intégrée à notre méthode de zonage et où cette validation n'est que visuelle, les résultats seront présentés dans la section 18 au même titre qu'un exemple d'application.

12.1. La validation statistique spatiale des cartes de zonage

Cette partie de notre travail est capitale car c'est par le biais de la validation spatiale, et donc des différents indices calculés, que nous allons pouvoir juger de la performance de la méthode de zonage dans la restitution des structures spatiales des données initiales.

Notre méthode de validation spatiale consiste à calculer différents indicateurs statistiques qui permettent de juger du gain en homogénéité apporté par le zonage par rapport aux cartes sur lesquelles il est fondé. Elle se rapproche d'une méthode de jugement des partitions préconisée par Lebart, Morineau et Piron [Lebart, Morineau, Piron 2000]. Ces auteurs décrivent plusieurs méthodes permettant de valider une classification. Mais peu d'entre elles sont automatiques et la grande majorité relève de l'appréciation empirique.

Notre méthode, s'appuie sur les coefficients de variation (C.V.) qui sont des indicateurs d'hétérogénéité. Leur utilisation pour le calcul de nos propres indicateurs présente l'avantage de réduire l'effet des ordres de grandeur (le C.V. étant une mesure de variabilité relative) [Dumolard, Dubus, Charleux, 2003] [Minvielle, Souiah 2003]. Le principal défaut de cet indicateur est qu'il nécessite une distribution à peu près symétrique.

L'ensemble des calculs est réalisé sur les valeurs d'origine des cartes, c'est-à-dire sans tenir compte d'une éventuelle pondération. L'objectif est ici de juger le zonage par rapport à l'information initiale disponible et non par rapport au point de vue de l'exploitant agricole.

12.1.1. Calcul d'indicateurs locaux pour chaque carte initiale

Les coefficients de variation sont calculés à partir de l'écart-type σ et de la moyenne arithmétique $\bar{x}$ de la façon suivante :

$$C.V. = \frac{\sigma}{\bar{x}} \quad (\text{Équation 4})$$

Un coefficient de variation $C.V._{C_i}$ est ainsi calculé pour chaque carte C_i (des n_c cartes initiales) et un autre $C.V._{C_i Z_j}$ pour chaque zone Z_j (des n_z zones finales) superposée à chaque carte C_i (voir Figure 41).

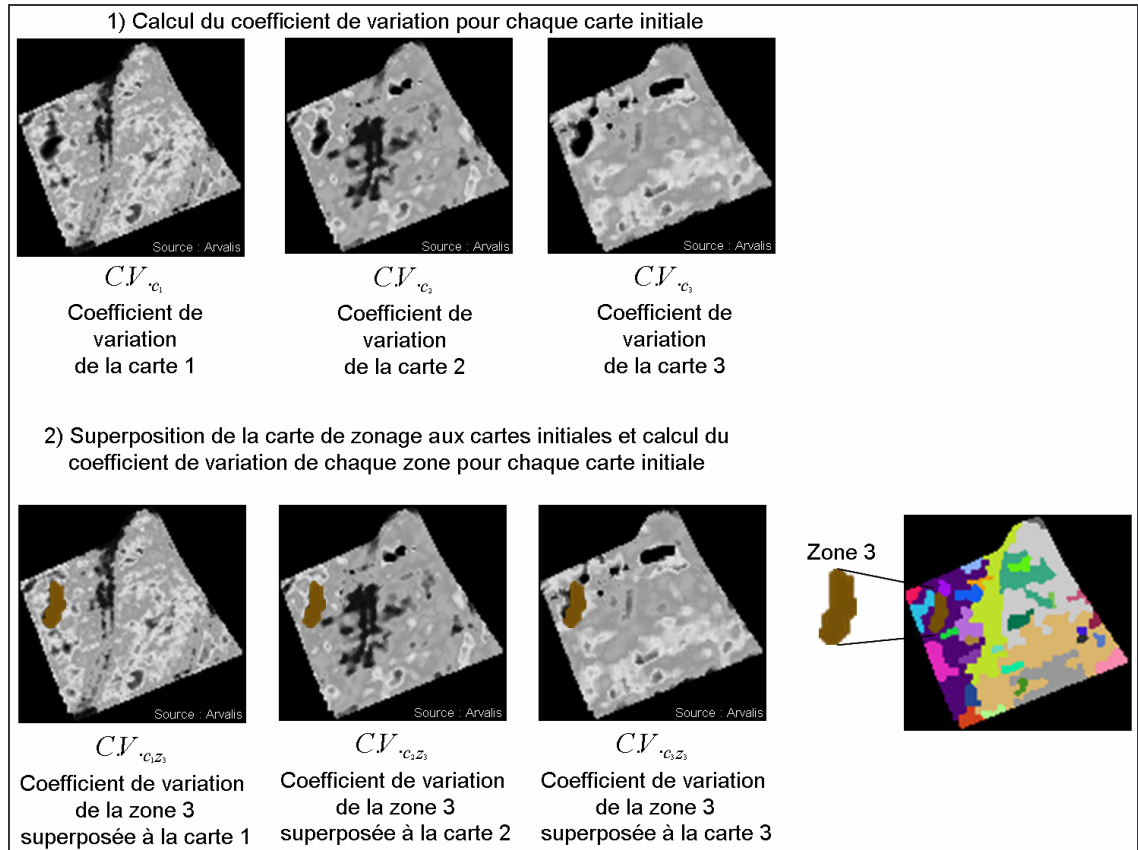


Figure 41 : Calcul des différents coefficients de variation

Pour chaque carte c_i et pour chaque zone $c_i z_j$, on calcule un quotient $Q.C.V_{c_i z_j}$ entre le coefficient de variation $C.V_{c_i z_j}$ de la zone superposée à la carte initiale et le coefficient de variation $C.V_{c_i}$ de cette carte initiale (voir Figure 42). Le quotient de coefficients de variation d'une zone $c_i z_j$ par rapport à une carte initiale c_i vaut donc :

$$Q.C.V_{c_i z_j} = \frac{C.V_{c_i z_j}}{C.V_{c_i}} \quad (\text{Équation 5})$$

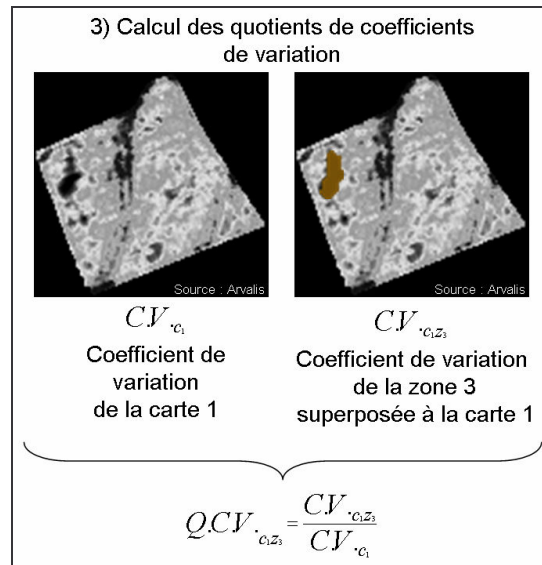


Figure 42 : Calcul des quotients de coefficients de variation

Etablir un quotient entre le coefficient de variation d'une zone d'une carte initiale et celui de cette carte initiale revient à comparer l'hétérogénéité globale de la carte et l'hétérogénéité locale en fonction du zonage. Il est donc possible de juger du gain d'homogénéité (par le biais de la perte d'hétérogénéité) de cette zone pour cette carte. L'interprétation se fait comme suit :

- Si $Q.C.V._{c_1 z_j} < 1$, la zone est moins hétérogène que la carte de référence. Plus le quotient est proche de 0, plus la perte d'hétérogénéité est importante.
- Si $Q.C.V._{c_1 z_j} = 1$, il n'y a ni perte ni gain d'hétérogénéité.
- Si $Q.C.V._{c_1 z_j} > 1$, la zone est plus hétérogène que la carte de référence. Plus le quotient est supérieur à 1, plus le gain d'hétérogénéité est important.

12.1.2. Calcul des Indices d'Hétérogénéité

A partir des indicateurs locaux que sont les quotients de coefficients de variation, nous allons calculer plusieurs indices d'hétérogénéité qui vont permettre de juger de la crédibilité du zonage obtenu.

Un indice d'hétérogénéité (I.H.) est une combinaison de quotients de coefficients de variation. Ce faisant, sa valeur étalon est 1, cas où il n'y a ni perte ni gain d'hétérogénéité. Et plus un I.H. est faible, meilleur il est. Par conséquent, en soustrayant l'I.H. obtenu de l'I.H. étalon de 1, on obtient la part d'hétérogénéité initiale supprimée par l'application du zonage. Cette part d'hétérogénéité supprimée peut être traduite en pourcentage. Ainsi, un I.H. de 0.75 signifie que l'application du zonage permet de réduire de 25% l'hétérogénéité initiale des valeurs.

Les différents Indices d'Hétérogénéité calculés sont les suivants :

- Pour chacune des cartes initiales, un I.H. indique dans quelles proportions l'application du zonage permet de réduire l'hétérogénéité des valeurs de cette carte initiale. Nous appelons cet I.H. "Indice d'Hétérogénéité d'une carte après zonage" (voir section 12.1.2.1).

- Pour chaque zone de la carte de zonage, un I.H. indique dans quelles proportions l'application de cette zone permet de réduire l'hétérogénéité des valeurs de l'ensemble des cartes initiales. Nous appelons cet I.H. "Indice d'Hétérogénéité d'une zone" (voir section 12.1.2.1).
- Pour la carte de zonage, un I.H. indique dans quelles proportions l'application du zonage permet de réduire l'hétérogénéité des valeurs de l'ensemble des cartes initiales. Nous appelons cet I.H. "Indice d'Hétérogénéité de la carte de zonage" (voir section 12.1.2.2).

12.1.2.1. I.H. des cartes initiales après zonage

L'I.H. d'une carte initiale C_i après zonage est la moyenne de ses différents quotients de coefficients de variation $Q.C.V._{C_i Z_j}$ pondérée par les pourcentages des surfaces des différentes zones p_{Z_j} (des n_z zones finales).

$$I.H._{C_i} = \frac{\sum_{j=1}^{n_z} Q.C.V._{C_i Z_j} \times p_{Z_j}}{\sum_{j=1}^{n_z} p_{Z_j}} \quad (\text{Équation 6})$$

Cet I.H. indique dans quelles proportions l'application du zonage permet de réduire l'hétérogénéité des valeurs de carte initiale C_i .

I.H. des différentes zones de la carte de zonage

L'I.H. d'une zone Z_j est la moyenne arithmétique des différents quotients de coefficients de variation $Q.C.V._{C_i Z_j}$ de cette zone pour chaque carte initiale C_i (voir Figure 43).

$$I.H._{Z_j} = \frac{\sum_{i=1}^{n_c} Q.C.V._{C_i Z_j}}{n_c} \quad (\text{Équation 7})$$

Où n_c est le nombre de cartes initiales et n_z le nombre de zones final sur la carte de zonage.

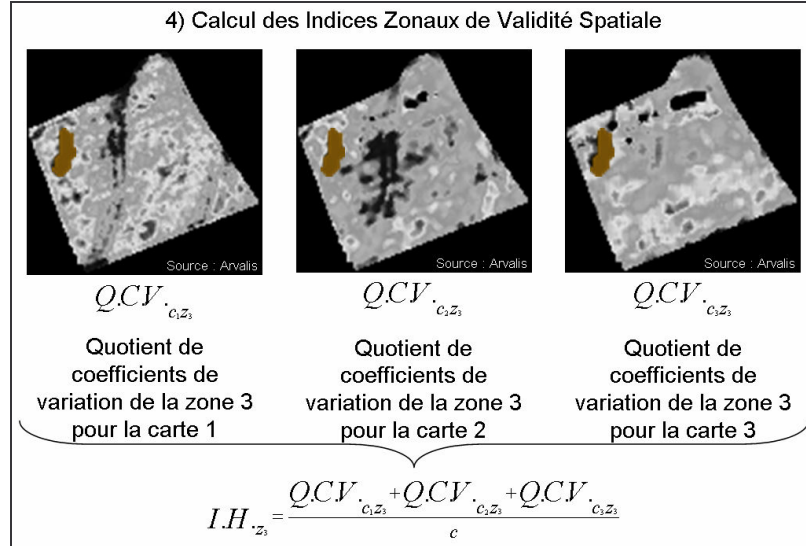


Figure 43 : Calcul des indicateurs zonaux de validité spatiale

Cet I.H. indique dans quelles proportions l'application de cette zone z_j permet de réduire l'hétérogénéité des valeurs des c cartes initiales.

12.1.2.2. I.H. de la carte de zonage

L'I.H. de la carte de zonage est la moyenne des I.H. des différentes zones z_j pondérée par les pourcentages des surfaces des différentes zones p_{z_j} (des n_z zones finales).

$$I.H._{\text{zonage}} = \frac{\sum_1^{n_z} I.H._{z_j} \times p_{z_j}}{\sum_1^{n_z} p_{z_j}} \quad (\text{Équation 8})$$

Cet I.H. indique dans quelles proportions l'application du zonage permet de réduire l'hétérogénéité des valeurs de l'ensemble des c cartes initiales.

12.1.3. Autres informations fournies

D'autres informations sont fournies en plus des I.H. Celles-ci n'ont pas pour but de valider le zonage (même si elles servent au calcul des I.H.) mais de permettre à l'exploitant agricole d'appréhender les caractéristiques des zones.

Ainsi, pour chaque carte initiale, les informations suivantes sont fournies (voir Figure 44) :

- La moyenne et l'écart-type de la carte.
- La moyenne et l'écart-type de chaque zone par rapport à cette carte.
- Le pourcentage de la carte ayant un coefficient de variation amélioré : il s'agit du pourcentage de surface de la carte initiale qui obtient des quotients de coefficients de variation inférieurs à la valeur étalon de 1.

- L'indice d'intensité R^2 : Il s'agit véritablement du pourcentage de variance expliquée mais il est appelé ici R^2 par analogie avec le R^2 de la régression. Il peut ici être considérée comme un indice de qualité globale du zonage. Néanmoins, dès que l'on considère qu'il y a des erreurs aléatoires dans le jeu de données et que l'on souhaite effectuer des tests, l'approche probabiliste s'appuie sur des hypothèses qui sont rarement vérifiées dans notre cadre d'utilisation. Si certaines distributions peuvent être multivariées, il est rare que les zones aient des nombres d'individus ou des variances comparables. Ce sont principalement ces raisons qui nous ont fait préférer le coefficient de variation au pourcentage de variance expliqué.

IH de la carte de zonage : 0.68					
	Aire(%) :	IH des zones :			
Zone 1 :	13.4	0.549			
Zone 2 :	2.45	0.673			
Zone 3 :	9.08	0.673			
Zone 4 :	57.9	0.736			
Zone 5 :	9.28	0.521			
Zone 6 :	7.82	0.692			
Carte : res1					
Poids : 10					
IH après zonage : 0.496					
Carte :	CV :	CV/Cvx :	Aire(%) :	Moyenne :	Ecart-type :
Zone 1 :	5.51			64.8	8.22
Zone 2 :	2.69	0.489	13.4	68.9	4.13
Zone 3 :	1.92	0.349	2.45	68.8	2.94
Zone 4 :	1.76	0.32	9.08	71.6	2.75
Zone 5 :	3.28	0.595	57.9	59.1	4.7
Zone 6 :	1.81	0.328	9.28	77.3	2.92
	1.2	0.219	7.82	76.3	1.94
Pourcentage de la carte ayant un coefficient de variation ameliore : 100					
Indice d'intensite (R²) : 0.747					
Carte : res2					
Poids : 10					
IH après zonage : 0.367					
Carte :	CV :	CV/Cvx :	Aire(%) :	Moyenne :	Ecart-type :
Zone 1 :	10.8			149	54.5
Zone 2 :	3.06	0.283	13.4	155	15.6
Zone 3 :	2.89	0.267	2.45	145	14.4
Zone 4 :	4.77	0.442	9.08	202	26.6
Zone 5 :	3.99	0.369	57.9	114	18.7
Zone 6 :	3.39	0.314	9.28	202	18.9
	5.43	0.503	7.82	278	34.4
Pourcentage de la carte ayant un coefficient de variation ameliore : 100					
Indice d'intensite (R²) : 0.855					

Figure 44 : Exemple de contenu du fichier d'informations fourni avec la carte de zonage

Par le biais de l'I.H. des différentes zones et de ces quelques informations, l'exploitant agricole doit donc être en mesure de juger chaque zone. Il lui est ainsi possible de décider où placer des analyses de sol. Mais il peut aussi, au vu de l'hétérogénéité interne de chaque zone, décider de faire plus d'analyses dans les zones plus hétérogènes... Par le biais de ces indicateurs, notre méthode de zonage apporte donc certaines solutions au problème de la localisation des analyses de sol évoqué en section 5.1.1.2.

Enfin, les différents Indices d'Hétérogénéité peuvent aussi être calculés par rapport aux valeurs d'une ou de plusieurs cartes de référence. Dans ce cas, les calculs sur les cartes initiales et les cartes de référence sont faits distinctement et les indicateurs sont fournis dans deux fichiers différents. Ce calcul d'indicateurs par rapport à des cartes de référence peut permettre d'étudier la reproductibilité de la structure spatiale d'une carte de référence à partir de différentes autres données.

12.2. L'analyse discriminante, une autre méthode de validation envisageable

Le terme "analyse discriminante" désigne couramment deux approches distinctes du problème de la classification supervisée :

- Une approche descriptive, l'analyse factorielle discriminante, qui crée de nouvelles variables permettant de séparer au mieux les classes. Ces nouvelles variables, appelées axes discriminants, sont des combinaisons linéaires des variables initiales.
- Une approche prédictive visant à déterminer les principales différences entre les classes identifiées avec les variables initiales. C'est cette approche qui nous intéresse plus particulièrement.

En effet, cette seconde approche permet de valider une classification puisqu'elle permet de s'assurer de l'optimalité de la partition. Les éléments affectés, selon l'analyse discriminante, à une "mauvaise" classe lors de la classification sont identifiés par le biais des scores des fonctions discriminantes. Il est ensuite possible d'isoler, au sein d'une même classe, les individus qui diffèrent sensiblement des caractéristiques générales de la classe grâce aux poids de discrimination.

Bien que cette méthode puisse être appliquée à notre cas, nous lui avons préféré le calcul des Indices d'Hétérogénéité présenté dans la section précédente. En effet, les fonctions discriminantes reposent sur l'hypothèse restrictive voulant que les classes aient des distributions normales. Or, si cette hypothèse n'est déjà pas systématiquement vérifiée lors d'une classification uniquement statistique, le recours à plusieurs contraintes spatiales lors de la classification la rend encore plus improbable. De même, l'analyse discriminante n'aborde la question de l'optimalité de la partition que d'un point de vue statistique. Or, dans notre cas, la prise en compte de contraintes spatiales est justement une condition de l'optimalité du zonage.

13. Exemple d'application

13.1. Objectifs

En section 9.1, nous avons défini des entités homogènes par la méthode d'Aide MultiCritère à la Décision (A.M.C.D.) E.L.E.C.T.R.E. TRI. Dans cet exemple d'application, nous avons utilisé notre propre méthode de zonage et avons comparé les résultats obtenus avec les deux méthodes. L'objectif principal était de vérifier si les cartes de zonage finales étaient comparables malgré les importantes différences de renseignements de paramètres existantes entre les méthodes.

La parcelle retenue est donc la même que pour l'application d'E.L.E.C.T.R.E. TRI, à savoir la parcelle expérimentale "Baziège" du réseau variabilité intraparcellaire d'Arvalis Institut du Végétal. Afin de pouvoir effectuer des comparaisons, les cartes utilisées pour tenter d'obtenir un zonage sont également les mêmes, c'est-à-dire des cartes de K₂O, de CaO et de pourcentage d'argile (voir Figure 46).

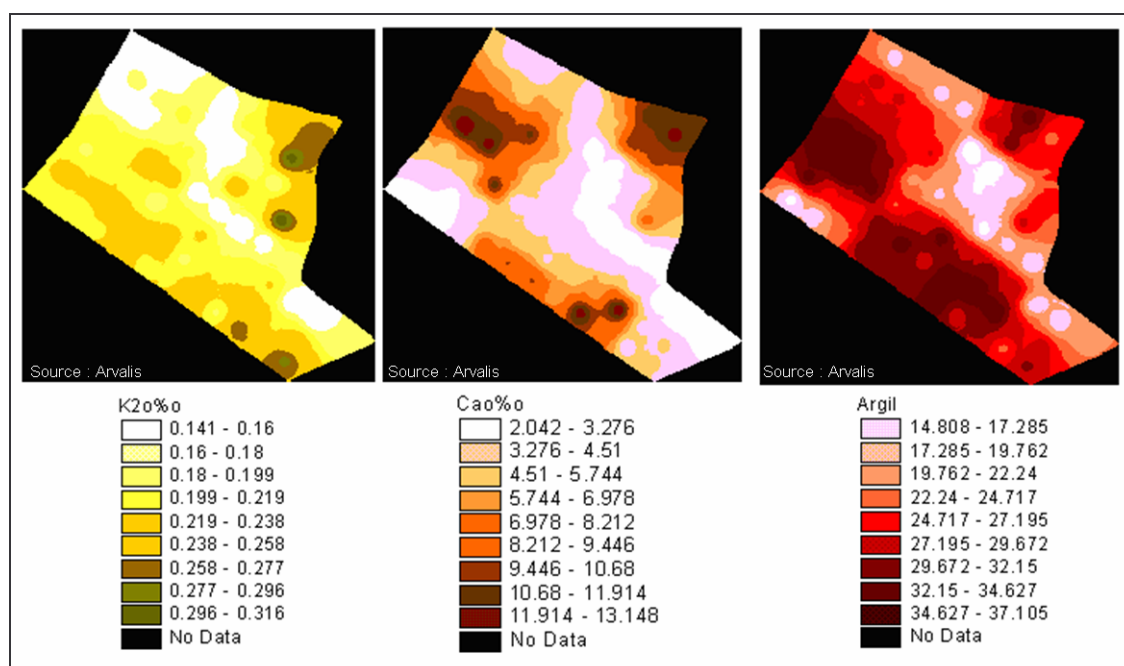


Figure 45 : Cartes de K2O, de CaO et de pourcentage d'argile sur la parcelle "Baziège"

13.2. Renseignement des paramètres de notre méthode de zonage

Conformément aux objectifs présentés en section 6.1, notre méthode de zonage est en grande partie automatisée. Les seuls paramètres à renseigner sont donc les cartes sur lesquelles fonder le zonage et les poids affectés à ces cartes. Pour cet exemple d'application, nous avons choisi de ne pas recourir à la pondération. Cela peut sembler paradoxal dans une optique de comparaison à une méthode d'A.M.C.D. ayant, elle, recours à la pondération. Mais nous voulons surtout juger ici de la ressemblance des cartes obtenues avec les deux méthodes par rapport aux renseignements de paramètres qu'elles exigent. Le fait de n'affecter aucune pondération aux cartes initiales contribue donc à réduire encore le peu de paramètres qu'exige notre méthode et donc à éloigner encore plus les logiques de renseignements de paramètres entre celle-ci et E.L.E.C.T.R.E. TRI.

Les trois cartes utilisées pour cet exemple d'application ont donc toutes reçu le même poids : 1.

De plus, même en ne définissant que trois catégories sous E.L.E.C.T.R.E. TRI comme nous l'avons fait en section 9.1.3, on ne peut non seulement pas contrôler le nombre de zones de la carte finale, mais on ne peut pas non plus savoir si ce nombre correspond ou non au nombre de zones intrinsèque aux données. Afin de nous assurer que les cartes obtenues avec les deux méthodes aient des nombres de zones comparables, nous avons décidé de retenir une vingtaine de zones lors de l'application de notre méthode.

13.3. Application de la méthode

La carte de zonage obtenue avec notre méthode est présentée en Figure 46. Pour mémoire, la carte de zonage obtenue avec E.L.E.C.T.R.E. TRI est présentée en Figure 47.

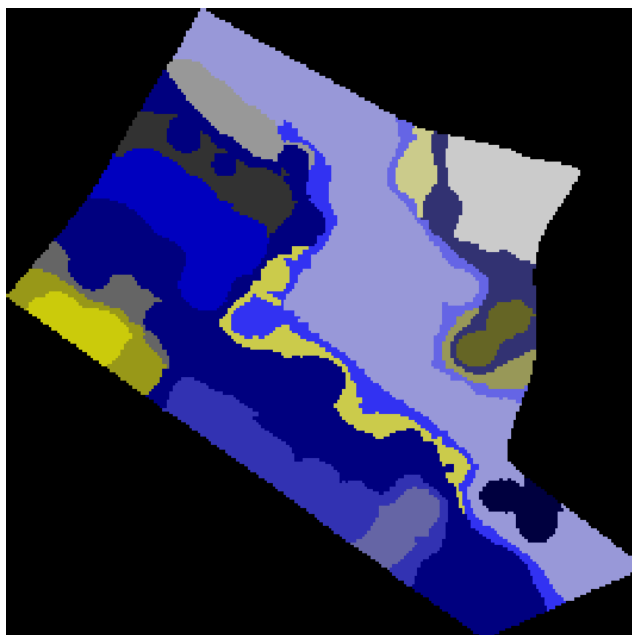


Figure 46 : Carte de zonage obtenue sur la parcelle "Baziège" avec notre méthode de zonage

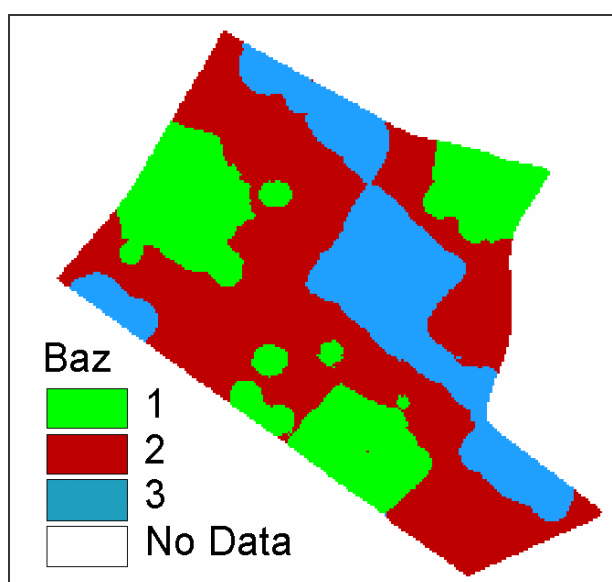


Figure 47 : Carte de zonage obtenue sur la parcelle "Baziège" avec la méthode E.L.E.C.T.R.E. TRI

Le premier constat concerne le nombre de zones des deux cartes. La carte issue de notre méthode de zonage ne compte pas moins de 19 zones. La fragmentation spatiale des catégories définies sous E.L.E.C.T.R.E. TRI fournit 11 zones. C'est la première catégorie qui est la plus fragmentée. Pourtant, plusieurs petites zones issues de cette catégorie pourraient être regroupées avec une zone plus importante voisine sans réelle perte d'information, en particulier à l'ouest et au sud de la parcelle.

Le second constat concerne la disposition des zones et la reconnaissance des tendances fortes de la parcelle. Sur la carte obtenue avec notre méthode de zonage, les zones voisines prises deux à deux ont souvent des formes de dimension et de forme assez proches, avec un effet d'emboîtement. On peut donc considérer que la grande majorité de ces zones constitue des transitions. Sur la carte obtenue avec E.L.E.C.T.R.E. TRI, nous avons vu que plusieurs petites zones de la première catégorie étaient en fait des excroissances de zones plus importantes de la même catégorie. Il existe donc une importante différence concernant les raisons d'existence des zones : zones de transition dans le premier cas, excroissances d'une zone plus importante dans le second cas. En dépit de cette différence, les tendances fortes mises en évidence sont les mêmes pour les deux zonages. Elles sont plus lisibles sur la carte obtenue avec E.L.E.C.T.R.E. TRI, celle-ci comptant moins de zones dont certaines peuvent encore être regroupées. Mais elles sont néanmoins bien identifiables sur la carte obtenue avec notre méthode de zonage. Dans le détail, certaines différences existent mais les deux cartes sont bel et bien comparables.

13.4. Conclusion

Les cartes obtenues avec E.L.E.C.T.R.E. TRI et notre méthode de zonage sont globalement comparables. C'est dans le nombre de paramètres à renseigner, et pour E.L.E.C.T.R.E. TRI dans la complexité des notions à appréhender pour ce faire, que se trouve la principale différence entre les deux méthodes. De plus, notre méthode permet de rechercher, par le biais de la C.A.H., un zonage avec un nombre de zones propice à une bonne restitution des données initiales, ce qu'E.L.E.C.T.R.E. TRI ne fait pas.

Nous avons développé notre méthode de définition d'entités homogènes sur la base d'une classification mixte à laquelle ont été intégrées plusieurs contraintes spatiales. Le principal intérêt d'une telle combinaison de techniques est l'adaptabilité de la méthode finale. Cette adaptabilité est à l'origine de l'adéquation de la méthode à nos contraintes de travail, qu'il s'agisse de la simplicité d'utilisation de la méthode ou de l'opérationnalité de la carte de zonage. A l'inverse, il est difficile de trouver un moyen réellement efficace et fiable de valider la méthode de zonage, tant les techniques spatiales et statistiques sont imbriquées.

Pour autant, les résultats fournis dans l'exemple d'application sont satisfaisants. La méthode a d'ailleurs été utilisée dans le cadre d'une étude bien plus vaste demandée par Arvalis Institut du Végétal. Cette étude sera présentée dans la partie suivante.

Cinquième partie :

Etude des configurations de données

Dans cette partie, nous allons exposer la façon dont nous avons exploité notre méthode de zonage dans le cadre d'une étude réalisée pour Arvalis Institut du Végétal. Cette étude, qui repose directement sur l'exploitation de certaines fonctionnalités offertes par notre méthode de zonage, visait à modéliser la structure spatiale de certaines informations coûteuses à partir d'autres informations l'étant moins.

Nous présenterons dans un premier temps le contexte de l'étude, puis son mode opératoire avant d'en exposer les résultats et d'en tirer des conclusions.

Enfin, nous exposerons les comparaisons qui ont été opérées entre certaines cartes de zonage obtenues lors de cette étude et celles obtenues avec le logiciel eCognition.

14. Contexte de l'étude

Aujourd'hui, beaucoup de méthodes de zonage nécessitent, pour obtenir des résultats de qualité, le recours à des données fiables et donc souvent chères à acquérir (voir section 6.2).

Nous avons vu en section 5.1.1 que les informations fournies par les analyses de sol sont à la fois primordiales pour la pratique de l'agriculture de précision mais aussi très onéreuses.

A raison de quelques dizaines d'euros par point d'analyse de sol, acquérir une information de qualité est donc d'un coût non négligeable [Fleming, Westfall, Bausch 2000]. Dillon, Mueller et Shearer considèrent que respecter l'information spatiale est nécessaire, mais que la précision requise pour les données induit une échelle fine. Les coûts d'acquisition sont alors très élevés, au point parfois de ne pas justifier les coûts supplémentaires pour atteindre l'échelle fine de gestion [Dillon, Mueller, Shearer 2003]. Cela constitue donc un des verrous socio-économiques relevés en section 4.1. Des données sol précises sont donc rarement disponibles sur les parcelles agricoles mais il est possible d'utiliser d'autres données, moins chères, sous certaines conditions [Domsch, Kaiser, Witzke, Zauer 2003] [Frogbook, Oliver, Derricourt 2003].

14.1. Objectif de l'étude

L'objectif de cette étude est de tenter d'obtenir des zonages fiables agronomiquement à partir de données relativement peu onéreuses, restituant au mieux la structure spatiale de données de référence plus chères. Pour ce faire, il est indispensable de disposer de deux jeux de données : un premier constitué de données sur lesquelles fonder les zonages ; et un second constitué de données agronomiques d'intérêt qui vont permettre de valider ou non les zonages obtenus.

14.2. Les données disponibles

L'ensemble des données utilisées dans le cadre de cette étude a été fourni par Arvalis Institut du Végétal.

14.2.1. Les données tests

Les données tests sont constituées de trois groupes de données :

- Des cartes de résistivité électrique.

- Des cartes de rendement.
- Des images satellitaires.

14.2.1.1. Les cartes de résistivité électrique

La résistivité électrique (voir section 5.1.4) a un coût de 100€ par hectare. C'est la donnée test la plus chère dont nous disposons.

Les trois voies de la résistivité électrique seront abrégées respectivement res1, res2 et res3.

14.2.1.2. Les cartes de rendement

Les semis de points utilisés dans le cadre de cette étude ont été corrigés selon la méthode développée par Mérour pour Arvalis Institut du Végétal (voir section 5.1.2.2) [Mérour 2002]. Les interpolations résultant des semis de points corrigés fournissent des cartes de rendement de meilleure qualité qu'une carte de rendement non corrigée.

Pour autant, ces cartes de rendement contiennent toujours des erreurs, au sens statistique et spatial du terme. Les rangs de collecte et les tournières, en particulier, sont encore visibles. Et le temps de montée en charge de la machine induit un léger décalage des valeurs de rendement entre deux rangs voisins moissonnés dans des directions opposées. Il a donc été décidé d'ajouter un second traitement aux données de rendement corrigées afin de tenter de réduire encore ces erreurs. Il s'agit d'effectuer un premier zonage, conservant un très grand nombre de zones (entre sept et dix), sur la base d'une carte de rendement corrigée. Chaque zone prend ensuite sa valeur moyenne de rendement. La carte de rendement moyenné ainsi obtenue est alors considérée comme une nouvelle carte de donnée test.

Tout au long de cette étude, les tests ont été réalisés en parallèle sur les cartes de rendement et sur les cartes de rendement moyenné. L'objectif de ces tests en parallèle est de vérifier l'utilité du pré-zonage moyenné des cartes de rendement corrigées. Autrement dit, pour le groupe de données tests concernant les cartes de rendement, nous disposons à la fois des cartes de rendement corrigées et des cartes de rendement moyenné.

Nous considérons les cartes de rendement comme une donnée test relativement peu chère. Si l'investissement initial est conséquent (15000€ pour l'acquisition d'un dispositif comprenant le G.P.S., le logiciel de cartographie du rendement et le capteur de rendement), chaque carte de rendement obtenue ensuite ne nécessite pas d'investissement supplémentaire.

Les cartes de rendement corrigées seront abrégées rdXX où XX sera l'année de mesure de la carte (rd99 pour la carte de rendement corrigée mesurée en 1999). Les cartes de rendement moyenné seront abrégées rdXXmoyYz où XX sera l'année de mesure de la carte et Y le nombre de zones de la carte (rd99moy8z pour la carte de rendement moyenné sur 8 zones mesurée en 1999).

14.2.1.3. Les images satellitaires

Elles proviennent de deux satellites distincts, Landsat et TERRA ASTER, en fonction des disponibilités sur les parcelles étudiées.

Les canaux retenus sur les différentes parcelles sont les canaux vert, rouge et Proche Infra-Rouge (P.I.R.), ceux-ci correspondant respectivement au bleu, au vert et au rouge d'une composition colorée "classique".

La seule différence notable entre les capteurs des deux satellites concerne la résolution des images : trente mètres pour Landsat ; quinze mètres pour TERRA ASTER.

Les images satellitaires utilisées dans le cadre de cette étude étaient disponibles gratuitement sur Internet.

Dans ce document, le canal vert sera abrégé c2, le canal rouge c3 et le canal P.I.R. c4.

14.2.2. Les données de référence

Les données de référence sont obtenues par interpolation des semis de points fournis lors d'analyses de sol, à raison de cinq à dix points par hectare, selon les parcelles. Il y a onze données de référence.

Ces analyses de sol sont réalisées sur les vingt cinq premiers centimètres du sol. Chaque analyse de sol coûte environ 23€.

On peut distinguer quatre groupes parmi les onze variables retenues :

- Les données de référence concernant la granulométrie : pourcentage d'argile (abrégé argi), pourcentage de limon fin (abrégé lifn), pourcentage de limon grossier (abrégé ligr), pourcentage de sable fin (abrégé safn) et pourcentage de sable grossier (abrégé sagr).
- Les données de référence concernant les teneurs physico-chimiques : teneur en CaO (abrégé cao), teneur en K₂O (abrégé k2o) et teneur en P₂O₅ (abrégé p2o5).
- Les données de référence concernant les pH : pH H₂O (abrégé phh2o) et pH KCl (abrégé phkcl).
- La profondeur de sol, mesurée à la tarière lors des analyses de sol (abrégé prof).

14.3. Résultats attendus

A l'issue de cette étude, nous recommandons, pour tout ou partie des données de référence agronomique en fonction des résultats, une ou plusieurs configurations de données tests (données tests à utiliser et pondération à leur appliquer) permettant d'obtenir un zonage respectant la structure spatiale de la donnée de référence concernée.

Pour une donnée de référence précise, les configurations de données recommandées pourront concerner des groupes ou des combinaisons de groupes de données tests différents. Cela permet de s'assurer que, pour cette donnée de référence, il y ait au moins une configuration de données qui corresponde aux données dont disposera l'exploitant agricole. En particulier, nous chercherons à proposer, pour chaque référence, une configuration de données n'utilisant pas la résistivité électrique, cette donnée étant la plus chère de nos données tests et donc la moins accessible aux exploitants agricoles.

Au cours de cette étude, nous tentons, pour chaque référence, de définir des configurations de données pour un petit nombre de zones (deux ou trois) et pour un grand nombre de zones (quatre à six). Au-delà, on considère que la complexité du zonage devient difficile à gérer agronomiquement.

Au final, les résultats attendus sont, pour chaque référence, une ou plusieurs configurations de données concernant des données tests suffisamment variées pour concerner la plupart des exploitants agricoles.

Bien que ceux-ci ne fassent pas directement partie des résultats attendus, plusieurs constats ont été faits concernant l'exploitabilité des différents groupes de données tests et sur l'intérêt des différentes données de référence. Ces constats sont également exposés.

14.4. Parcelles étudiées

L'étude a porté sur quatre parcelles du réseau "variabilité intraparcellaire" d'Arvalis Institut du Végétal : La Jaillière, Imbault, Magneraud et Ouzouer. L'ensemble des données utilisées dans le cadre de cette étude a donc été fourni par Arvalis Institut du Végétal.

Dans un souci de clarté et de concision, les données et résultats de cette étude de configurations de données ne sont illustrés qu'avec les cartes concernant la parcelle de "La Jaillière".

14.4.1. Parcelle "La Jaillière"

La parcelle "La Jaillière" se situe en Loire-Atlantique, dans le bocage Angevin, sur la commune de la Chapelle St Sauveur.

Trois types de sol sont rencontrés sur la parcelle :

- Sol d'apport alluvial hydromorphe sur altérite de schiste profonde d'une épaisseur supérieure à un mètre
- Sol brun faiblement hydromorphe sur schiste altéré à faible profondeur (trente à soixante centimètres)
- Sol brun hydromorphe sur matériau à dominante sableuse

Le relief est légèrement vallonné avec des pentes de 3 à 4 %.

Nous disposons des données suivantes sur la parcelle :

- Les voies 1 et 2 de la résistivité électrique (voir Figure 48). Les mesures ayant fourni des résultats aberrants pour la troisième voie, nous ne l'avons pas utilisée.

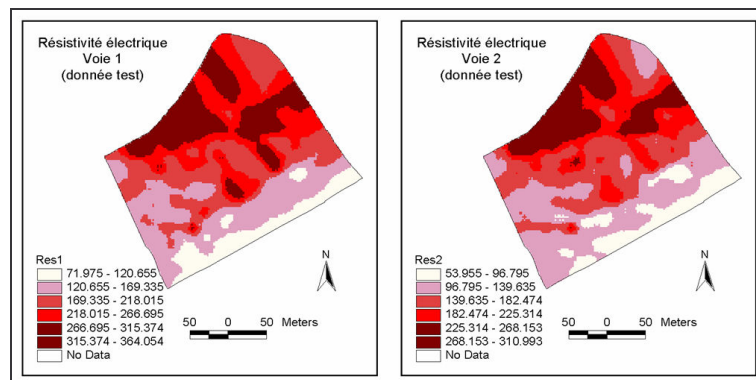


Figure 48 : Cartes des données tests de résistivité électrique sur la parcelle "La Jaillière"

- Deux cartes de rendement pour les années 1999 et 2000 et les deux cartes de rendement moyenné correspondantes (voir Figure 49).

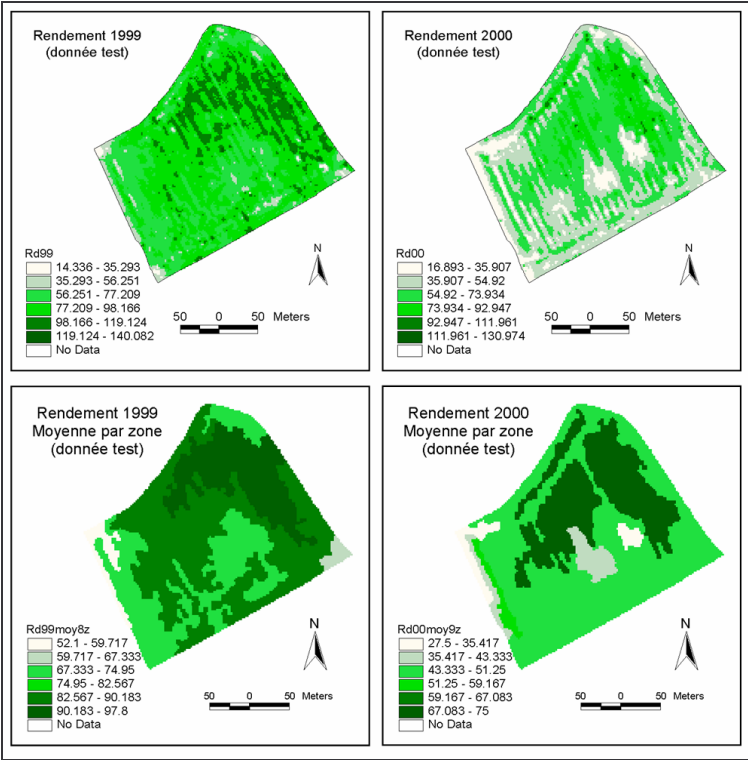


Figure 49 : Cartes des données tests de rendement et rendement moyenné sur la parcelle "La Jaillière"

➤ Les trois canaux c2, c3 et c4 d'une image Landsat à trente mètres de résolution (voir Figure 50 et Figure 51). L'image originale date du mois d'août 2000.

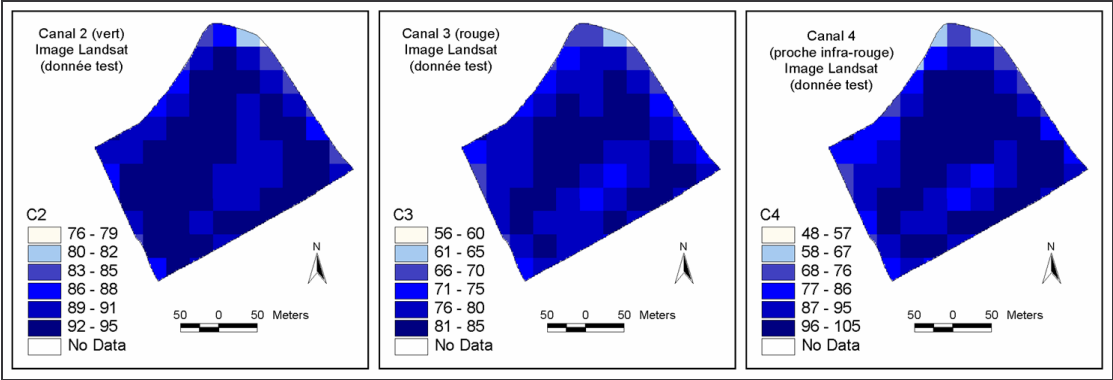


Figure 50 : Cartes des données tests des canaux de l'image satellitaire sur la parcelle "La Jaillière"

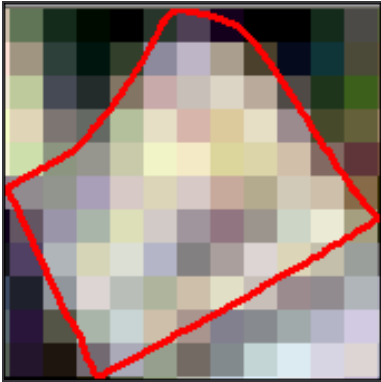


Figure 51 : Composition colorée correspondant aux trois canaux de l'image Landsat sur la parcelle "La Jaillière"

- Les onze cartes correspondant aux données de référence décrites en section 14.2.2, à savoir, les références concernant la granulométrie (voir Figure 52), les teneurs physico-chimiques (voir Figure 53), les pH (voir Figure 54) et la profondeur de sol (voir Figure 55).

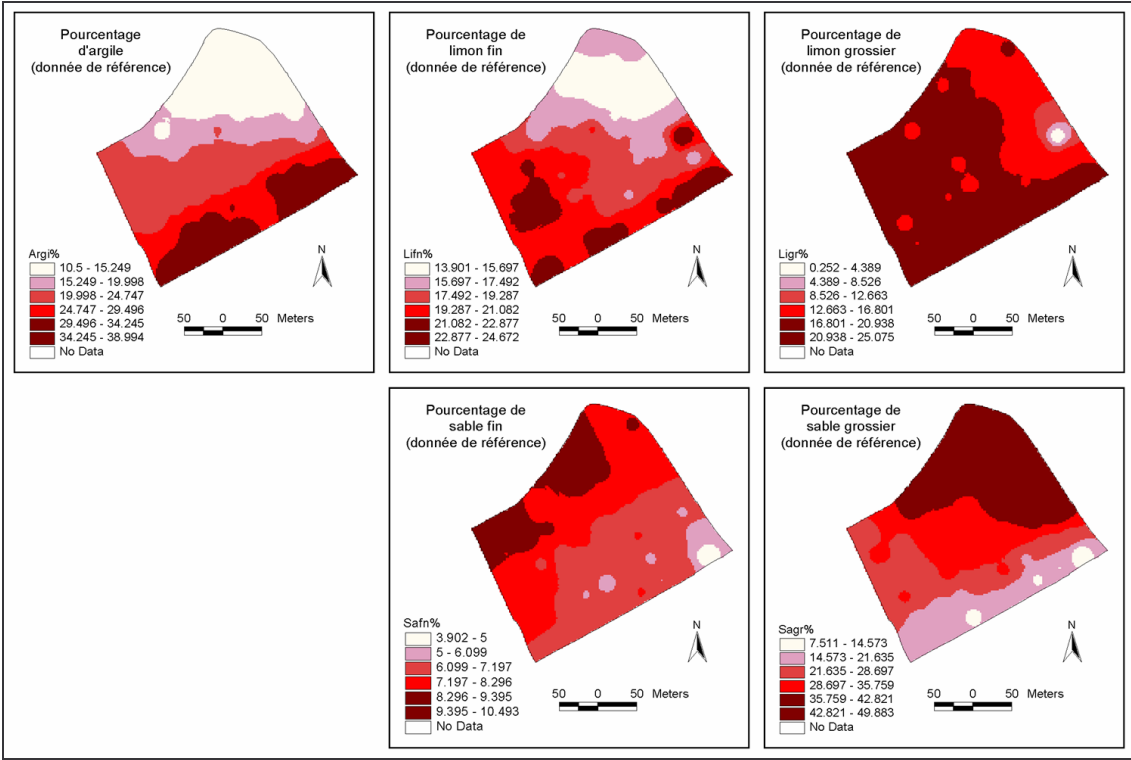


Figure 52 : Cartes des données de référence concernant la granulométrie sur la parcelle "La Jaillière"

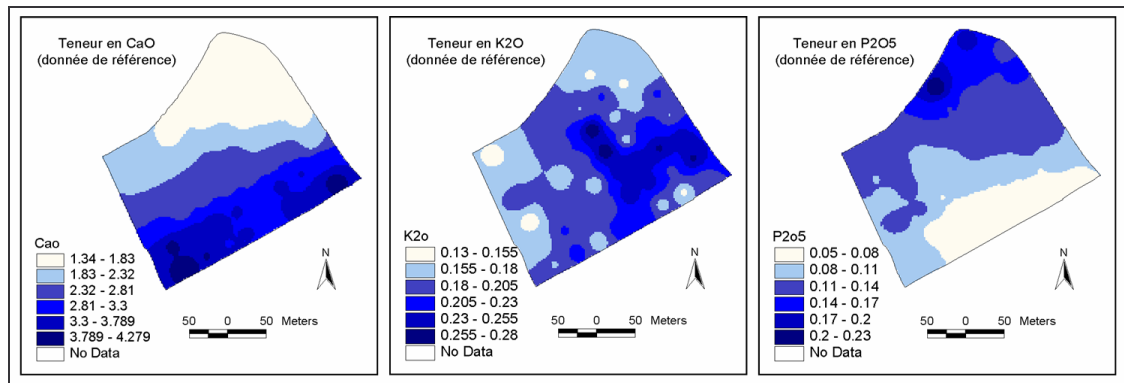


Figure 53 : Cartes des données de référence concernant les teneurs physico-chimiques sur la parcelle "La Jaillière"

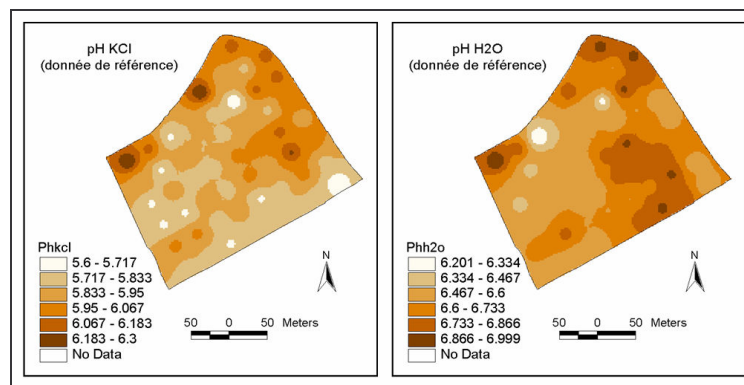


Figure 54 : Cartes des données de référence concernant les pH sur la parcelle "La Jaillière"

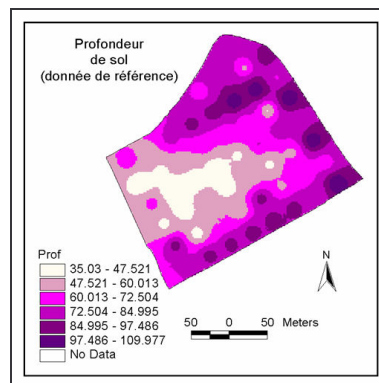


Figure 55 : Cartes des données de référence concernant la profondeur de sol sur la parcelle "La Jaillière"

14.4.2. Parcelle "Imbault"

La parcelle Imbault est localisée dans le Loiret, à proximité de Malesherbes. Située dans une zone de transition assez vaste constituant un passage progressif entre la Beauce et le Gâtinais, on y rencontre donc certaines caractéristiques respectives des deux régions :

- Le sous-sol est formé par le calcaire du Gâtinais (étage inférieur du calcaire de Beauce) très répandu dans le Gâtinais Beauceron.
- Le relief est plus marqué qu'en Beauce, les plateaux sont plus étroits.
- Les caractères généraux des sols sont assez comparables à ceux de Beauce, tout au moins sur les plateaux.

La parcelle Imbault a une texture d'argile limoneuse variant peu dans l'épaisseur du limon jusqu'au substratum. Ce dernier apparaît entre 40 et 80 centimètres de profondeur soit sous forme de calcaire dur, soit sous forme de calcaire meuble d'altération. La parcelle est aussi caractérisée par un fond d'écoulement qui la traverse de son angle nord-ouest à son angle sud-est.

Nous disposons des données suivantes sur la parcelle :

- Les voies 1, 2 et 3 de la résistivité électrique.
- Trois cartes de rendement pour les années 1999, 2000 et 2001 et les trois cartes de rendement moyenné correspondantes.
- Les deux canaux c3 et c4 d'une image Landsat à trente mètres de résolution. Le canal c2 était indisponible pour cette parcelle. L'image originale date du mois d'août 2000.
- Dix cartes correspondant aux données de référence décrites en section 14.2.2. La référence non représentée est le pH KCl.

14.4.3. Parcelle "Magneraud"

La parcelle "Magneraud" se situe en Charente-maritime au nord-est de la commune de Saint Georges du Bois.

Le sol est de nature argilo-calcaire typique du Poitou-Charentes. Sa teneur en argile est voisine de 30%. Le sous-sol est un calcaire marneux tendre plus ou moins fissuré. La réserve utile est d'environ cent millimètres. C'est un sol à ressuyage facile et à réchauffement rapide ayant dans l'ensemble une bonne fertilité.

La dénomination locale de ce type de sol est "Groie". Ce type de sol concerne 700 000 à 800 000 Hectares en Poitou-Charentes et se prolonge dans d'autres régions : Champagne Berrichonne, Plateaux de Bourgogne.

Dans son orientation nord-sud, le nord de la parcelle est plat. A partir du milieu de la parcelle intervient une faible pente orientée au sud.

Nous disposons des données suivantes sur la parcelle :

- Les voies 1, 2 et 3 de la résistivité électrique.
- Deux cartes de rendement pour les années 2000 et 2001 et les trois cartes de rendement moyenné correspondantes.
- Les trois canaux d'une image TERRA ASTER à quinze mètres de résolution renommés c2, c3 et c4. L'image originale date du mois d'août 2000.
- Les onze cartes correspondant aux données de référence décrites en section 14.2.2.

14.4.4. Parcelle "Ouzouer"

La parcelle "Ouzouer" est localisée dans la Loiret, dans la région naturelle appelée "Petite Beauce". Elle est proche du lieu dit "Poisseaux", à six kilomètres d'Epieds en Beauce.

Le trait marquant le paysage est la présence quasi uniforme de très grandes parcelles agricoles à champs ouverts. Dans toute la région, le paysage présente une morphologie légèrement ondulée, l'altitude variant entre 120 et 150 mètres. La Petite Beauce est située dans la partie occidentale de la formation des calcaires de Beauce. Ce calcaire, qui forme le substratum de la région, correspond à l'ensemble des faciès lacustres à dominante calcaire qui se sont déposés du Lutécien à l'Aquitaniens dans le sud-ouest du Bassin Parisien. Sur ce calcaire, un dépôt de limon des plateaux a été déposé par le vent et par le ruissellement. Il recouvre de façon quasi continue les calcaires aquitaniens. Son épaisseur varie de 150 centimètres à moins de 60 centimètres.

La parcelle "Ouzouer" est caractéristique de ce milieu, tant par sa profondeur de limon que son relief. Du fait de ses caractéristiques, la composante principale du drainage dans ces sols est verticale (absence de ruissellement).

Nous disposons des données suivantes sur la parcelle :

- Les voies 1, 2 et 3 de la résistivité électrique.
- Deux cartes de rendement pour les années 1999 et 2001 et les trois cartes de rendement moyenné correspondantes.
- Les trois canaux c2, c3 et c4 d'une image Landsat à trente mètres de résolution. L'image originale date du mois d'août 2000.
- Onze cartes correspondant aux données de référence décrites en section 14.2.2. Les résultats concernant la référence "pourcentage de sable grossier" ne sont pas disponibles pour la combinaison de groupes de données {résistivité électrique + rendement moyenné + image satellitaire}.

14.5. Etudes similaires déjà entreprises

14.5.1. Par Dampney, King, Lark, Wheeler, Bradley et Mayr

Dampney, King, Lark, Wheeler, Bradley et Mayr ont tenté de modéliser des informations sol à partir d'autres données sur quatre parcelles d'essai [Dampney, King, Lark, Wheeler, Bradley, Mayr 2003]. La méthode de zonage utilisée dans le cadre de cette étude s'appuyait sur une classification floue sans contrainte de contiguïté.

Ces auteurs considèrent que l'induction électromagnétique est une donnée intéressante économiquement pour identifier et cartographier les caractéristiques intraparcélaires du sol. Ils ont mis en évidence des relations significatives entre la conductivité électrique et des propriétés du sol telles que la texture, le carbone organique et les ressources en eau.

De plus, ils montrent que les zones de hauts rendements correspondent à des sols contenant beaucoup d'argile et peu de sable alors que les zones de bas rendements correspondent à des sols contenant peu d'argile et beaucoup de sable. Or, la conductivité électrique est fortement liée à la teneur en argile et à la teneur en sable.

A l'issue de leur étude, il s'avère que le rendement et la conductivité électrique apportent tous deux des informations sur des propriétés importantes du sol. Mais il demeure impossible de choisir l'une ou l'autre de ces données pour prédire correctement les propriétés du sol sur l'ensemble des parcelles d'essai. Les auteurs préconisent donc d'utiliser le rendement et la conductivité électrique simultanément.

14.5.2. Par Domsch, Kaiser, Witzke et Zauer

Domsch, Kaiser, Witzke et Zauer travaillent à partir de cartes de rendement et de conductivité électrique (méthode E.M. 38). Selon eux, lorsque des données sol précises ne sont pas disponibles, la conductivité électrique peut être utilisée pour réaliser une carte de modulation.

Ces auteurs s'appuient sur le concept de potentiel de rendement qui permet de moyenner les conditions du site en quelques zones. Le potentiel de rendement reprend des conditions naturelles (relief, réserve utile, climat...), les espèces cultivées et la stratégie de l'exploitant agricole. Cette approche peut être mise en relation avec notre propre traitement des cartes de rendement (voir section 14.2.1.2). Les cartes de potentiel de rendement servent ensuite de base pour la modulation.

Domsch, Kaiser, Witzke et Zauer ont montré qu'une faible conductivité correspondait à des zones de faibles rendements et qu'une forte conductivité correspondait à des zones de forts rendements [Domsch, Kaiser, Witzke, Zauer 2003].

14.5.3. Par Frogbook, Oliver et Derricourt

Frogbook, Oliver et Derricourt considèrent que la gestion de la variabilité intraparcellaire nécessite une information spatialisée sur les propriétés du sol. Mais pour identifier cette variabilité, il faut une densité d'échantillons très importante, ce qui entraîne des coûts d'acquisition de l'information très élevés. Ces auteurs préconisent donc de recourir à d'autres données moins chères à condition qu'elles soient corrélées avec certains paramètres du sol pertinents agronomiquement.

Leur étude, visant à mettre en évidence ces corrélations, a montré que la conductivité électrique est en relation avec l'humidité et la texture du sol. Ainsi, il existe de fortes corrélations entre la conductivité électrique et la granulométrie. Les relations entre la conductivité électrique et les propriétés du sol telles que la teneur en argile et les ressources en eau sont plus complexes et Frogbook, Oliver et Derricourt considèrent que la conductivité électrique ne peut pas se substituer à une analyse de sol pour ces informations [Frogbook, Oliver, Derricourt 2003].

15. Mode opératoire de l'étude

Le mode opératoire appliqué pour atteindre notre objectif se déroule en quatre phases :

- Pour chaque parcelle, l'analyse des données tests et des données de référence.
- Pour chaque parcelle, l'application de la méthode de zonage sur les données tests par le biais de multiples simulations.
- Pour chaque parcelle, le calcul d'indices d'hétérogénéité permettant de quantifier la restitution des structures spatiales des données de référence par les zonages obtenus lors des simulations.
- Pour l'ensemble des parcelles, l'exploitation des indices pour définir des configurations de données "types".

L'utilisation conjointe des méthodes d'analyse des données, de validation des cartes de zonage obtenues et d'interprétation des résultats permet, à l'issue des simulations, de tirer des conclusions quant aux données tests à traiter et aux poids à leur affecter pour une préconisation précise (voir Figure 56).

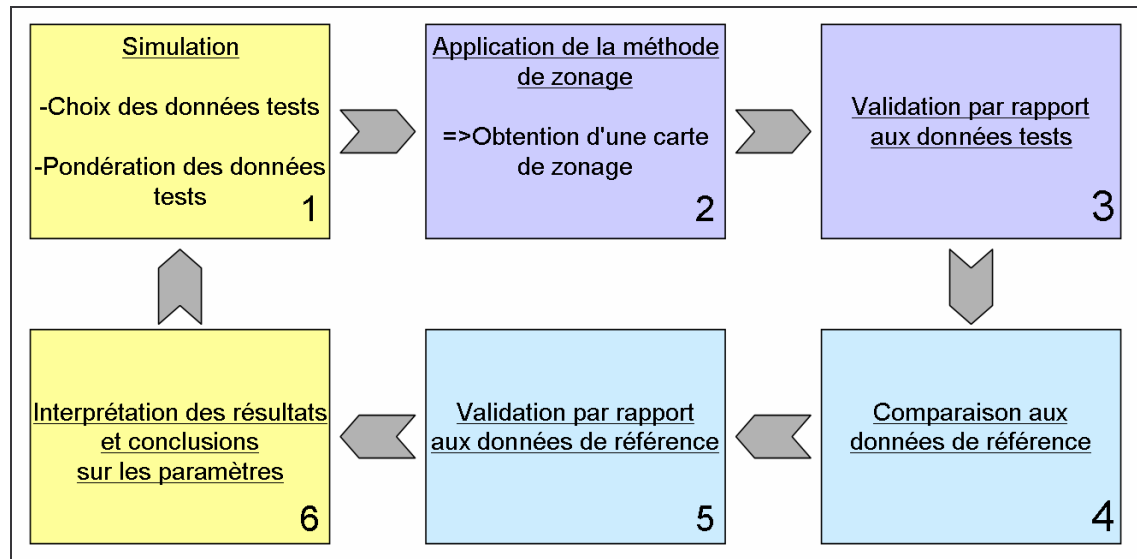


Figure 56 : Méthodologie de configuration et de validation des simulations

15.1. Analyse des données

La principale méthode d'analyse des données à laquelle nous avons eu recours pour cette étude est l'A.C.P. (voir section 11.1).

L'autocorrélation spatiale de toutes les données tests a également été calculée pour l'ensemble des parcelles afin d'estimer la facilité avec laquelle ces données nous fourniraient un zonage.

Mais nous avons également tiré parti des corrélations. Indicateurs statistiques élémentaires d'utilisation aisée [Groupe Chadule 1987], les corrélations vont principalement nous permettre de juger d'une éventuelle relation entre une donnée test et une donnée de référence. Pour chaque parcelle étudiée, nous avons calculé l'ensemble des corrélations entre chaque donnée test et chaque donnée de référence. Moins synthétique et lisible qu'une A.C.P. en début d'étude, les corrélations vont toutefois nous permettre d'affiner nos interprétations finales visant à définir les configurations de données.

On notera qu'en toute rigueur, ce sont les pourcentages de variance expliquée (r^2), mesurant l'intensité de la relation, qui aurait dû ici être utilisé à la place des corrélations (r). Cependant, ce ne sont pas les valeurs absolues des indicateurs statistiques qui nous intéressent ici, mais leurs importances relatives et leurs évolutions d'une parcelle à l'autre. Par conséquent, dans notre cas précis, l'utilisation des corrélations ou des pourcentages de variance expliquée ne modifie pas l'importance relative des relations et donc nos observations.

15.2. Application de la méthode de zonage

15.2.1. Mise en forme préalable des données

Les différentes données utilisées dans le cadre de cette étude ont été créées et/ou modifiées afin d'être conformes aux conditions de superposition des cartes initiales exposées en section 10.2.

Les semis de points concernant la résistivité électrique, le rendement et toutes les données de référence sont interpolés avec la méthode Inverse Distance Weight (I.D.W.) sous Arcview. Toutes les cartes sont paramétrées pour avoir une résolution de deux mètres ainsi qu'une hauteur et une largeur identiques.

Les cartes des différents canaux de l'image satellitaire, d'une résolution initiale de quinze ou trente mètres, sont rééchantillonnées à deux mètres avec la méthode du plus proche voisin sous Arcview.

15.2.2. Choix des cartes initiales et pondération

Selon la conception initiale de notre méthode de zonage, cette étape doit permettre d'intégrer la connaissance et les préférences de l'exploitant agricole en lui offrant la possibilité d'opérer certains choix (voir section 10.2.1). Elle est toutefois utilisée dans une autre optique dans le cadre de cette étude.

En effet, notre objectif est ici de tenter de définir des configurations de données respectant les structures spatiales de références agronomiques à partir de données peu onéreuses. Nous devons donc effectuer, pour chaque parcelle, de nombreuses simulations avec des combinaisons de données tests et de poids différentes. De multiples possibilités ont été explorées et les simulations suivantes ont donc été effectuées pour chaque parcelle :

- Chacune des cartes de résistivité électrique (voie 1, 2 et 3) prise individuellement avec un poids de 10,
- Chacune des cartes de rendement prise individuellement avec un poids de 10,
- Chacun des canaux de l'image satellitaire pris individuellement avec un poids de 10,
- L'ensemble des cartes du groupe "résistivité électrique" avec un poids variant de 1 et 10 pour chaque voie.
- L'ensemble des cartes du groupe "rendement" avec un poids variant de 1 et 10 pour chaque carte.
- L'ensemble des cartes du groupe "image satellitaire" avec un poids variant de 1 et 10, pour chaque canal.
- Les cartes des groupes "résistivité électrique" et "rendement" avec des poids variant de 1 à 10 (les cartes d'un même groupe recevant le même poids).
- L'ensemble des cartes des groupes "résistivité électrique" et "image satellitaire" avec des poids variant de 1 à 10 (les cartes d'un même groupe recevant le même poids).
- L'ensemble des cartes des groupes "rendement" et image satellitaire" avec des poids variant de 1 et 10 (les cartes d'un même groupe recevant le même poids).
- Conformément aux tests sur le rendement moyenné (voir section 14.2.1.2), l'ensemble des cartes des groupes "rendement moyenné" et "résistivité électrique" avec des poids variant de 1 à 10 (les cartes d'un même groupe recevant le même poids).
- Conformément aux tests sur le rendement moyenné (voir section 14.2.1.2), l'ensemble des cartes des groupes "rendement moyenné" et "image satellitaire" avec des poids variant entre 1 à 10 (les cartes d'un même groupe recevant le même poids).
- L'ensemble des cartes des groupes "résistivité électrique", "rendement" et "image satellitaire" avec des poids variant de 1 à 10 (les cartes d'un même groupe recevant le même poids).
- Conformément aux tests sur le rendement moyenné (voir section 14.2.1.2), l'ensemble des cartes des groupes "résistivité électrique", "rendement moyenné" et "image satellitaire" avec des poids variant de 1 à 10 (les cartes d'un même groupe recevant le même poids).

La multiplicité des simulations réalisées devait nous permettre d'éviter tout à priori sur la capacité des données tests à restituer certaines données de référence. Au total, plus de 4500 simulations¹ ont été réalisées sur chacune des parcelles, sachant qu'une même simulation peut fournir plusieurs cartes de zonage satisfaisantes (voir sections 10.2.8 et 14.3).

15.3. Validation spatiale des résultats

L'objectif de cette étude étant de modéliser au mieux la structure spatiale des données de référence à partir des données tests, nous nous sommes appuyés sur les indices d'hétérogénéité des cartes de référence après zonage (voir section 12.1.2.1).

Au moment où a été réalisée cette étude, l'appellation des I.H. (Indices d'Hétérogénéité, voir section 12.1) de cartes après zonage était Indice Global de Validité Spatiale (I.G.V.S.). Dans le reste de cette partie concernant l'étude des configurations de données, c'est donc le terme I.G.V.S. qui sera utilisé.

15.4. Mise en forme des résultats et définition des configurations de données

Le choix des configurations de données, pour chaque référence agronomique et pour chaque nombre de zones, repose sur la démarche suivante :

- Pour chaque parcelle, l'ensemble des cartes de zonage réalisées est classé en fonction de leur I.G.V.S.,
- Pour chaque parcelle, on sélectionne les cartes ayant obtenu les meilleurs I.G.V.S. pour un petit et un grand nombre de zones, et ce pour chaque groupe ou combinaison de groupes de données tests,
- Pour l'ensemble des parcelles, on réalise ensuite une moyenne des différents I.G.V.S. obtenus, pour un petit et un grand nombre de zones, pour chaque groupe ou combinaison de groupes de données tests,
- On sélectionne les groupes ou combinaisons de groupes de données tests qui fournissent les meilleurs I.G.V.S. moyens, en s'assurant d'en conserver au moins un ne concernant pas la résistivité électrique (voir section 14.3),
- On recherche d'éventuelles règles de pondérations au sein de chacun des groupes ou combinaisons de groupes de données tests retenus,
- On vérifie s'il existe des liens entre les pondérations de chacun de ces groupes ou combinaisons de groupes de données tests retenus et leurs corrélations avec la référence concernée.

Pour chaque référence, les configurations de données seront présentées de la manière suivante :

¹ La réalisation concrète de ces nombreuses simulations n'aurait pas été possible sans la mise à disposition par la M.S.H.-Alpes (Maison des Sciences de l'Homme – Alpes) de "juliette", puissante machine de calcul parallèle, en l'occurrence un S.G.I. Altix 350.

- Un tableau présente les meilleurs I.G.V.S. obtenus pour chaque groupe ou combinaison de groupes, pour un petit et un grand nombre de zones. Les pondérations des simulations retenues sont également présentées. Conformément aux modalités de pondération définies en 15.2.1, les poids sont indiqués : pour toutes les cartes lorsqu'un seul groupe de données tests est concerné (pour le groupe {résistivité électrique}, une pondération "1 4 3" signifie que la voie 1 a été pondérée à 1, la voie 2 à 4 et la voie 3 à 3) ; pour l'ensemble des cartes d'un groupe lorsque plusieurs groupes de données tests sont concernés (pour la combinaison de groupes {résistivité électrique + rendement + image satellitaire}, une pondération "1 4 3" signifie que les trois voies de la résistivité ont reçu un poids de 1, l'ensemble des cartes de rendement un poids de 4 et les trois canaux de l'image satellitaire un poids de 3). Le tableau des I.G.V.S. est dérivé de plusieurs tableaux de mise en forme des résultats qui sont présentés en section 21 en annexe.
- Pour un petit nombre de zones, description des configurations de données retenues (dont au moins une excluant la résistivité électrique, voir section 14.3), avec leurs I.G.V.S. moyens et les remarques sur d'éventuelles règles de pondérations et relations avec les corrélations.
- Pour un grand nombre de zones, description des configurations de données retenues (dont au moins une excluant la résistivité électrique, voir section 14.3), avec leurs I.G.V.S. moyens et les remarques sur d'éventuelles règles de pondérations et relations avec les corrélations.
- Quelques remarques sur la possibilité de reproduire la structure spatiale de la donnée de référence concernée avec les données tests dont nous disposons.
- Le gain d'I.G.V.S. lorsque l'on passe d'un petit à un grand nombre de zones,
- La recherche d'éventuelles règles de pondération générales, quel que soit le nombre de zones, pour une même configuration de données.

Rappelons que, dans le cadre de cette étude, nous cherchons à maximiser l'homogénéité des zones, donc à minimiser l'hétérogénéité des dites zones, donc à obtenir les I.G.V.S. les plus faibles possibles.

16. Présentation des résultats

L'ensemble des résultats présentés ici concerne les quatre parcelles étudiées.

Cependant, dans un souci de clarté, ces résultats seront illustrés uniquement pour la parcelle "La Jaillière".

De même, par souci de concision, les résultats ne seront présentés que pour trois des onze références agronomiques étudiées. Ces références ont été choisies pour leur représentativité par rapport aux résultats globaux de l'étude. Toutefois, en section 16.2.4, un tableau récapitulatif reprendra l'ensemble des I.G.V.S. obtenu pour chacune des données de référence.

16.1. Analyse des données

16.1.1. Considération sur l'hétérogénéité des parcelles étudiées

C'est le tout premier point à considérer. L'objet de cette étude est de définir des configurations de données qui vont permettre de reproduire la structure spatiale d'une donnée de référence par le biais d'un zonage, celui-ci visant à réduire l'hétérogénéité initiale de la donnée de référence.

Cette réduction de l'hétérogénéité est mesurée par l'I.G.V.S. Or il y a deux cas pour lesquels les I.G.V.S. peuvent être médiocres :

- Nous ne disposons pas des données tests adéquats pour reproduire la structure spatiale de la donnée de référence concernée sur la parcelle concernée.
- Ou alors (encore que les deux cas pourraient être rencontrés en même temps sur certaines parcelles) la parcelle est globalement déjà très homogène et le zonage ne peut donc réduire que modérément une hétérogénéité déjà faible.

Afin d'éviter toute confusion entre les deux causes possibles identifiées ici, nous avons donc examiné l'hétérogénéité des différentes données de référence sur les différentes parcelles. Ceci doit nous permettre de juger de l'hétérogénéité globale des différentes parcelles les unes par rapport aux autres.

Le Tableau 7 reprend l'ensemble des coefficients de variation (C.V.) pour chaque donnée de référence sur l'ensemble des données étudiées.

Donnée de référence	C.V. pour la parcelle "La Jaillière"	C.V. pour la parcelle "Imbault"	C.V. pour la parcelle "Magneraud"	C.V. pour la parcelle "Ouzouer"
Pourcentage d'argile	0.320	0.143	0.081	0.043
Pourcentage de limon fin	0.124	0.226	0.081	0.051
Pourcentage de limon grossier	0.140	0.201	0.124	0.035
Pourcentage de sable fin	0.127	0.168	0.091	0.410
Pourcentage de sable grossier	0.315	0.529	0.642	1.118
Teneur en CaO	0.316	0.151	0.030	0.202
Teneur en K ₂ O	1.327	1.047	0.120	2.399
Teneur en P ₂ O ₅	2.306	1.987	0.233	6.884
pH KCl	0.040	0.028	0.008	0.092
pH H ₂ O	0.035	0.042	0.007	0.078
Profondeur de sol	0.225	0.345	0.162	0.142

Tableau 7 : Coefficients de variation des onze données de référence pour chaque parcelle étudiée.

Par rapport à l'ensemble des parcelles de cette étude, les parcelles "La Jaillière" et "Imbault" constituent des parcelles globalement hétérogènes.

La parcelle "Magneraud" obtient des C.V. très faibles par rapport à ceux des autres parcelles. C'est sans aucun doute la parcelle la plus homogène traitée dans le cadre de cette étude.

On remarque que la parcelle "Ouzouer" est singulière par rapport aux autres parcelles. Pour quatre des onze références (pourcentage d'argile, de limon fin, de limon grossier et profondeur de sol), elle obtient les plus faibles C.V. Pour cinq autres (pourcentage de sable fin, de sable grossier, teneur en K₂O, en P₂O₅, pH KCl et pH H₂O), elle obtient les plus importants C.V. La parcelle "Ouzouer" est donc, par rapport à l'ensemble des parcelles étudiées ici, soit très homogène, soit très hétérogène en fonction de la référence considérée.

Il est donc possible que nous obtenions des I.G.V.S. médiocres pour de nombreuses données de référence sur la parcelle "Magneraud" et pour certaines données de référence sur la parcelle "Ouzouer".

16.1.2. Autocorrélation spatiale des parcelles

L'autocorrélation spatiale d'une carte peut permettre d'estimer la facilité avec laquelle la carte permettra d'obtenir un zonage (voir section 11.3.1). Le calcul de l'autocorrélation spatiale ne présente donc d'intérêt que pour les données tests, celles-ci seules étant utilisées avec notre méthode de zonage. Les résultats fournis pour l'indice de MORAN I sont présentés dans le Tableau 8.

Donnée test	MORAN I pour la parcelle "La Jaillière"	MORAN I pour la parcelle "Imbault"	MORAN I pour la parcelle "Magneraud"	MORAN I pour la parcelle "Ouzouer"
rd99	0.864	0.93	-	0.899
rd00	0.887	0.866	0.937	-
rd01	-	0.917	0.899	0.89
res1	0.987	0.994	0.993	0.938
res2	0.988	0.991	0.991	0.961
res3	-	0.994	0.99	0.969
c2	0.91	-	0.945	0.882
c3	0.924	0.971	0.925	0.859
c4	0.923	0.941	0.957	0.934

Tableau 8 : Indices d'autocorrélation de MORAN I des données tests sur les parcelles étudiées.

On remarque que les indices sont tous proches, voire très proches de +1, ce qui indique une forte autocorrélation spatiale. Au vu de ces résultats, il doit être possible d'obtenir aisément des zonages en s'appuyant sur les cartes concernées.

Sur les quatre parcelles étudiées, ce sont toujours les cartes de résistivité électrique qui obtiennent les meilleurs indices (aucune valeur inférieure à 0.938 avec une moyenne de 0.981). Viennent ensuite les images satellitaires (valeurs allant de 0.859 à 0.971 avec une moyenne de 0.924) et les cartes de rendement corrigées (valeurs allant de 0.864 à 0.937 avec une moyenne de 0.899).

Il faut ici rappeler que les cartes de rendement et de résistivité électrique ont été obtenues par interpolation de semis de points avec la méthode I.D.W. (Inverse Distance Weight). Une méthode d'interpolation estimant la valeur d'un point donné à partir des valeurs des points voisins, il est logique que les cartes interpolées soient sujettes à l'autocorrélation spatiale. Les indices obtenus par les cartes de rendement et de résistivité électrique n'ont donc rien de surprenant.

Toutefois, il est très peu probable que le recours à une interpolation soit le seul facteur expliquant les très forts indices d'autocorrélation spatiales obtenus. En effet, les interpolations fournissant les cartes de rendement et de résistivité électrique s'appuient sur des semis de points très denses (voir sections 5.1.2 et 5.1.4). Ainsi, ceux-ci restituent déjà assez fidèlement la structure spatiale du paramètre sur la parcelle, et donc les continuités et discontinuités influençant l'autocorrélation spatiale. C'est particulièrement vrai pour les cartes de résistivité électrique dont les semis de points initiaux sont à la fois denses et très bien géolocalisés. Ce constat reste valable dans une moindre mesure pour les cartes de rendement qui doivent être corrigées pour être exploitables et qui, même alors, restent moins fiables que les cartes de résistivité électrique. Au final, dans notre cas, les paramètres représentés par les différentes cartes considérées sont sans aucun doute bien autocorrélés spatialement, même en faisant abstraction de l'effet induit par l'interpolation.

A l'inverse, le mode d'acquisition des images satellitaires rend inutile tout recours à une interpolation et restitue fidèlement, le cas échéant, les discontinuités spatiales. Cela doit donc se ressentir sur les indices d'autocorrélation spatiale. Les images satellitaires n'obtiennent cependant pas de moins bons indices que les cartes de rendement ou de résistivité. On peut donc conclure qu'elles sont, elles aussi bien autocorrélées spatialement.

16.1.3. Les Analyses en Composantes Principales (A.C.P.)

Ces analyses ont été réalisées pour chaque parcelle afin de constater d'éventuelles attractions fortes entre certaines données tests et certaines données de référence.

La Figure 57 reprend le plan factoriel de l'A.C.P. réalisée sur les données tests et les données de référence sur la parcelle "La Jaillière". Les trois premières composantes regroupent plus de 73% de l'information totale.

Le groupe le plus net concerne les variables pH KCl, pH H₂O, teneur en K₂O et teneur en P₂O₅, quatre données de référence. Les teneurs physico-chimiques et les pH sont logiquement groupés. Un autre groupe est constitué des pourcentages d'argile, de limon fin et de limon grossier, ce qui n'a rien d'inattendu. Ce groupe intègre également la teneur en CaO. On retrouve aussi les trois canaux de l'image satellite Landsat. Les deux cartes de rendement et les deux voies de la résistivité sont placées à proximité des pourcentages de sable fin et grossier et de la profondeur de sol. Au final, il apparaît que les différentes variables des différents groupes de données tests forment des ensembles nets mais il est difficile d'établir des liens flagrants avec des données de référence, d'autant plus que les groupes de données tests sont proches du centre du diagramme alors que celui-ci regroupe pourtant près des trois quarts de l'information initiale.

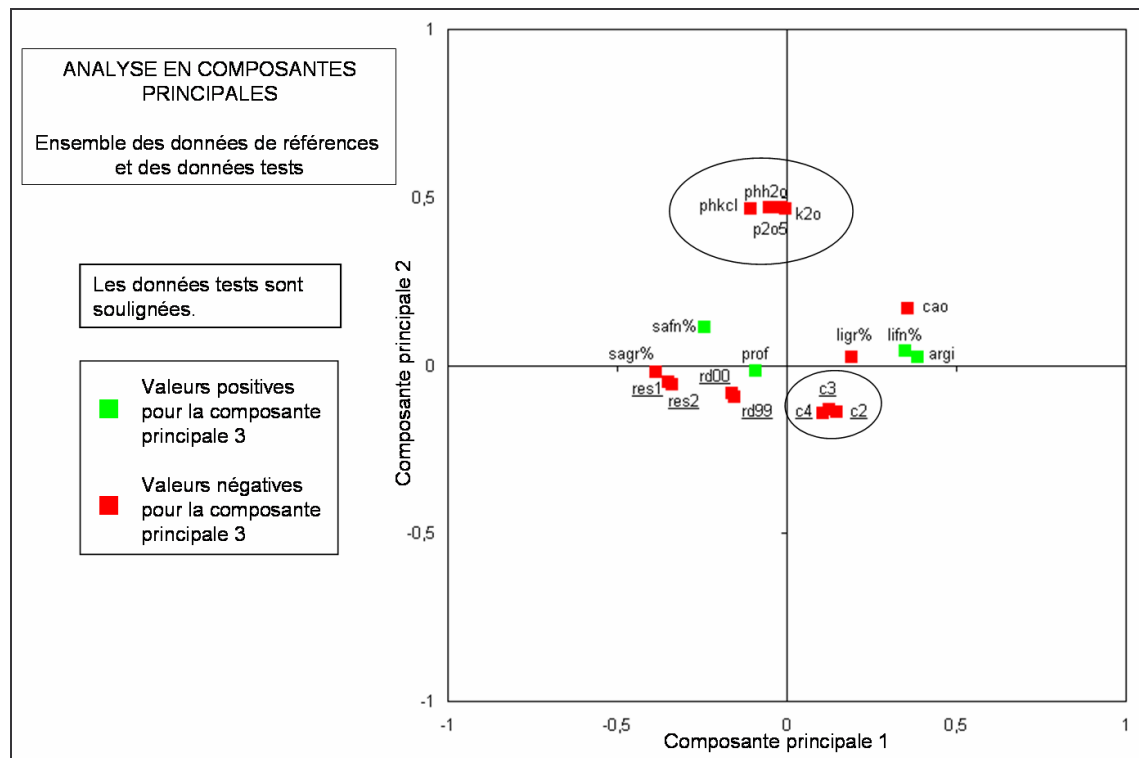


Figure 57 : Analyse en Composantes Principales des données tests et des données de référence sur la parcelle "La Jaillièrre"

Au final, sur l'ensemble des parcelles étudiées, aucune attraction récurrente entre une donnée test précise et une donnée de référence précise n'a pu être clairement établi pour l'ensemble des parcelles étudiées.

16.1.4. Les corrélations

Comme cela l'a été précisé en section 15.1, les corrélations ont été calculées entre chaque donnée test et chaque donnée de référence, et ce, pour chacune des parcelles. Les corrélations vont nous permettre d'affiner nos interprétations finales visant à définir les configurations de données.

En effet, si, pour une référence donnée, on obtient sur les quatre parcelles des pondérations comparables et que, parallèlement, les corrélations entre les données tests et la donnée de référence concernée sont aussi comparables ; alors, on peut émettre l'hypothèse que, pour cette référence, ce sont les corrélations qui permettent d'estimer la pondération à affecter aux données tests.

Bien entendu, il est impensable, même si nous devons faire ce constat pour plusieurs données de référence, de considérer cette relation pondération-corrélations comme acquise. En effet, le faible nombre de cas traité, en l'occurrence quatre parcelles, empêche toute généralisation de ce genre. Néanmoins, rien ne dit que ce type d'études ne se généralisera pas, le problème abordé ici étant un des verrous économiques majeurs de l'agriculture de précision et d'autres études du même genre ayant déjà été entreprises (voir section 14.5). Au final, si ce type de constat est effectué sur un nombre conséquent de parcelles, il pourra peut-être être généralisé.

	argi	lifn	ligr	safn	sagr	cao	k2o	p2o5	phkc l	phh2 o	prof
rd99	-0.32	-0.4	-0.28	0.02	0.39	-0.33	0.01	0.03	0.02	-0.07	0.21
rd00	-0.38	-0.42	-0.23	0.07	0.43	-0.37	0.03	0.05	0.05	-0.04	0.12
rd99 moy 8z	-0.41	-0.52	-0.32	0.03	0.49	-0.41	-0.02	0.01	0.02	-0.08	0.33
rd00 moy 9z	-0.42	-0.42	-0.29	0.08	0.46	-0.4	0	0.04	0.03	-0.06	0.21
res1	-0.8	-0.73	-0.34	0.53	0.8	-0.77	-0.02	0.07	0.16	-0.07	0.05
res2	-0.79	-0.69	-0.27	0.57	0.76	-0.74	-0.02	0.08	0.14	-0.12	0.08
c2	0.28	0.26	0.22	-0.14	-0.3	0.24	-0.02	-0.06	-0.2	-0.14	-0.17
c3	0.26	0.19	0.18	-0.2	-0.25	0.21	0.03	-0.03	-0.15	-0.1	-0.1
c4	0.21	0.12	0.13	-0.23	-0.18	0.16	0.02	-0.04	-0.16	-0.09	-0.05

Tableau 9 : Corrélations entre les données de référence et les données tests sur la parcelle "La Jaillière".

En ce qui concerne la parcelle "La Jaillière" (voir Tableau 9), plusieurs données de référence ont des corrélations significatives avec un ou plusieurs groupes de données tests. C'est particulièrement vrai pour les informations concernant la granulométrie et la teneur en K2O pour lesquelles des corrélations existent avec les trois groupes de données tests. Les références ne présentant pas de réelles corrélations avec les données tests sont la teneur en K2O, la teneur en P2O5, le groupe des pH. La profondeur de sol est faiblement corrélée avec le groupe de données tests concernant le rendement.

Les corrélations sur la parcelle "La Jaillière" sont suffisamment significatives pour que l'on puisse espérer obtenir des résultats satisfaisants pour plusieurs données de référence.

	argi	lifn	ligr	safn	sagr	cao	k2o	p2o5	phkc l	phh2 o	prof
rd99	0,29	0,43	0,3	0,33	-0,16	-0,41	-0,01	-0,03	-	-0,11	0,51
rd00	-0,12	-0,23	-0,18	-0,12	0,03	0,37	0,03	0,04	-	0,01	-0,29
rd01	0,43	0,58	0,34	0,34	-0,25	-0,55	-0,04	-0,07	-	-0,12	0,62
rd99 moy 10z	0,26	0,42	0,35	0,33	-0,14	-0,47	-0,02	-0,03	-	-0,12	0,53
rd00 moy 8z	-0,08	-0,23	-0,25	-0,07	0,06	0,44	0,05	0,58	-	0,12	-0,32
rd01 moy 9z	0,47	0,64	0,38	0,37	-0,29	-0,59	-0,05	-0,08	-	-0,13	0,66
res1	-0,63	-0,8	-0,52	-0,5	0,49	0,56	0,06	0,09	-	0,13	-0,78
res2	-0,62	-0,79	-0,41	-0,48	0,47	0,58	0,03	0,06	-	0,13	-0,74
res3	-0,62	-0,8	-0,49	-0,51	0,48	0,63	0,04	0,07	-	0,12	-0,82
c3	0,39	0,53	0,16	0,26	-0,45	-0,3	0,01	-0,03	-	-0,1	0,38

Tableau 10 : Corrélations entre les données de référence et les données tests sur la parcelle "Imbault".

Pour la parcelle "Imbault" (voir Tableau 10), comme pour la parcelle "La Jaillière", la quasi-totalité des données de référence a des corrélations significatives avec la plupart des groupes de données tests. Pour cette parcelle, la référence "profondeur de sol" est elle aussi corrélée avec les trois groupes de données tests. Les trois données de référence non corrélées aux données tests sont la teneur en K₂O, la teneur en P₂O₅ et le pH H₂O.

	argi	lifn	ligr	safn	sagr	cao	k2o	p2o5	phkc l	phh2 o	prof
rd00	-0,25	-0,15	0,13	-0,25	0,24	-0,05	-0,08	0,13	0,08	-0,1	-0,19
rd01	-0,19	-0,15	0,15	-0,24	0,12	0,05	-0,15	0,01	0,07	-0,04	-0,01
rd00 moy 10z	-0,28	-0,16	0,04	-0,25	0,29	-0,05	-0,17	0,03	0,12	-0,1	-0,17
rd01 moy 10z	-0,16	-0,13	0,15	-0,27	0,15	0,09	-0,18	-0,04	0,09	-0,01	-0,07
res1	-0,39	-0,26	-0,22	-0,2	0,13	-0,16	-0,26	0,06	0,19	-0,21	-0,37
res2	-0,29	-0,24	-0,17	-0,18	0,07	-0,05	-0,21	0,11	0,11	-0,23	-0,4
res3	-0,21	-0,29	-0,14	-0,18	-0,01	0,09	-0,22	0,1	0,05	-0,23	-0,38
c2	-0,32	-0,38	-0,22	-0,15	0,01	-0,05	-0,25	-0,03	0,2	-0,1	-0,19
c3	-0,45	-0,43	-0,32	-0,25	0,11	-0,17	-0,42	-0,24	0,31	-0,05	-0,16
c4	-0,12	-0,21	-0,04	0,03	-0,1	0,07	0,04	0,23	0,02	-0,13	-0,16

Tableau 11 : Corrélations entre les données de référence et les données tests sur la parcelle "Magneraud".

La parcelle "Magneraud" est un cas particulier (voir Tableau 11). Aucune corrélation significative n'existe pour les groupes de données de référence concernant les teneurs en CaO et en P2O5 et les pH. Mais, même pour les autres données de référence, les corrélations existantes sont bien moins prononcées que celles des parcelles "La Jaillière" et "Imbault". Seule la référence "profondeur de sol" présente des corrélations significatives avec le groupe de données "résistivité électrique".

	argi	lifn	ligr	safn	sagr	cao	k2o	p2o5	phkc l	phh2 o	prof
rd99	0,22	-0,01	0,36	0,36	0,29	-0,11	0,34	0,35	0,34	0,3	0,13
rd99 moy 9z	0,19	-0,01	0,48	0,45	0,37	-0,1	0,42	0,44	0,44	0,38	0,19
rd01	0,05	0,21	0,07	0,34	0,35	0,06	0,34	0,34	0,31	0,32	0,09
rd01 moy 10z	0,08	0,33	0,13	0,49	0,49	0,03	0,49	0,49	0,42	0,44	0,19
res1	-0,04	-0,14	-0,14	-0,02	-0,01	0,23	-0,01	-0,01	0,06	0,05	-0,3
res2	-0,06	-0,25	-0,35	-0,07	-0,04	0,42	-0,05	-0,04	0,06	0,05	-0,47
res3	-0,02	-0,23	-0,32	0,01	0,04	0,44	0,03	0,03	0,08	0,11	-0,44
c2	0,12	-0,14	0,05	0,25	0,27	0,21	0,26	0,26	0,31	0,29	-0,15
c3	0,11	-0,09	0,14	0,31	0,29	0,14	0,3	0,31	0,35	0,32	-0,01
c4	0,07	0,02	-0,38	0,02	0,09	0,3	0,03	0,03	0,05	0,09	-0,21

Tableau 12 : Corrélations entre les données de référence et les données tests sur la parcelle "Ouzouer".

La parcelle "Ouzouer" est singulière (voir Tableau 12). Des corrélations existent entre chacune des données de référence et un ou plusieurs groupes de données. Toutefois, ces corrélations sont rarement très prononcées. Ainsi, ce sont les cartes de rendement qui obtiennent les meilleures corrélations avec les données de référence concernant la granulométrie, la résistivité électrique n'étant vraiment corrélée qu'avec les pourcentages de limon fin et grossier. Concernant les teneurs physico-chimiques, il existe toujours une corrélation avec l'image satellitaire. De plus, la teneur en Cao est corrélée avec la résistivité électrique, la teneur en K2O et en P2O5 avec le rendement. Les références concernant le pH sont corrélées avec le rendement et dans une moindre mesure avec l'image satellitaire. Enfin, la profondeur de sol est bien corrélée avec la résistivité électrique.

On remarquera qu'il est possible d'établir une certaine forme de relations entre l'homogénéité des différentes parcelles (voir section 16.1.1) et les corrélations relevées sur celles-ci. Les parcelles "La Jaillière" et "Imbault", toutes deux relativement hétérogènes, obtiennent des corrélations significatives pour les données de référence concernant la granulométrie, la teneur en CaO et, dans une moindre mesure, la profondeur de sol. La parcelle "Magneraud", qui est la plus homogène, n'obtient de corrélations significatives que pour la référence "profondeur de sol". Enfin, la parcelle "Ouzouer", qui est soit la plus homogène soit la plus hétérogène en fonction des références, obtient des corrélations pour chacune des données de référence, mais celles-ci concernent souvent un nombre restreint de données tests.

Il est difficile à ce stade de l'étude de pouvoir s'assurer de ces relations, ni même de les expliquer. La poursuite de l'étude sur d'autres parcelles devra permettre d'entériner ou non les relations ici mises en évidence.

On peut avancer, avec beaucoup de prudence, une hypothèse sur les éventuelles raisons des relations entre l'homogénéité des parcelles et les corrélations. Il existe en effet certaines analogies entre les opérations mises en œuvre dans les calculs des C.V. et des corrélations, même si les premiers concernent une variable et les secondes deux. Dans les deux cas, l'écart-type et la moyenne interviennent et l'indicateur s'appuie sur un quotient. Le caractère généralisable de cette piste reste toutefois à démontrer.

16.2. Présentation des configurations de données

16.2.1. Pour la référence "pourcentage d'argile"

16.2.1.1. Tableau des I.G.V.S.

Concernant la référence "pourcentage d'argile", les simulations retenues pour chaque parcelle, pour chaque groupe ou combinaison de groupes de données tests sont présentées dans le Tableau 13.

Groupe de données	Parcelle	2-3 zones		4-6 zones	
		IGVS	Pondération	IGVS	Pondération
Résistivité électrique	La Jaillière	0.79	4res1 5res2	0.64	5res1 8res2
	Imbault		6 4 8		5 3 2
	Magneraud		1 6 8		10 10 5
	Ouzouer		8 2 5		8 1 8
Rendement	La Jaillière	0.84	1 6	0.75	4 9
	Imbault		9 8 10		5 3 10
	Magneraud		7 8		7 8
	Ouzouer		10rd99		1 10
Image satellitaire	La Jaillière	0.91	1 6 2	0.85	1 6 1
	Imbault		10c3		2c3 4c4
	Magneraud		6 3 1		2 8 1
	Ouzouer		1 10 8		1 4 3
Résistivité électrique + rendement	La Jaillière	0.78	5 4	0.69	7 7
	Imbault		4 4		8 7
	Magneraud		2 10		2 7
	Ouzouer		8 8		6 1
Résistivité électrique + rendement moyenné	La Jaillière	0.8	9 3	0.68	5 1
	Imbault		2 1		10 8
	Magneraud		4 10		3 8
	Ouzouer		1 8		1 10
Résistivité électrique + image satellitaire	La Jaillière	0.79	9 3	0.66	7 4
	Imbault		3 7		4 3
	Magneraud		1 10		4 6
	Ouzouer		7 1		8 1
Rendement + image satellitaire	La Jaillière	0.83	8 7	0.76	8 4
	Imbault		3 7		6 4
	Magneraud		7 2		3 9
	Ouzouer		2 6		3 6
Rendement moyenné + image satellitaire	La Jaillière	0.84	10 7	0.77	5 2
	Imbault		3 9		3 10
	Magneraud		7 2		7 2
	Ouzouer		3 7		3 8
Résistivité électrique + rendement + image satellitaire	La Jaillière	0.76	9 4 10	0.66	8 7 3
	Imbault		7 1 6		9 3 9
	Magneraud		1 1 2		1 8 2
	Ouzouer		1 3 8		5 1 1
Résistivité électrique + rendement moyenné + image satellitaire	La Jaillière	0.74	7 3 4	0.66	9 6 1
	Imbault		7 2 8		8 5 1
	Magneraud		1 10 10		1 5 10
	Ouzouer		10 7 5		10 7 4

Tableau 13 : Simulations retenues pour la référence "pourcentage d'argile"

16.2.1.2. Pour un petit nombre de zones

Les configurations de données sont présentées dans le Tableau 14, la carte de zonage correspondant à la meilleure configuration de données sur la parcelle "La Jaillière" est présentée dans la Figure 58.

La référence "pourcentage d'argile" est une information qu'il est possible de représenter correctement pour un petit nombre de zones. Avec 0.74 pour la meilleure configuration de données {résistivité électrique + rendement moyenné + image satellitaire}, la réduction de l'hétérogénéité n'est pas négligeable. Qui plus est, cette configuration de données {résistivité électrique + rendement moyenné + image satellitaire} obtient sur chaque parcelle le premier ou le second meilleur I.G.V.S.

On note toutefois qu'en l'absence de la donnée test "résistivité électrique", l'I.G.V.S. moyen se dégrade et remonte à 0,83 pour la configuration de données {rendement + image satellitaire}. Cela s'explique par les très bonnes corrélations entre la résistivité électrique et le pourcentage d'argile (sauf sur la parcelle "Ouzouer" pour les raisons identifiées en section 16.1.4), même s'il est impossible de mettre clairement en évidence des relations entre les pondérations et les corrélations lorsque plusieurs groupes de données tests sont combinés. Autrement dit, si l'exploitant agricole ne possède pas la résistivité électrique, le zonage à petit nombre de zones perd beaucoup de son intérêt car l'I.G.V.S. est alors mauvais.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement moyenné + image satellitaire	0.74	La résistivité est plus fortement pondérée sauf sur la parcelle "Magneraud".	Aucune
Rendement + image satellitaire	0.83	Aucune	Aucune

Tableau 14 : Configurations de données recommandées pour la référence "Pourcentage d'argile" pour un petit nombre de zones

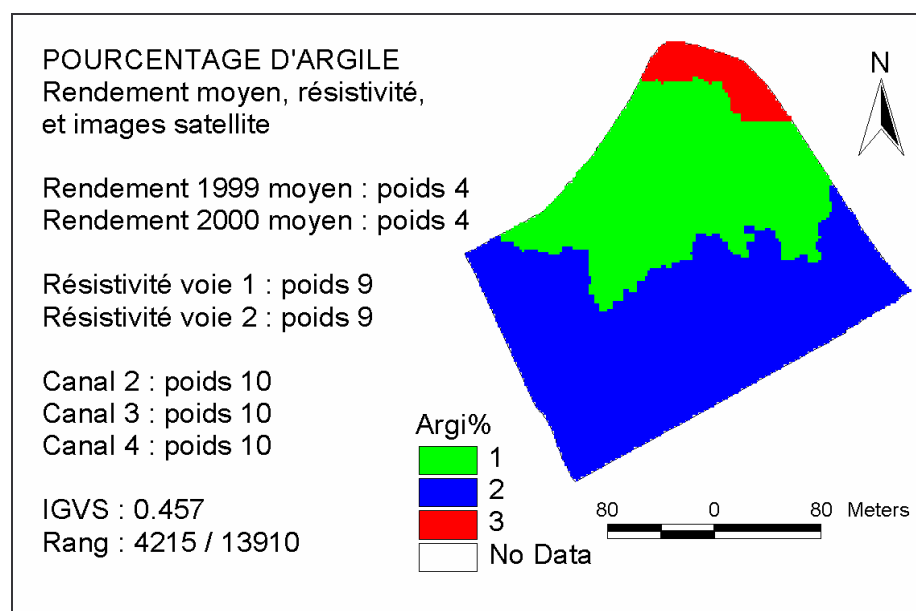


Figure 58 : Zonage correspondant à la meilleure configuration de données pour la référence "pourcentage d'argile" pour un petit nombre de zones sur la parcelle "La Jaillière"

16.2.1.3. Pour un grand nombre de zones

Les configurations de données sont présentées dans le Tableau 15, la carte de zonage correspondant à la meilleure configuration de données sur la parcelle "La Jaillière" est présentée dans la Figure 59.

En ce qui concerne la première configuration de donnée recommandée, à savoir la résistivité électrique, on notera que cette donnée obtient sur l'ensemble des parcelles le premier ou le second meilleur I.G.V.S. pour un grand nombre de zones. Il est toutefois impossible de dégager des règles de pondération type et aucune relation claire n'a pu être établie entre les pondérations et les corrélations des différentes voies. Ces bons résultats s'expliquent par les corrélations.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique	0.64	Aucune	Aucune
Résistivité électrique + rendement moyenné + image satellitaire	0.66	Pondérer fortement la résistivité électrique, moyennement le rendement moyenné et faiblement l'image satellitaire	Aucune
Rendement	0.75	Aucune	Aucune

Tableau 15 : Configurations de données recommandées pour la référence "pourcentage d'argile" pour un grand nombre de zones

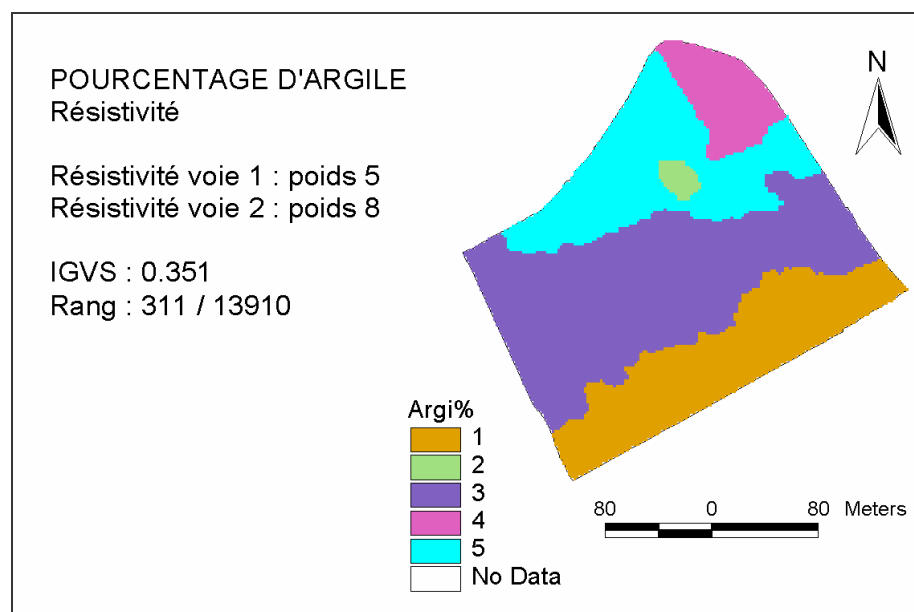


Figure 59 : Zonage correspondant à la meilleure configuration de données pour la référence "pourcentage d'argile" pour un grand nombre de zones sur la parcelle "La Jaillière"

16.2.1.4. Synthèse

Il est relativement aisé de reproduire la structure spatiale de la référence "pourcentage d'argile" avec les données tests dont nous disposons. Quelque soit le nombre de zones concerné, la perte d'hétérogénéité est conséquente, en particulier pour les cartes avec un grand nombre de zones. Toutefois, ces résultats de qualité restent liés aux fortes corrélations entre le pourcentage d'argile et la résistivité électrique. En effet, dès que cette information n'est plus prise en compte, les I.G.V.S. augmentent sensiblement.

On notera que les meilleures configurations de données pour un petit et un grand nombre de zones obtiennent toujours les premiers ou seconds meilleurs I.G.V.S. sur l'ensemble des parcelles. Toutefois, une seule règle de pondérations a pu être mise en évidence. Elle concerne le groupe de données tests {résistivité électrique + rendement moyenné + image satellitaire} pour une carte de zonage à grand nombre de zones. Cette règle nous incite à pondérer plus fortement la résistivité électrique, moyennement le rendement moyenné et faiblement l'image satellitaire. Il n'existe aucune relation entre les pondérations des différentes configurations de données retenues et les corrélations.

La configuration de données la mieux adaptée pour la référence "pourcentage d'argile" est {résistivité électrique + rendement moyenné + image satellitaire}. Elle fait partie des meilleures configurations de données quelque soit le nombre de zones considéré. De plus, des règles de pondération ont pu être mises en évidence. La résistivité électrique, en particulier, est plus fortement pondérée que les autres données tests.

16.2.2. Pour la référence "teneur en P2O5"

16.2.2.1. Tableau des I.G.V.S.

Concernant la référence "teneur en P2O5", les simulations retenues pour chaque parcelle, pour chaque groupe ou combinaison de groupes de données tests sont présentées dans le Tableau 16.

Groupe de données	Parcelle	2-3 zones		4-6 zones	
		IGVS	Pondération	IGVS	Pondération
Résistivité électrique	La Jaillière	0.77	10res2	0.65	3res1 6res2
	Imbault		6 3 5		1 3 5
	Magneraud		1 4 3		9 4 7
	Ouzouer		8 5 9		2 10 2
Rendement	La Jaillière	0.83	1 6	0.74	4 9
	Imbault		2 5 3		1 1 9
	Magneraud		1 1		3 5
	Ouzouer		2 4		1 5
Image satellitaire	La Jaillière	0.91	1 6 2	0.84	1 9 7
	Imbault		10c3		10c3
	Magneraud		5 2 7		1 9 4
	Ouzouer		1 8 1		9 8 9
Résistivité électrique + rendement	La Jaillière	0.8	3 1	0.67	9 1
	Imbault		6 6		7 1
	Magneraud		3 10		3 10
	Ouzouer		4 4		6 5
Résistivité électrique + rendement moyenné	La Jaillière	0.85	6 2	0.69	5 1
	Imbault		2 1		10 1
	Magneraud		4 10		10 8
	Ouzouer		9 1		1 8
Résistivité électrique + image satellitaire	La Jaillière	0.77	10 2	0.64	9 2
	Imbault		3 7		7 7
	Magneraud		7 6		8 1
	Ouzouer		7 1		5 10
Rendement + image satellitaire	La Jaillière	0.82	8 7	0.74	7 2
	Imbault		8 4		8 4
	Magneraud		9 3		6 1
	Ouzouer		4 8		2 10
Rendement moyenné + image satellitaire	La Jaillière	0.83	10 7	0.76	4 2
	Imbault		3 9		7 1
	Magneraud		3 5		5 7
	Ouzouer		1 3		7 6
Résistivité électrique + rendement + image satellitaire	La Jaillière	0.74	7 3 4	0.66	5 2 1
	Imbault		4 2 4		4 2 6
	Magneraud		9 1 5		6 1 4
	Ouzouer		6 7 8		3 1 4
Résistivité électrique + rendement moyenné + image satellitaire	La Jaillière	0.75	7 2 2	0.67	10 7 1
	Imbault		8 3 5		9 1 9
	Magneraud		1 5 6		3 4 10
	Ouzouer		5 7 7		5 1 9

Tableau 16 : Simulations retenues pour la référence "teneur en P2O5"

16.2.2.2. Pour un petit nombre de zones

Les configurations de données sont présentées dans le Tableau 17, la carte de zonage correspondant à la meilleure configuration de données sur la parcelle "La Jaillière" est présentée dans la Figure 60.

Les I.G.V.S. obtenus sont médiocres, mais néanmoins satisfaisants pour un petit nombre de zones. C'est la résistivité électrique qui permet, ici encore, d'obtenir ces résultats satisfaisants. En effet, si l'on supprime ce groupe de données, la meilleure configuration de données est alors {rendement + image satellitaire} avec un mauvais I.G.V.S. de 0.82.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.74	Aucune	Aucune
Rendement + image satellitaire	0.82	Aucune	Aucune sauf pour la parcelle "Magneraud"

Tableau 17 : Configurations de données recommandées pour la référence "teneur en P2O5" pour un petit nombre de zones

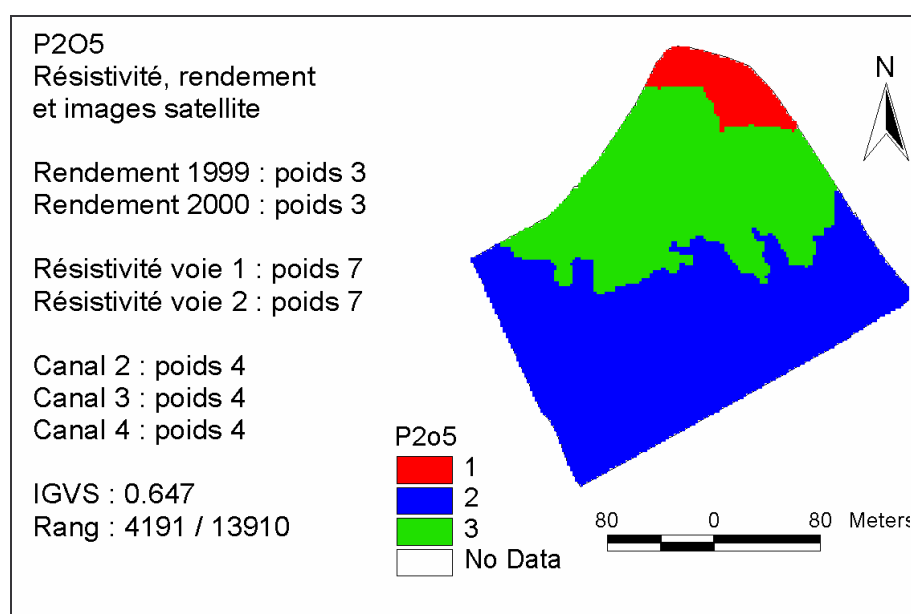


Figure 60 : Zonage correspondant à la meilleure configuration de données pour la référence "teneur en P2O5" pour un petit nombre de zones sur la parcelle "La Jaillière"

16.2.2.3. Pour un grand nombre de zones

Les configurations de données sont présentées dans le Tableau 18, la carte de zonage correspondant à la meilleure configuration de données sur la parcelle "La Jaillière" est présentée dans la Figure 61.

Ici encore, la résistivité électrique est une composante incontournable pour obtenir des cartes de zonages de bonne qualité. Avec 0.64, la configuration de données {résistivité électrique + image satellitaire} obtient un bon I.G.V.S. La première configuration excluant la résistivité électrique n'obtient qu'un I.G.V.S. médiocre de 0.74.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + image satellitaire	0.64	Aucune	Quelques relations concernant les parcelles "Magneraud" et "Ouzouer"
Rendement	0.74	Aucune	Aucune

Tableau 18 : Configurations de données recommandées pour la référence "teneur en P2O5" pour un grand nombre de zones

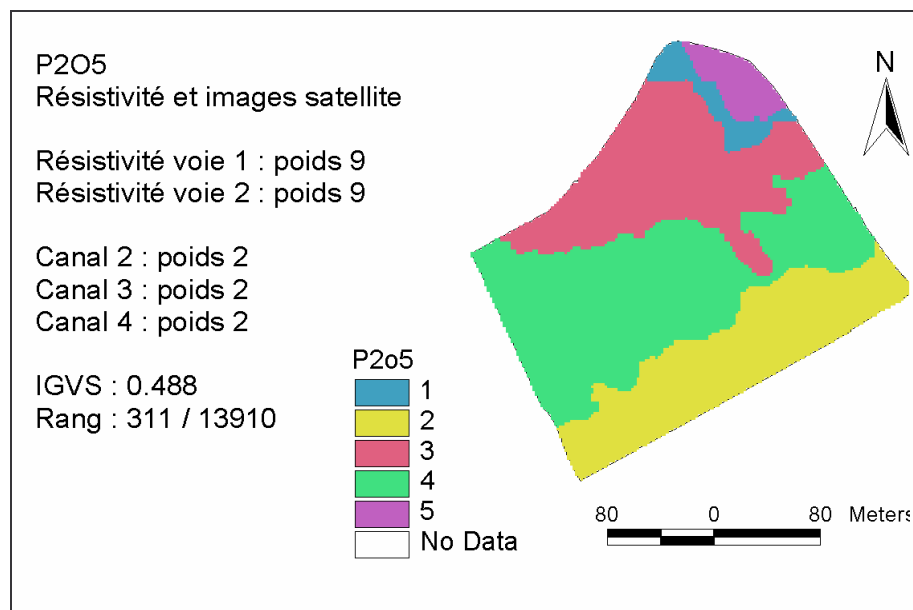


Figure 61 : Zonage correspondant à la meilleure configuration de données pour la référence "teneur en P2O5" pour un grand nombre de zones sur la parcelle "La Jaillière"

16.2.2.4. Synthèse

La référence "teneur en P2O5" est une référence qu'il est possible de modéliser correctement. Toutefois, pour obtenir une carte de zonage à la fois exploitable et de bonne qualité, disposer de la résistivité électrique est indispensable. Pourtant, il n'existe aucune corrélation significative entre ce groupe de données tests et la donnée de référence. Mais cela est globalement vrai aussi pour les autres données tests, à l'exception du rendement (moyenné ou non) pour la parcelle "Ouzouer".

La référence "teneur en P2O5" bénéficie d'une forte baisse des I.G.V.S. entre les cartes à petit nombre de zones et les cartes à grand nombre de zones. Cette baisse est toutefois moins nette pour les configurations de données excluant la résistivité électrique.

La configuration de données la mieux adaptée pour la référence "teneur en P2O5" est {résistivité électrique + rendement + image satellitaire}. Elle obtient le meilleur I.G.V.S. quelque soit le nombre de zones retenu.

16.2.3. Pour la référence "pH H2O"

16.2.3.1. Tableau des I.G.V.S.

Concernant la référence "pH H2O", les simulations retenues pour chaque parcelle pour chaque groupe ou combinaison de groupes de données tests sont présentées dans le Tableau 19.

Groupe de données	Parcelle	2-3 zones		4-6 zones	
		IGVS	Pondération	IGVS	Pondération
Résistivité électrique	La Jaillière	0.91	10res2	0.83	7 res2 8res2
	Imbault		10 1 1		10 2 5
	Magneraud		1 4 3		9 1 5
	Ouzouer		8 5 9		8 1 4
Rendement	La Jaillière	0.9	4 6	0.85	10rd00
	Imbault		9 5 8		10 2 5
	Magneraud		7 8		1 5
	Ouzouer		3 7		1 3
Image satellitaire	La Jaillière	0.91	5 1 1	0.86	10 3 2
	Imbault		10c3		10c3
	Magneraud		5 3 2		3 6 7
	Ouzouer		1 10 7		1 4 9
Résistivité électrique + rendement	La Jaillière	0.92	2 4	0.87	4 5
	Imbault		6 6		7 1
	Magneraud		4 6		7 2
	Ouzouer		8 8		6 5
Résistivité électrique + rendement moyenné	La Jaillière	0.93	2 2	0.86	9 2
	Imbault		2 9		10 8
	Magneraud		4 10		3 10
	Ouzouer		9 1		3 1
Résistivité électrique + image satellitaire	La Jaillière	0.9	4 4	0.83	10 2
	Imbault		3 7		5 7
	Magneraud		3 4		5 6
	Ouzouer		7 1		3 3
Rendement + image satellitaire	La Jaillière	0.89	5 4	0.84	6 6
	Imbault		6 2		5 1
	Magneraud		7 1		9 2
	Ouzouer		6 3		2 10
Rendement moyenné + image satellitaire	La Jaillière	0.92	7 9	0.85	7 6
	Imbault		3 9		4 1
	Magneraud		3 4		3 4
	Ouzouer		1 3		4 1
Résistivité électrique + rendement + image satellitaire	La Jaillière	0.88	2 1 3	0.82	8 7 3
	Imbault		6 2 9		1 9 2
	Magneraud		1 1 2		3 1 5
	Ouzouer		8 1 10		6 6 3
Résistivité électrique + rendement moyenné + image satellitaire	La Jaillière	0.89	6 3 10	0.83	1 10 7
	Imbault		5 2 10		7 1 7
	Magneraud		4 1 7		3 1 4
	Ouzouer		8 1 7		10 1 1

Tableau 19 : Simulations retenues pour la référence "pH H2O"

16.2.3.2. Pour un petit nombre de zones

Les configurations de données sont présentées dans le Tableau 20, la carte de zonage correspondant à la meilleure configuration de données sur la parcelle "La Jaillière" est présentée dans la Figure 62.

Le meilleur résultat est obtenu par la configuration de données {résistivité électrique + rendement + image satellitaire}. L'I.G.V.S. correspondant est toutefois mauvais puisqu'il atteint 0.88.

Pour un petit nombre de zones, aucune des données tests dont nous disposons n'est en mesure d'approcher la structure spatiale de la carte de référence.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.88	La résistivité électrique et l'image satellitaire ont des poids comparables. Le rendement est toujours plus faiblement pondéré que les deux autres groupes de données tests.	Aucune
Rendement + image satellitaire	0.89	Le rendement est toujours plus fortement pondéré que l'image satellitaire.	Aucune

Tableau 20 : Configurations de données recommandées pour la référence "pH H₂O" pour un petit nombre de zones

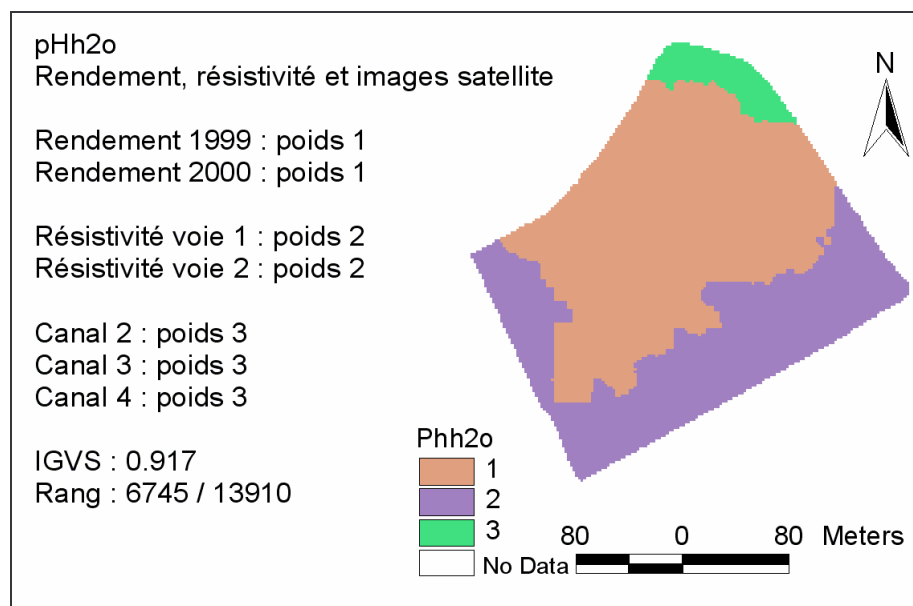


Figure 62 : Zonage correspondant à la meilleure configuration de données pour la référence "pH H₂O" pour un petit nombre de zones sur la parcelle "La Jaillière"

16.2.3.3. Pour un grand nombre de zones

Les configurations de données sont présentées dans le Tableau 21, la carte de zonage correspondant à la meilleure configuration de données sur la parcelle "La Jaillière" est présentée dans la Figure 63.

Comme pour les cartes de zonage à petit nombre de zones, les résultats sont mauvais, aucune configuration de données ne permettant de réduire significativement les I.G.V.S.

Des règles de pondération ont été mises en évidence pour le groupe de données tests concernant la résistivité électrique.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique+ rendement + image satellitaire	0.82	Aucune	Aucune
Résistivité électrique	0.83	La voie 1 doit être pondérée fortement, la voie 2 faiblement, la voie 3 moyennement.	Aucune
Rendement + image satellitaire	0.84	Aucune	Aucune

Tableau 21 : Configurations de données recommandées pour la référence "pH H₂O" pour un grand nombre de zones

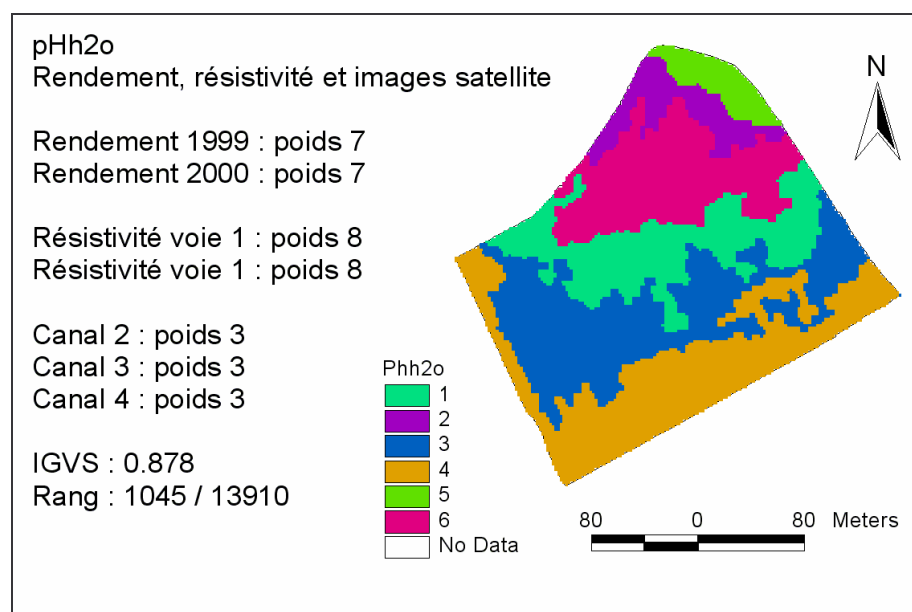


Figure 63 : Zonage correspondant à la meilleure configuration de données pour la référence "pH H₂O" pour un grand nombre de zones sur la parcelle "La Jaillière"

16.2.3.4. Synthèse

Les données tests dont nous disposons ne nous permettent pas de modéliser, même médiocrement, la structure spatiale de la donnée de référence "pH H₂O". Même les cartes de zonage à grand nombre de zones obtiennent de mauvais I.G.V.S. En effet, la baisse des I.G.V.S. en passant d'un petit à un grand nombre de zones est quasiment nulle.

La configuration de données la mieux adaptée pour la référence "pH H₂O" est {résistivité électrique + rendement + image satellitaire}. Elle obtient le meilleur I.G.V.S. quelque soit le nombre de zones retenu. Des règles de pondération existent pour les zonages à petit nombre de zones.

D'autres explications à ces mauvais résultats seront fournies en section 17.1.3.

16.2.4. Résultats obtenus pour les autres données de référence

Les tableaux suivants reprennent les configurations de données retenues pour les références agronomiques étudiées mais non présentées dans cette section.

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.72	Aucune	Aucune
Résistivité électrique + rendement	0.75	Pondérer de manière assez proche les deux groupes (pas plus de 2 de différence sur une échelle entre 1 et 10)	Aucune
Rendement	0,79	Aucune	Aucune

Tableau 22 : Configurations de données recommandées pour la référence "Pourcentage de limon fin" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.63	Aucune	Aucune
Rendement + image satellitaire	0.72	Les poids entre les deux groupes doivent être similaires.	Les pondérations suivent globalement les corrélations, sauf pour la parcelle "Ouzouer".

Tableau 23 : Configurations de données recommandées pour la référence "pourcentage de limon fin" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Rendement + image satellitaire	0.85	Aucune	Aucune
{résistivité électrique + rendement + image satellitaire}	0.86	Aucune	Aucune

Tableau 24 : Configurations de données recommandées pour la référence "Pourcentage de limon grossier" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.76	La résistivité électrique doit être plus fortement pondérée que les deux autres groupes de données tests. Cela n'est toutefois pas valable sur la parcelle "Magneraud".	Aucune
Rendement	0.79	Aucune	Aucune

Tableau 25 : Configurations de données recommandées pour la référence "pourcentage de limon grossier" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Rendement + image satellitaire	0.82	Aucune	Aucune
Résistivité électrique + rendement + image satellitaire	0.84	Aucune	Aucune

Tableau 26 : Configurations de données recommandées pour la référence "Pourcentage de sable fin" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + image satellitaire	0.76	Les poids sont à peu près les mêmes entre les deux groupes de données (pas plus de 4 de différence sur une échelle de 1 à 10).	Aucune
Rendement + image satellitaire	0.76	Les poids des deux groupes de données sont relativement proches sauf sur la parcelle "La Jaillière", avec une différence de 5 sur une échelle de 1 à 10.	Aucune
Résistivité électrique + rendement + image satellitaire	0.76	Aucune	Aucune sauf sur la parcelle "Magneraud"

Tableau 27 : Configurations de données recommandées pour la référence "pourcentage de sable fin" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement moyenné+ image satellitaire	0.71	Aucune	Aucune
Résistivité électrique + image satellitaire	0.75	Les poids des deux groupes de données sont proches sauf pour la parcelle "Ouzouer".	Aucune sauf pour la parcelle "Magneraud"
Rendement	0.82	Aucune	Aucune

Tableau 28 : Configurations de données recommandées pour la référence "Pourcentage de sable grossier" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique+ rendement moyenné + image satellitaire	0.62	Aucune	Aucune sauf pour la parcelle "La Jaillière"
Résistivité électrique	0.66	La voie 1 est toujours fortement pondérée (10 ou 9 sur une échelle de 1 à 10), les autres voies n'ont jamais un poids supérieur à 5.	Aucune sauf pour la parcelle "Magneraud"
Rendement + image satellitaire	0.78	Les poids sont à peu près les mêmes entre les deux groupes de données (pas plus de 3 de différence sur une échelle de 1 à 10).	Aucune

Tableau 29 : Configurations de données recommandées pour la référence "pourcentage de sable grossier" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.74	Aucune	Aucune
Rendement	0.76	Les poids des différentes cartes de rendement sont relativement proches.	Aucune

Tableau 30 : Configurations de données recommandées pour la référence "teneur en CaO" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique	0.58	Aucune	Aucune
Résistivité électrique + rendement + image satellitaire	0.6	Aucune	Aucune
Rendement	0.7	Aucune	Aucune

Tableau 31 : Configurations de données recommandées pour la référence "teneur en CaO" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.69	Aucune	Aucune sauf sur la parcelle "Ouzouer"
Rendement	0.81	Aucune	

Tableau 32 : Configurations de données recommandées pour la référence "teneur en K₂O" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique+ rendement + image satellitaire	0.61	Aucune	Aucune
Rendement + image satellitaire	0.67	Aucune	Aucune sauf sur la parcelle "Magneraud"

Tableau 33 : Configurations de données recommandées pour la référence "teneur en K₂O" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.82	Aucune	Aucune
Rendement + image satellitaire	0.83	Aucune	Aucune

Tableau 34 : Configurations de données recommandées pour la référence "pH KCl" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.73	Aucune	Aucune
Rendement moyenné + image satellitaire	0.76	Le rendement moyenné est toujours plus fortement pondéré sauf sur la parcelle "Ouzouer".	Aucune

Tableau 35 : Configurations de données recommandées pour la référence "pH KCl" pour un grand nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique + rendement + image satellitaire	0.82	Les groupes de la résistance électrique et de l'image satellite sont pondérés de façon similaire. Le rendement reçoit un poids très faible. Ces règles ne sont pas valables sur la parcelle de La Jaillière.	Aucune
Rendement + image satellitaire	0.84	Aucune	Aucune

Tableau 36 : Configurations de données recommandées pour la référence "profondeur de sol" pour un petit nombre de zones

Configuration de données	I.G.V.S. moyen pour les quatre parcelles	Règles de pondération	Relations entre la pondération et les corrélations
Résistivité électrique	0.69	Aucune	Oui.
Résistivité électrique + rendement + image satellite	0.72	Aucune	Aucune pour les "La Jaillière" et "Imbault", relations nettes pour les parcelles "Magneraud" et "Ouzouer".
Rendement	0.81	Aucune	Aucune

Tableau 37 : Configurations de données recommandées pour la référence "profondeur de sol" pour un grand nombre de zones

17. Conclusions de l'étude

17.1. Concernant les groupes de données de référence

17.1.1. La granulométrie

Nous avons obtenu des résultats satisfaisants pour trois des cinq cartes de référence liées à la granulométrie (pourcentage d'argile, pourcentage de limon fin, pourcentage de limon grossier, pourcentage de sable fin, pourcentage de sable grossier). Les mauvais résultats peuvent s'expliquer par le fait que nous travaillons sur cinq données de référence concernant la granulométrie, ces données étant exprimées en pourcentages. Par conséquent, il y a des relations directes entre les différents pourcentages, comme l'ont montré les A.C.P (voir section 16.1.2). Dès lors, si l'on obtient d'excellents résultats pour certaines références, de par les fortes corrélations entre ces références et certaines données tests, il est probable que d'autres références soient nettement moins corrélées avec ces mêmes données tests, en tout cas sur certaines parcelles. On dispose alors de données tests permettant une restitution très fragmentaire de la structure spatiale de ces autres références, ce qui se traduit par de mauvais I.G.V.S.

Bien qu'il soit possible de travailler sur chacune de ces cinq données séparément, il serait sans aucun doute plus pratique de travailler sur une information synthétique. Gillet et Bodilis nous ont proposé de travailler sur une carte de texture de sol et/ou sur une carte d'indice de battance [Gillet, Bodilis 2004]¹.

Lors de pluies intenses, les gouttes font éclater les agrégats du sol, particulièrement si celui-ci est limoneux ou argileux. Une partie des fines particules est alors emportée par ruissellement et, au retrait de l'eau, les particules très fines restantes se solidarisent pour finalement former une couche dure à la surface. Ce phénomène, appelé "battance", est défavorable aux cultures au sens où la couche dure empêche ensuite l'eau et l'air de bien s'infiltrer dans le sol, ce qui nuit aux racines, à l'activité de la flore et de la faune du sol et au développement des plantes.

La carte de texture du sol devra être obtenue à partir du diagramme triangulaire présenté dans la Figure 64. L'inconvénient de cette méthode est de fournir une carte à l'information qualitative nominale qui sera donc très difficilement exploitable dans notre cas.

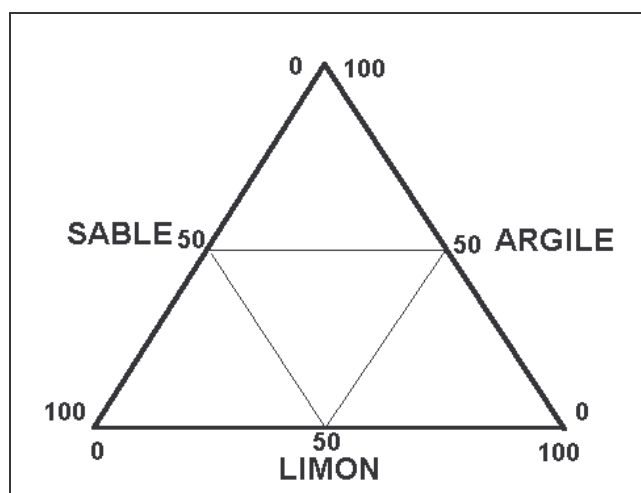


Figure 64 : Triangle de définition de la texture d'un sol à partir des différents pourcentages granulométriques

¹ [Gillet, Bodilis 2004] : Communication privée. Jean-Paul Gillet et Anne-Monique Bodilis sont respectivement chargé de la gestion de l'exploitation et ingénieur régional sur la station expérimentale Arvalis Institut du Végétal de La Jaillière.

La solution visant à exploiter l'indice de battance (I.B.) semble plus adaptée. La formule de calcul de cet indice est la suivante :

$$I.B. = \frac{(1.5 \times \text{lifn}) + (0.75 \times \text{ligr})}{\text{argi} + (10 \times \text{teneur en matière organique})} - c \quad (\text{Équation 9})$$

Avec :

$c = 0.2 \times (\text{pH} - 7)$ si le pH est supérieur à 7,

$c = 0$ si le pH est inférieur à 7.

L'I.B. présente plusieurs avantages. Il prend la forme d'une information quantitative, ce qui le rend parfaitement adapté à notre travail. Il est plus commode à manier que cinq cartes de granulométrie. Il intègre la teneur en matière organique et le pH H₂O. Il s'agit d'un indice fréquemment utilisé par les agronomes, que l'on peut donc considérer comme pertinent agronomiquement.

Nous avons donc décidé d'effectuer quelques tests sur l'indice de battance. La carte de référence "indice de battance" sur la parcelle "La Jaillière" est présentée dans la Figure 65.

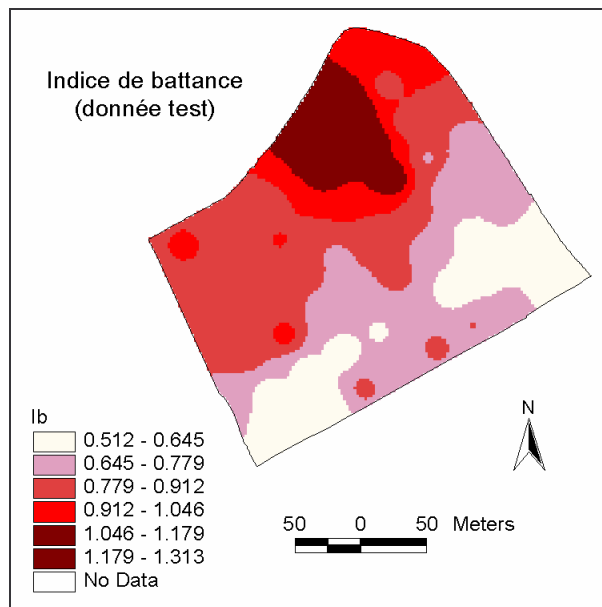


Figure 65 : Carte de la donnée de référence "indice de battance" sur la parcelle "La Jaillière"

Nous avons consulté les résultats des différents groupes ou combinaisons de groupes de données tests obtenus pour les cinq cartes de référence concernant la granulométrie. Nous avons décidé d'effectuer les simulations sur les parcelles "La Jaillière" et "Imbault", l'une étant soumise au phénomène de battance, l'autre moins. Nous avons retenu les deux configurations de données suivantes pour réaliser ces simulations :

- {résistivité électrique + rendement + image satellitaire} avec des poids respectivement de 10, 5 et 5.
- {résistivité électrique + image satellitaire} avec des poids respectivement de 8 et 5.

Afin de nous assurer de la possibilité de substituer l'indice de battance aux différentes données concernant la granulométrie, nous avons calculé les I.G.V.S. des zonages obtenus non seulement par rapport à la carte de l'indice de battance, mais aussi par rapport aux cinq données de référence concernant la granulométrie.

Les résultats des différentes simulations sont présentés dans le Tableau 38. Les cartes de zonage obtenues sur la parcelle "La Jaillière" sont présentées dans la Figure 66. L'absence de zonages à petit nombre de zones sur la parcelle "Imbault" est due, comme nous l'exposerons en section 17.3, au fonctionnement de notre méthode de zonage, en particulier de la C.A.H. avec contrainte de contiguïté.

Simulation	I.G.V.S. pour la parcelle "La Jaillière"		I.G.V.S. pour la parcelle "Imbault"	
	Petit nombre de zones	Grand nombre de zones	Petit nombre de zones	Grand nombre de zones
Résistivité électrique : 10 Rendement : 5 Image satellitaire : 5	ib : 0.81 argi% : 0.6 lifn% : 0.7 ligr% : 0.93 safn% : 0.88 sagr% : 0.61	ib : 0.69 argi% : 0.44 lifn% : 0.54 ligr% : 0.79 safn% : 0.79 sagr% : 0.49	Pas de carte générée	ib : 0.74 argi% : 0.73 lifn% : 0.58 ligr% : 0.83 safn% : 0.76 sagr% : 0.73
Résistivité électrique : 8 Image satellitaire : 5	ib : 0.82 argi% : 0.62 lifn% : 0.7 ligr% : 0.93 safn% : 0.89 sagr% : 0.62	ib : 0.8 argi% : 0.58 lifn% : 0.66 ligr% : 0.91 safn% : 0.86 sagr% : 0.6	Pas de carte générée	ib : 0.76 argi% : 0.72 lifn% : 0.59 ligr% : 0.86 safn% : 0.75 sagr% : 0.75

Tableau 38 : Résultats des quelques simulations effectuées pour la référence "indice de battance" sur les parcelles "La Jaillière" et "Imbault"

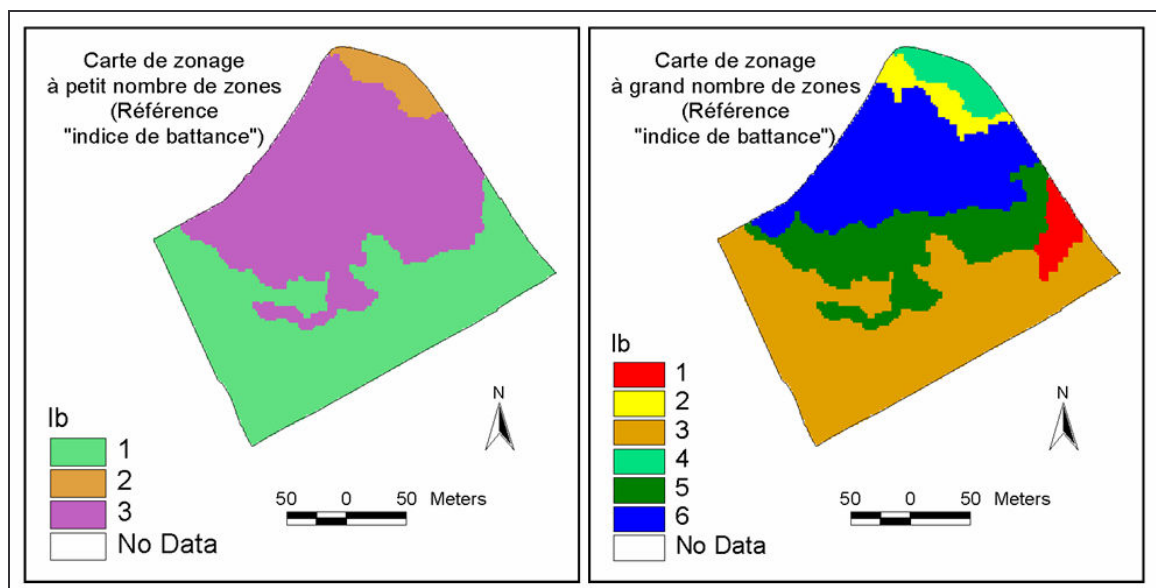


Figure 66 : Zonages correspondants à la simulation {résistivité électrique + rendement + image satellitaire} pour la référence "indice de battance" pour un petit et un grand nombre de zones sur la parcelle "La Jaillière"

Avec quelques simulations, nous obtenons des I.G.V.S. médiocres, mais acceptables, pour la référence "indice de battance". En ce qui concerne les cinq données de référence concernant la granulométrie, les résultats obtenus ici sont en accord avec les résultats obtenus précédemment (voir section 16.2.1 et 16.2.4). Nous restituons mieux les structures spatiales, et donc obtenons de meilleurs I.G.V.S. (de 0.44 à 0.75), pour les références "pourcentage d'argile", "pourcentage de limon fin" et "pourcentage de sable grossier" que pour les références "pourcentage de limon grossier" et "pourcentage de sable fin" (de 0.75 à 0.93). Les différences d'I.G.V.S. entre les informations concernant la granulométrie sont moins nettes sur la parcelle "Imbault" que sur la parcelle "La Jaillière".

Les résultats obtenus devraient être améliorés dans le cadre de tests systématiques tels que ceux pratiqués lors de cette étude.

Au vu des avantages liés à l'utilisation de l'indice de battance et des résultats que nous avons obtenu concernant la restitution de la structure spatiale de cette information et de celles des différentes informations de granulométrie, nous considérons qu'il est possible de remplacer les cinq données de référence concernant la granulométrie par une unique donnée de référence telle que l'indice de battance en obtenant des résultats satisfaisants.

17.1.2. Les teneurs physico-chimiques

Nous avons globalement obtenu des résultats satisfaisants pour les trois cartes de référence de ce groupe, même si ceux-ci varient d'une référence à l'autre et d'une parcelle à l'autre.

Bodilis et Gillet ont attiré notre attention sur l'absence de références concernant les teneurs en matière organique (M.O.) et en MgO. Des discussions devraient donc être engagées sur ce point.

17.1.3. Les pH

Les résultats obtenus pour les deux références concernant les pH sont franchement mauvais. Cela s'explique par le fait que l'information pH est très variable. Elle peut subir des variations de 0.7 à 1 point sur une année selon les parcelles.

17.1.4. La profondeur de sol

Disposer d'une information sur la profondeur de sol permet, entre autres, d'estimer la réserve utile (R.U.). Le calcul de cette dernière nécessite toutefois de connaître globalement la texture du sol de la parcelle, ce qui nous ramène au problème évoqué en 17.1.1 sur la nature qualitative de cette information. Par conséquent, travailler sur la référence "profondeur de sol" s'avère plus aisé que d'élaborer une carte de R.U. et permet, dans une certaine mesure et tout en restant conscient des limites d'une telle attitude, de s'en approcher.

L'obtention d'un zonage de qualité pour la référence "profondeur de sol" n'a rien eu de systématique dans le cadre de cette étude. Les résultats obtenus pour la référence "profondeur de sol" sont mauvais pour les cartes de zonage à petit nombre de zones, le meilleur I.G.V.S. étant de 0.82. Pour les cartes à grand nombre de zones, les résultats sont médiocres, le meilleur I.G.V.S. étant de 0.69.

17.2. Concernant les groupes de données tests

17.2.1. La résistivité électrique

Si la résistivité électrique est la donnée test la plus chère dont nous disposions, elle est aussi la plus discriminante. En effet, rares sont les références pour lesquelles les I.G.V.S. ne sont pas sensiblement dégradés si l'on exclut la résistivité électrique des simulations. Qui plus est, lorsqu'une configuration de données concerne plusieurs groupes de données tests, dont la résistivité électrique, et que des règles de pondération existent, la résistivité électrique est souvent fortement pondérée. De même, pour plusieurs références, la résistivité électrique seule suffit à obtenir un zonage satisfaisant lorsqu'on conserve un grand nombre de zones.

C'est la qualité de la mesure de cette donnée test (semis de points extrêmement dense et localisé de manière très précise) et le soin apporté au calibrage des données issues de ces mesures qui explique son exploitabilité (voir section 5.1.4). Cela fait de la résistivité électrique une donnée particulièrement fiable (au sens exposé en section 10.2.1). Mais c'est aussi la seule de nos données tests qui soient effectivement mesurée dans le sol. Or nous cherchons ici à reproduire des structures spatiales fondées initialement sur des analyses de sol. Le mode de mesure même de la résistivité électrique en fait, autant que la précision de cette mesure, une donnée test particulièrement adaptée à nos besoins dans le cadre de cette étude.

Toutefois, dans notre cas précis, cette donnée a quand même un défaut, conséquent, directement lié à sa qualité. En effet, par la précision de sa mesure, la résistivité électrique permet de restituer avec beaucoup de finesse les transitions entre les tendances fortes d'une structure spatiale. Mais lorsque la résistivité électrique est utilisée seule, il arrive fréquemment (en moyenne une simulation sur trois lors de cette étude) que notre méthode de zonage ne parvienne pas à définir de zonages à petit nombre de zones, le premier zonage défini comptant souvent 5 zones ou plus. Cela explique le fait que le groupe de données tests {résistivité électrique} seul soit rarement proposé comme configuration de données pour un zonage à petit nombre de zones alors qu'il l'est régulièrement lorsqu'on conserve un grand nombre de zones.

17.2.2. Le rendement

Le rendement, par sa disponibilité, constitue une donnée test incontournable. Si l'on peut s'interroger sur la justesse de traiter cette information fondamentalement synthétique comme une donnée test (voir section 5.1.2), il n'est pas possible pour autant de l'ignorer. De plus, l'information relevée par le rendement concerne bien entendu la plante. L'utilisation de cette information synthétique pour tenter de modéliser plusieurs informations précises fondées sur des analyses réalisées sur les vingt cinq premiers centimètres de sol peut donc poser des problèmes.

L'utilisation des cartes de rendement devrait donc s'accompagner de précautions. Conserver une trace du type de culture année par année nous semble être un minimum. Cela permettra d'utiliser des cartes de rendement concernant des cultures aux itinéraires techniques similaires à celle en cours (la culture de la pomme de terre diffère de la culture du blé). De même, lister les accidents (météorologiques ou autres) permettra d'éviter d'utiliser les cartes correspondantes. Par exemple, en 2000, un problème de traitement phytosanitaire a contribué à inverser les régions habituelles de bons et de mauvais rendement sur la parcelle "Imbault". Dans ce cas, le problème n'a rien de naturel et devrait suffire à considérer la carte de rendement finale avec prudence.

Hélas, dans le cadre de cette étude, nous n'avons pas pu respecter ces recommandations, les cartes de rendement que nous avons utilisées datant d'au moins trois ans et les informations concernant les cultures n'étant pas disponibles.

Les données de rendement contiennent souvent des erreurs de diverses natures comme nous l'avons exposé en section 14.2.1.2. Même après correction, les cartes de rendement peuvent encore présenter des singularités (effets de rang, tournières...) qui peuvent se répercuter sur les cartes de zonage. Ainsi, plusieurs des cartes de zonage présentées ici contiennent des artefacts issus des cartes de rendement.

En ce qui concerne le recours au pré-zonage des cartes de rendement pour obtenir des cartes de rendement moyenné, les résultats montrent que dans vingt et un cas sur vingt deux (onze références avec zonages à petit et à grand nombre de zones proposés pour chacune), des simulations incluant les trois groupes de données tests ont obtenus des résultats parmi les meilleurs disponibles. La différence moyenne d'I.G.V.S. entre les configurations utilisant le rendement et les configurations utilisant le rendement moyenné est de 0.013, autrement dit négligeable. Qui plus est, lorsqu'une ou plusieurs simulations incluant le rendement et un seul autre groupe de données tests ont obtenu des résultats intéressants, c'est quasiment toujours le rendement qui est concerné, et non le rendement moyenné (une seule et unique exception dans cette étude). Par conséquent, avoir recours au rendement moyenné n'a strictement rien apporté dans le cadre de cette étude.

Ainsi, pour la donnée de référence "pourcentage de sable grossier", il est possible de remplacer la configuration de données recommandée {résistivité électrique + rendement moyenné + image satellitaire} par la configuration de données {résistivité électrique + rendement + image satellitaire}, la différence d'I.G.V.S. étant de 0.03, soit négligeable.

Les cartes de rendement, si elles doivent être manipulées avec précaution, n'en sont pas moins des données tests utiles. Même si elles sont souvent moins fortement pondérées que les cartes de résistivité électrique, la quasi totalité des configurations de données les prennent en compte, ce qui prouve leur utilité dans le cadre de cette étude.

17.2.3. L'image satellitaire

Si l'image satellitaire seule n'est jamais recommandée comme configuration de données, elle fait partie de la totalité des configurations de données recommandées, la majorité de celles-ci concernant les trois groupes de données tests. Pour ces configurations, lorsque des règles de pondération existent, le groupe de données "image satellitaire" n'est pas souvent le plus fortement pondéré. Mais, pour secondaire qu'il soit, le rôle de l'image satellitaire dans cette étude n'en demeure pas moins réel.

17.3. Concernant le nombre de zones des cartes de zonage

17.3.1. Nombre de zones et qualité de la carte de zonage

Comme nous l'avons exposé en section 17.2.1, il arrive souvent que nos simulations concernant la résistivité électrique seule ne fournissent pas de cartes de zonage à petit nombre de zones, et ce quelle que soit la parcelle concernée. Cela peut aussi survenir, pour certaines parcelles, lorsque d'autres groupes de données tests sont pris en compte et sont nettement moins fortement pondérés que la résistivité électrique (voir section 17.1.1).

En effet, conserver deux ou trois zones revient à s'approcher des limites du mode de fonctionnement de la Classification Ascendante Hiérarchique (C.A.H.) (voir section 10.2.8). Car à la fin de la C.A.H., lorsqu'il ne reste plus que quelques zones à fusionner, la fusion concerne souvent des zones identifiées depuis de nombreuses itérations comme étant très différentes. Autrement dit, on fusionne certaines tendances fortes de la structure spatiale de la parcelle en dépit de leurs différences. Et comme notre C.A.H. à contraintes spatiales a été conçue pour déterminer le meilleur nombre de zones d'un point de vue statistique et spatial, il n'est pas rare qu'elle exclue les cartes à trop faible nombre de zones car celles-ci restituent trop partiellement la structure spatiale de la parcelle.

C'est particulièrement net sur la Figure 36 (section 10.2.8, page 117) où l'on peut constater que le premier minimum local après la fusion de toutes les zones (soit en partant de la gauche de la figure) concerne une carte de zonage avec une dizaine de zones.

Par conséquent, les cartes à faible nombre de zones (deux ou trois), si elles sont lisibles et exploitables sur le terrain, n'en restent pas moins des généralisations très simples, voire simplistes, de la structure spatiale de la parcelle.

Conserver un grand nombre de zones (entre quatre et six) présente deux avantages et un inconvénient. Tout d'abord, il est très rare, dans ce cas, qu'une simulation ne fournisse pas de carte de zonage. Ensuite, les I.G.V.S. sont souvent nettement inférieurs à ceux des cartes de zonage à petit nombre de zones, la restitution de la structure spatiale de la carte de référence étant plus précise. Les différences d'I.G.V.S. entre les cartes à petit et à grand nombre de zones varient entre 0.06 et 0.12 en fonction des références avec une moyenne de 0.09. L'inconvénient des cartes de zonage à grand nombre de zones est que celles-ci sont moins faciles à lire, et donc à exploiter.

Le choix du nombre de zones à conserver dépend donc du compromis recherché entre la lisibilité de la carte de zonage et sa précision statistique et spatiale.

17.3.2. Nombre de zones et adéquation des données tests au sujet de cette étude

Au cours de cette étude, nous avons également généré des cartes de zonage disposant d'un très grand nombre de zones (entre sept et dix). Tout comme pour les cartes à petit et grand nombre de zones, les performances de ces cartes à très grand nombre de zones ont été mises en forme (voir sections 15.2.1 et 15.4), combinaison de données tests par combinaison de données tests, donnée de référence par donnée de référence.

Conformément au mode de fonctionnement de la C.A.H. (voir section 10.2.8), les résultats obtenus avec les cartes de zonage à très grand nombre de zones sont meilleurs que ceux obtenus avec les cartes à grand nombre de zones.

Ainsi, la donnée de référence "pH H₂O", référence la plus difficile à modéliser, obtient une configuration de données avec un I.G.V.S. médiocre, mais acceptable, de 0.76. L'ensemble des autres données de référence obtient des bons I.G.V.S. inférieurs ou égaux à 0.7, plus souvent inférieurs ou égaux à 0.65.

Pour un très grand nombre de zones, le seul groupe de données tests fournissant de mauvais I.G.V.S. pour quasiment toutes les données de référence (par rapport aux résultats des autres groupes de données tests) est l'image satellitaire prise seule. Cela s'explique aisément par la différence de résolution initiale entre les groupes de données tests. Bien que la résolution finale de toutes les cartes soient de deux mètres, celle de l'image satellitaire reste nécessairement liée à sa résolution initiale (quinze ou trente mètres en fonction du capteur) même si elle a été rééchantillonnée. Alors que les cartes de résistivité électrique et de rendement ont été interpolées directement dans la résolution de deux mètres. Lorsqu'on conserve un très grand nombre de zones, la finesse de résolution des cartes de résistivité électrique et de rendement est donc pleinement exploitée alors que l'exploitabilité des canaux de l'image satellitaire reste liée à leur résolution initiale. Ce constat n'est toutefois valable que lorsqu'on considère chaque groupe de données tests individuellement. Dès que l'on combine les groupes, l'image satellitaire joue à nouveau un rôle et elle n'influe pas négativement sur les résultats.

La principale conclusion que l'on peut tirer de ces différents constats est la suivante : quoique nous ayons pu en dire auparavant, toutes les données tests dont nous disposons dans le cadre de cette étude sont à même de restituer, même partiellement, la structure spatiale des différentes données de référence. C'est le nombre de zones retenu pour la carte de zonage qui va faire que cette capacité de restitution va s'exprimer ou pas.

17.4. Concernant les configurations de données

Pour l'ensemble des données de référence, une configuration de données fournit très souvent des résultats satisfaisants, ou tout au moins des résultats parmi les meilleurs que nous ayons obtenus. Il s'agit de la combinaison des trois groupes de données tests {résistivité électrique + rendement + image satellitaire}. Cette configuration de données est recommandée pour dix des onze données de référence et ce quelque soit le nombre de zones retenu. Pour la donnée de référence "pourcentage d'argile", c'est la configuration de données {résistivité électrique + rendement moyenné + image satellitaire} qui a été recommandée mais la configuration de données {résistivité électrique + rendement + image satellitaire} obtient les seconds meilleurs résultats avec une différence négligeable de l'I.G.V.S. de 0.02.

La totalité des vingt-deux configurations de données recommandées dans cette étude (deux zonages différents, à petit et à grand nombre de zones, pour chacune des onze données de référence), concerne une combinaison des trois groupes de données tests. Nous recommandons donc l'emploi de la configuration de données {résistivité électrique + rendement + image satellitaire} pour la modélisation de la structure spatiale, par application de notre méthode de zonage, de l'ensemble des données de référence concernée par cette étude.

En ce qui concerne l'importance des différents groupes de données, on notera le rôle quasi indispensable de la résistivité électrique à l'obtention de bons résultats, viennent ensuite respectivement le rendement et l'image satellitaire. De plus, la résistivité électrique est souvent plus fortement pondérée que le rendement et l'image satellitaire.

Pour chaque donnée de référence, et pour chaque nombre de zones, au moins une configuration de données excluant la résistivité électrique a été proposée afin d'éviter d'imposer l'achat de cette donnée test à un exploitant agricole. Néanmoins, dans la très grande majorité des cas, les résultats fournis par ces configurations de données sont nettement moins bons que ceux fournis par les premières configurations de données retenues, qui incluent quasiment toujours la résistivité électrique.

Enfin, le pré-zonage du rendement afin d'obtenir une carte de rendement moyenné n'a été d'aucune utilité dans le cadre de cette étude. Les résultats obtenus avec le rendement moyenné n'ont jamais été significativement meilleurs que ceux obtenus avec le rendement corrigé. Ils se sont souvent avérés de qualité équivalente, voire même parfois inférieure.

17.5. Pour une éventuelle poursuite de l'étude

Si l'étude devait être reproduite sur d'autres parcelles, il serait opportun de considérer les points suivants concernant les données de référence :

- Remplacer les cinq données de référence concernant la granulométrie par une unique référence concernant l'indice de battance, information synthétique plus utilisable par les agronomes.
- La teneur en MgO et la teneur en matière organique devraient être prises en compte car elles sont fréquemment utilisées par les agronomes.
- A l'inverse, la référence "pH KCl", qui obtient de mauvais résultats et qui n'est pas utilisée par les agronomes, pourrait ne plus être utilisée. La référence "pH H₂O", fréquemment utilisée par les agronomes, et même si elle obtient des résultats pires que ceux de la référence "pH KCl", devrait être conservée afin de confirmer la difficulté à la modéliser que nous avons mis en évidence.

En ce qui concerne les échelles de pondération des données tests, il est conseillé de conserver une échelle de poids de 1 à 10 car celle-ci est très intuitive pour l'utilisateur. Toutefois, dans la mesure où les différences de un ou deux points ne sont pas avérées particulièrement discriminantes lors de cette étude, surtout lorsque de nombreuses données tests sont considérées, nous recommanderions une échelle de pondération prenant les valeurs 1, 4, 7 et 10 pour les éventuelles simulations systématiques à venir.

18. Comparaison des zonages obtenus avec le logiciel eCognition et notre méthode de zonage

Antunes a comparé les zonages obtenus avec notre méthode de zonage lors de l'étude des configurations de données à ceux obtenus avec le logiciel eCognition (présenté en section 6.2.2.4) [Antunes 2004]. Ces comparaisons devaient permettre de valider d'une manière générale les résultats fournis par notre méthode de zonage (et non les résultats de l'étude des configurations de données) et de juger des différences entre les deux techniques afin d'envisager d'éventuelles modifications de notre méthode.

18.1.1. Modalités des comparaisons

Nous avons décidé d'effectuer les comparaisons par rapport aux zonages d'eCognition à partir des cartes de zonage issues de notre étude sur les configurations de données, présentée en section 14. Toutefois, devant la multitude de cartes générées durant cette étude, nous avons, dans un souci de clarté, restreint les comparaisons à quelques cartes.

Les cartes proposées ici concernent toutes la parcelle "La Jaillière". Les cartes retenues, conformément aux conditions de l'étude sur les configurations de données, comptent un petit ou un grand nombre de zones (respectivement entre deux et trois et entre quatre et six) obtenues à partir de différents groupes de données. Les zonages présentés concernent des données de référence différentes en fonction du ou des groupes de données tests considérés. En effet, les I.H. varient pour un même zonage d'une donnée de référence à l'autre et nous avons tenu à effectuer les comparaisons avec eCognition sur la base de nos meilleurs résultats.

Les cartes de zonage présentées ici sont issues des groupes de données suivants :

- Les canaux vert (c2), rouge (c3) et proche infrarouge (c4) d'une image Landsat datant du mois d'août 2000. L'information sol de référence est le pourcentage d'argile.
- Deux cartes de rendement concernant les années 1999 et 2000. L'information sol de référence est le pourcentage d'argile.
- Les voies 1 et 2 de la résistivité électrique. L'information sol de référence est la profondeur de sol.
- L'ensemble de ces données simultanément. L'information sol de référence est le pourcentage de sable fin.

Pour générer des cartes de zonage avec eCognition, Antunes utilise une méthode un peu différente de celle que le manuel d'utilisation préconise. En effet, ce manuel suggère d'exécuter une segmentation puis de vérifier sa qualité visuellement mais aussi grâce à la fonction permettant d'éditer les données statistiques résultantes. Si les résultats ne conviennent pas, il faut alors segmenter à nouveau jusqu'à obtenir un résultat satisfaisant pour l'utilisateur. Malheureusement, la fonction "statistiques" n'était pas disponible sur notre version d'essai d'eCognition. Seule la validation visuelle a donc pu être opérée.

Les cartes de zonage retenues à l'issue de l'étude sur les configurations de données l'ont été après de très nombreux tests de pondération. Faute de temps pour réaliser les mêmes tests sous eCognition, Antunes a appliqué les mêmes pondérations que celles des tests avec notre méthode de zonage. Les tests de paramétrage d'eCognition ont donc surtout porté sur les paramètres d'échelle, de couleur et de forme pour obtenir le même nombre de zones que les cartes de zonage issues de notre méthode.

18.1.2. Comparaison des cartes de zonage issues de l'image satellitaire

Les cartes obtenues avec les deux méthodes de zonage sont présentées Figure 67 et Figure 68.

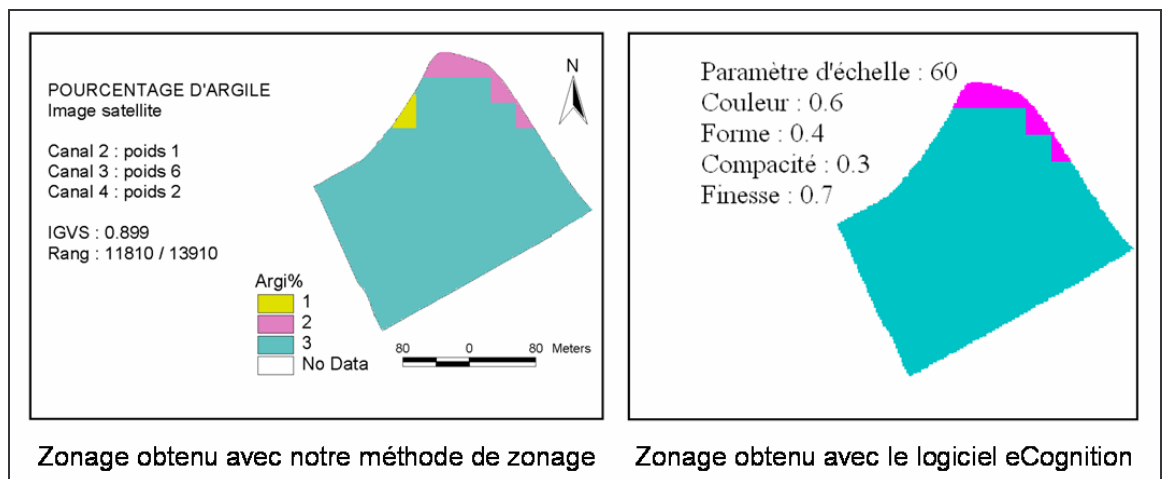


Figure 67 : Cartes de zonage à petit nombre de zones obtenues avec l'image satellitaire

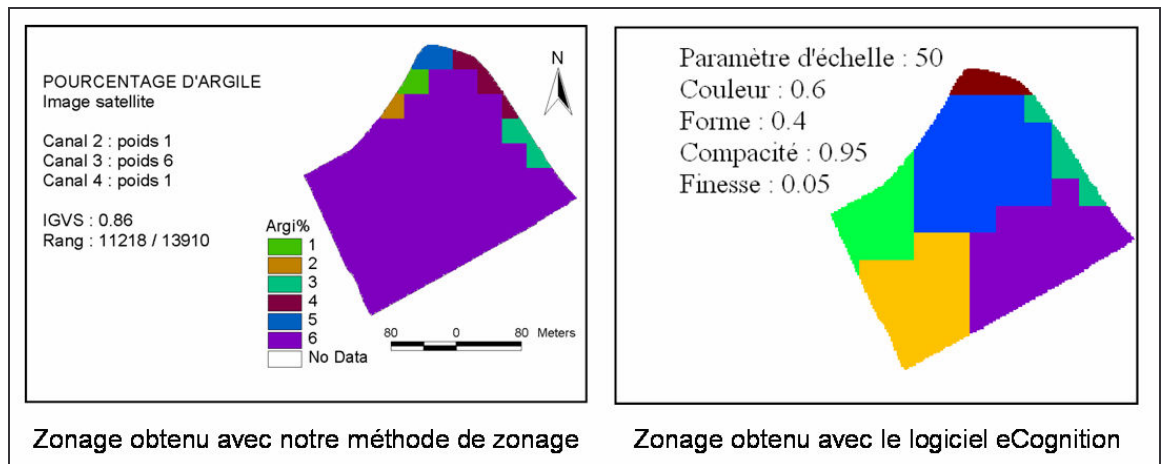


Figure 68 : Cartes de zonage à grand nombre de zones obtenues avec l'image satellitaire

Pour un petit nombre de zones, les deux cartes de zonage ont des aspects quasiment identiques, à l'exception d'une seule petite zone supplémentaire au nord de la carte pour notre méthode.

Pour un grand nombre de zones, les résultats sont très différents. Les petites zones identifiées par notre méthode sur le pourtour de la parcelle sont également souvent présentes pour le zonage eCognition. En revanche, cette dernière présente des grandes zones sur l'ensemble de la surface de la parcelle là où notre méthode n'en identifie aucune de comparable.

Les résultats obtenus avec l'image satellitaire sont donc ambivalents. On notera que les paramètres d'eCognition varient sensiblement entre les zonages à petit et à grand nombre de zones.

18.1.3. Comparaison des cartes de zonage issues des cartes de rendement

Les cartes obtenues avec les deux méthodes de zonage sont présentées Figure 69 et Figure 70.

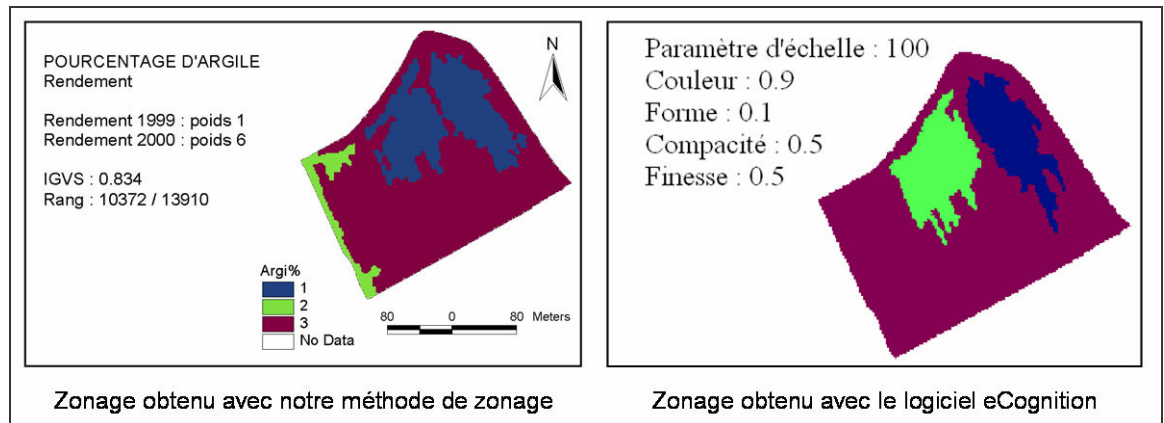


Figure 69 : Cartes de zonage à petit nombre de zones obtenues avec les cartes de rendement

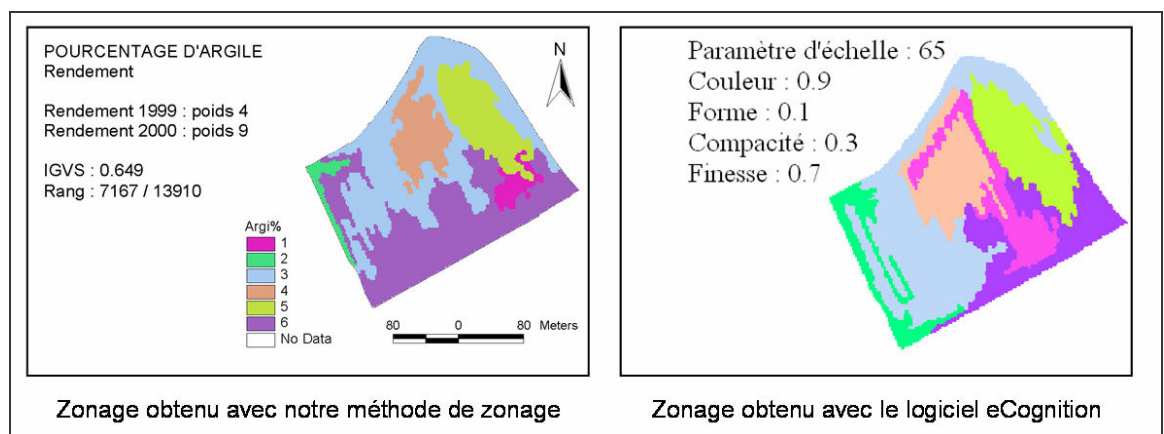


Figure 70 : Cartes de zonage à grand nombre de zones obtenues avec les cartes de rendement

Avec les cartes de rendement, les résultats sont comparables quelque soit le nombre de zones considéré. Si les zones sont différemment regroupées ou scindées et si des nuances apparaissent en fonction de la méthode, les grandes tendances identifiées sont les mêmes.

Exception faite du paramètre d'échelle, les paramètres d'eCognition sont à peu près les mêmes quel que soit le nombre de zones obtenu.

18.1.4. Comparaison des cartes de zonage issues des cartes de résistivité électrique

Les cartes obtenues avec les deux méthodes de zonage sont présentées Figure 71 et Figure 72.

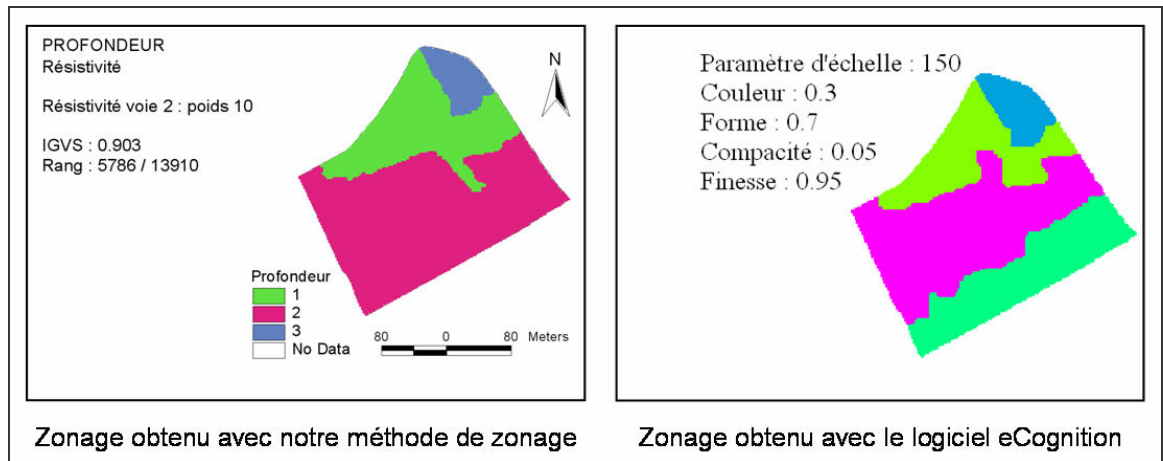


Figure 71 : Cartes de zonage à petit nombre de zones obtenues avec les cartes de résistivité électrique

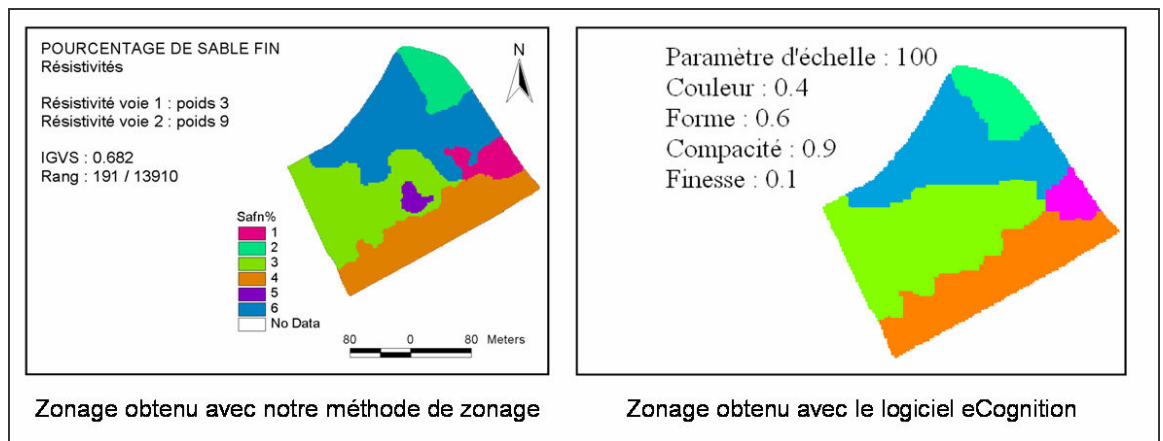


Figure 72 : Cartes de zonage à grand nombre de zones obtenues avec les cartes de résistivité électrique

Pour un petit nombre de zones, le zonage eCognition identifie une zone de plus que notre propre méthode. Les autres zones identifiées par les deux méthodes sont globalement similaires. La modification du paramètre d'échelle sous eCognition afin de n'obtenir que trois zones amène à un regroupement des deux zones au nord. La zone bleue du zonage eCognition n'est peut-être pas suffisamment vaste pour être maintenue avec un autre paramètre d'échelle.

Pour un grand nombre de zones, les deux cartes de zonage présentent de grandes similitudes. Seule la zone 5 de notre méthode n'apparaît pas sur le zonage eCognition, sans doute à cause de sa petite taille par rapport aux paramètres de la segmentation.

On notera les différences sensibles des valeurs des paramètres d'échelle et de compacité entre les deux cartes de zonage issues d'eCognition.

18.1.5. Comparaison des cartes de zonage issues de l'ensemble des données

Les cartes obtenues avec les deux méthodes de zonage sont présentées Figure 73 et Figure 74.

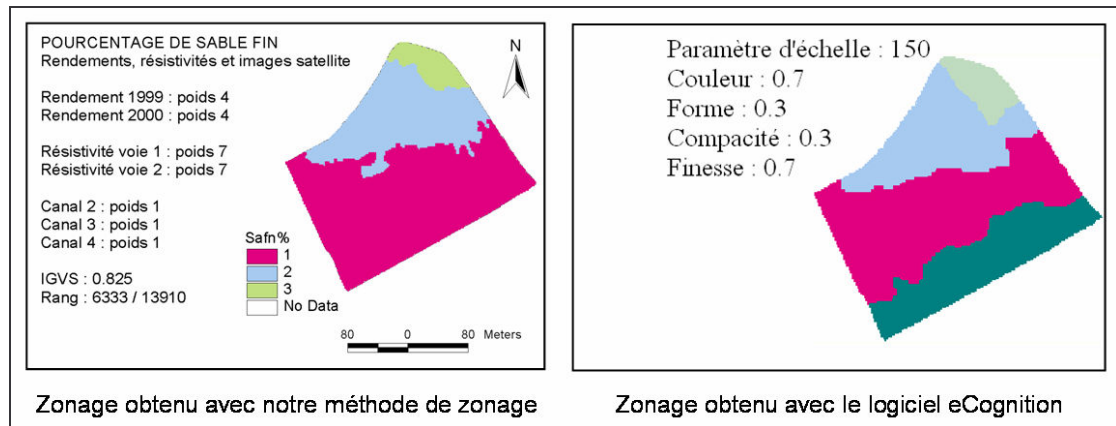


Figure 73 : Cartes de zonage à grand nombre de zones obtenues avec l'ensemble des données

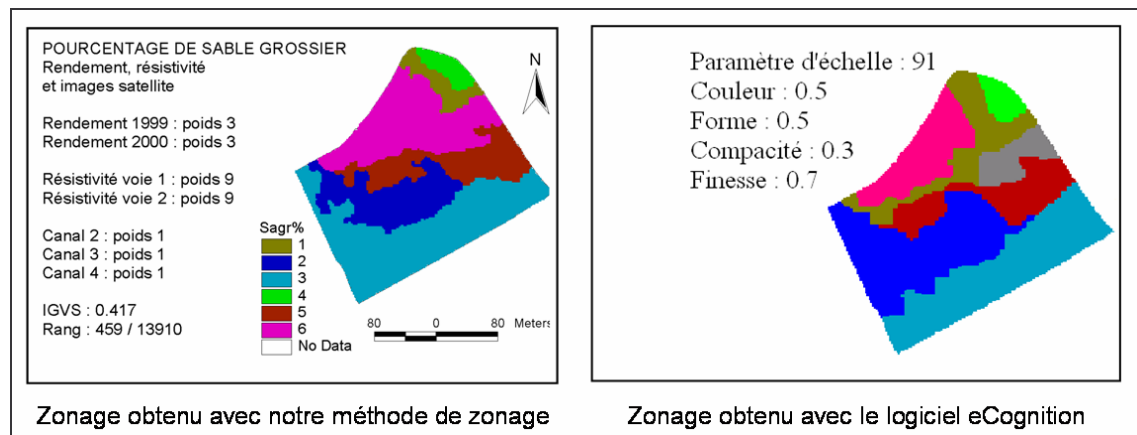


Figure 74 : Cartes de zonage à grand nombre de zones obtenues avec l'ensemble des données

Pour un petit nombre de zones, les résultats sont similaires à ceux obtenus avec la résistivité électrique seule, à savoir une zone supplémentaire pour le zonage eCognition, les autres zones étant globalement les mêmes.

Pour un grand nombre de zones, les cartes sont également très semblables. Le zonage eCognition scinde toutefois la zone 6 de notre méthode en trois zones, dont l'une équivaut à notre zone 1.

Ici encore, le paramètre d'échelle est sensiblement différent en fonction du nombre de zones.

18.1.6. Conclusion

Exception faite des zonages uniquement fondés sur l'image satellitaire, la grande majorité des cartes obtenues avec notre méthode est visuellement comparable à celles obtenues avec eCognition. Les différences entre les deux techniques n'interviennent donc pas au niveau des résultats, mais au niveau du mode de fonctionnement :

- Tout d'abord, eCognition nécessite le renseignement de nombreux paramètres. Outre la pondération, paramètre également présent dans notre méthode, les paramètres d'échelle, de couleur, de forme, de compacité et de finesse doivent être renseignés. Autant de concepts qui nécessitent d'être appréhendés avant de pouvoir utiliser le logiciel.

- Les valeurs renseignées pour les paramètres d'eCognition diffèrent, pour un même groupe de données, d'un nombre de zones à l'autre. Si les poids issus de notre méthode de zonage restent comparables entre un petit et un grand nombre de zones (exception faite des tests sur les cartes de rendement seules, bien que la carte de l'année 2000 soit toujours plus fortement pondérée que la carte de l'année 1999), certains paramètres spécifiques à eCognition varient sensiblement. Toutefois, c'est le paramètre d'échelle qui est principalement affecté. Cela est logique puisque c'est ce paramètre qui permet justement d'influencer le nombre de zones final.
- eCognition, en tant qu'outil d'analyse d'image, accorde beaucoup plus d'influence aux contraintes spatiales que notre propre méthode. Là où notre méthode intègre des contraintes spatiales mais s'attache à préserver au mieux l'information attributaire, eCognition permet aux contraintes spatiales (pas de zones trop petites, surface régulièrement segmentée, compacité des zones...) de modifier sensiblement la carte finale. Cette différence entre les deux techniques a été mise en évidence de façon très nette lors des tests sur l'image satellitaire seule (voir 18.1.2).

Ainsi, s'il est possible d'obtenir des cartes de zonage comparables avec les deux méthodes, leurs objectifs respectifs diffèrent sensiblement.

eCognition s'adresse à des utilisateurs avertis maîtrisant les concepts sous-jacents aux différentes techniques auxquelles il a recours. Notre méthode de zonage n'induit le renseignement que d'un seul paramètre : la pondération ; et peut donc être utilisée par un public plus vaste.

eCognition permet d'obtenir un compromis entre une carte fiable statistiquement et "agréable" visuellement alors que notre méthode offre une carte très fiable statistiquement et ne laisse intervenir les contraintes spatiales que si l'information attributaire et la structure spatiale inhérente sont scrupuleusement respectées.

L'étude des configurations de données a présenté, dans le cadre de notre travail, un double avantage.

Elle a tout d'abord permis de s'assurer de l'exploitabilité de la méthode de zonage. Automatisée au point qu'il n'y a qu'à renseigner le nom des cartes initiales et leur poids et travaillant à partir de fichiers issus d'Arcview 3.2, la méthode a pu être appréhendée et mise en œuvre rapidement par les quelques personnes ayant participé à l'étude. De plus, l'influence des poids sur la carte de zonage obtenue, bien que déjà vérifiée par ailleurs, a été ici mise clairement en évidence.

L'exercice nous a également permis d'exploiter notre méthode de zonage dans le cadre d'une véritable étude aux objectifs et aux moyens concrets et bien formalisés pour Arvalis Institut du Végétal.

Enfin, sur la multitude de simulations réalisées, nous avons sélectionné les meilleurs résultats afin de comparer les cartes correspondantes à celles fournies par un logiciel professionnel réputé performant, eCognition. Bien que les comparaisons n'aient été que visuelles et que notre méthode soit beaucoup plus spécialisée que ce logiciel, nous avons pu constater la similitude des cartes de zonage obtenues avec les deux méthodes.

Devant le souci croissant de nos sociétés modernes de préserver l'environnement, l'agriculture productiviste a montré ses limites. En intégrant les préceptes du développement durable, l'agriculture durable a ouvert la voie à plusieurs approches. Parmi celles-ci, l'agriculture de précision se distingue des autres modes de conduite des cultures par ses besoins spécifiques en information géographique et par l'utilisation qu'elle fait des S.I.G. et de l'analyse spatiale. Pour le géographe, c'est une occasion de s'investir et de contribuer au développement d'une agriculture plus respectueuse de la nature.

Parmi les divers besoins de l'agriculture de précision, notre attention s'est portée sur la définition d'entités homogènes à partir d'une méthode à la fois simple d'utilisation et réalisant des cartes fiables et opérationnelles.

Après avoir entrepris une première série de tests fondés sur l'Aide MultiCritère à la Décision, nous nous sommes rendus compte que cette approche, bien que fournissant des résultats satisfaisants, reposait sur des méthodes complexes. Cette complexité, gage de qualité, rend peu aisée leur utilisation dans le cadre de notre travail. Nous avons donc décidé d'abandonner cette voie. On notera qu'il existe une méthode plus récente qu'E.L.E.C.T.R.E. TRI, nommée I.R.I.S. (Interactive Robustness analysis and parameters Inference for multicriteria Sorting problems), qui tout en s'appuyant sur certaines fonctionnalités de son aînée, est nettement plus simple d'utilisation. Nous n'avons hélas pas pu tester cette méthode.

Dans un second temps, nous avons mis la statistique multivariée à contribution pour la conception de notre méthode de zonage. Mais cet ensemble de techniques seul ne suffit pas pour atteindre notre objectif. Nous avons donc intégré à notre chaîne de traitement statistique des données (en l'occurrence une classification mixte) un ensemble de contraintes spatiales. Au final, notre méthode respecte l'ensemble des contraintes initiales tout en fournissant des résultats satisfaisants. De plus, quelques indicateurs statistiques d'interprétation simple permettent de juger de la qualité des cartes de zonage obtenues.

Ce sont ces indicateurs, et en particulier la possibilité de les calculer à partir de cartes de référence différentes des cartes utilisées pour établir le zonage, qui ont été la clé de voûte de l'étude des configurations de données. Lors de cette étude, nous avons tenté de reproduire la structure spatiale de données onéreuses à partir d'autres données l'étant nettement moins. A l'issue des nombreuses simulations réalisées sur quatre parcelles, nous avons montré que la combinaison des cartes de résistivité électrique, de rendement, et des canaux vert, rouge et proche infra-rouge d'une image satellitaire permettaient systématiquement d'obtenir la meilleure approximation de nos cartes de référence. De plus, pour certaines références, des règles de pondération ont été mises en évidence. Elles consistent presque toujours à pondérer fortement les cartes de résistivité électrique, moyennement les cartes de rendement corrigées et faiblement les cartes issues de l'image satellitaire. D'autres simulations ont été réalisées, dans un second temps, sur quatre autres parcelles. Nous avons manqué de temps pour en analyser les résultats.

Plusieurs pistes pourraient permettre d'améliorer notre méthode de zonage. La première concerne l'utilisation, lors de l'A.C.P., d'une matrice d'autocorrélation spatiale [Gangloff, Reich, Westfall, Koshla 2001]. L'A.C.P. garantit l'orthogonalité des composantes à une distance 0 car les composantes principales sont basées sur une matrice de corrélation qui ne prend pas en compte l'autocorrélation spatiale. En utilisant une matrice d'autocorrélation, on peut prendre en compte une unité de temps ou d'espace entre les variables. Cela pourrait s'avérer très utile pour effectuer des zonages sur des parcelles de taille très différentes et/ou pour s'assurer de la validité d'un zonage dans le temps.

Cette validité du zonage dans le temps pourrait également être surveillée en utilisant un corrélogramme, en particulier si le zonage est fondé sur plusieurs cartes de rendement. On établit un corrélogramme avec des données quantitatives utilisées pour fonder le zonage et variant de façon importante dans le temps, tel le rendement. Une rupture de pente dans le corrélogramme signifie une modification dans le temps de la répartition spatiale des aires de bons et de mauvais rendements. Il faut alors s'interroger sur l'intérêt de prendre en compte la totalité des cartes de rendement pour fonder le zonage.

D'autres notions pourraient également s'avérer utiles. Ainsi, l'entropie conditionnelle permettrait, si l'on dispose de quelques zonages, d'estimer le zonage de l'année suivante sans avoir recours à des données extérieures telles que les cartes de rendement et autres. La notion d'information mutuelle pourrait servir de critère lors des classifications à la place des traditionnels critères métriques.

Enfin, il est envisageable d'établir, sous Arcview, une légende type attribuant à chaque zone de la carte de zonage une couleur en fonction de son Indice d'Hétérogénéité. En plus de l'I.H. global de la carte, l'utilisateur disposerait ainsi, rapidement et de façon très lisible, d'une information supplémentaire sur la qualité du zonage établi.

Au final, notre méthode de définitions d'entités homogènes, fondée sur une classification mixte intégrant des contraintes spatiales, permet d'établir rapidement et simplement des cartes de zonage opérationnelles. Cette simplicité d'utilisation laisse néanmoins à l'exploitant agricole la possibilité de faire valoir ses préférences. En effet, celui-ci garde la main sur la pondération des cartes initiales et il peut modifier la carte de zonage finale s'il le souhaite en s'appuyant sur les indicateurs de validité proposés.

Annexes

19. Vocabulaire commun en A.M.C.D.

Les quelques définitions données ici sont extraites du livre de Bernard Roy : "Méthodologie d'aide multicritère à la décision" [ROY 1985].

19.1. Acteur

Un individu ou un groupe d'individus est acteur d'un processus de décision si, par son système de valeurs, que ce soit au premier degré du fait des intentions de cet individu ou groupe d'individus ou au second degré par la manière dont il fait intervenir celles d'autres individus, il influence directement ou indirectement la décision. De plus, pour qu'un groupe d'individus soit identifié comme un seul et même acteur, il faut que, relativement au processus, les systèmes de valeurs, systèmes informationnels et réseaux relationnels des divers membres du groupe n'aient pas à être différenciés.

19.2. Action

Une action a est la représentation d'une éventuelle contribution à la décision globale susceptible, eu égard à l'état d'avancement du processus de décision, d'être envisagée de façon autonome et de servir de point d'application à l'aide à la décision (ce point d'application pouvant suffire à caractériser a).

19.3. Aide à la décision

Activité de celui qui, prenant appui sur des modèles clairement explicités mais non nécessairement complètement formalisés, aide à obtenir des éléments de réponse aux questions que se pose un intervenant dans un processus de décision, éléments concourant à éclairer la décision et normalement à prescrire, ou simplement favoriser, un comportement de nature à accroître la cohérence entre l'évolution du processus d'une part, les objectifs et le système de valeurs au service desquels cet intervenant se trouve placé d'autre part.

On notera que Schärliig recommande d'utiliser avec précaution le terme "aide à la décision". En effet, il considère qu'il est parfois plus question d'aide au choix ou à l'évaluation [Schärliig 1996].

19.4. Conséquence

Tout effet ou tout attribut de l'action a susceptible d'interférer avec les objectifs ou avec le système de valeurs d'un acteur du processus de décision en tant qu'élément primaire à partir duquel il élabore, justifie ou transforme ses préférences est appelé une conséquence de a .

19.5. Critère

Une fonction g à valeurs réelles définie sur A est, pour un acteur Z , une fonction-critère ou un critère appréhendant le sous-nuage des conséquences $v_g(A)$ si :

Le nombre $g(a)$ est déterminé si et seulement si une évaluation $\Gamma g(a)$ de $vg(a)$ est disponible ; le modèle $\Gamma g(A)$ qui fournit cette évaluation est appelé support de la fonction-critère g .

L'acteur Z (ou l'homme d'étude jugeant au nom de Z) reconnaît l'existence d'un axe de signification sur lequel deux actions potentielles quelconques a et a' peuvent être comparées relativement aux seuls aspects des conséquences qui recouvre $vg(A)$ et il accepte de modéliser cette comparaison conformément à

$$g(a') \geq g(a) \Rightarrow a' S_g a \quad (\text{Équation 10})$$

S_g désignant une relation de surclassement restreint à l'axe de signification du critère g (faisant en particulier abstraction de tous les aspects des conséquences non modélisées dans le support de g).

19.6. Echelle de préférence

Une échelle de préférence E est un ensemble d'états rangés selon un préordre complet noté $\geq$ jouissant de la propriété suivante : raisonnant sur deux actions idéales a et a' dont la comparaison correspond à celle de deux états e et e' de E , tout acteur Z admet :

la situation d'indifférence aIa' (a indifférent à a') lorsque e et e' sont ex æquo dans le préordre ($e = e'$) ;

la situation de préférence $a > a'$ lorsque e précède e' dans le préordre ($e > e'$).

On appelle échelon de E toute classe d'équivalence du préordre ; l'échelle peut être assimilée à l'ensemble de ces échelons. Elle est dotée de l'ordre complet $>$.

19.7. Tableau de performances

Soit un ensemble B d'actions potentielles (B pouvant être égal à A) et $g_1, \dots, g_n$, n critères formant une famille F . On appelle tableau des performances de B sur F le tableau contenant les valeurs $g_j(a)$ pour tout j appartenant à F et pour tout a appartenant à B ainsi que les caractéristiques des éventuelles fonctions-seuils $q_j(g_j(a))$ et $p_j(g_j(a))$.

20. Prise en compte des informations initiales dans une stratégie de définition d'entités homogènes

Doerge est à l'origine de considérations sur la stratégie de définitions d'entités homogènes [Doerge 2001]. Une partie de ces considérations concerne l'utilisation des informations initiales.

Dans un premier temps, il distingue quatre catégories d'informations permettant d'établir des entités homogènes (voir Tableau 39).

Type de caractéristiques de parcelle	Exemples
Quantitatives, stables	Altitude/topographie, matière organique dans le sol, pH ou carbonate de calcium (CaCO_3), conductivité électrique du sol, cartes de lever topographique à très grande échelle, courbure de la surface et propriétés hydrologiques
Quantitatives, dynamiques	Données de capteur de rendement, densité et distribution des adventices, structure de la canopée ou température de la culture, humidité ou salinité du sol, situation en azote du sol ou de la plante
Qualitatives, stables	Couleur du sol, cartes de lever topographique à grande échelle, éléments nutritifs (potassium et phosphore), éléments pathogènes et nuisibles, profondeur au sous-sol, aération du sol/situation du drainage
Intuitives/historiques	Connaissance des caractéristiques de la parcelle de l'exploitant, structure d'ensemble du rendement et pratiques historiques, couche arable et qualité de la couche arable, rotations de cultures passées, limites anciennes des champs, niveau du terrain et structure du drainage, caractéristiques du sous-sol

Tableau 39 : Types de caractéristiques sur lesquelles fonder des entités homogènes en agriculture de précision.

Doerge aborde ensuite des questions pratiques afin de sélectionner des informations conformes aux moyens et aux objectifs de l'exploitant agricole :

Relations de l'information avec le rendement : Doerge insiste sur le fait que les zones de bon et mauvais rendement peuvent changer sensiblement d'une année à l'autre.

Coût de l'information : certaines informations sont très peu chères voire gratuites (connaissance de l'exploitant, carte, de sol, images satellitaires disponibles sur Internet).

Information quantitative et durable dans le temps : topographie, conductivité électrique, propriétés physiques... Si ces informations sont corrélées au rendement, elles peuvent être utilisées pour définir des entités homogènes.

Densité de l'information : les données densément réparties sur la parcelle donneront des résultats plus robustes à l'issue d'une interpolation. Les cartes de rendement, les Modèles Numériques de Terrain (M.N.T.), les photographies aériennes numérisées sont des données denses.

Echelle de l'information : les données devraient être collectées à une échelle au moins aussi fine que celle à laquelle les entités homogènes seront définies.

A partir de ces considérations pratiques et en fonction de l'opération culturale, il est possible de sélectionner les informations conformes à la stratégie de l'exploitant agricole afin de définir des entités homogènes (voir Tableau 40).

Opération culturale	Informations utilisables
Eléments nutritifs (potassium et phosphore)	Topographie, données d'analyses de sol, photographie de sol nu, carte de lever topographique, conductivité électrique
Azote et engrais (fumure)	Texture du sol, matière organique, zones de rendement, photographie du sol nu, reliquats azotés, réflectance de la canopée
Chaux	pH du sol, texture du sol
Gypse	Connaissance de l'exploitant, caractéristiques du rendement, conductivité électrique, analyses de sol pour le pH et le sodium
Densité de semis	Historique des niveaux de rendement, profondeur de la couche arable
Variété	Topographie, caractéristiques du rendement, connaissance de l'exploitant, photographie du sol nu, échantillons de nuisibles géoréférencés
Herbicides	Cartes d'adventices, matière organique du sol, texture du sol
Pesticides	Propriétés du sol, analyses de sol géoréférencées et rapports de reconnaissance sur le terrain
Eau	Texture du sol, topographie, matière organique du sol, zones de rendement

Tableau 40 : Opérations culturales ayant fréquemment recours à la modulation et informations couramment utilisés pour définir les zones homogènes.

21. Mise en forme des tableaux d'I.G.V.S.

Entre l'obtention des cartes de zonage après application de notre méthode de zonage et la recommandation de configurations de données, les résultats subissent plusieurs mises en forme successives afin d'accroître leur lisibilité.

A l'issue des simulations, pour chaque parcelle, on obtient onze fichiers (un par référence) contenant l'ensemble des I.G.V.S. obtenus par l'ensemble des simulations. Le Tableau 41 reprend une partie des 14000 résultats de simulations du tableau original pour la référence "pourcentage d'argile" sur la parcelle "La Jaillière".

IGVS obtenu	Nombre de zones	Nom de la Simulation
...	...	...
0,321	6	5res1_5res2_1rd99moy8Z_1rd00moy9Z_6z
0,322	10	2rd99_2rd00_4res1_4res2_2c2_2c3_2c4_10z
...	...	...
0,351	5	5res1_8res2_5z
0,352	9	4rd99_4rd00_6res1_6res2_2c2_2c3_2c4_9z
0,352	10	5rd99_5rd00_7res1_7res2_2c2_2c3_2c4_10z
0,352	10	2rd99_2rd00_6res1_6res2_3c2_3c3_3c4_10z
0,352	9	10res1_10res2_2c2_2c3_2c4_9z
0,352	5	3res1_6res2_5z
0,352	6	3res1_6res2_6z
0,352	6	9res1_9res2_1rd99moy8Z_1rd00moy9Z_6z
0,353	10	1rd99_1rd00_5res1_5res2_4c2_4c3_4c4_10z
0,353	10	2res1_9res2_10z
0,353	8	5res1_5res2_3rd99moy8Z_3rd00moy9Z_8z
0,354	10	1rd99moy8z_1rd00moy9z_8res1_8res2_3c2_3c3_3c4_10z
0,354	9	3rd99_3rd00_7res1_7res2_2c2_2c3_2c4_9z
0,354	9	3rd99_3rd00_8res1_8res2_3c2_3c3_3c4_9z
0,354	9	5rd99_5rd00_8res1_8res2_1c2_1c3_1c4_9z
0,354	8	9res1_9res2_5rd99_5rd00_8z
0,354	10	1res1_5res2_10z
0,354	7	2res1_4res2_7z
0,355	10	1rd99moy8z_1rd00moy9z_5res1_5res2_1c2_1c3_1c4_10z
0,355	8	1rd99_1rd00_10res1_10res2_6c2_6c3_6c4_8z
0,355	10	6rd99_6rd00_7res1_7res2_1c2_1c3_1c4_10z
0,355	10	4res1_4res2_5rd99_5rd00_10z
0,355	10	5res1_5res2_6rd99_6rd00_10z
0,355	6	7res1_7res2_7rd99_7rd00_6z
...	...	...

Tableau 41 : Partie du tableau de résultats des simulations pour la référence "pourcentage d'argile" sur la parcelle "La Jaillière"

Les simulations sont donc classées par I.G.V.S. et nous recherchons le meilleur I.G.V.S. pour chaque groupe et chaque combinaison de groupes de données tests, et ce pour chaque donnée de référence et pour chaque parcelle.

Une fois cette recherche effectuée pour l'ensemble des références sur l'ensemble des parcelles, nous regroupons les informations des différentes parcelles dans onze tableaux, un par référence. Le Tableau 42, le Tableau 43 et le Tableau 44 représentent les tableaux concernant respectivement le pourcentage d'argile (pour lequel on obtient de bons résultats), la teneur en K₂O (pour laquelle on obtient des résultats mitigés) et le pH H₂O (pour lequel on obtient de mauvais résultats). Pour chaque référence, ces tableaux contiennent les informations suivantes pour un petit et un grand nombre de zones :

Le meilleur I.G.V.S. obtenu par chaque groupe et combinaison de groupes de données tests sur chaque parcelle.

Le rang de la simulation concernée : cette information permet de situer la simulation retenue parmi les milliers réalisées sur chaque parcelle.

La pondération de la simulation concernée : cette information est indispensable pour pouvoir déceler d'éventuelles règles de pondération d'une parcelle à l'autre pour un même groupe ou une même combinaison de groupes de données tests.

Le classement des quatre premiers I.G.V.S. sur chaque parcelle, pour un petit et un grand nombre de zones. Ce classement se distingue par les couleurs des cellules : vert pour le meilleur I.G.V.S. obtenu pour cette parcelle pour cette référence, jaune, pour le second, orange pour le troisième et rouge pour le quatrième. Cette coloration des meilleures simulations permet d'appréhender rapidement les groupes ou combinaisons de groupes de données ayant fourni des bons résultats sur l'ensemble des parcelles.

Groupe ou combinaison de groupes de données tests	Parcelle	2-3 zones		4-6 zones	
		IGVS	Rang / Pondération	IGVS	Rang / Pondération
Résistivité électrique	La Jaillière	0,537	3847 / 4res1_5res2	0,351	311 / 5res1_8res2
	Imbault	0,716	1877 / 6 4 8	0,688	460 / 5 3 2
	Magneraud	0,96	13218 / 1 6 8	0,693	254 / 10 10 5
	Ouzouer	0,942	10466 / 8 2 5	0,833	515 / 8 1 8
Rendement	La Jaillière	0,834	10370 / 1 6	0,649	7167 / 4 9
	Imbault	0,758	5361 / 9 8 10	0,714	1761 / 5 3 10
	Magneraud	0,814	7330 / 7 8	0,731	637 / 7 8
	Ouzouer	0,958	12793 / 10rd99	0,926	8233 / 1 10
Image satellitaire	La Jaillière	0,899	11810 / 1 6 2	0,86	11218 / 1 6 1
	Imbault	0,95	14418 / 10c3	0,896	11889 / 2c3_4c4
	Magneraud	0,857	10043 / 6 3 1	0,763	1410 / 2 8 1
	Ouzouer	0,925	8013 / 1 10 8	0,878	3011 / 1 4 3
Résistivité électrique + rendement	La Jaillière	0,552	4480 / 5 4	0,355	334 / 7 7
	Imbault	0,757	5203 / 4 4	0,716	1855 / 8 7
	Magneraud	0,854	9949 / 2 10	0,761	1295 / 2 7
	Ouzouer	0,959	12939 / 8 8	0,921	7571 / 6 1
Résistivité électrique + rendement moyenné	La Jaillière	0,55	4398 / 9 3	0,321	72 / 5 1
	Imbault	0,866	11098 / 2 1	0,709	1414 / 10 8
	Magneraud	0,834	8976 / 4 10	0,772	1921 / 3 8
	Ouzouer	0,946	11133 / 1 8	0,909	5986 / 1 10
Résistivité électrique + image satellitaire	La Jaillière	0,542	4041 / 9 3	0,37	479 / 7 4
	Imbault	0,744	3916 / 3 7	0,7	946 / 4 3
	Magneraud	0,926	12101 / 1 10	0,72	539 / 4 6
	Ouzouer	0,937	9768 / 7 1	0,844	1018 / 8 1
Rendement + image satellitaire	La Jaillière	0,823	10035 / 8 7	0,693	7635 / 8 4
	Imbault	0,757	5207 / 3 7	0,707	1297 / 6 4
	Magneraud	0,821	8021 / 7 2	0,757	1130 / 3 9
	Ouzouer	0,923	7802 / 2 6	0,886	3610 / 3 6
Rendement moyenné + image satellitaire	La Jaillière	0,822	9919 / 10 7	0,714	7840 / 5 2
	Imbault	0,773	6642 / 3 9	0,744	3923 / 3 10
	Magneraud	0,808	6654 / 7 2	0,714	468 / 7 2
	Ouzouer	0,951	11786 / 3 7	0,894	4349 / 3 8
Résistivité électrique + rendement + image satellitaire	La Jaillière	0,547	4215 / 9 4 10	0,374	514 / 8 7 3
	Imbault	0,752	4751 / 7 1 6	0,705	1228 / 9 3 9
	Magneraud	0,808	6720 / 1 1 2	0,723	569 / 1 8 2
	Ouzouer	0,922	7735 / 1 3 8	0,838	730 / 5 1 1
Résistivité électrique + rendement moyenné + image satellitaire	La Jaillière	0,537	3832 / 7 3 4	0,397	800 / 9 6 1
	Imbault	0,743	3899 / 7 2 8	0,708	1402 / 8 5 1
	Magneraud	0,808	6701 / 1 10 10	0,73	630 / 1 5 10
	Ouzouer	0,889	3847 / 10 7 5	0,825	271 / 10 7 4

Tableau 42 : Tableau de synthèse des résultats des parcelles pour la référence "pourcentage d'argile"

Groupe ou combinaison de groupes de données tests	Parcelle	2-3 zones		4-6 zones	
		IGVS	Rang / Pondération	IGVS	Rang / Pondération
Résistivité électrique	La Jaillière	0,894	5028 / 10res2	0,809	1349 / 3res1_9res2
	Imbault	0,62	2908 / 6 3 5	0,575	296 / 1 3 5
	Magneraud	0,919	12856 / 1 4 3	0,691	231 / 9 4 7
	Ouzouer	0,936	12530 / 8 2 5	0,724	3141 / 1 7 10
Rendement	La Jaillière	0,941	10316 / 4 6	0,851	2332 / 10rd99
	Imbault	0,636	4082 / 2 5 3	0,567	142 / 1 1 9
	Magneraud	0,838	9883 / 7 8	0,799	7923 / 9 10
	Ouzouer	0,848	8512 / 4 8	0,764	5013 / 1 10
Image satellitaire	La Jaillière	0,953	11327 / 4 1 1	0,924	8395 / 10 1 8
	Imbault	0,893	13502 / 10c3	0,868	13021 / 10c3
	Magneraud	0,799	7920 / 10c2	0,737	2515 / 9 10 1
	Ouzouer	0,767	5149 / 7 8 5	0,561	131 / 7 3 5
Résistivité électrique + rendement	La Jaillière	0,931	9219 / 9 2	0,855	2484 / 4 5
	Imbault	0,796	11232 / 6 6	0,611	2156 / 7 1
	Magneraud	0,855	10932 / 3 4	0,747	3883 / 8 2
	Ouzouer	0,82	7589 / 4 4	0,711	2692 / 2 2
Résistivité électrique + rendement moyenné	La Jaillière	0,929	9026 / 6 2	0,801	1243 / 9 2
	Imbault	0,899	13616 / 2 1	0,617	2658 / 8 2
	Magneraud	0,865	11464 / 4 5	0,743	3330 / 2 1
	Ouzouer	0,851	8610 / 1 9	0,777	5680 / 2 10
Résistivité électrique + image satellitaire	La Jaillière	0,91	6576 / 9 5	0,826	1575 / 4 1
	Imbault	0,647	4775 / 2 8	0,584	666 / 2 7
	Magneraud	0,865	11472 / 5 10	0,714	709 / 8 1
	Ouzouer	0,797	6645 / 1 7	0,578	267 / 2 6
Rendement + image satellitaire	La Jaillière	0,869	3097 / 4 5	0,791	1164 / 7 2
	Imbault	0,654	5255 / 8 4	0,604	1662 / 8 4
	Magneraud	0,793	7691 / 1 10	0,753	4690 / 3 9
	Ouzouer	0,765	5058 / 2 6	0,554	75 / 3 7
Rendement moyenné + image satellitaire	La Jaillière	0,916	7205 / 7 9	0,817	1437 / 3 8
	Imbault	0,663	5864 / 7 4	0,632	3794 / 4 1
	Magneraud	0,81	8283 / 4 8	0,766	5766 / 5 7
	Ouzouer	0,757	4658 / 1 2	0,559	117 / 2 7
Résistivité électrique + rendement + image satellitaire	La Jaillière	0,608	365 / 7 4 1	0,527	70 / 9 3 1
	Imbault	0,641	4452 / 9 3 10	0,591	1026 / 2 1 7
	Magneraud	0,801	8028 / 4 2 10	0,747	4001 / 6 2 8
	Ouzouer	0,732	3454 / 1 8 7	0,576	246 / 2 1 6
Résistivité électrique + rendement moyenné + image satellitaire	La Jaillière	0,917	7229 / 1 3 8	0,843	2019 / 9 6 1
	Imbault	0,648	4904 / 10 4 9	0,6	1444 / 9 1 9
	Magneraud	0,852	10788 / 2 4 10	0,728	1571 / 2 1 4
	Ouzouer	0,751	4322 / 3 8 7	0,603	605 / 1 3 8

Tableau 43 : Tableau de synthèse des résultats des parcelles pour la référence "teneur en K2O"

Groupe ou combinaison de groupes de données tests	Parcelle	2-3 zones		4-6 zones	
		IGVS	Rang / Pondération	IGVS	Rang / Pondération
Résistivité électrique	La Jaillière	0,927	8630 / 10res2	0,887	1936 / 7res2_8res2
	Imbault	0,833	7726 / 10 1 1	0,771	2068 / 10 2 5
	Magneraud	0,962	11739 / 1 4 3	0,881	4379 / 9 1 5
	Ouzouer	0,902	8052 / 8 5 9	0,772	69 / 8 1 4
Rendement	La Jaillière	0,948	11332 / 4 6	0,9	3937 / 10rd00
	Imbault	0,748	858 / 9 5 8	0,698	59 / 10 2 5
	Magneraud	0,967	12106 / 7 8	0,911	8254 / 1 5
	Ouzouer	0,929	10911 / 3 7	0,889	6494 / 1 3
Image satellitaire	La Jaillière	0,94	10235 / 5 1 1	0,923	7728 / 10 3 2
	Imbault	0,876	10484 / 10c3	0,861	9700 / 10c3
	Magneraud	0,909	8083 / 5 3 2	0,866	2047 / 3 6 7
	Ouzouer	0,93	10981 / 1 10 7	0,811	364 / 1 4 9
Résistivité électrique + rendement	La Jaillière	0,935	9696 / 2 4	0,884	1589 / 4 5
	Imbault	0,88	10668 / 6 6	0,814	5925 / 7 1
	Magneraud	0,957	11450 / 4 6	0,882	4501 / 7 2
	Ouzouer	0,919	9796 / 8 8	0,884	5899 / 6 5
Résistivité électrique + rendement moyenné	La Jaillière	0,93	8892 / 2 2	0,887	1943 / 9 2
	Imbault	0,896	11520 / 2 9	0,803	4722 / 10 8
	Magneraud	0,947	10937 / 4 10	0,884	4844 / 3 10
	Ouzouer	0,937	11784 / 9 1	0,868	4049 / 3 1
Résistivité électrique + image satellitaire	La Jaillière	0,918	6909 / 4 4	0,876	921 / 10 2
	Imbault	0,828	7311 / 3 7	0,767	1826 / 5 7
	Magneraud	0,93	9875 / 3 4	0,863	1662 / 5 6
	Ouzouer	0,942	12434 / 7 1	0,805	259 / 3 3
Rendement + image satellitaire	La Jaillière	0,921	7429 / 5 4	0,888	2067 / 6 6
	Imbault	0,784	2953 / 6 2	0,748	832 / 5 1
	Magneraud	0,92	9102 / 7 1	0,869	2467 / 9 2
	Ouzouer	0,93	10996 / 6 3	0,872	4495 / 2 10
Rendement moyenné + image satellitaire	La Jaillière	0,932	9428 / 7 9	0,854	159 / 7 6
	Imbault	0,907	12071 / 3 9	0,814	5935 / 4 1
	Magneraud	0,895	6280 / 3 4	0,868	2314 / 3 4
	Ouzouer	0,938	11900 / 1 3	0,874	4714 / 4 1
Résistivité électrique + rendement + image satellitaire	La Jaillière	0,917	6745 / 2 1 3	0,878	1045 / 8 7 3
	Imbault	0,815	6136 / 6 2 9	0,769	1970 / 1 9 2
	Magneraud	0,907	7848 / 1 1 2	0,856	907 / 3 1 5
	Ouzouer	0,884	5953 / 8 1 10	0,794	153 / 6 6 3
Résistivité électrique + rendement moyenné + image satellitaire	La Jaillière	0,907	4882 / 6 3 10	0,871	630 / 1 10 7
	Imbault	0,842	8505 / 5 2 10	0,767	1863 / 7 1 7
	Magneraud	0,913	8581 / 4 1 7	0,855	793 / 3 1 4
	Ouzouer	0,895	7153 / 8 1 7	0,817	500 / 10 1 1

Tableau 44 : Tableau de synthèse des résultats des parcelles pour la référence "pH H2O"

Pour chaque référence, on réalise ensuite la moyenne des différents I.G.V.S. obtenus par les différents groupes ou combinaisons de groupes de données tests sur les différentes parcelles. Après avoir supprimé les rangs des simulations retenues, on obtient les tableaux d'I.G.V.S. présentés dans ce document pour chacune des références et sur la base desquels les configurations de données à recommander sont définies.

Bibliographie

- [Adinarayana, Maitra, Venkataraman 2001] ADINARAYANA J., MAITRA S., VENKATARAMAN G., "Spatial Decision support system for rural land use planning", In Steffe J., *Third European Conference of the European Federation for Information Technology in Agriculture, Food and the Environment*, tome 1, p. 233-237, agro Montpellier, Montpellier, 2001.
- [AIRE/CTS 2001] Association Inter-Régionale Européenne pour la Promotion de la Cartographie Tridimensionnelle des Sols Agricoles, *Pédographie des parcelles agricoles appliquée à l'agriculture de précision et aux problématiques agri-environnementales : utilisation conjointe de l'induction électromagnétique et de la télédétection*, AIRE/CTS 2001.
- [Antunes 2004] ANTUNES C., La classification : une méthode proposée par le logiciel eCognition, Arvalis Institut du Végétal, 2004.
- [Ariaux 2000] ARIAUX B., "Les S.I.G. utilisés en agriculture de précision", In Zwaenepoel P., *Actes du colloque Agriculture de précision : avancées de la recherche technologique et industrielle*, p. 55-65, Educagri Editions, Dijon, 2000.
- [Bailey 1994] BAILEY T.C., "G.I.S. and spatial analysis : a review of statistical spatial analysis in geographical information systems", In Fotheringham S., Rogerson P., *Spatial Analysis and G.I.S.*, p. 13-44, Taylor and Francis, Bristol, 1994.
- [Baratte 2002] BARATTE E., "L'exploitation céréalière de précision : une technologie qui respecte l'environnement", In *Grains Magazine*, Octobre 2002.
- [Bavoux, Chapuis, Delmer, Mannone, Passegué, Volpoët 1998] BAVOUX J.J., CHAPUIS R., DELMER S., MANNONE V., PASSEGUE S., VOLPOET P., *Introduction à l'analyse spatiale*, collection "Synthèse", série "Géographie", Armand Colin, Paris, 1998.
- [Bichat 1999] BICHAT H.-H., "A propos de l'agriculture de précision", In *Bulletin du Conseil Général du Génie Rural, des Eaux et des Forêts*, n° 53, p. 93-96, avril 1999.
- [Boisgontier 1997a] BOISGONTIER D., "Hétérogénéités : comment les déceler ?", In *Agriculture de précision, cap sur le XXIème siècle*, supplément à *Perspectives Agricoles*, n° 222, p. 14-19, mars 1997.
- [Boisgontier 1997b] BOISGONTIER D., "Prendre en compte l'hétérogénéité des parcelles", In *Agriculture de précision, cap sur le XXIème siècle*, supplément à *Perspectives Agricoles*, n° 222, p. 20-22, mars 1997.
- [Boisgontier 1999] BOISGONTIER D., "Quels enjeux environnementaux pour l'agriculture de précision ?", In *Ingénieurs de la Vie*, n° 450, p. 10-11, juillet-août-septembre 1999.
- [Boisgontier 2000] BOISGONTIER D., "La mise en œuvre de l'agriculture de précision", In Zwaenepoel P., *Actes du colloque Agriculture de précision : avancées de la recherche technologique et industrielle*, p. 365-374, Educagri Editions, Dijon, 2000.
- [Boisgontier, Ariaux 1998] BOISGONTIER D., ARIAUX B., "Agriculture de précision : que faire aujourd'hui et avec quels moyens ?", In *Perspectives Agricoles*, n° 235, p. 14-20, mai 1998.
- [Boisgontier, Hébrard 2001] BOISGONTIER D., HEBRARD J.-P., "S.I.G. : Une base de données à l'échelle intraparcellaire", In *Perspectives Agricoles*, n° 271, p. 40-41, septembre 2001.
- [Bouroche, Saporta 1989] BOUROCHE J. M., SAPORTA G., *L'analyse des données*, Presses Universitaires de France, Paris, France, 1989.

[Bourret, Reggia, Samuelides 1991] BOURRET P., REGGIA J., SAMUELIDES M., *Réseaux Neuronaux, une approche connexionniste de l'intelligence artificielle*, Teknea, Toulouse, France, 1991.

[Boydell, McBratney 1999] BOYDELL B., McBRATNEY A.B., "Identifying potential within-field management zones from cotton yield estimates", In Stafford J.V., *Precision Agriculture '99*, tome 1, p. 331-341, Sheffield Academic Press Limited, Sheffield, 1999.

[Chaléard, Charvet 2004] CHALEARD J.L., CHARVET J.P., *Géographie agricole et rurale*, Collection Atouts Géographie, Belin, Paris, 2004.

[Chang, Clay, Carlson, Clay, Reese 2000] CHANG J., CLAY D.E., CARLSON C.G., CLAY S.A., REESE C.L., "The influence of different approaches for identifying inorganic N and P management zones on fertilizer recommendation", In *Proceedings of the Fifth International Conference on Precision Agriculture*, ASA-CSSA-SSSA, Madison, Etats-Unis, 2000.

[Chung, Sudduth, Drummond, Kitchen 2001] CHUNG S.O., SUDDUTH K.A., DRUMMOND S.T., KITCHEN N.R., "Analysis of soil, topographic and yield data using nested variograms", *Third International Conference on Geospatial Information in Agriculture and Forestry*, 2001.

[Culligan Dunsmore, Salt 1998] CULLIGAN DUNSMORE M., SALT C.A., "A G.I.S.-based decision support system using multicriteria decision making methodology for radiological countermeasure suitability assessment", In *Proceedings of the First International Conference on Geospatial Information in Agriculture and Forestry*, tome 1, p. 645-652, ERIM International, Ann Arbor, 1998.

[Cupitt, Whelan 2001] CUPITT J., WHELAN B.M., "Determining potential within-field crop management zones", In Grenier G., Blackmore S., *Third European Conference on Precision Agriculture*, tome 1, p. 7-12, agro Montpellier, Montpellier, 2001.

[Dabas, Tabbagh, Boigontier 2001] DABAS M., TABBAGH J., BOISGONTIER D., "Multi-depth continuous electrical profiling (MuCEP) for characterization of in-field variability", In Grenier G., Blackmore S., *Third European Conference on Precision Agriculture*, tome 1, p. 361-366, agro Montpellier, Montpellier, 2001.

[Dampney, King, Lark, Wheeler, Bradley, Mayr 2003] DAMPNEY P.M.R., KING J.A., LARK R.M., WHEELER H.C., BRADLEY R.I., MAYR T.R., "Automated methods for mapping pattern of soil physical properties as a basis for variable management of crops within fields", In Stafford J.V., Werner A. *Precision Agriculture*, p. 135-140, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Delpech, Biagini 2000] DELPECH C., BIAGGINI F., "Progresser dans la gestion agronomiques des parcelles", In *Agrodistribution*, p. 16-19, juin 2000.

[De Montgolfier, Bertier 1978] DE MONTGOLFIER J., BERTIER P., *Approche multicritère des problèmes de décision*, Editions Hommes et Techniques, Suresnes, France, 1978.

[Dias, Mousseau, Figueira, Climaco, Silva ?] DIAS L., MOUSSEAU V., FIGUEIRA J., CLIMACO J., SILVA C.G., *IRIS - Interactive Robustness analysis and parameters' Inference for multicriteria Sorting problems*, ?. <http://www4.fe.uc.pt/lmcdias/iris.htm>

[Dillon, Mueller, Shearer 2003] DILLON C.R., MUELLER T., SHEARER S., "An economic optimization model for management zone configuration", In Stafford J.V., Werner A. *Precision Agriculture*, p. 165-169, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Dobermann, Ping, Simbahan, Adamchuk 2003] DOBERMANN A., PING J.L., SIMBAHAN G.C., ADAMCHUK V.I., "Processing of yield map data for delineating yield zones", In Stafford J.V., Werner A. *Precision Agriculture*, p. 177-185, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Doerge 2001] DOERGE T.A., "Management Zone Concepts", In *Site-Specific Management Guidelines*, Potash & Phosphate Institute, 2001. www.ppi-far.org/ssmg

[Domsch, Kaiser, Witzke, Zauer 2003] DOMSCH H., KAISER T., WITZKE K., ZAUER O., "Empirical methods to detect management zones with respect to yield", In Stafford J.V., Werner A. *Precision Agriculture*, p. 187-192, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Dumolard 1981] DUMOLARD P., *L'espace différencié*, collection "Géographia", Economica, Paris, 1981.

[Dumolard, Dubus, Charleux, 2003] DUMOLARD P., DUBUS N., CHARLEUX L., *Les statistiques en géographie*, Collection Atouts Géographie, Editions Belin, Paris, 2003.

[Fleming, Westfall, Bausch 2000] FLEMING K.L., WESTFALL D.G., BAUSCH W.C., "Evaluating management zone technology and grid soil sampling for variable rate nitrogen application", In *Proceedings of the Fifth International Conference on Precision Agriculture*, ASA-CSSA-SSSA, Madison, Etats-Unis, 2000.

[Fourty 1996] FOURTY T., Méthodes d'estimation du contenu biochimique d'un couvert végétal à partir de mesures haute résolution spectrale acquises au niveau satellitaire, Thèse de troisième cycle, P. 132-139, 1996.

[Franzen, Halvorson, Hofman 2000] FRANZEN D.W., HALVORSON A.D., HOFMAN V.L., "Management zones for soil N and P levels in the northern great plains", In *Proceedings of the Fifth International Conference on Precision Agriculture*, ASA-CSSA-SSSA, Madison, Etats-Unis, 2000.

[Fridgen, Fraisse, Kitchen, Sudduth 2000] FRIDGEN J.J., FRAISSE C.W., KITCHEN N.R., SUDDUTH K.A., "Delineation and analysis of site-specific management zones", In *Proceedings of the Second International Conference of Geospatial Information in Agriculture and Forestry*, tome 2, p. 402-411, ERIM International, Ann Arbor, 2000.

[Frogbook, Oliver, Derricourt 2003] FROGBOOK Z.L., OLIVER M.A., DERRICOURT K.E., "Exploring the spatial relations between soil properties and Electro-Magnetic Induction (EMI) and the implications for management", In Stafford J.V., Werner A. *Precision Agriculture*, p. 217-222, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Gaillard 2004] GAILLARD D., *Gestion concertée du ruissellement-érosif dans les espaces agricoles des plateaux de grande culture*, Thèse de doctorat de Géographie, Université de Rouen, U.F.R. de lettres et sciences humaines, 2004.

[Gangloff, Reich, Westfall, Koshla 2001b] GANGLOFF W.J., REICH R.M., WESTFALL D.G., KOSHLA R., "Principal component kriging utilizing an autocorrelation matrix", In *Proceedings of the Third International Conference on Geospatial Information in Agriculture and Forestry*, Veridian, Ann Arbor, Etats-Unis, 2001.

[Gay 1995] GAY J.C., *Les discontinuités spatiales*, collection "Poche Géographie", Economica, Paris, 1995.

[Gay 1997] GAY M., "Agriculture de précision : un nouvel enjeu pour la céréaliculture européenne", In *Revue Purpan*, n° 184, p. 125-148, juillet-septembre 1997.

[Gerwig, Sadler, Evans 2000] GERWIG B.K., SADLER E.J., EVANS D.E., "Evaluating techniques for defining management zones in the coastal plain", In *Proceedings of the Fifth International Conference on Precision Agriculture*, ASA-CSSA-SSSA, Madison, Etats-Unis, 2000.

[Girard, Girard 1999] GIRARD M-C, GIRARD C.M., *Traitement des données de télédétection*, Dunod, Paris, France, 1999.

[Godard 2001] GODARD A., Production de modèles prédictifs de paramètres pédologiques à partir de la conductivité électro-magnétique, Mémoire d'études, Institut National Agronomique Paris-Grignon, 2001.

[Graillot, Londiche, Dechomets, Batton-Hubert ?] GRAILLOT D., LONDICHE H., DECHOMETTS R., BATTON-HUBERT M., *Analyse multicritère spatiale pour l'identification de la vulnérabilité aux inondations*, ?.

[Grenier 1997] GRENIER G., "La maîtrise informatique des données nécessaires à l'agriculture de précision", In 5^{ème} rencontres internationales de l'A.F.C.O.M.E., 1997.

[Groupe Chadule 1987] GROUPE CHADULE, *Initiation aux pratiques statistiques en géographie*, Masson, Paris, France, 1987.

[Haining 1994] HAINING R., "Designing spatial data analysis modules for geographical information systems", In Fotheringham S., Rogerson P., *Spatial Analysis and G.I.S.*, p. 45-63, Taylor and Francis, Bristol, 1994.

[Hornung, Koshla, Reich, Westfall 2003] HORNUNG A., KOSHLA R., REICH R., WESTFALL D.G., "Evaluation of site-specific management zones : grain yield and nitrogen use efficiency", In Stafford J.V., Werner A. *Precision Agriculture*, p. 297-302, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Hoskinson, Hess, Hempstead 1998] HOSKINSON R.L., HESS J.R., HEMPSTEAD D.W., "Precision farming results from using the decision support system for agriculture (DSS4AG)", In *Proceedings of the First International Conference on Geospatial Information in Agriculture and Forestry*, tome 1, p. 206-210, ERIM International, Ann Arbor, 1998.

[Joërin 1997] JOERIN F., Décider sur le territoire, proposition d'une approche par utilisation de S.I.G. et de méthodes d'analyse multicritère, Thèse de doctorat, Ecole Polytechnique Fédérale de Lausanne, 1997.

[Joërin, Thériault, Musy 2001] JOËRIN F., THERIAULT M., MUSY A., "Using G.I.S. and outranking multicriteria analysis for land-use suitability assessment", In *International Journal of Geographical Information Science*, vol. 15(2), p. 153-174, 2001.

[Johannsen 1998] JOHANNSEN C.J., "Glossary of Terms for Precision Farming", In *Modern Agriculture*, p. 44-46, été 1998.

[Jullien, Huet 2005] JULLIEN A., HUET P., "Agriculture de précision" In Laurent C., Thinon P., *Agricultures et territoires*, Collection Information géographique et aménagement du territoire, Hermès, Paris, 2005.

[Kachanoski, O'halloran, Aspinall, Von Bertoldi 1996] KACHANOSKI R.G., O'HALLORAN I.P., ASPINALL D., VON BERTOLDI P., "Delta Yield : Mapping Fertilizer Nitrogen Requirement for Crops", In *Better Crops*, vol. 80, n° 3, p. 20-23, 1996.

[Kravchenko 2003] KRAVCHENKO A., "The effects of spatial structure on accuracy of map and performance of interpolation methods", In Stafford J.V., Werner A. *Precision Agriculture*, p. 329-333, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

[Laaribi 2000] LAARIBI A., *S.I.G. et analyse multicritère*, Hermès Sciences, Paris, 2000.

[Lark, Stafford 1998] LARK R.M., STAFFORD J.V., "Information on within-field variability from sequences of yield maps : multivariate classification as a first step of interpretation", In *Nutrient Cycling in Agroecosystems*, n° 50, p. 277-281, 1998.

[Laurent 2005] LAURENT C., "Agricultures et territoires : changements de perspective, méthodes d'analyse et nouvelles recherches" In Laurent C., Thinon P., *Agricultures et territoires*, Collection Information géographique et aménagement du territoire, Hermès, Paris, 2005.

[Laurent Thinon 2005] LAURENT C., THINON P., "*Agricultures et territoires : changements de perspective, méthodes d'analyse et nouvelles recherches*" In Laurent C., Thinon P., *Agricultures et territoires*, Collection Information géographique et aménagement du territoire, Hermès, Paris, 2005.

[Lebart, Morineau, Piron 2000] LEBART L., MORINEAU A., PIRON M., *Statistiques exploratoire et multidimensionnelle*, Dunod, Paris, France, 2000.

[Lefaucheur 1999] LEFAUCHEUR R., "L'agriculture de précision : introduction et tendances d'adoption", 1999. <http://www.agrimine.com>

[Luchiari, Shanahan, Francis, Schlemmer, Schepers, Liebig, Schepers Payton 2000] LUCHIARI A.Jr., SHANAHAN J., FRANCIS J., SCHLEMMER M., SCHEPERS J., LIEBIG M., SCHEPERS A., PAYTON S., "Strategies for establishing management zones for site specific nutrient management", In *Proceedings of the Fifth International Conference on Precision Agriculture*, ASA-CSSA-SSSA, Madison, Etats-Unis, 2000.

[Mainguenaud 2002] MAINGUENAUD M. (sous la direction de), *Langages pour les S.I.G. : conception, développement et I.H.M.*, Hermès Sciences, Paris, 2002.

[Mallarino, Hinz, Oyarzabal 1996] MALLARINO A.P., HINZ P.N., OYARZABAL E.S., "Multivariate Analysis as a Tool for Interpreting Relationships Between Site Variables and Crop Yields", In *Proceedings of the Third International Conference on Precision Agriculture*, p. 151-157, ASA-CSSA-SSSA, Madison, 1996.

[Malzer 1996] MALZER L.G., "Corn Yield Response Variability and Potential Profitability of Site-Specific Nitrogen Management", In *Better Crops*, vol. 80, n° 3, p. 6-8, 1996.

[Masson 2000] MASSON B., "l'agriculture de précision se met sur orbite", In *Bima*, n° 1484, p. 22-23, juillet-août 2000.

[Maystre, Pictet, Simos 1994] MAYSTRE L.-Y., PICTET J., SIMOS J., *Méthodes multicritères ELECTRE : description, conseils pratiques et cas d'application à la gestion environnementale*, Collection Gérer l'Environnement, Presses Polytechniques et Universitaires Romandes, Lausanne, Suisse.

[McKinion, Olson, Sequeira, Whisler, Varco, Williford 1997] McKINION J.M., OLSON R.L., SEQUEIRA R.A., WHISLER F.D., VARCO J.J., WILLIFORD J.R., "Model-based decision support systems for precision agriculture in cotton production", In Stafford J.V., *Precision Agriculture '97*, tome 2, p. 769-775, BIOS Scientific Publishers Limited, Oxford, 1997.

[Mérour 2002] MEROUR X., *La cartographie de rendement : traitement et valorisation de l'information*, Mémoire d'études, Ecole Supérieure d'Agriculture d'Angers, 2002.

[Minvielle, Souiah 2003] MINVIELLE E., SOUIAH S-A., *L'analyse statistique et spatiale : statistiques cartographie, télédétection, S.I.G.*, Collection Outils et Méthodes en Géographie, Editions du Temps, Nantes, 2003.

[Molines 2003] MOLINES N., *Méthodes et outils pour la planification des grandes infrastructures linéaires et leur évaluation environnementale*, Thèse de doctorat, Département de Géographie, Faculté de Sciences Humaines et Sociales, Université Jean Monnet de Saint-Etienne, 2003.

[Molines, Chevallier 2001] MOLINES N., CHEVALLIER J.J., "Système d'information géographique et analyse multicritère : une association novatrice au service du processus d'évaluation des grandes infrastructures linéaires", In Colorni A., Paruccini M., Roy B., *A-MCD-A*, p. 195-212, Office for Official Publications of the European Communities, Luxembourg, 2001.

[Openshaw 1991] OPENSHAW S., "Developing Appropriate Spatial Analysis Methods for G.I.S.", In Maguire D.J., Goodchild M.F., Rhind D.W., *Geographical Information Systems, Principles and Applications*, tome 1, p. 389-402, 1991.

- [Pavard 1996] PAVARD P., "L'agriculture de précision : pour quoi faire ?", In *La France Agricole*, p. 22, 27 septembre 1996.
- [Pawlak 2003] PAWLAK J., "Precision agriculture – economic aspects", In Stafford J.V., Werner A. *Precision Agriculture*, p. 527-532, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.
- [Pilesjö, Thylén, Persson 2000] PILESJÖ P., THYLÉN L., PERSSON A., "Digital elevation models for delineation of agricultural management zones", In *Proceedings of the Second International Conference of Geospatial Information in Agriculture and Forestry*, tome 1, p. 224-231, ERIM International, Ann Arbor, 2000.
- [R.A.C.A.L. 1999] R.A.C.A.L., "Explication du D.G.P.S., Guide d'agriculture de précision destiné aux fermiers", 1999.
- [Robert 2000] ROBERT C.R., "L'agriculture de précision : les verrous liés à la technologie et à la gestion agronomique", In Zwaenepoel P., *Actes du colloque Agriculture de précision : avancées de la recherche technologique et industrielle*, p. 11-29, Educagri Editions, Dijon, 2000.
- [Robin 1995] ROBIN M., *La télédétection*, Collection Fac Géographie, Editions Nathan, Paris 1995.
- [Rouiller, Dabas, Tabbagh, Perrard 2001] ROUILLER D., DABAS M., TABBAGH J., PERRARD C., "Definition of homogeneous agricultural areas with the MuCEP system", In Grenier G., Blackmore S., *Third European Conference on Precision Agriculture*, tome 1, p. 103-108, agro Montpellier, Montpellier, 2001.
- [Rousseau, Martel 1994] ROUSSEAU A., MARTEL J.M., "Environmental Assessment of an electric transmission line project : a MCDA method", In Paruccini M., *Applying Multiple Criteria Aid for Decision to Environmental Mangement*, p. 163-185, Kluwer Academic Publisher, Londres, Angleterre 1994.
- [Roy 1985] ROY B., *Méthodologie multicritère d'aide à la décision*, Economica, Paris, 1985.
- [Ruefenacht, Vanderzanden, Morrison, Golden 2002] RUEFENACHT B., VANDERZANDEN D., MORRISON M., GOLDEN M., *New Techniques for Segmenting Images*, 2002.
- [Schärlig 1985] SCHÄRLIG A., *Décider sur plusieurs critères : panorama de l'aide à la décision multicritère*, Presses Polytechniques et Universitaires Romandes, Lausanne, Suisse, 1985.
- [Stafford, Lark 1997] STAFFORD J.V., LARK R.M., "Classification as a first step in the interpretation of temporal and spatial variation of crop yield", In *Annals of Applied Biology*, n° 130, p. 111-121, 1997.
- [Stafford, Lark, Bolam, ?] STAFFORD J.V., LARK R.M., BOLAM H.C., "Using Yield Maps to Regionalize Fields into Potential Managements Units", In ?, ?.
- [Sudduth, Drummond, Birrell, Kitchen 1996] SUDDUTH K.A., DRUMMOND S.T., BIRRELL S.J., KITCHEN N.R., "Analysis of Spatial Factors Influencing Crop Yield", In Robert P.C., Rust R.H., Larson W.E., *Proceedings of the Third International Conference on Precision Agriculture*, p. 129-140, ASA-CSSA-SSSA, Madison, 1996.
- [Suhaedi, Metternicht, Lodwick 2001] SUHAEDI E., METTERNICHT G., LODWICK G., "Integrated multicriteria decision making and geographic information systems for spatial planning of rural areas", In *Proceedings of the Third International Conference on Geospatial Information in Agriculture and Forestry*, Veridian, Ann Arbor, Etats-Unis, 2001.
- [Thévenet 1997] THEVENET G., "Quels enjeux pour l'agriculture de précision ?", In *Agriculture de précision, cap sur le XXIème siècle*, supplément à *Perspectives Agricoles*, n° 222, p. 5-7, mars 1997.

[Thévenet 1999] THEVENET G., "L'agriculture de précision en France : état de la situation et questions posées", In *Conférence-débat du S.I.M.A. sur L'enjeu français de l'agriculture de précision*, p. 15-17, 1999.

[Triboulet, Lardon, Langlet 1996] TRIBOULET P., LARDON S., LANGLET A., "Segmentation de textures pour l'identification d'espaces ruraux différenciés", In *Etude des phénomènes spatiaux en agriculture*, p. 47-64, I.N.R.A. Editions Les Colloques, Paris, 1996.

[Van Alphen, Stoorvogel 1999] VAN ALPHEN B.J., STOORVOGEL J.J., A Methodology to Define Management Units in Support of an Integrated, Model-Based Approach to Precision Agriculture, 1999.

[Varma, Ferguson, Wild 1998] VARMA V.K., FERGUSON I.S., WILD I., "Decision support system for the sustainable forest management", In *Proceedings of the First International Conference on Geospatial Information in Agriculture and Forestry*, tome 1, p. 170-177, ERIM International, Ann Arbor, 1998.

[Young, Gerbrands, Van Vliet 1998] YOUNG I.T., GERBRANDS J.J., VAN VLIET L.J., *Image Processing Fundamentals*, Delft University of Technology, Pays-Bas, 1998.

[Zimmermann, Plöchl, Luckhaus, Domsch 2003] ZIMMERMANN H.M., PLOCHL M., LUCKHAUS C., DOMSCH H., "Selecting the optimum locations for soil investigations", In Stafford J.V., Werner A. *Precision Agriculture*, p. 759-764, Wageningen Academic Publishers, Wageningen, Pays-Bas, 2003.

RESUME

Dans le domaine agricole, l'application du principe du développement durable s'est faite en même temps que la prise de conscience des limites de l'agriculture productiviste pour donner naissance à l'agriculture durable. Cette dernière englobe toute une série d'approche de la conduite des cultures dont les plus connues sont l'agriculture biologique et l'agriculture raisonnée.

Une autre approche, l'agriculture de précision, s'appuie sur de l'information géographique à très grande échelle et des outils adaptés (Systèmes d'Information Géographique et analyse spatiale) pour gérer la variabilité intraparcellaire afin d'optimiser les coûts d'exploitation et de limiter les impacts environnementaux des opérations culturales. La pratique de l'agriculture de précision, en particulier la mise en oeuvre opérationnelle de la modulation des intrants, nécessite le découpage de la parcelle en quelques grandes entités homogènes.

Une méthode de zonage a donc été développée. Dans un premier temps, des tests ont été effectués sur la base de l'Aide MultiCritère à la Décision. Mais cette approche a été abandonnée pour des raisons d'applicabilité. La méthode développée au final s'appuie sur une classification mixte intégrant plusieurs contraintes spatiales et permettant la prise en compte du savoir de l'exploitant agricole. Elle permet d'obtenir rapidement et simplement des zonages satisfaisants tout en offrant quelques indicateurs de validité.

Une étude visant à modéliser la structure spatiale de données onéreuses en utilisant d'autres l'étant nettement moins a été menée. Nous avons montré qu'avec notre méthode de zonage, il était possible d'approcher correctement la structure spatiale de plusieurs variables agronomiques pertinentes en se basant sur une combinaison de cartes de résistivité électrique, de rendement et d'images satellitaires.

MOTS-CLÉS : Analyse spatiale, zonage, agriculture durable, agriculture de précision, entité homogène.

ABSTRACT

In the agricultural field, the principle of sustainable development began to be applied when productivist agriculture showed its limits. This gave rise to sustainable agriculture which includes a whole series of approach to control the cultures among which biological agriculture and reasoned agriculture are the most known.

Another approach, precision agriculture, relies on very large scale geographical information and adapted tools (Geographical Information Systems and spatial analysis) to manage in-field variability in order to optimize exploitation costs and to limit the environmental impacts of the farming operations. The practice of precision agriculture, in particular the operational implementation of the modulation of inputs, requires the division of the parcel into a few big management zones.

A zoning method thus has been developed. Initially, tests were carried out on the basis of MultiCriteria Decision Making. But this approach was abandoned for reasons of applicability. The method finally developed is based on a hybrid clustering integrating several spatial constraints and allowing to take into account the farmer's knowledge. It makes it possible to obtain quickly and simply satisfactory zonings while offering some validity indicators.

A study aiming at modelling the spatial structure of expensive data while using others cheaper was conducted. We have demonstrated that with our zoning method, it was possible to correctly approach the spatial structure of several relevant agronomic variables using a combination of electrical resistivity, yield and satellite imagery maps.

KEYWORDS : Spatial analysis, zoning, sustainable agriculture, precision agriculture, management zone.