



HAL
open science

Utilisation de l'information incertaine dérivée d'une base de données sols. Application à la cartographie des propriétés hydriques à l'échelle régionale

D.R. Cazemier

► To cite this version:

D.R. Cazemier. Utilisation de l'information incertaine dérivée d'une base de données sols. Application à la cartographie des propriétés hydriques à l'échelle régionale. Sciences du Vivant [q-bio]. Ecole Nationale Supérieure Agronomique, 1999. Français. NNT: . tel-02834337

HAL Id: tel-02834337

<https://hal.inrae.fr/tel-02834337v1>

Submitted on 7 Jun 2020

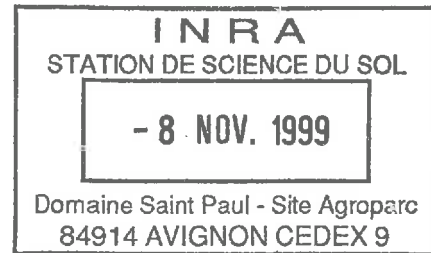
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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

TH-HB50

MINISTERE DE L'AGRICULTURE
ECOLE NATIONALE SUPERIEURE AGRONOMIQUE DE MONTPELLIER

T H E S E



présentée à l'Ecole Nationale Supérieure Agronomique de Montpellier
pour obtenir le diplôme de DOCTORAT en SCIENCE DU SOL

Ecole Doctorale Biologie des Systèmes Intégrés –Agronomie – Environnement

Utilisation de l'information incertaine dérivée d'une base de données sols

Application à la cartographie des propriétés hydriques à l'échelle régionale

par

Durk R. CAZEMIER

Soutenue le 12 février 1999 devant le jury composé de:

Prof. Dr.	J.C. REMY	ENSA Montpellier	Président
Dr.	M. BORNAND	INRA Montpellier	Dir. de thèse
Prof. Dr.	P.A. BURROUGH	Universiteit Utrecht (Pays-Bas)	Rapporteur
Prof. Dr.	J. DESACHY	Université P. Sabatier Toulouse	Rapporteur
Dr.	D. KING	INRA Orléans	Examineur
Dr.	P. LAGACHERIE	INRA Montpellier	Examineur
Dr.	R. MARTIN-CLOUAIRE	INRA Toulouse	Examineur

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
LABORATOIRE DE SCIENCE DU SOL
MONTPELLIER

Avant propos

Le travail de recherche présenté dans ce document n'aurait pu être effectué sans le soutien et l'aide d'un certain nombre de personnes que je tiens à remercier ici.

En premier lieu, j'exprime ma gratitude à Philippe Lagacherie et à Roger Martin-Clouaire qui ont assumé l'encadrement quotidien de cette thèse. J'ai pu bénéficier pendant ces trois dernières années de leur grande disponibilité, leur volonté de coopération et leurs critiques, conseils et encouragements, particulièrement pendant la période de rédaction. Cette thèse me semble être le fruit d'une parfaite collaboration entre chercheurs de laboratoires INRA différents. Encore merci.

Ensuite, je remercie Michel Bornand, directeur de recherche INRA, qui a dirigé cette thèse. Son soutien moral ainsi que ses corrections minutieuses de l'écrit ont été de grande valeur pour l'achèvement de cette thèse.

J'exprime également ma reconnaissance à Jean-Claude Rémy, professeur science du sol à l'ENSA de Montpellier, qui a bien voulu présider le jury de la thèse. Je le remercie également pour toutes les démarches qu'il a entreprises afin d'obtenir la dérogation de DEA, qui m'a permis de débiter la thèse directement après mes études à Utrecht.

"Een bijzonder woord van dank aan" Peter Burrough professeur en géographie physique à l'université d'Utrecht (NL) et rapporteur de cette thèse. Je le remercie beaucoup d'avoir accepté d'être rapporteur et d'avoir jugé le travail avec autant d'attention. En envoyant un de ses étudiants à Montpellier, il a contribué aux échanges franco néerlandais dans le domaine de la "fuzzy" science du sol. Merci également à Jacky Desachy, professeur à l'université Paul Sabatier à Toulouse, pour avoir accepté de lire et de juger ce document. Etant donné sa position externe à la science du sol, en général, et au laboratoire en particulier, j'estime ses critiques et ses remarques de grande valeur.

Je suis également reconnaissant à Dominique King, directeur de recherche à l'INRA d'Orléans, pour avoir suivi le travail de l'extérieur, au cours des réunions de pilotage, et d'avoir accepté de participer au jury de la soutenance. Il a témoigné de son grand intérêt sur le sujet de la thèse et il a pu influencer le travail grâce à quelques conseils pertinents.

Après les remerciements envers les membres de jury, je tiens à remercier un certain nombre de personnes avec lesquelles j'ai eu le plaisir de collaborer.

Je tiens à remercier Jean-Claude Arvieu de m'avoir accueilli au laboratoire de Science du Sol de l'INRA de Montpellier où j'ai pu travailler dans des conditions favorables. Je re-

mercie également Marc Voltz, capitaine de l'équipe "spatialisation", qui m'a fait bénéficier de son attitude critique tout au long de mon travail et de ses conseils au cours de quelques réunions.

Ensuite, je remercie Gwenn Trotoux pour le "coup de main" durant la collecte des échantillons, pendant les semaines chaudes de '96. Les siestes à l'ombre, au bord de l'Hérault, étaient méritées et indispensables pour un déroulement efficace de ce travail musclé.

Je remercie Mark Rounsevell, coordinateur du projet IMPEL, qui, pendant les réunions du projet, m'a fait bénéficier de son expérience sur les projets de recherches européens et de ses conseils en ce qui concerne mon travail de thèse. Il a d'ailleurs témoigné d'une grande connaissance de l'art de vivre méditerranéen et d'un intérêt particulier envers le vin languedocien. Je remercie également les autres membres du projet pour les discussions intéressantes et constructives pendant les réunions.

Les problèmes informatiques ont été nombreux pendant ma thèse. Plusieurs personnes m'ont aidé sur des aspects différents. En premier lieu, je remercie Jean-Pierre Rellier et Robert Faivre du laboratoire de biométrie et d'intelligence artificielle de l'INRA à Toulouse. Le premier pour les modifications de certains bugs dans Conflex et le deuxième pour l'aide à l'écriture du modèle de la simulation de Monte Carlo.

Ensuite, je remercie Christian Floure pour la mise en page de certaines illustrations de la thèse et Pierrot Falipou pour les aides aux problèmes informatiques et pour l'introduction dans le monde des bases de données sols. Je suis également très reconnaissant envers Jean-Christophe Fabre, informaticien au laboratoire: Conflex sous Linux était effectivement la solution. Je n'oublie pas de remercier Catherine Mazzoni, Jean-Marc Robbez-Masson. Je les remercie pour l'assistance aux problèmes liés à l'informatique et liés au SIG.

En venant d'un autre pays d'Europe, les affaires administratives sont nombreuses. Je tiens à remercier Jeannette Salze, sans qui je n'aurais pas pu faire cette thèse, car elle a assumé toutes les affaires administratives nécessaires à mon séjour au laboratoire.

Le séjour au laboratoire a été rendu agréable par l'ambiance chaleureuse et jeune. Je pense en particulier ici aux stagiaires: Sebastien, Armel, Guillaume, Gaël, Rosa-Anna, Hugues, François, Philippe, Nicolas (j'en oublie sûrement), aux thésards: Xavier, Isabelle, Sylvie, Walter, Safar, Valérie et Franck et aux post-docs: Maria et Bernd. Je fais également un coucou au reste du personnel du laboratoire et je les remercie pour leur gentillesse et la collégialité au cours des trois dernières années.

Pour finir, je tiens à remercier quelques personnes classées "extragrades". Tout d'abord, je remercie Cesare Tassinari. Les nombreuses discussions romagnoles-bataves sur les selles de vélo ou autour d'une table couverte de petits délices italiens ont été précieuses dans ma vie de thésard. J'ai beaucoup apprécié cette hospitalité, ce soutien moral et cette volonté de donner un coup de main en période surchargée. Je tiens également à remercier Ca-

mille et Fabrice pour leur amitié et leur hospitalité vis à vis d'un étranger. Les parties de cartes et de grillades dans la garrigue constituaient la relaxation indispensable pendant les week-ends. Tom, compatriote, merci pour les pauses café, "de koekjes", et les aides diverses durant la dernière partie de la thèse. Ensuite, je remercie Jean-François, Coralie, Cécile Philippe, William, Odile, Serge, Les Kadors, les "footeurs". Tous ces gens ont rendu le séjour à Montpellier plus qu'agréable.

Bijna aan het eind van deze "avant propos", een woord van dank aan mijn ouders, Anne, Otto, Kirsten en André. Ook jullie hebben door de morele steun op afstand een steentje bijgedragen aan deze intellectuele inspanning.

Enfin, la personne qui m'a supporté, aidé et encouragé le plus, Vanessa, merci!

Montpellier, printemps 1999

Durk Cazemier

Sommaire

Avant propos	iii
Sommaire	vii
Introduction générale.....	1
1 Cadre général de l'étude.....	5
1.1 Applicabilité des différents types de cartographie de sol pour la présente étude.....	6
1.1.1 Estimation spatiale des propriétés hydriques en utilisant l'échantillonnage systématique	6
1.1.2 Dériver les propriétés hydriques du sol à partir des cartes pédologiques existantes.....	7
1.1.2.1 Concepts cartographiques utilisés.....	7
1.1.2.2 Bases de données sols et systèmes d'information géographique.....	9
1.1.2.3 Détermination des propriétés hydriques du sol à l'aide des fonctions de pédotransfert.....	9
1.1.3 Utilisation de cartes pédologiques à petite échelle	11
1.2 La stratégie proposée pour la cartographie des propriétés hydriques à l'échelle régionale.....	13
2 Analyse de la procédure d'estimation des propriétés hydriques.....	17
2.1 Incertitude et imprécision	18
2.2 Origines de l'incertitude dans l'information pédologique.....	19
2.2.1 Les caractéristiques naturelles du sol.....	19
2.2.2 Méthodes d'observation du sol.....	20
2.2.3 Quantification de l'incertitude dans l'information pédologique	21
2.3 Analyse de l'information contenue dans la base de données régionale.....	22
2.3.1 Représentation de l'imprécision sur la délimitation des unités cartographiques de sols.....	22
2.3.2 Incertitude liée à la délimitation géographique des UTS	23
2.3.2.1 Les données caractérisant l'environnement des UTS	24
2.3.2.2 Incertitude associée aux données géographiques supplémentaires	24
2.3.3 Différents modes de description des propriétés de sol d'une UTS	27
2.3.3.1 Représentation des UTS au moyen d'un profil représentatif	27
2.3.3.2 Description textuelle	28
2.3.3.3 Description des propriétés du sol par profil de référence	29
2.3.4 Incertitude liée à l'estimation des propriétés hydriques	30
2.3.4.1 Incertitude liée à une régression multiple linéaire	31
2.3.4.2 Incertitude liée à la représentativité des FPT	31
2.4 Propagation d'incertitude	32
2.4.1 Propagation de l'incertitude dans l'estimation de la réserve utile en eau.....	32
2.4.2 Evaluation de l'incertitude dans la cartographie des UTS	33
2.5 Conclusion.....	34

3	Représentation et traitement de l'information incertaine.....	35
3.1	Deux cadres théoriques possibles	36
3.2	Application d'une approche statistique et probabiliste.....	37
3.2.1	Représentation de la connaissance en distributions de probabilité	38
3.2.2	Propagation de l'incertitude avec la simulation Monte Carlo.....	39
3.3	Les sous-ensembles flous et la théorie des possibilités	40
3.3.1	Définitions des sous-ensembles flous	40
3.3.2	Opérations élémentaires avec les sous-ensembles flous	42
3.3.3	Relations floues	43
3.3.4	La théorie de possibilité	44
3.3.5	Propagation de l'imprécision à travers des relations floues.....	47
3.3.5.1	Calcul d'intervalles flous	47
3.3.5.2	Calcul d'intervalles flous avec des variables interactives.....	48
3.4	Interaction probabilité - possibilité.....	50
3.4.1	Comparaison des deux mesures.....	50
3.4.2	Transformations de distributions de probabilité et de distributions de possibilité.....	50
3.5	Application des sous-ensembles flous dans les domaines pédologiques et géographiques.....	52
3.6	Conclusion.....	54
4	La région d'étude et les données utilisées	55
4.1	Présentation de la zone d'étude.....	56
4.1.1	Les vallées de l'Hérault, de l'Orb et du Libron	56
4.1.1.1	Localisation et généralités géographiques	56
4.1.1.2	Principales entités géologiques.....	58
4.1.1.3	Principaux types de sol	59
4.1.2	Le secteur de référence Adissan-Cazouls-Montagnac	62
4.2	Présentation des données utilisées.....	63
4.2.1	L'information pédologique régionale	63
4.2.2	Les fonctions de pédotransfert.....	64
4.2.3	Information disponible exprimant d'autres relations entre les propriétés du sol.....	65
4.2.4	Les données géographiques supplémentaires utilisées	66
4.2.4.1	Les cartes géologiques	66
4.2.4.2	L'information topographique.....	67
4.2.4.3	Le modèle numérique d'altitude (MNA).....	67
4.2.5	Les données disponibles pour la validation	67
4.2.5.1	Observations ponctuelles des propriétés hydriques du sol	68
4.2.5.3	Carte contenant la localisation exacte des UTS	69

5	Représentation de la connaissance	71
5.1	Représentation de la connaissance intervenant dans l'estimation de la réserve utile	72
5.1.1	Interprétation et représentation de la description textuelle	72
5.1.1.1	<i>Texture</i>	74
5.1.1.2	<i>Profondeur du sol</i>	76
5.1.1.3	<i>Pierrosité</i>	77
5.1.1.4	<i>Matière Organique</i>	77
5.1.2	Interprétation des descriptions par profils de référence	78
5.1.3	Densité apparente.....	78
5.1.4	Contraintes exprimant les interactions entre variables dans l'estimation des propriétés hydriques	80
5.1.4.1	<i>Interactions entre propriétés du sol appartenant à un même horizon</i>	80
5.1.4.2	<i>Interactions entre propriétés appartenant à deux horizon voisins</i>	82
5.1.5	Représentation de l'incertitude liée aux régressions	84
5.2	Représentation de la connaissance dans la cartographie des UTS.....	85
5.2.1	Interprétation et représentation de l'information issue de la base de données sols	86
5.2.2	Détermination des limites floues des cartes choroplèthes utilisées	87
5.2.3	Quantification de l'incertitude sur la pente	90
5.3	Conclusion.....	91
6	Estimation des propriétés hydriques par UTS.....	93
6.1	Mise en œuvre des modèles.....	94
6.1.1	Résolution de CSP avec CON'FLEX.....	94
6.1.1.1	<i>Données d'entrées du problème</i>	94
6.1.1.2	<i>Formulation des requêtes</i>	96
6.1.1.3	<i>Paramétrage de la recherche des solutions avec CON'FLEX</i>	97
6.1.1.4	<i>Résolution des CSP et temps de calcul</i>	99
6.1.2	Calcul d'intervalles flous	100
6.1.3	Simulation Monte Carlo.....	100
6.1.2.1	<i>Création de la fonction de répartition à partir d'une fonction de densité de probabilité pour le tirage aléatoire</i>	101
6.1.2.2	<i>Prise en compte des contraintes</i>	102
6.2	Protocoles d'évaluation.....	103
6.2.1	Expérimentations effectuées	103
6.2.2	Méthodes d'évaluation.....	104
6.3	Comparaison des différentes méthodes d'estimation des propriétés hydriques du sol	105
6.3.1	Présentation et interprétation des résultats bruts	106
6.3.2	Comparaison des estimations avec des mesures de terrain	108
6.3.3	Conclusion	111
6.4	Comparaison des estimations issues de différentes données	112
6.4.1	Analyse graphique des résultats bruts	112
6.4.2	Analyse des résultats à l'aide des deux critères d'évaluation.....	115
6.4.3	Différences de précision des estimations issues des cartes pédologiques à différentes échelles...	118
6.5	Conclusion.....	120

7	Cartographie des unités typologiques de sols	121
7.1	Le problème de la localisation des UTS au sein des UCS.....	122
7.2	Méthodes et techniques appliquées pour la localisation des UTS	122
7.2.1	Localisation par appariement partiel et opération ensembliste	123
7.2.2	L'utilisation d'un SIG en mode raster.....	124
7.3	Protocole d'évaluation	127
7.3.1	Données de validation disponibles.....	128
7.3.2	Série d'expérimentations à effectuer.....	128
7.3.3	Trois mesures quantitatives pour évaluer la qualité de la localisation des UTS	129
7.3.3.1	<i>Proportion de pixels bien classés</i>	<i>129</i>
7.3.3.2	<i>Imprécision de la localisation des UTS.....</i>	<i>130</i>
7.3.3.3	<i>Cohérence de l'estimation.....</i>	<i>131</i>
7.4	Résultats obtenus pour les trois mesures d'évaluation.....	131
7.4.1	Proportion d'estimations correctes	131
7.4.2	Evaluation de l'imprécision sur la localisation	133
7.4.3	L'évaluation de la cohérence de l'information pour chaque combinaison.	134
7.4.4	Evaluation globale en combinant les critères d'évaluation	134
7.5	Conclusion et discussion sur la localisation des UTS	136
8	Cartographie des propriétés hydriques	139
8.1	Estimation spatiale des propriétés hydriques	140
8.1.1	Intégration de la première et deuxième étape par combinaison disjonctive.....	140
8.1.2	Protocole d'évaluation des estimations spatiales	141
8.1.3	Résultats de l'évaluation des estimations spatiales.....	141
8.2	Représentation cartographique de l'estimation spatiale.....	143
8.2.1	Détermination de la position des résultats par rapport à une valeur seuil.....	143
8.2.2	Représentation de l'intervalle par α -coupe dans un seul pixel	145
8.3	Conclusion.....	146
	Conclusion générale	151
	Bibliographie	157
	Glossaire.....	165
	Liste des figures	167
	Liste des tableaux	169
	Annexes	171

Introduction générale

La connaissance des propriétés hydriques du sol est une information essentielle pour les agriculteurs et pour les aménageurs, ainsi que pour les scientifiques dans différents domaines de recherche agronomique et environnementale comme, par exemple, l'évaluation des impacts d'un changement climatique sur le rendement des cultures agricoles (Rounsevell *et al.*, 1998) et la modélisation hydrologique (Moussa *et al.*, 1998). Ces utilisateurs sont impliqués dans des projets et des études à caractère géographique. L'information recherchée ne concerne généralement pas la valeur d'une propriété hydrique en un point précis, mais plutôt la répartition de cette propriété au sein d'un ensemble géographique assez vaste. C'est pourquoi, il est important de pouvoir restituer cette information avec une dimension spatiale, c'est-à-dire sous la forme d'une cartographie de propriétés hydriques du sol.

La surface sur laquelle porte cette cartographie influence la disponibilité des données et les méthodes cartographiques à appliquer. La cartographie des propriétés hydriques pour une surface réduite (par exemple une parcelle) permet d'effectuer un échantillonnage systématique suivi d'une analyse géostatistique (Voltz *et al.*, 1997). Ces méthodes s'avèrent moins adaptées pour des surfaces plus importantes, du fait des contraintes méthodologiques, de la lourdeur des analyses et du coût qu'elles entraînent. L'alternative à la cartographie systématique, présentée par de nombreux auteurs, réside dans l'utilisation de l'information sur les sols issue des études pédologiques existantes. L'estimation des propriétés hydriques pour chaque unité de sol définie sur la carte s'effectue généralement par l'application des fonctions de pédotransfert à l'information issue d'un profil de sol qui représente au mieux cette unité de sol (i.e. le profil représentatif).

Lorsque la région d'intérêt correspond à un vaste territoire (région, pays), la seule information disponible sur la variabilité des sols est constituée par les cartes pédologiques à petite échelle. L'utilisation de ces cartes selon l'approche mentionnée ci-dessus devient problématique. D'une part, les profils caractérisant les unités de sol ont, à cette échelle, une représentativité douteuse. De ce fait, les estimations de propriétés hydriques utilisant ces profils deviennent très incertaines et ne représentent pas la variabilité interne des unités de sol. D'autre part, les unités de sol des cartes pédologiques à petite échelle sont généralement complexes, c'est à dire qu'elles regroupent plusieurs unités de sol. En conséquence, l'appartenance d'un site donné à une unité de sol n'est pas connue avec précision à partir de

la carte à petite échelle ce qui constitue une nouvelle source d'incertitude sur la cartographie des propriétés hydriques.

Pour estimer les propriétés hydriques des unités de sol en tenant compte de leur variabilité interne et pour localiser ces unités de sol au sein d'unités cartographiques complexes, il est possible d'exploiter d'autres informations qui décrivent globalement les propriétés (ex: texture, profondeur,...) et l'environnement (ex: pente, géologie,...) des unités de sol. Ces informations se présentent sous la forme de descriptions qualitatives de l'unité de sol ou sous la forme d'intervalles de valeurs donnant une estimation d'une propriété. Il s'agit donc d'informations imprécises qui sont plus complexes à manipuler qu'un ensemble de propriétés attachées à un profil représentatif. L'absence d'approches permettant de manipuler cette imprécision a, jusqu'à présent, fortement limité l'utilisation de ces informations pour cartographier les propriétés des sols sur une région.

L'objectif du travail de recherche qui va être présenté dans ce mémoire est d'utiliser les descriptions imprécises des propriétés et de l'environnement des unités de sol d'une carte à petite échelle pour cartographier les propriétés hydriques des sols sur une région. L'approche envisagée s'appuie sur la théorie des possibilités pour (i) représenter l'information imprécise d'une carte pédologique à petite échelle, (ii) cartographier les unités de sol au sein des unités cartographiques complexes et (iii) estimer les propriétés hydriques de chaque unité de sol. L'approche est appliquée à la cartographie de l'humidité au point de flétrissement du sol et à la réserve utile en eau du sol sur la région "vallées de l'Hérault-Orb-Libron" située dans la plaine languedocienne, au sud de la France.

L'organisation du mémoire est la suivante :

Le **premier chapitre** expose le cadre général de l'étude, avec définition des objectifs et de la démarche générale proposée pour une cartographie régionale des propriétés hydriques. Le **deuxième chapitre** analyse en détail les différentes étapes de la cartographie des propriétés hydriques, afin de déterminer les différentes sources d'incertitude pouvant intervenir. Le **troisième chapitre** aborde le choix des cadres théoriques possibles: l'approche statistique dans laquelle l'incertitude est exprimée par la mesure de probabilité ou la théorie des possibilités basée sur la théorie des sous-ensembles flous.

Le **quatrième chapitre** présente la zone d'étude choisie ainsi que les données réellement disponibles dans cette zone. Le **cinquième chapitre** présente ensuite l'interprétation et la représentation de la connaissance en accord avec la théorie des possibilités qui a été choisie.

La cartographie des propriétés hydriques est donc réalisée selon les trois étapes principales de la démarche générale adoptée dans la présente étude. Les trois derniers chapitres de ce mémoire exposent successivement les résultats obtenus et une évaluation de ces résultats.

1 Cadre général de l'étude

Il existe un besoin important d'information spatiale concernant le sol dans les disciplines touchant à la recherche environnementale et dans les domaines de l'aménagement des territoires agricoles. Les cartes pédologiques fournissent suffisamment d'information si l'on s'intéresse à l'organisation spatiale du sol en général. Cependant, dans la plupart des cas, l'information demandée concerne plutôt une propriété du sol spécifique. Dans ces cas, les cartes pédologiques classiques et les bases de données associées ne fournissent pas directement l'information demandée, mais elles servent comme source de données, à partir de laquelle l'information recherchée peut-être dérivée.

La présente étude s'intéresse à la cartographie des propriétés hydriques du sol à l'échelle régionale. La principale contrainte de ce type de cartographie est l'importance de la surface sur laquelle porte la cartographie. Cette ambition concernant la surface à couvrir limite, pour des raisons évidentes de coût, l'échelle des prospections et la disponibilité des données nécessaires à la cartographie. Elle détermine également les méthodes et les techniques appropriées à l'analyse des données et à la représentation cartographique. C'est pour ces raisons que la cartographie d'une propriété du sol à l'échelle régionale constitue un sujet qui nécessite une étude approfondie concernant les données disponibles et les méthodes de cartographie à utiliser. Dans l'étude qui suit, il est donc important d'avoir une vue globale sur les potentialités des différents types de cartographie avant d'élaborer la démarche à suivre pour la cartographie des propriétés hydriques du sol à l'échelle régionale.

Ce chapitre a pour objectif de justifier le choix de la démarche adoptée dans ce travail pour cartographier les propriétés hydriques du sol à l'échelle régionale. On présentera les différentes alternatives pour cartographier les propriétés hydriques du sol: applicabilité de la cartographie systématique des propriétés hydriques et possibilité d'exploiter l'information de cartes déjà existantes. Cette analyse permet d'établir la stratégie adoptée dans la suite du travail pour cette cartographie.

1.1 Applicabilité des différents types de cartographie de sol pour la présente étude

Les développements techniques et méthodologiques de ces dernières décennies dans le domaine de la cartographie du sol ont entraîné une diversité importante de types de cartographie du sol (Legros, 1996). Chaque type de cartographie pédologique, général ou thématique, a pour objectif de représenter la variabilité spatiale du phénomène étudié de la façon la plus réaliste possible sur la base des données disponibles, tout cela en tenant compte de l'échelle cartographique. Il est ainsi possible de caractériser les différents types de cartographie selon leur utilisation finale, l'échelle cartographique et les données disponibles.

La cartographie d'une propriété hydrique à l'échelle régionale est fortement problématique compte tenu de la difficulté à obtenir, sur une aussi vaste surface, un nombre suffisant de mesures permettant de caractériser la variabilité régionale. Le choix du type de cartographie des propriétés hydriques à l'échelle régionale sera donc déterminé par la quantité et la qualité des données disponibles à cette échelle. L'analyse qui suit permet de préciser les choix retenus.

1.1.1 Estimation spatiale des propriétés hydriques en utilisant l'échantillonnage systématique

L'utilisation des méthodes cartographiques systématiques pour l'estimation spatiale des propriétés hydriques a été présentée par de nombreux auteurs (Mallants *et al.*, 1996; Scheinost *et al.*, 1997a,b; Voltz & Goulard, 1994; Voltz *et al.*, 1997). Ces exemples utilisent des mesures de terrain des propriétés hydriques pour la quantification d'une propriété hydrique sur un ensemble de points dans la zone d'étude. La prise des échantillons et l'analyse au laboratoire sont des étapes délicates et doivent respecter certains protocoles pour éviter l'introduction d'erreurs importantes (Soil Survey Division Staff, 1993; Baize & Jabiol, 1995). L'intérêt de l'échantillonnage sur le terrain est la production de données objectives sur le comportement hydrique du sol. Le grand inconvénient de l'échantillonnage et de la mesure des propriétés hydriques du sol est la lourdeur et le coût des travaux de terrain et des analyses au laboratoire.

Les analyses statistiques et géostatistiques (i.e. méthode d'interpolation comme le krigeage) sont ensuite utilisées pour estimer les propriétés hydriques aux endroits non échantillonnés. Les méthodes géostatistiques ont été décrites dans le cadre de la théorie des variables régionalisées (Matheron, 1965 ; Journel & Huijbregts, 1978 ; Webster & Oliver, 1990). Cette théorie est basée sur le fait que la variation d'une propriété du sol dans l'espace peut être modélisée par une fonction continue et aléatoire exprimant la dépendance spatiale.

La possibilité d'appliquer les méthodes géostatistiques à l'échelle régionale est fortement réduite à cause de deux facteurs. Premièrement, l'hypothèse fondamentale de stationnarité est discutable à cette échelle. Cette hypothèse repose en effet sur le principe que, à partir du moment où la structure globale de la variation spatiale est prise en compte, la variation restante est homogène et la différence entre deux points géographiques est seulement fonction de la distance entre ces deux points (exprimé par le variogramme expérimental¹). Or, le milieu naturel (géologie, sol et végétation) à l'échelle régionale montre généralement trop d'hétérogénéité pour que la condition de stationnarité de la variation spatiale soit acceptable. Deuxièmement, un découpage de la zone en question en plusieurs sous-régions, satisfaisant chacune l'hypothèse de stationnarité, serait une solution. Mais, ceci nécessite un grand nombre d'observations nécessaires pour pouvoir bien estimer le variogramme expérimental pour chaque sous région. La lourdeur des analyses des propriétés du sol et le coût financier de celle-ci rendent quasi impossible le recours à un grand nombre d'échantillons.

Pour ces deux raisons (hypothèse de stationnarité non valide globalement et lourdeur des analyses), nous n'avons pas retenu la voie d'une analyse statistique suivie d'une approche géostatistique pour la cartographie des propriétés hydriques du sol à l'échelle régionale.

1.1.2 Dériver les propriétés hydriques du sol à partir des cartes pédologiques existantes

L'alternative choisie dans le cadre de ce travail consiste à estimer les propriétés hydriques recherchées à partir de l'information contenue dans les cartes pédologiques. Dans la suite sont développés, les concepts cartographiques utilisés dans cette approche, la formalisation de l'information issue d'une carte pédologique au moyen de systèmes d'information géographiques et au moyen de base de données et le recours à de fonctions de pédotransfert utilisées pour estimer des propriétés du sol non couramment déterminées dans les études de sol classiques.

1.1.2.1 Concepts cartographiques utilisés

Les cartes pédologiques présentent des échelles cartographiques très variables représentant le sol de manière différente. Pour faire la distinction entre les différents types de cartes pédologiques, il est tout d'abord nécessaire de définir quelques termes ou concepts de

¹ Un variogramme expérimental exprime la variance d'une propriété du sol entre les points d'échantillonnage en fonction de la distance entre ces points. Un nombre important d'observations est nécessaire pour pouvoir bien déterminer le variogramme pour une propriété du sol dans une zone d'étude (Webster & Oliver, 1990).

cartographie pédologique. La connaissance pédologique présente dans les cartes et les bases de données est structurée selon une certaine logique. Dans la littérature, il existe différentes façons de représenter et de hiérarchiser l'information, c'est-à-dire les différents types d'unités de sols et les volumes caractéristiques distinguables au sein de la couverture pédologique (Baize, 1986).

La notion d'horizon est connue et utilisée depuis très longtemps. Il s'agit d'une subdivision de la couverture pédologique en volumes sensiblement homogènes, constituant des couches le plus souvent parallèles à la surface (Girard, 1983; Girard, 1984; Jamagne, 1993). Le découpage du sol en un ou plusieurs horizons est le résultat d'une observation du sol à l'aide d'une fosse ou d'un sondage à la tarière. Ce découpage s'opère par combinaison d'un certain nombre de critères sélectionnés. Dès qu'un caractère jugé suffisamment important se modifie d'une façon notable, un horizon peut être distingué. L'ensemble des horizons décrits dans une fosse constitue le profil pédologique.

L'horizon constitue l'unité de base de la caractérisation verticale de la couverture pédologique. La cartographie pédologique consiste à déterminer l'extension spatiale des horizons (Baize, 1986). Pour pouvoir déterminer et représenter cette organisation tri-dimensionnelle il est nécessaire de simplifier la réalité pédologique. Cette simplification s'effectue grâce au concept des unités typologiques de sols (UTS). Une unité typologique de sols est l'objet conceptuel qui rend compte d'un ensemble d'observations de sol considérées comme homogènes. Une UTS est alors constituée par la superposition ou l'agencement d'horizons aux caractéristiques bien définies. Sa définition repose sur le rattachement à un classement le plus souvent de type référentiel comme une classification taxonomique (AFES-INRA, 1992; Jamagne, 1993). Les unités typologiques de sols sont principalement caractérisées par les horizons et leur superposition au sein de l'unité. D'autres propriétés concernent les éléments paysagers comme la géologie et la topographie souvent utilisés dans la définition des UTS.

Dans une étude pédologique classique, la description des unités de sol d'une carte s'effectue de deux manières différentes:

- a) dans la légende et la notice associée à la carte. L'UTS est décrite qualitativement. Cette description qualitative représente le synthèse des observations (sondages, profils et analyses) réalisées sur une UTS.
- b) un profil représentatif est choisi au sein de l'UTS. C'est à partir de ce profil que sont estimées quantitativement les propriétés du sols. Lorsque l'UTS est suffisamment importante, plusieurs profils peuvent la caractériser ce qui donne la possibilité de décrire ces propriétés sous forme d'intervalle de valeurs.

1.1.2.2 Bases de données sols et systèmes d'information géographique

Une étude pédologique résulte généralement du recueil d'une quantité d'information importante et de nature variable (description des horizons et des UTS, analyses au laboratoire des échantillons de sol, carte). Pour éviter une perte de l'information recueillie et pour optimiser l'exploitation de cette information, il est important de pouvoir la stocker de façon optimale et de faciliter son utilisation à l'aide d'un support adéquat (Gaultier *et al.*, 1992). Traditionnellement, la carte pédologique est dessinée sur support papier accompagnée d'une légende et d'une notice. Depuis l'existence des bases de données informatisées et des systèmes d'information géographique, les données pédologiques sont généralement représentées et stockées sous un format informatisé, facilement accessible à l'utilisateur (Bornand *et al.*, 1994; Brown *et al.*, 1987; Gaultier *et al.*, 1993; King *et al.*, 1995).

Ces outils informatiques permettent de stocker les données de manière efficace et de représenter l'information cartographique de manière flexible. Généralement, l'information codée en machine existe en parallèle avec l'information conventionnelle (carte papier + notice). Elle est généralement composée d'une partie géométrique et d'une partie sémantique. La partie géométrique, c'est-à-dire l'organisation géographique des unités cartographiques de sols, est intégrée dans un SIG et elle est représentée en mode vecteur. Les unités cartographiques de sols sont alors représentées sous forme de polygones. L'information sémantique, c'est-à-dire, la description des horizons, des UTS est stockée dans une SGBDR (Système de Gestion de Bases de Données Relationnelles). Cet outil fournit une organisation de fichiers qui permet de recueillir les descriptions des entités pédologiques quelle que soit l'échelle de travail considérée et la nature des descriptions (description qualitative, intervalle de valeurs, profil représentatif).

Il existe un fort développement des bases de données sols nationales et continentales (Le Bas & Jamagne, 1996; Heineke *et al.*, 1998). L'application de ces bases de données se trouve principalement dans le domaine de l'évaluation de terre aux niveaux régional, national et continental (Hubrechts *et al.*, 1998; Proctor *et al.*, 1998; Vossen & Meyer-Roux, 1995).

1.1.2.3 Détermination des propriétés hydriques du sol à l'aide des fonctions de pédotransfert.

La plupart des propriétés hydriques, nécessaires à l'utilisation de modèles mécanistes de bilan hydrique, demandent des déterminations lourdes qui ne sont généralement pas effectuées dans une étude pédologique classique. Pour avoir une estimation de ces propriétés à partir des informations usuellement contenues dans une carte pédologique, plusieurs auteurs (Bastet *et al.*, 1998; Bouma & Van Lanen, 1987; Van Genuchten & Leij, 1992; Tomasella & Hodnett, 1998) ont proposé l'utilisation de fonctions de pédotransfert à partir de sites-

références sur lesquels sont précisées les propriétés recherchées.

Ces fonctions expriment la relation mathématique existant entre une propriété ou un comportement du sol et des caractéristiques de constitution de ce sol. De nombreux travaux concernant l'établissement des fonctions de pédotransfert ont été menés depuis le début des années 80. Ils ont abouti à la création de plusieurs types de fonctions de pédotransfert différents (Bastet *et al.*, 1998). Dans la littérature, on distingue deux types principaux (Van Genuchten & Leij, 1992; Bastet *et al.*, 1998). Le premier type exprime la relation entre les propriétés facilement déterminables et les paramètres d'une équation mathématique. C'est par exemple le cas des fonctions de pédotransfert qui déterminent les paramètres d'une courbe de rétention en eau (Vereecken *et al.*, 1989; Wösten *et al.*, 1995). Le deuxième type de fonctions de pédotransfert exprime la relation entre les mêmes propriétés du sol et les valeurs de rétention en eau pour différents niveaux de potentiel matriciel (Gupta & Larson, 1979; Rawls *et al.*, 1982).

Les fonctions de pédotransfert sont généralement établies à partir d'un ensemble de données quantitatives. Ces données sont généralement obtenues par les analyses d'une série d'échantillons provenant des horizons de certains types de sols. Dans ce cas, les données contiennent, d'un côté des mesures de la variable dépendante (la propriété hydrique pour laquelle on cherche à établir la fonction de pédotransfert) et de l'autre côté, des mesures des variables explicatives (granulométrie, densité apparente, taux de matière organique). La relation mathématique entre les variables à estimer et les variables indépendantes a, pour les deux types, généralement la forme d'une régression multiple linéaire (Bastet *et al.*, 1998). C'est à ce niveau que les données présentes dans une base de données sols peuvent s'avérer utiles. En effet, ces bases contiennent généralement de multiples exemples d'associations entre les variables explicatives et la variable dépendante.

Il existe un certain nombre d'avantages et d'inconvénients attachés à l'application de fonctions de pédotransfert pour l'estimation des propriétés hydriques du sol. L'avantage principal est la réduction du nombre d'observations et d'analyses à faire dans le cas où on dispose d'une fonction de pédotransfert précise et fiable correspondant bien au type de sol en question. Les inconvénients sont liés d'une part aux limites de représentativité de la fonction en dehors de la zone pour laquelle la fonction a été établie et d'autre part, à l'incertitude entachant les paramètres. L'application des fonctions de pédotransfert est donc contraignante du fait de ces facteurs et requiert une utilisation prudente.

Lorsque les données disponibles ne permettent pas de calculer des fonctions de pédotransfert, l'alternative est dans l'établissement des règles de pédotransfert (Van Ranst *et al.*, 1995). Ces règles correspondent à une approche principalement qualitative et combinent de manière booléenne des informations pédologiques classiques comme par exemple le type de roche mère et la classe texturale pour la détermination de certaines propriétés du sol peu disponibles, par exemple, la densité apparente du sol, la CEC ou la réserve utile. La prédiction

de la propriété du sol recherchée s'effectue par l'affectation d'une classe prédéfinie (faible, moyenne, ou forte) dans le domaine de la variable concernée; ce qui correspond à une représentation très grossière de l'information. La présente étude a pour objectif de prédire les propriétés hydriques avec plus de précision à l'aide des fonctions de pédotransfert qui sont disponibles pour la zone d'étude.

1.1.3 Utilisation de cartes pédologiques à petite échelle

L'objectif étant de cartographier les propriétés hydriques des sols sur de vastes surfaces en Europe, il est indispensable de s'assurer que les informations pédologiques utilisées dans la démarche sont disponibles. Ceci amène à écarter à priori les cartes pédologiques à moyenne échelle car seuls certains petits pays Européens ont une couverture systématique de leur territoire à cette échelle (Magaldi, 1995). La seule alternative demeure l'utilisation de cartes à petite échelle.

La cartographie du sol à petite échelle présente différents niveaux de rendus possibles. Il est nécessaire de faire une distinction entre les cartes à très petites échelles (1:1.000.000 – 1:5.000.000) et les cartes à échelles régionales (1:100.000 – 1:500.000). Les cartes à très petites échelles (CEC, 1985; FAO-UNESCO, 1974) couvrent des continents entiers et représentent un découpage pédologique très grossier, ayant surtout une valeur pédagogique ou didactique. L'information pédologique est très imprécise et pas du tout adaptée à une estimation des propriétés hydriques à l'échelle régionale. Ce type de cartes pédologiques ne peut donc pas, à priori, être pris en considération pour notre travail.

Les cartes pédologiques à l'échelle régionale contiennent en général une information pédologique beaucoup plus précieuse. Il existe un nombre important de programmes de cartographie à l'échelle régionale et en particulier à l'échelle du 1:250.000 (Le Bas & Jamagne, 1996; Bornand *et al.*, 1989; Favrot *et al.*, 1994; Heineke *et al.*, 1998). La cartographie à cette échelle utilise des unités cartographiques complexes où les associations de sols sont représentées par des unités qui sont le résultat de l'intégration d'un ensemble d'horizons pédologiques et qui sont homogènes au niveau de certains éléments paysagers (substrat géologique, hydrologie, géomorphologie, occupation végétale). Ces associations de sols ont été appelées des Unités de Pédopaysages (Bornand *et al.*, 1989). Elles constituent les unités cartographiques de sol (UCS).

Une unité cartographique de sols (UCS) est le regroupement d'une ou plusieurs UTS de façon à pouvoir en faire une représentation cartographique à une échelle choisie (Jamagne, 1993). Elle constitue donc une unité sémantique rassemblant un ensemble de plages cartographiques. Dans tous les cas, une UCS est définie géographiquement par la délimitation d'une ou plusieurs plages cartographiques (surface fermée dont le contenu sémantique est

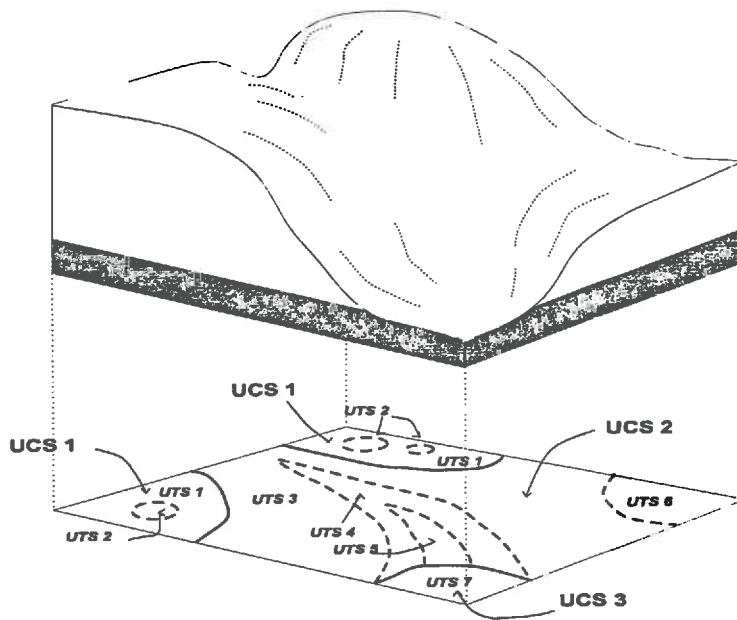


Figure 1.1: Schéma d'organisation des entités cartographiques pour une carte pédologique à l'échelle régionale.

considéré comme homogène).

La Figure 1.1 expose les caractéristiques cartographiques de la base de données sol régionale et elle montre la hiérarchie dans l'organisation des données. Dans cet exemple fictif, trois UCS ont été distinguées par le pédologue [les sommets (UCS 1), les pentes (UCS 2) et la vallée qui est creusée dans un matériau géologique différent (UCS 3)]. Au sein des UCS, le pédologue a pu distinguer les différences pédologiques en créant des UTS (indiqué par les lignes pointillées dans la figure 3). Ces UTS ne sont pas représentées sur la carte (Bornand *et al.*, 1989; Jamagne *et al.*, 1995). Ce type de cartographie permet de représenter le sol au niveau des régions européennes et elle augmente ainsi considérablement la disponibilité des données pédologiques en général, sur une région européenne donnée. La disponibilité des données constitue l'avantage principal de l'utilisation des cartes pédologiques à l'échelle régionale.

L'application de tels types de cartes n'est cependant pas sans contraintes. Deux problèmes principaux sont attachés à l'utilisation de ces cartes: (i) la difficulté de représenter des UTS par des profils représentatifs et (ii) leur absence de localisation géographique.

L'influence de l'échelle cartographique sur la qualité de l'estimation de propriétés hydriques à partir des profils représentatifs des cartes à grande et moyenne échelle a été étudié par Bregt & Beemster (1989) et Leenhardt *et al.* (1995). Ces deux études démontrent une croissance d'erreur d'estimation lorsque l'échelle cartographique diminue (de 1:10.000 au 1:100.000). Leenhardt *et al.* (1995) montrent que l'erreur d'estimation augmente pour les cartes pédologiques d'une échelle inférieure à 1:25.000 dans le cas des sols de la région Languedoc-Roussillon. Ceci est dû à (i) l'importance de la variabilité intra-unité, ce qui

augmente l'erreur commise lorsqu'on utilise une valeur moyenne de propriété pour toute l'unité et (ii) la difficulté de pouvoir estimer une valeur moyenne à partir d'un profil représentatif car le choix du profil représentatif s'appuie sur un nombre d'observations de terrain plus limité lié à l'échelle d'observation imposée. Le profil représentatif n'est donc pas une forme de description qui semble adapté aux caractéristiques de l'UTS définies à cette échelle.

Comme il est indiqué plus haut, la plupart des unités cartographiques de sols sont complexes. Ceci implique une indétermination, en un point donné, quant aux valeurs estimées pour chaque unité typologique de sols potentiellement présente. Ngongo *et al.* (1995) et Le Bas *et al.* (1998) ont essayé de déterminer l'erreur provoquée par cette confusion. Ngongo *et al.* (1995) ont essayé d'agréger l'information de plusieurs unités cartographiques (d'une carte pédologique au 1:1.000.000) pour une surface régulière définie par la cellule d'une grille (20x20 km et 40x40 km). Trois pondérations différentes sont proposées et la moyenne des valeurs de toutes les UTS présentes dans la cellule est considérée comme la meilleure estimation de la variable à cartographier. Cette estimation est peu réaliste, car elle gomme toute l'hétérogénéité existante à l'intérieur de la cellule.

Les deux contraintes évoquées ci-dessus nous amènent à effectuer les choix suivants.

- a) Les profils représentatifs ne peuvent pas être utilisés comme information descriptive de l'unité typologique de sols à cette échelle. Il est nécessaire de faire appel à d'autres sources d'information pour l'estimation des propriétés hydriques comme (i) la description textuelle basée sur l'expérience de prospection du pédologue et (ii) les données exprimant les propriétés du sol sous forme d'intervalles de valeurs basés sur tous les profils disponibles pour l'UTS.
- b) En l'absence de localisation des UTS fournie par la carte, il est nécessaire d'estimer leur localisation au sein des UCS à partir de la description de leur environnement (géologie, occupation du sol, relief,...)

Le point commun entre ces deux aspects est qu'ils nécessitent un développement d'approches permettant la manipulation d'une information imprécise associée à chaque UTS.

1.2 La stratégie proposée pour la cartographie des propriétés hydriques à l'échelle régionale

Le problème à résoudre consiste à estimer une propriété hydrique du sol à un endroit donné, dans une région donnée. Ce problème peut se découper en trois étapes (cf. Figure 2.1). La première étape (a) consiste à estimer les propriétés hydriques pour chaque unité typologique du sol décrite dans la zone d'étude. L'estimation des propriétés hydriques

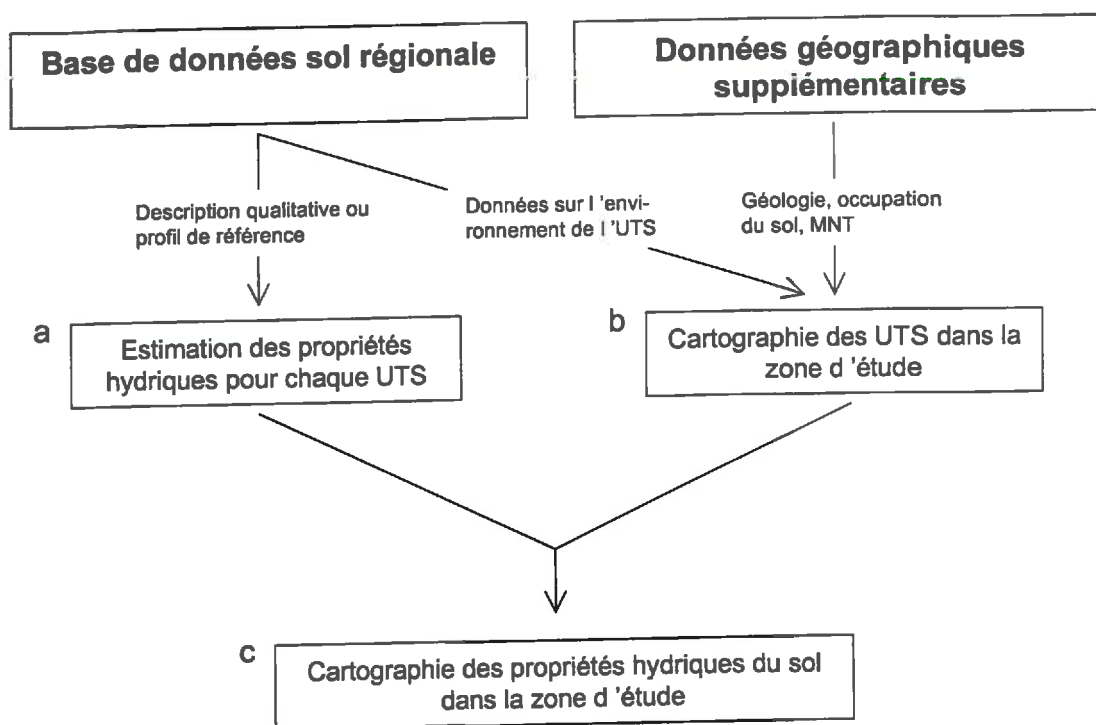


Figure 1.2: La démarche proposée pour l'estimation des propriétés hydriques à partir d'une carte pédologique régionale. Trois étapes principales ont été définies (a, b et c). Voir texte pour plus d'explications.

s'effectue au niveau de l'unité typologique car c'est à ce niveau qu'est associée la description des propriétés morphologiques des sols rencontrés. Cependant, l'information attachée aux unités typologiques de sols contient rarement des données concernant les propriétés hydriques du sol. Mais en utilisant des fonctions de pédotransfert, il est possible de dériver une estimation des propriétés souhaitées.

La deuxième étape (b) est nécessaire pour déterminer à quelle unité typologique du sol on peut rattacher un point dans l'espace. Étant donné que les unités cartographiques sont des unités complexes où la localisation des UTS n'est pas connue, ce point peut appartenir à un ensemble d'UTS. En utilisant l'information géographique supplémentaire, il sera possible de mieux positionner les UTS au sein des UCS. Cette analyse est basée sur des relations entre le sol et le paysage (Hudson, 1992). Il s'agit de prendre en compte ici les facteurs paysagers pour lesquels une information supplémentaire est disponible et qui interviennent dans la formation des sols et leur répartition (Jenny, 1941). Ces éléments sont principalement la lithologie, la morphologie et l'occupation du sol. Un problème essentiel consiste à exprimer cette relation quantitativement. Un nombre important d'auteurs a essayé d'utiliser les éléments paysagers pour la cartographie du sol en général à différentes échelles cartographiques avec des méthodes quantitatives différentes (De Bruin & Stein, 1998; Dalal-Clayton & Robinson, 1992; Lagacherie, 1992; Odeh *et al.*, 1992; Robbez-Masson *et al.*, 1996;

Webster & Burrough, 1972).

Dans la présente étude, on essaie d'exploiter des relations entre l'occurrence d'une unité typologique de sols et les facteurs qui composent et structurent les paysages. Ces relations sont dérivées exclusivement de l'information contenue dans la base régionale des données sols où chaque description des unités typologiques de sols contient l'information concernant les variables environnementales. Le positionnement des UTS au sein des UCS s'effectue ensuite grâce à la comparaison d'éléments donnés par les relations logiques avec l'information géographique supplémentaire, généralement plus précise.

Les deux étapes a et b étant réalisées, l'étape c consiste à combiner les deux résultats produits pour cartographier la propriété hydrique du sol dans la zone d'étude.

L'étape b du schéma présenté dans la Figure 1.2 est nécessaire pour les cartes pédologiques dont les unités cartographiques de sols sont constituées d'associations de sols. Cette étape peut être éliminée dans le cas où l'on dispose d'une carte pédologique dont les unités cartographiques de sols correspondent à des unités typologiques de sols. La cartographie des propriétés hydriques devient alors une estimation fondée sur une fonction de pédotransfert et sur les données extraites de la base de données sols comme proposé dans l'étape a.

2 Analyse de la procédure d'estimation des propriétés hydriques

Le précédent chapitre a exposé l'intérêt de l'utilisation d'une carte pédologique à l'échelle régionale pour l'estimation des propriétés hydriques. Une démarche générale a été proposée pour effectuer cette estimation en tenant compte des caractéristiques principales de la carte pédologique à l'échelle régionale. L'information disponible intervenant dans cette estimation est bien souvent imparfaite à l'égard de l'objectif. Il est essentiel d'en tenir compte dans l'estimation et il est nécessaire pour cela de caractériser la qualité de l'information.

Il est donc important d'analyser en détail les sources d'incertitude pesant sur l'estimation des propriétés hydriques. La présence d'incertitude affecte le mode de représentation des données dans la base de données ainsi que les traitements appliqués à ces données pour obtenir le résultat final.

Ce chapitre a ainsi pour objectif d'examiner les données nécessaires et disponibles pour chaque étape de la procédure d'estimation proposée, en insistant sur les caractères imprécis ou incertain de l'information. Un paragraphe introductif sera consacré aux notions d'imprécision et d'incertitude. Les implications de ces notions dans le domaine de la pédologie en général seront présentés dans le paragraphe 2.2. Le paragraphe 2.3 exposera successivement l'imprécision associée à l'aspect cartographique, l'organisation hiérarchique des données, les différents modes de description des UTS et l'incertitude liée à la prédiction des propriétés hydriques des UTS. Cette analyse permettra d'exposer dans le paragraphe 2.4 les problèmes rencontrés au moment de la mise en œuvre de la cartographie des propriétés hydriques à l'échelle régionale.

2.1 Incertitude et imprécision

La connaissance dont dispose une personne sur une situation donnée est généralement imparfaite. Pour exprimer le manque de connaissance, la personne utilise un certain vocabulaire. Le langage courant et celui des sciences de la représentation de la connaissance utilisent les termes "incertains" et "imprécis" pour exprimer la qualité d'une information donnée. Il est important, dans le cadre de cette étude, de connaître exactement la signification des termes "imprécision et incertitude". Les ouvrages (Dubois & Prade, 1988) et (Klir & Yuan, 1995) constituent des références adéquates sur le sujet.

Tout d'abord, il est nécessaire d'introduire la notion de proposition. Une proposition est une spécification de la valeur d'une variable. Une proposition est imprécise quand la spécification est incomplète et insuffisante pour l'usage qu'on voudrait en faire. Généralement, l'imprécision concerne le contenu de la proposition et se manifeste par une hésitation face à plusieurs alternatives entre lesquelles on ne peut pas trancher. La proposition n'est alors pas restreinte à un seul élément. Par exemple, si on sait que l'âge de la voiture est entre 10 et 14 ans, la variable *âge de la voiture* peut prendre toutes les valeurs entre 10 et 14. On ne pourra donc pas répondre à toutes les questions relatives à cet âge de la voiture (par exemple, "est-ce que la voiture a plus de 12 ans?").

Il existe des cas où la proposition est imprécise, et où l'ensemble des alternatives possibles pour la valeur de la variable n'est pas bien définie comme par exemple, "l'âge de la voiture est environ 13 ans". Les propositions de ce type sont qualifiées de floues ou vagues et sont donc des propositions imprécises particulières.

La certitude d'une proposition qualifie la vérité ou la fausseté de la proposition. Une proposition incertaine est alors une proposition pour laquelle la vérité ou fausseté n'est pas complètement établie. La notion de vérité fait référence à ce qui est connu de la réalité (basé sur des observations réelles ou des données supposées) (Badia & Martin-Clouaire, 1989). L'expression de l'incertitude d'une proposition précise, c'est-à-dire, une proposition dont la variable sous-jacente n'a qu'une valeur possible, s'effectue par le rattachement d'un degré de confiance exprimant jusqu'où il est cru que la proposition est fausse ou vraie.

Présenté plus haut comme deux notions différentes il est important d'illustrer la relation entre incertitude et imprécision. Une propriété s'exprime à travers la valeur d'une variable. Une proposition imprécise peut ainsi être vue comme une disjonction de propositions précises mais incertaines. Par exemple, la proposition "la voiture de Pierre est récente" est équivalente à la disjonction de "la voiture a 1 an", "la voiture a 2 ans", ..., "la voiture a 6 ans", chacune étant pondérée par un degré d'incertitude qualifiant la plus ou moins grande compatibilité avec le terme "récente". En conséquence, parler d'imprécision revient à parler de l'incertitude induite sur un ensemble d'alternatives.

Ceci implique, pour notre étude, que les formulations imprécises des connaissances pédologiques au moyen d'intervalles de valeurs, peuvent être considérées comme l'expression de ce que le pédologue connaît de l'incertitude des informations qu'il fournit.

2.2 Origines de l'incertitude dans l'information pédologique

L'utilisation de l'information pédologique à petite échelle nécessite une analyse préalable des différentes sources d'incertitude qui pèsent sur les données et sur les connaissances à manipuler. La qualité de l'information pédologique est liée aux deux composantes principales communément admises lorsqu'on aborde des données géographiques. Il s'agit de la composante géométrique et de la composante sémantique. De nombreux auteurs ont étudié la qualité de ces deux composantes dans le cadre de l'information spatiale en général et dans le cadre des SIG en particulier (Goodchild & Gopal, 1989; Burrough & McDonnell, 1998; Frank, 1998). L'information spatiale est assujettie à un grand nombre de sources d'incertitude. Ces sources varient de l'observation des phénomènes géographiques jusqu'à la représentation de l'information cartographique. L'information présentée sur les cartes pédologiques et dans les bases de données est également affectée par ces mêmes sources d'incertitude.

De façon générale, la cartographie pédologique à grande échelle permet de représenter le sol de manière plus détaillée qu'une carte à petite échelle. Dans le langage courant, on dit que ces cartes à grande échelle sont plus *précises* que les cartes à petite échelle pour dire que les cartes à grande échelle s'approchent plus de la réalité que les cartes à petite échelle. L'analyse de l'incertitude et des sources d'incertitude dans l'information pédologique constitue un sujet délicat (Beckett & Webster, 1971; Legros, 1978a,b; Marsman & De Gruijter, 1986). Les sources d'incertitude dans l'information pédologique peuvent être analysées selon l'effet qu'elles induisent sur la pureté et l'homogénéité (Bascomb & Jarvis, 1967). La pureté exprime jusqu'à quel niveau la définition des plages cartographiques correspond à la définition donnée dans la légende de la carte. L'homogénéité exprime la variation des caractères du sol à l'intérieur d'une unité cartographique de sols. Ce paragraphe expose donc les facteurs principaux affectant la pureté et l'homogénéité, c'est-à-dire, l'incertitude des cartes pédologiques.

2.2.1 Les caractéristiques naturelles du sol

L'information pédologique disponible dans une carte de sol et dans une base de données sol est une représentation d'une connaissance concernant un phénomène géographique

tri-dimensionnel. Cette information ne reflète jamais la réalité exacte parce que le pédologue est obligé de schématiser la variabilité spatiale et verticale respectivement en unités typologiques de sols et en horizons. L'incertitude affectant les horizons et les unités cartographiques de sols dépend, à densité d'observation égale, des caractéristiques naturelles de la zone d'étude. Ceci peut être illustré par deux exemples. Premièrement, une variabilité spatiale importante d'une caractéristique naturelle du sol provoque généralement une faible homogénéité des unités typologiques de sols définies (Beckett & Webster, 1971). Ceci provoque une incertitude élevée concernant l'information représentée sur la carte. Deuxièmement, le pédologue utilise des relations entre des phénomènes de surface (i.e. végétation, topographie) et la présence de certains types de sol pour interpoler entre ces observations de terrain et localiser les limites des unités cartographiques de sols (Hudson, 1990). L'utilisation de cette information associée aux caractéristiques naturelles du sol augmente généralement la certitude, avec laquelle on décrit les unités cartographiques de sols et elle réduit l'incertitude sur le positionnement des limites de la cartes.

Un dernier aspect naturel du sol affectant l'incertitude sur l'information pédologique concerne la transition entre deux unités typologiques de sols voisines ou un horizon et son horizon sous-jacent. Cette transition est souvent de caractère graduel (Burrough, 1989; Lagacherie *et al.*, 1996; McBratney & Odeh, 1997). Ceci provoque des difficultés pour la détermination des limites des unités cartographiques sur la carte et les limites des horizons dans les profils. L'information pédologique contient alors une incertitude sur le positionnement exact des limites sur la carte et sur la définition des horizons. La pureté de la carte sera alors d'autant plus faible que cette incertitude sera élevée.

2.2.2 Méthodes d'observation du sol

L'incertitude de l'information cartographique est influencée par la densité des observations faites sur le terrain. Une densité d'observation élevée permet de déterminer de façon plus détaillée l'organisation spatiale des unités typologiques de sols et l'homogénéité des propriétés du sol au sein des UTS. Toutefois, les méthodes d'observation ne sont pas infaillibles et ont une capacité limitée à informer sur ce qui est recherché. Elles ne renvoient donc jamais la vérité complète et introduisent souvent une erreur dans la localisation d'un objet ou une incertitude sur la vraie valeur de l'attribut (Burrough & McDonnell, 1998). Les principales sources d'incertitude proviennent des erreurs des instruments de mesure utilisés sur le terrain pour mesurer certaines propriétés de sol ou pour localiser l'endroit de l'observation. Les mesures effectuées au laboratoire sur les échantillons pris sur le terrain introduisent également une incertitude. L'incertitude liée à ces méthodes d'observation s'exprime généralement par un intervalle de confiance établi à l'aide des données statistiques.

L'observation par expert constitue une autre façon d'observer le sol. La procédure à suivre par l'expert pour observer le sol est généralement bien définie dans les guides de description de sol (Baize & Jabiol, 1995; Hodgson, 1974; Soil Survey Division Staff, 1993). Pourtant, le facteur humain est important dans ce type d'observations où l'intuition personnelle joue un grand rôle. Les représentations cartographiques des mêmes sols effectuées par différents experts montrent généralement de fortes différences (Bie & Beckett, 1971; Legros, 1996). Une incertitude importante est alors introduite dans l'information cartographique du sol.

2.2.3 Quantification de l'incertitude dans l'information pédologique

Les deux paragraphes précédents ont montré la sensibilité de la cartographie de sol en général à l'introduction d'incertitude. Une information s'avère inutile si le degré de vérité attaché à cette information est inconnu. Il est donc nécessaire de pouvoir quantifier l'incertitude sur l'information cartographique disponible. Le nombre d'auteurs qui ont essayé de quantifier l'incertitude dans les cartes pédologiques classiques est relativement réduit (Beckett & Webster, 1971; Marsman & De Gruijter, 1986; Walter, 1990; Leenhardt *et al.*, 1994; Rogowski, 1996). Ces auteurs ont essayé de déterminer la pureté et l'homogénéité des cartes de manière statistique au moyen des méthodes statistiques disponibles comme l'analyse de variance, par exemple (Webster & Oliver, 1990). Ceci permet, par exemple, de quantifier l'homogénéité d'une propriété du sol grâce à la variabilité intra-unité ou grâce au rapport entre la variabilité globale et la variabilité intra-unité.

Les travaux menés sur la quantification de l'incertitude concernent, du fait du grand nombre d'observations de terrain nécessaire, les cartes pédologiques à grande et moyenne échelle sur des surfaces réduites. La quantification de l'incertitude des cartes à petites échelles en utilisant une analyse statistique n'est pas envisageable. Il convient alors de trouver d'autres moyens pour quantifier l'incertitude associée à la carte pédologique à petite échelle.

D'autres caractérisations se sont limitées à l'appréciation d'une propriété de sol. Ainsi, il existe quelques études sur la vérification de l'appréciation de la texture (De Leenheer *et al.*, 1954; Hodgson *et al.*, 1976). Pour vérifier la qualité d'une appréciation, il suffit de reporter dans le triangle de texture, le point fourni par les analyses sur la composition granulométrique de l'échantillon sur lequel le test a porté (Legros, 1996). Par classe texturale prédite sur le terrain, il est ensuite possible, dans le cas où il existe suffisamment de données, d'établir un nuage de points. Ce nuage montre ensuite si les points sont bien situés à l'intérieur de la classe texturale ou si les points sont dispersés à l'extérieur de la classe.

2.3 Analyse de l'information contenue dans la base de données régionale

Nous avons vu que l'information pédologique était affectée par plusieurs sources d'incertitude. Maintenant nous examinons, au travers de l'exemple de la base de données régionale, comment cette incertitude se matérialise sur les données qui sont nécessaires pour estimer les propriétés hydriques. Pour chaque ensemble de données, on recherche systématiquement les 'méta-données' permettant d'estimer, même de façon approximative, leur incertitude. En particulier, nous considérons les formulations imprécises de l'information qui, traduites en intervalles de valeurs, expriment la connaissance incomplète du pédologue et permettent de quantifier l'incertitude sur ce qui en est extrait.

Cette analyse est divisée en quatre parties. Premièrement, l'information nécessaire est une information cartographique. En cela, l'information à traiter est concernée par l'imprécision sur le positionnement des limites. Il est donc important d'identifier ici ce qui permet de borner cette imprécision. La deuxième partie concerne l'organisation hiérarchique des données dans la base de données. Cette organisation introduit le problème de la non localisation précise des unités typologiques de sols. Le paragraphe 2.3.2 expose l'organisation des données dans une base de données régionale et les données nécessaires pour la localisation des UTS à partir des informations secondaires. La troisième partie de l'analyse expose les différents modes de description des unités typologiques de sols. La quatrième et dernière partie concerne l'incertitude associée à l'estimation des propriétés hydriques pour chaque unité typologique de sols à l'aide des fonctions de pédotransfert.

Ce chapitre utilise des exemples tirés de la base de données du Languedoc-Roussillon (Bornand *et al.*, 1989; Bornand *et al.*, 1994). Les caractéristiques et les problèmes rencontrés ici sont cependant transférables à d'autres situations et à d'autres bases de données régionales (King *et al.*, 1994; Le Bas & Jamagne, 1996).

2.3.1 Représentation de l'imprécision sur la délimitation des unités cartographiques de sols

L'information concernant la distribution spatiale du sol est représentée dans les cartes pédologiques à petite échelle par des UCS (cf. § 1.1.2). Les unités cartographiques de sols sont représentées sous la forme d'une carte choroplèthe, c'est-à-dire, sous la forme d'un ensemble de plages cartographiques donnant une seule valeur pour une plage cartographique (Burrough & McDonnell, 1998). Les plages cartographiques sont différenciées à l'aide d'une série de couleurs et symboles et les limites entre les plages cartographique se présentent sous la forme d'un trait en principe infiniment étroit. L'utilisation d'un tel type de carte provoque

une incertitude liée à la densité d'observations caractérisant la carte. Ici, on s'intéresse d'abord à la précision du positionnement des limites.

Deux facteurs sont en jeu: l'échelle de représentation cartographique et la nature de la transition naturelle entre les unités de pédopaysages. En ce qui concerne l'influence de l'échelle cartographique, l'imprécision sur le positionnement des limites est plus importante sur une carte à petite échelle que sur une carte à grande échelle (Beckett & Webster, 1971; Marsman & De Gruijter, 1986). Par exemple, un millimètre d'écart sur une carte au 1:250.000 entre la vraie limite et la limite dessinée sur la carte correspond à 250 m sur le terrain comparé avec 25 m pour une carte au 1:25.000. L'imprécision de la première carte est alors plus importante. Cette imprécision est néanmoins quantifiable. Il suffit d'avoir, pour une partie de la zone d'étude, une carte détaillée avec le positionnement précis des limites. La distance moyenne entre ces limites précises et celles données sur la carte à petite échelle peut être une mesure de l'erreur (Dunn *et al.*, 1990; Bolstadt *et al.*, 1990).

La transition souvent graduelle entre les unités cartographiques de sols des cartes utilisées est également à prendre en compte dans la détermination de l'imprécision du positionnement des limites sur la carte. La quantification de cette imprécision est néanmoins difficile. Lagacherie *et al.* (1996) utilisent pour la largeur maximale de la transition une valeur donnée par le pédologue. Cette valeur maximale permet une représentation imprécise des limites de sol sous forme de "zones-tampons". Dans la suite de ce travail, on propose un tel type d'approche dans lequel une représentation imprécise de la limite permettra de représenter les incertitudes liées d'une part, à la transition naturelle entre les UCS et d'autre part, à l'effet d'échelle cartographique.

2.3.2 Incertitude liée à la délimitation géographique des UTS

Compte tenu de l'extrême imprécision pesant sur les limites des UTS à l'échelle régionale, la cartographie du sol à cette échelle régionale utilise une représentation géographique imprécise en appliquant le principe des associations de sol: les unités cartographiques de sols (UCS) sont composées d'une ou plusieurs unités typologiques de sols (UTS). Pour localiser les UTS au sein des UCS, il est cependant possible d'utiliser l'information caractérisant l'environnement de l'UTS dans la mesure où les données d'environnement sont souvent disponibles sur l'ensemble de la région. Dans la suite nous étudions ces données et nous analysons l'incertitude liée à leur utilisation.

UTS	Description
102	Sol formé dans les grands méandres et les berges de l'Orb et de l'Hérault. Ripisylve, vignes, aspergeraies. D'amont en aval, le pH des sols se modifie (les sables quartzeux cévenols s'enrichissant progressivement en éléments calcaires).
303	Sol développé sur les terrasses moyennes de l'Hérault et de l'Orb, dans son cours médian. Vignes dominantes.
345	Ce sol est observé sur des reliquats d'anciens placages pliocènes ou villafranchiens; vestige d'un ancien appareil fluvial très développé.
974	Sol développé sur molasse, grès et calcaire gréseux à passées de lumachelles de l'Eo-Oligo-Miocène. Pente faible. Vigne. Carcassès, Biterrois.

Tableau 2.1: La description textuelle concernant les données de l'environnement pour 4 UTS dans la base de données sol au 1/250.000 du Languedoc-Roussillon (Bornand et al. 1992).

2.3.2.1 Les données caractérisant l'environnement des UTS

La nécessité d'introduire des données caractérisant l'environnement des UTS est fondée sur l'existence de relations entre type de sol et éléments de paysage. Ces données se localisent au niveau de la description textuelle des UTS. Le Tableau 2.1 montre quelques exemples de ce type d'information qui renseigne sur la formation géologique, la lithologie, l'occupation du sol, la végétation, la morphologie et la pente. Ces exemples illustrent deux aspects importants de cette information. Premièrement, les propositions sont imprécises (par ex. pente faible). Elles sont globales et relativement peu informatives. Deuxièmement, ces propositions sont incomplètes, c'est-à-dire, qu'il existe des variables renseignées pour une UTS mais qui ne le sont pas pour d'autres.

2.3.2.2 Incertitude associée aux données géographiques supplémentaires

Pour pouvoir exploiter les caractéristiques d'environnement de chaque UTS présentées ci-dessus, il convient d'avoir des données supplémentaires correspondant aux variables géographiques considérées. Elles concernent:

- a) L'information géologique disponible dans les cartes géologiques à moyenne échelle.
- b) L'information topographique (végétation et occupation de sol) disponible dans les cartes topographiques à grandes et moyennes échelles.
- c) Des données sur l'altitude et la pente: modèle numérique d'altitude (MNA) dérivé de l'information topographique.

Voyons comment la qualité de ces données diffère pour chaque type d'information.

a. *Information géologique*

La formation géologique de surface constitue un facteur important dans le développement du sol. La nature de la roche est ainsi un facteur discriminant dans la définition des unités typologiques du sol. Une information exacte sur la géologie pourrait servir à déterminer l'organisation spatiale des UTS. Une carte géologique classique contient de l'information sur l'organisation spatiale des formations géologiques de surface et pourrait donc être utilisée pour cette analyse. Contrairement aux cartes pédologiques, les cartes géologiques à moyenne échelle sont beaucoup plus disponibles. La carte géologique au 1:50:000 (BRGM) est disponible pour la majeure partie du territoire français.

L'utilisation des cartes géologiques à moyenne échelle introduit de l'incertitude liée, en particulier, à l'imprécision de la carte elle-même. Elle concerne par exemple le positionnement des limites entre les unités géologiques dessinées sur la carte. La transition naturelle entre deux unités géologiques peut être graduelle tout comme la transition entre deux unités de sol. D'autres sources d'incertitude sont: l'incertitude sur le contenu des plages cartographiques et les erreurs liées au traitement informatique (digitalisation des cartes etc.).

Pour pouvoir utiliser les cartes géologiques dans la localisation des unités typologiques de sols, il est nécessaire d'avoir la correspondance entre l'information géologique disponible dans la description de l'UTS et l'information géologique de la carte géologique. La taxonomie utilisée par le pédologue pour décrire les formations géologiques associées à l'UTS ne correspond pas obligatoirement à celle de la légende de la carte géologique. Ceci nécessite de faire des appariements entre les deux types d'informations qui ne peuvent être que partiels. Par ailleurs, il est également nécessaire de rendre compte de l'imprécision sur les limites de la même façon que ce qui a été proposé pour les cartes des pédopaysages.

b. *Information topographique et occupation végétale*

Le sol et la végétation sont des phénomènes naturels interdépendants et cette relation est utilisée dans la cartographie du sol. La définition des UTS contient alors souvent un aspect concernant la végétation naturelle ou agricole. Une information exacte sur l'occupation du sol ou sur la végétation naturelle pourrait aider à la détermination de l'organisation spatiale des UTS. Cette information est principalement présentée sur des cartes de végétation et sur les cartes topographiques. En France, on dispose seulement d'une couverture systématique du territoire exploitant le concept de végétation potentielle. Mais, la disponibilité des cartes de végétation réelle est réduite, ce qui rend ce type d'information peu ou non exploitable pour cette étude.

La présente étude utilise des cartes topographiques à grande échelle (1:25.000) pour la dérivation des données sur l'occupation du sol et sur la végétation naturelle. La disponibilité de ce type d'information est grande du fait des travaux cartographiques importants effectués par les services de cartographie topographique nationaux comme l'Institut Géographique National (IGN). La qualité de cette information est élevée par rapport aux autres sources d'informations géographiques supplémentaires. L'erreur affectant les limites de ces cartes topographiques au 1:25.000 est de l'ordre de quelques mètres. Ceci est négligeable par rapport à l'incertitude des cartes géologiques à moyenne échelle et des cartes des pédopaysages à l'échelle régionale.

L'utilisation des cartes topographiques nécessite, comme pour l'information géologie, la traduction de l'information issue de la description textuelle en unité de légende de la carte. Ceci introduit une source d'incertitude, lorsque l'appariement entre les deux sources d'information n'est que partiel.

c. *Modèle numérique d'altitude*

Tout comme la géologie et la végétation naturelle, le relief est lié au développement du sol et vice versa. Le relief peut être exprimé en termes d'altitude, de pente et d'exposition. Il est possible d'obtenir l'information concernant ces variables à l'aide d'un modèle numérique d'altitude (MNA). Un MNA donne, pour une grille régulière, l'altitude en chaque point. La pente et l'exposition peuvent ensuite être calculées à partir de cette information. Le modèle numérique d'altitude est calculé à partir d'une série de points pour lesquels on connaît précisément l'altitude. Une interpolation est ensuite effectuée pour calculer l'altitude sur chaque point de la grille régulière.

Un modèle numérique d'altitude est une approximation de la réalité qui fournit une valeur précise pour l'altitude en chaque point. Cette valeur est plus ou moins fiable (i.e. certaine) pour deux raisons. Premièrement, l'incertitude est provoquée par la qualité des données avec lesquelles le MNA a été élaboré. L'altitude peut être mesurée avec une certaine précision en fonction des méthodes et instruments utilisés. La densité des points d'observation est également un aspect de la qualité des données d'entrée. Généralement, une faible densité d'observations provoque une forte incertitude du MNA. Deuxièmement, le calcul de l'altitude sur les points non observés (interpolation) provoque également une incertitude. Il existe plusieurs méthodes d'interpolation avec des caractéristiques diverses et une qualité du résultat différente. Les méthodes les plus souvent appliquées sont le krigeage, la triangulation et la distance inverse (Burrough & McDonnell, 1998).

La qualité d'un MNA est quantifiable avec une série de données de validation. Une série de données suffisamment grande permet de calculer une distribution statistique de la diffé-

rence entre la valeur mesurée et la valeur estimée. La quantification de l'incertitude permet ainsi de donner une distribution au lieu d'une valeur précise pour l'altitude.

2.3.3 Différents modes de description des propriétés de sol d'une UTS

Les unités typologiques de sols sont des objets complexes auxquels est associée une incertitude non négligeable, en particulier dans les bases de données régionales. Plusieurs alternatives existent pour décrire ces UTS. Elles diffèrent par la source d'information qu'elles utilisent et par leur degré de précision. Ici, on distingue trois types de données:

- a. La description au moyen d'un profil représentatif.
- b. La description textuelle de l'unité de sol.
- c. La description des propriétés du sol par profil de référence.

Les trois sections suivantes explicitent par type de données, les caractéristiques des données, l'origine des données, la possibilité de quantifier l'incertitude et l'applicabilité des données dans la présente étude.

2.3.3.1 *Représentation des UTS au moyen d'un profil représentatif*

La description des UTS par profil représentatif est fréquemment utilisée dans les bases de données sol. Ce type de représentation consiste à choisir un profil de sol parmi l'ensemble des profils étudiés comme le plus représentatif (Baize, 1986 ; Legros, 1996). Ce profil est donc un profil réellement observé à l'intérieur de l'unité typologique du sol et considéré par le pédologue pour représenter au mieux les caractéristiques moyennes de l'UTS.

Les informations relatives aux profils représentatifs sont généralement précises. Les horizons du sol sont décrits avec beaucoup de soin et les propriétés du sol à l'intérieur des horizons ont également été déterminées avec beaucoup de précision (surtout après analyse des échantillons au laboratoire). Cette précision des données s'accompagne cependant d'une incertitude importante, puisque la représentativité de l'information donnée est douteuse au regard de la variabilité spatiale à l'intérieur des UTS et ce d'autant plus que l'échelle de la carte considérée est petite.

La quantification de l'incertitude mentionnée ci-dessus est difficile pour les unités typologiques du sol à l'échelle régionale à cause de l'absence de données statistiques : le nombre d'observations de terrain par unité typologique de sols est réduite, ce qui ne permet pas d'estimer statistiquement une mesure de variabilité (Webster & Oliver, 1990). Dans ces cas, il

UTS	Description
102	<i>Sol très profond, à texture grossière (la fraction sableuse atteint 60 %), avec parfois, apparition à la base de galets de quartz. Sol brun, peu organique, faible capacité d'échange.</i>
303	<i>Sol profond, à forte charge caillouteuse dans une matrice limono-sablo-argileux à limono-argilo-sableux rouge. Teneur en carbonate total faible à forte (0 à 30 %) avec toutefois une teneur en calcaire actif faible à très faible. Bien pourvu en cations alcalino-ferreux ainsi qu'en phosphore et potassium.</i>
345	<i>Sol profond, argilo-sableux à limono-argileux dès la surface, à charge caillouteuse souvent importante (galets cristallins). Structure polyédrique à prismatique, rouge, pH acide à neutre. Forte capacité d'échange.</i>
974	<i>Sol moyennement profond (40-60 cm), limono-sableux, très calcaire (30-50%), brun jaune, structure polyédrique sub-anguleuse dominante.</i>

Tableau 2.2: La description textuelle concernant les données purement pédologiques pour 4 UTS dans la base de données sol au 1/250.000 du Languedoc-Roussillon (Borand *et al.* 1992).

est parfois nécessaire de déterminer une valeur exprimant la variabilité spatiale globale pour la région étudiée (Beckett & Webster, 1971).

2.3.3.2 Description textuelle

La description textuelle est une description globale de l'unité typologique de sols qui correspond à la légende associée à la carte des sols. L'expert représente l'ensemble de sa connaissance concernant les UTS et les UCS (profils et sondages) à l'aide de termes linguistiques spécifiques. Le Tableau 2.2 donne la description textuelle pour 4 unités typologiques de sols.

Pour les exemples donnés, le sol est décrit dans sa globalité, c'est-à-dire, qu'il n'existe pas de distinction entre les horizons du sol. La description textuelle contient une richesse importante en informations. Néanmoins, l'application de cette description imprécise dans l'estimation des propriétés hydriques pose le problème de l'exploitation rigoureuse et cohérente d'une information essentiellement qualitative. En d'autres termes, il convient de donner un contenu sémantique clair à cette description.

Pour bien pouvoir interpréter la description textuelle, il faudra connaître l'origine de cette description. Pour pouvoir décrire ce qu'il observe, le pédologue a besoin d'une échelle de mesure ou d'une référence sur laquelle il peut baser la formulation de ses observations du sol et du paysage. Cette référence ou mesure d'échelle est souvent de caractère continu, comme par exemple la profondeur. Cependant, pour pouvoir rendre compte de l'incertitude de l'observation de terrain et de la variabilité à l'intérieur d'une unité de sol, le pédologue utilise

une subdivision du domaine continu en un nombre limité d'intervalles de valeurs ou de classes (profondeur du sol entre 25 et 50 cm, profondeur entre 100 et 150 cm.). Pour chaque classe, un nom approprié a été choisi sous la forme d'un terme linguistique (sol peu profond, sol moyennement profond et sol très profond, etc.). L'ensemble des références pour chaque variable constitue alors le vocabulaire nécessaire pour pouvoir décrire le sol et le paysage (Legros & Nortcliff, 1990). Au moment de l'interprétation d'une description textuelle, il faudra alors connaître la signification des termes linguistiques utilisés. La signification des termes est décrite dans des guides de description de sols (Baize et Jabiol, 1995; Soil Survey Staff, 1993; Hodgson, 1974).

L'information donnée par un pédologue sous forme de description textuelle est imprécise puisque les propriétés de l'unité typologique de sols sont données sous forme de classes qui correspondent à un intervalle de valeurs donné par le système de codage employé. Le pédologue utilise ces systèmes de codages pour exprimer l'incertitude sur les propriétés du sol au sein des unités typologiques. D'un autre côté, ces systèmes de codages peuvent parfois limiter artificiellement la précision des descriptions, des subdivisions de classes n'étant que rarement possible (par ex. Sableux à sable fin vs. Sableux à sable grossier). Il faut également noter qu'une nouvelle incertitude est à prendre en compte lorsqu'on utilise une description textuelle. Il s'agit de celle associée à l'erreur d'affectation de classes. Dans quelques cas, principalement pour l'appréciation de la texture, ces erreurs ont été mesurées à partir des profils analysés (De Leenheer *et al.*, 1954; Hodgson *et al.*, 1976).

En résumé, chaque terme correspond à un intervalle de valeurs dont il convient de définir les bornes avec une certaine flexibilité pour représenter le plus fidèlement le sens du terme et pour éviter le plus possible les différences d'interprétation d'un pédologue à l'autre.

2.3.3.3 Description des propriétés du sol par profil de référence

Les profils associés à une étude de sol font l'objet d'observations précises et d'analyses de laboratoire. L'ensemble de ces données constitue une source d'information importante pour la définition des unités typologiques du sol. Dans ce cas, l'UTS peut être décrite par un profil de référence (Gaultier *et al.*, 1992) dans lequel les propriétés sont décrites sous forme d'intervalle de valeur en distinguant des horizons. Le Tableau 2.3 donne un extrait de la base de données pour illustrer ce type de représentation pour les mêmes unités de sol décrites dans les Tableaux 2.1 et 2.2.

L'intervalle de valeurs pour chaque variable est identifié par trois valeurs: la valeur minimale, la valeur moyenne et la valeur maximale. Les valeurs minimales et maximales sont des indications grossières pour l'intervalle dans lequel peut se trouver la propriété du sol pour l'horizon donné. La valeur moyenne indique la valeur dominante au sein de l'intervalle. Cependant, ces valeurs ne correspondent pas à des mesures statistiques souvent appliquées

UTS	Horizon	Nom	Epaisseur (cm)			Taux d'argile (%)			PH			Etc...
			Min	Moy	Max	Min	Moy	Max	Min	Moy	Max	
102	102001	Ap	10	20	60		5		7	8	9	
	102002	B	30	60	80	0	5	10	7	8.3	9	
	102003	C					0			8.3		
303	303001	Ap	15	50	75	10	25	45	6	7.9	8.8	
	303002	CB/Bt	35	50	75	15	30	65	6.5	7.9	8.8	
345	345001	Ap		40			30		7.2			
	345002	B		60			50		7.4			
	345003	C		20			35		8.3			
974	974001	Ap/A1	5	25	55	5	15	30	7.8	8.4	9	
	974002	B	10	20	40	5	15	30	8	8.6	9	
	974003	BC				1	10	20	8.5	8.6	9	

Tableau 2.3: Un extrait de la base de données sol au 1/250.000 du Languedoc-Roussillon. L'ensemble des horizons pour une UTS constitue le profil de référence.

pour la description d'un ensemble d'observations. Ainsi, les valeurs minimale et maximale ne correspondent pas aux valeurs minimale et maximale pour un degré de confiance d'une distribution statistique et la valeur moyenne n'est pas la moyenne d'un ensemble d'observations statistiquement représentatives.

Il existe également des variables non renseignées pour certains horizons, c'est-à-dire, des valeurs manquantes. Dans ces cas, le pédologue n'est pas capable de donner une valeur sur la base de ses observations de terrain. Ceci peut être lié aux cas où la variable n'a pas été déterminée ou aux situations où il n'existe pas suffisamment de profils mesurés pour pouvoir donner un intervalle fiable.

La description des unités typologiques de sol par profils de référence constitue la synthèse des données disponibles dans les études de profils. A priori, cette représentation, qui est plus détaillée, et moins globalisante que la description textuelle et donc moins imprécise (représentée par des intervalles plus étroits). Toutefois, cela ne suffit pas à la rendre plus utile dans notre problème (comme on le verra dans les résultats expérimentaux) car son intérêt dépend aussi de son adéquation aux autres connaissances mises en œuvre dans la résolution du problème.

2.3.4 Incertitude liée à l'estimation des propriétés hydriques

L'application des fonctions de pédotransfert introduit deux types d'incertitude. La première source d'incertitude est liée à la construction du modèle, i.e. l'établissement de la régression et la deuxième est liée à la représentativité du modèle.

2.3.4.1 *Incertitude liée à une régression multiple linéaire*

La régression multiple linéaire est une représentation mathématique relativement simple des données empiriques. Cette méthode permet de constituer rapidement un modèle multidimensionnel des données multivariées en prenant en compte la corrélation entre la variable dépendante et les variables indépendantes. Cependant, cette méthode est assujettie d'une incertitude pour deux raisons principales (Blalock, 1988). Premièrement, la qualité de la régression est influencée par le nombre de données utilisées pour l'analyse. Un nombre suffisamment grand est nécessaire pour déterminer précisément les distributions des variables d'entrées, les corrélations entre les variables dépendantes et indépendantes et la distribution du facteur d'erreur (e). Deuxièmement, la linéarité entre la variable dépendante et la variable indépendante est souvent douteuse. Même dans le cas des régressions non linéaires, le choix du type de régression reste arbitraire, vu le manque de connaissance sur la nature exacte des relations entre variables à partir des données empiriques.

Il existe deux méthodes pour analyser la qualité des équations de régression. La première est l'analyse de variance qui compare la variance prise en compte par la régression avec la variance totale. Une deuxième façon d'analyser la qualité des équations obtenues est par analyse du résidu. La distribution des résidus (valeur estimée – valeur observée) donne une indication sur la proximité générale entre les données et le modèle. Elle donne également une information sur un éventuel biais. Sans biais, la forme de la distribution des résidus est normale et la qualité de la régression peut être exprimée sous forme d'écart type des résidus (Blalock, 1988).

En conclusion, les fonctions de pédotransfert (FPT) sous la forme d'une régression multiple contiennent de l'incertitude. Il est ensuite important de pouvoir prendre en compte cette incertitude au moment où ces équations de régression sont appliquées pour l'estimation de propriétés hydriques. L'analyse des résidus et l'analyse de variance fournissent les métadonnées nécessaires pour la quantification de l'incertitude.

2.3.4.2 *Incertitude liée à la représentativité des FPT*

La représentativité de la fonction de pédotransfert est un aspect important en ce qui concerne l'applicabilité des fonctions. Les données utilisées pour l'établissement des FPT's sont généralement issues d'une zone géographique avec un type de sol spécifique. Il est alors important de savoir si ces FPT's sont valables, c'est-à-dire représentatives dans une autre région, pour un même type de sol. Une fonction de pédotransfert établie par exemple pour un sol très commun aux Etats-Unis n'est pas automatiquement représentative pour un même type de sol en Europe (Bastet *et al.*, 1998). Ce problème introduit une incertitude dans une estimation utilisant des FPT's. Ce type d'incertitude est de plus difficilement quantifiable. Une

bonne connaissance de l'origine de la FPT et du type de sol auquel on souhaite appliquer les FPT's est nécessaire.

Une régression multiple est établie pour un intervalle donné de la variable dépendante. Il n'existe pas de données en dehors de cet intervalle, ce qui implique un manque de connaissance sur l'extrapolation de la relation mathématique. L'utilisation des FPT se réduit donc à la détermination d'une valeur à l'intérieur de l'intervalle pour laquelle la FPT a été faite.

2.4 Propagation d'incertitude

L'analyse précédente a permis de mettre à plat les caractéristiques principales des données nécessaires pour la cartographie des propriétés hydriques à l'échelle régionale. Cette analyse a montré l'existence de nombreuses sources d'incertitude. La cartographie des propriétés hydriques, en utilisant une combinaison de différentes données incertaines, produit ensuite un résultat incertain: l'incertitude sur les variables d'entrées se propage dans l'analyse cartographique proposée.

Les aspects de propagation de l'incertitude seront analysés dans les deux paragraphes suivants pour les deux principales étapes de la démarche générale: estimation de la propriété hydrique pour chaque UTS et localisation des UTS dans la zone d'étude. Cette analyse permettra de définir les trois problèmes principaux à résoudre dans la présente étude.

2.4.1 Propagation de l'incertitude dans l'estimation de la réserve utile en eau

La propriété hydrique du sol à laquelle on s'intéresse plus particulièrement dans la présente étude, est la réserve utile du sol. La réserve utile du sol est définie comme le réservoir en eau disponible pour la végétation et joue ainsi un rôle crucial dans, par exemple, la modélisation du bilan hydrique du sol dans les modèles hydrologiques (Leenhardt *et al.*, 1995), éventuellement liée à la modélisation de croissance de cultures (Armstrong *et al.*, 1996). La réserve utile n'est pas une propriété du sol facilement mesurable sur le terrain. Elle est généralement estimée de manière indirecte à partir d'une série de propriétés du sol. L'équation, sous sa forme la plus simple, pour la détermination de la réserve utile est la suivante (Baize & Jabiol, 1995):

$$RU = (\theta_{cr} - \theta_{pf}) \times E \quad (2.1)$$

La réserve utile est calculée pour une profondeur E à partir de deux niveaux d'humidité: l'humidité à la capacité au champ, qui prend classiquement le nom de capacité de rétention (θ_{cr}) et l'humidité au point de flétrissement (θ_{pf}). Ces deux niveaux d'humidité

correspondent respectivement à l'humidité volumique de pF2 et pF4.2. La valeur pour ces deux variables est déterminée en utilisant des fonctions de pédotransfert. Ces fonctions ont la forme générale suivante:

$$\theta_x = f(\%argile, \%limon, \%sable, \text{densité apparente}, \%matière\ organique) \quad (2.2)$$

Il est donc possible d'évaluer la réserve utile pour chaque unité typologique de sol en utilisant les opérations arithmétiques des équations 2.1 et 2.2, si on connaît les fonctions de pédotransfert et si on dispose des données nécessaires pour chaque UTS.

Cependant, les données nécessaires pour chaque UTS sont principalement représentées sous la forme d'intervalles de valeurs. L'estimation de la réserve utile telle qu'elle a été présentée ci-dessus devient ainsi fortement imprécise si les variables d'entrée sont très imprécises elles-mêmes. Ceci permet de conclure sur deux aspects importants concernant l'estimation de la réserve utile pour chaque UTS:

- a) Il est essentiel de représenter de manière fiable la connaissance imprécise (les intervalles de valeurs) exprimant l'incertitude sur chaque propriété de sol afin de pouvoir évaluer l'incertitude sur l'estimation de la réserve utile.
- b) De plus, l'estimation de la réserve utile doit tenir compte des relations entre les variables. Ces relations entre les variables peuvent s'exprimer, par exemple, sous la forme de corrélations ou sous la forme d'égalités et d'inégalités (par exemple, la somme du taux d'argile, de limon et de sable est toujours égale à 100%). Estimer la réserve utile sans prendre en compte ces relations risque d'influencer fortement l'imprécision de cette estimation.

2.4.2 Evaluation de l'incertitude dans la cartographie des UTS

L'information de la base de données sol va nous permettre d'établir, pour chaque UTS, des règles logiques permettant de déterminer la présence ou l'absence d'une UTS en un point donné dans la zone d'étude. La prise en compte de l'incertitude attachée à chaque variable et la propagation de cette incertitude dans la détermination de la présence ou l'absence de l'UTS en un point donné est essentielle.

A présent, nous pouvons distinguer trois aspects indispensables à la résolution de ce problème cartographique. Premièrement, il est important de pouvoir représenter, par des classes plus ou moins strictes, la connaissance issue de la description textuelle concernant de propriétés environnementale de l'UTS. Deuxièmement, il convient de pouvoir comparer ou confronter ces classes à des valeurs observées ou estimées avec plus ou moins de précision. Troisièmement, une combinaison ou agrégation est nécessaire pour les confrontations sur chaque propriété.

2.5 Conclusion

Dans ce chapitre, nous avons analysé les données nécessaires pour cartographier des propriétés hydriques du sol sur l'ensemble d'une région. Il a été montré que ces données étaient affectées par une incertitude importante qu'il était indispensable de représenter et de propager dans l'analyse cartographique afin de pouvoir déterminer l'incertitude sur l'estimation de la propriété hydrique recherchée.

Pour cela, nous devons apporter des solutions méthodologiques à trois problèmes:

- a) **Comment représenter une information pédologique imprécise avec des données sur l'imprécision (méta-données) limitées?**
- b) **Comment estimer, par des opérations arithmétiques, la réserve utile pour chaque UTS en propageant l'incertitude sur chaque variable intervenant dans l'estimation et en tenant compte des interactions entre ces variables?**
- c) **Comment localiser, par des combinaisons logiques, les UTS au sein des UPP en prenant en compte l'imprécision de la description des UTS et l'incertitude sur l'information géographique supplémentaire utilisée?**

Dans la suite de ce travail nous allons, par différentes étapes, essayer de répondre à chaque problème soulevé.

3 Représentation et traitement de l'information incertaine

Le précédent chapitre montre que la cartographie des propriétés hydriques du sol à l'échelle régionale est sujette à de nombreuses sources d'imprécision et d'incertitude. Cette cartographie nécessite un cadre théorique et méthodologique permettant le traitement de l'information imprécise et incertaine, le traitement de l'information manquante et la représentation des dépendances entre variables.

Il est nécessaire de déterminer quelles méthodes et techniques sont à notre disposition et quelles sont les mieux adaptées à notre situation. Ce chapitre a pour but d'explorer les choix possibles parmi les différents cadres théoriques et méthodologiques. Nous donnons des exemples et des références bibliographiques de ce qui existe et a déjà été appliqué dans le domaine de la pédologie et des systèmes d'information géographique.

On présente successivement les deux approches théoriques pour le traitement de l'imprécision et de l'incertitude. La première approche s'appuie sur un cadre théorique statistique utilisant une mesure de probabilité. La deuxième approche met en jeu le cadre théorique basé sur la logique floue et utilisant une mesure de possibilité. Il est également nécessaire ici d'analyser les différences et similitudes des deux approches théoriques proposées.

3.1 Deux cadres théoriques possibles

Les connaissances dont nous disposons sur une situation quelconque sont généralement imparfaites, soit parce que nous avons un doute sur leur validité (incertitude), soit parce que nous éprouvons une difficulté à les exprimer clairement (imprécision) (Bouchon-Meunier, 1993). Les notions d'incertitude et d'imprécision ont été explicitées dans le précédent chapitre, mais il s'agit aussi de pouvoir les mathématiser afin de les manipuler rigoureusement.

Dans la littérature, on distingue deux types d'approches. La première concerne l'approche statistique classique. La statistique permet de décrire et de généraliser l'information concernant une population sur la base d'une série d'observations sur cette population ou sur la base de la formulation de lois générales établies à partir d'observations répétitives (induction ou inférence statistique). La statistique inductive est directement liée à la théorie des probabilités. La théorie des probabilités est une théorie mathématique de l'incertitude reposant sur un ensemble d'axiomes dont les plus importants spécifient que (i) l'incertitude peut être représentée par une mesure et une seule (la mesure de probabilité), (ii) la mesure attachée à un événement détermine complètement celle attachée à l'événement contraire et (iii) la mesure attachée à une disjonction d'événements indépendants est la somme des mesures de chaque événement.

Cependant, la notion de probabilité ne permet pas de traiter les connaissances vagues ou des connaissances sur lesquelles plane une part d'ignorance non liée au hasard (Bouchon-Meunier, 1993; Klir & Yuan, 1995). Une alternative à une approche statistique dans ces cas est constituée par la théorie des possibilités (Zadeh, 1978b). La théorie des possibilités est issue de la théorie des sous-ensembles flous, considérée comme une généralisation de la théorie classique des ensembles. La théorie des sous-ensembles flous est basée sur quelques d'idées fondamentales (Bouchon-Meunier, 1993): (i) l'idée de l'appartenance partielle à une classe, (ii) l'idée des catégories aux limites mal définies, (iii) l'idée de la gradualité dans le passage d'une catégorie à une autre admettant des situations intermédiaires entre le tout et le rien. Les sous-ensembles flous sont principalement utilisés dans les situations de connaissances imparfaitement décrites, vagues ou imprécises. Ils établissent alors une interface entre l'information brute (souvent sous forme d'expression linguistique) et l'information représentée numériquement et manipulable. Un ensemble flou caractérise un ensemble de valeurs possibles. A partir de celui-ci, on peut construire différentes mesures d'incertitude, dont les mesures de possibilité et de nécessité, qui sont non-additives. La théorie de possibilités offre un cadre mathématique pour traiter l'imprécision et l'incertitude de nature qualitative en utilisant des opérateurs de comparaison (max et min) plutôt que la somme et le produit (utilisées dans la théorie de probabilité). La notion de possibilité d'un événement correspond à l'idée d'absence de surprise. La notion duale de nécessité (ou certitude) correspond à l'impossibilité de l'événement contraire. Il existe plusieurs ouvrages qui explorent divers

aspects de la théorie des possibilités, c'est-à-dire, les bases théoriques et les applications. Quelques références intéressantes sur ce sujet sont Dubois & Prade (1988) et Klir & Yuan (1995).

En résumé, la théorie de probabilité est un outil idéal pour la formalisation de l'incertitude dans les cas où la connaissance est basée sur la distribution de fréquences mesurées¹ (incertitude liée à un phénomène aléatoire observé). En revanche, la théorie des possibilités est idéale pour la formalisation de l'information incomplète exprimée en terme de propositions imprécises et vagues (Klir & Yuan, 1995) (incertitude liée à une part d'ignorance qui empêche de déduire l'incertitude sur un événement à partir du celui du contraire). Néanmoins, dans bon nombre de cas pratiques, il existe différents types d'information à prendre en compte. Une partie des données disponibles peut être de caractère statistique et une autre partie des données sous la forme de propositions imprécises. Pour exploiter ces informations, il convient alors de se donner un cadre théorique unique. Dans le cas d'un problème de décision, on choisira de préférence une transformation qui ramène les possibilités à des probabilités pour pouvoir ensuite appliquer les outils de la théorie de la décision. Dans les problèmes de raisonnement, comme celui abordé dans cette thèse, on mettra toute l'information dans le cadre possibiliste. Ceci nécessite des techniques de transformation d'une information probabiliste en information possibiliste et inversement (Klir & Parviz, 1992; Dubois *et al.*, 1993), qui, compte tenu des différences entre les deux cadres, implique une perte d'information (pour le passage probabilité vers possibilité) ou l'introduction d'une part d'arbitraire (pour le passage des possibilités aux probabilités).

3.2 Application d'une approche statistique et probabiliste

L'intérêt général d'une approche statistique, dans la résolution d'un problème pratique donné, est la capacité d'évaluer l'incertitude sur la réalisation d'un événement en terme de probabilité, c'est-à-dire, de chance d'occurrence d'événements. Une analyse statistique nécessite d'abord une représentation de l'information disponible permettant l'évaluation de cette incertitude. Cette représentation sera brièvement exposée, en ce qui concerne les variables continues, dans le paragraphe suivant. Ensuite, nous présenterons également une méthode fréquemment utilisée pour la propagation de l'incertitude dans la modélisation: la simulation de Monte Carlo.

¹ Il existe aussi une interprétation non-fréquentiste des probabilités: les probabilités subjectives qui s'utilisent dans un cadre bayésien.

3.2.1 Représentation de la connaissance en distributions de probabilité

La statistique descriptive permet de représenter un ensemble d'observations sous la forme d'une distribution de fréquence. La distribution de fréquence pour une variable continue peut être exprimée sous la forme d'un histogramme. Dans un histogramme, on exprime la fréquence absolue ou relative (la proportion) pour chaque intervalle défini sur un domaine de référence continu X . Un histogramme est principalement une représentation visuelle des données permettant de dériver approximativement l'information sur la valeur la plus fréquente et sur la dispersion des valeurs. D'autres moyens permettant de caractériser une distribution de fréquence de façon plus quantitative sont le minimum, le maximum, la moyenne arithmétique, le médian, l'écart type et la variance. La moyenne arithmétique ou le médian sont des mesures exprimant la valeur la plus fréquente (i.e. la plus probable) alors que l'écart type et la variance sont des mesures de dispersion.

Un histogramme exprimant la fréquence relative peut-être interprété comme distribution de probabilité en classes discrètes. Considérons un domaine de référence X divisé en n classes, la distribution de probabilité p est alors $\{p_1, p_2, \dots, p_n\}$. La distribution de probabilité est normalisée de la manière suivante:

$$\sum_{i=1}^n p_i = 1 \quad (3.1)$$

Un histogramme est une distribution de probabilité réaliste si l'on tient compte des données brutes, mais l'utilisation de cette information est limitée à cause de la discrétisation du domaine de référence. Une distribution de probabilité continue est alors nécessaire pour pouvoir déterminer la probabilité pour chaque point ou pour chaque intervalle dans le domaine de référence X . Une distribution de probabilité continue est définie par la fonction de densité de la probabilité. La condition de normalisation s'exprime de la manière suivante:

$$\int_{-\infty}^{+\infty} p(x) dx = 1 \quad (3.2)$$

Il existe plusieurs types de fonctions théoriques de densité de probabilité. Le choix du type de fonction à utiliser dépend de la distribution des données brutes décrites par l'histogramme (unimodale, avec une dérive ou uniforme). Les fonctions unimodales généralement utilisées sont les fonctions normale et lognormale. Ces deux fonctions utilisent l'écart type et la moyenne arithmétique comme paramètres. L'utilisation de ces deux types de fonctions est limitée par le nombre d'observations nécessaires pour l'estimation de ces paramètres. Dans le cas où ces paramètres sont difficilement estimables, il est parfois préférable d'utiliser une fonction uniforme affectant une probabilité égale à chaque élément d'un sous-ensemble de X .

3.2.2 Propagation de l'incertitude avec la simulation Monte Carlo

Le paragraphe précédent a exposé la représentation de l'information sous la forme d'une distribution de probabilité qui peut-être définie par une fonction de densité de probabilité. Pour un problème quantitatif où plusieurs variables numériques entrent en jeu et où chaque variable est exprimée sous forme de distribution de probabilité, il est important de pouvoir propager l'incertitude sur chaque variable lors de la résolution du problème. Le but de la propagation d'incertitude est la détermination de la distribution de probabilité sur une variable particulière. Il existe plusieurs types de calcul de propagation d'incertitude. Un seul type de propagation sera exposé ici. Il s'agit de la simulation Monte Carlo (Shannon, 1977; Manly, 1991). La simulation Monte Carlo est basée sur un principe relativement simple qui consiste en la réalisation d'une série de simulations du modèle avec, chaque fois, un jeu de valeurs différent pour les variables d'entrée. Le jeu de valeurs est choisi de manière aléatoire en tenant compte de la densité de probabilité définie pour chaque variable. Un grand nombre de tirages aléatoires est nécessaire pour obtenir la distribution de probabilité du résultat avec suffisamment de fiabilité. Des références bibliographiques donnent souvent des valeurs autour de mille tirages pour un niveau de signification statistique de 5% (Manly, 1991; Heuvelink, 1993).

Dans le cas le plus simple de la simulation Monte Carlo, les variables sont considérées comme indépendantes. Or, ceci n'est souvent pas le cas. L'échantillonnage des valeurs aléatoires doit satisfaire certaines conditions quand les variables sont dépendantes (Johnson, 1987). Ces conditions peuvent être définies par la matrice de la covariance ou la matrice de corrélation. D'autres relations entre variables, par exemple des inégalités, peuvent être introduites comme contraintes sur l'échantillonnage des valeurs aléatoires.

L'utilisation pratique de cette méthode existe dans divers domaines de modélisation. Elle a été appliquée, par exemple, dans la modélisation des processus hydrologiques par Jones (1989), Smith & Freeze (1979), Smith & Hebbert (1979) et De Roo *et al.* (1992). La simulation Monte Carlo dans la modélisation à l'aide des systèmes d'information géographique a été utilisée par Fischer (1992), Heuvelink (1993) et Goodchild *et al.* (1992). Il existe également des exemples de propagation d'incertitude à l'aide de la simulation Monte Carlo dans le domaine de la cartographie d'une propriété du sol (Heuvelink & Bierkens, 1992; Heuvelink & Burrough, 1993).

La simulation Monte Carlo présente plusieurs avantages et désavantages. L'avantage principal est la création d'une distribution de probabilité de la variable de sortie du modèle par une procédure généralement facile à mettre en place. Un désavantage concerne la lourdeur en terme de temps de calcul dû au grand nombre de simulations nécessaires. Or, la capacité de calcul des ordinateurs d'aujourd'hui permet la réalisation de nombreuses simulations dans un délai acceptable. Toutefois, la lourdeur de calcul pose encore des problèmes dans les applica-

tions spatiales où une série de simulations est nécessaire en chaque point d'une grille régulière (Heuvelink, 1993). Le deuxième désavantage de la simulation Monte Carlo concerne le choix du type de fonction de densité de la probabilité. La connaissance sur la distribution exacte des variables d'entrée est généralement absente à cause d'un manque de données. Dans beaucoup d'applications pratiques de la simulation Monte Carlo, on suppose une fonction normale pour la distribution des variables. Ceci est un choix souvent arbitraire qui pèse sur la crédibilité des résultats.

3.3 Les sous-ensembles flous et la théorie des possibilités

Une alternative à une approche probabiliste existe dans la théorie des possibilités. Ce paragraphe exposera les principales caractéristiques des ensembles flous et de la théorie des possibilités.

3.3.1 Définitions des sous-ensembles flous

La notion d'ensemble flou a été introduite par Zadeh (1965). Cette notion met en évidence la difficulté qui existe à décrire certaines connaissances humaines en logique classique. La logique classique ne permet pas de représenter des expressions vagues telles que "profond", "très", et "environ", car elle est bivalente, i.e. une expression quelle qu'elle soit, est soit complètement vraie soit complètement fausse. Or, très souvent, l'information dont on dispose et que l'on voudrait représenter est entachée d'imprécision ("sol profond") ou d'incertitude ("il est presque certain que cette parcelle se situe dans UTS 975"). Pour résoudre ce problème de représentation, le seuil unique de décision pour l'appartenance à l'ensemble des sols "profonds" est remplacé par une zone de transition entre les éléments parfaitement adaptés aux sols "profonds" et les éléments pas du tout adaptés aux sols "profonds". A l'intérieur de la zone de transition, on définit des éléments plus ou moins adaptés. On introduit alors la notion de degré d'appartenance qui permet d'exprimer jusqu'à quel point il est acceptable de dire qu'un élément appartient à un ensemble.

Un sous-ensemble flou est ainsi une généralisation d'un sous-ensemble classique. Un sous-ensemble classique A de X peut-être défini par une fonction d'appartenance ($\mu_A(x)$) qui détermine si les éléments appartiennent à A (valeur 1) ou si les éléments n'appartiennent pas à A (valeur 0):

$$\mu_A(x) : X \rightarrow \{0,1\}. \quad (3.3)$$

Un sous-ensemble flou A de X est défini par une fonction d'appartenance ($\mu_A(x)$) qui associe à chaque élément x de X le degré d'appartenance compris entre 0 et 1:

$$\mu_A(x) : X \rightarrow [0,1]. \quad (3.4)$$

Dans la présente étude, nous interprétons les degrés d'appartenance comme un degré d'incertitude qui est alors la possibilité que l'élément concerné soit effectivement dans l'ensemble flou. Cet ensemble sert alors à délimiter la disjonction de valeurs possible que peut prendre la variable à laquelle on s'intéresse.

Par exemple, supposons que la variable d'intérêt est la profondeur du sol et que le pédologue a exprimé sa connaissance comme "sol profond". Les valeurs de profondeur du sol possibles sont représentables par l'ensemble flou de la Figure 3.1.

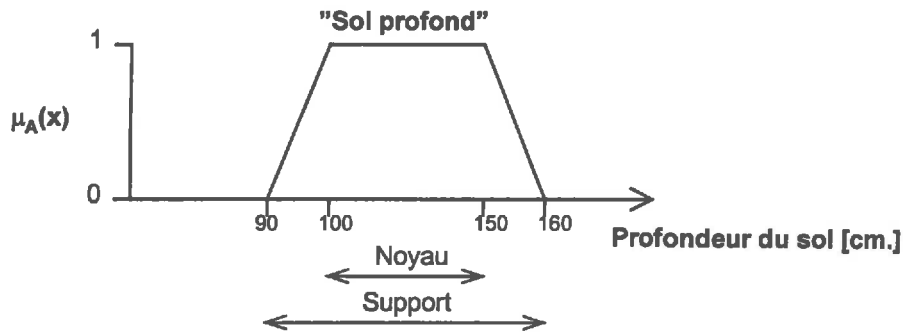


Figure 3.1: Sous-ensemble flou sous la forme d'une fonction trapézoïdale.

La Figure 3.1 est un exemple de fonction d'appartenance trapézoïdale. Il est nécessaire d'introduire ici certaines définitions qui permettent de décrire un sous-ensemble flou. La première concerne le support d'un sous-ensemble flou. Le support d'un sous-ensemble flou est la partie de X pour laquelle la fonction d'appartenance de A n'est pas nulle:

$$\text{Supp}(A) = \{x \in X / \mu_A(x) \neq 0\}. \quad (3.5)$$

La hauteur d'une ensemble flou est exprimée par:

$$h(A) = \sup_{x \in X} \mu_A(x). \quad (3.6)$$

où \sup indique ici le supremum, c'est-à-dire, le degré d'appartenance maximal. \inf exprime l'infimum, c'est-à-dire le plus petit d'entre eux. L'ensemble de tous les éléments appartenant de façon absolue (avec un degré 1) à A est appelé le noyau de A et noté $\text{noy}(A)$.

$$\text{Noy}(A) = \{x \in X / \mu_A(x) = 1\}. \quad (3.7)$$

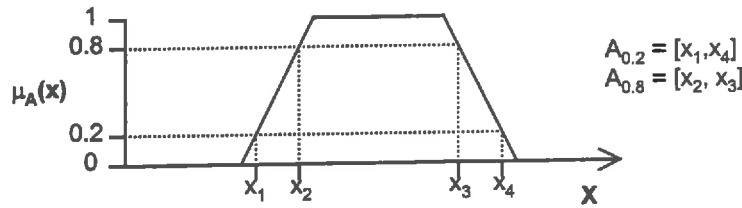


Figure 3.2: La définition des α -coupes.

Un sous-ensemble flou peut-être représenté par une série d'ensembles non-flous par ce qu'on appelle des α -coupes ou coupes de niveau α (voir la Figure 3.2). Une α -coupe A_α est l'ensemble des éléments ayant au moins le degré d'appartenance de niveau α :

$$A_\alpha = \{x; \mu_A(x) \geq \alpha\}. \quad (3.8)$$

Si le référentiel X est égal à l'ensemble des nombres réels, alors tout ensemble flou sur X est appelé quantité floue. Une quantité floue A dont la fonction d'appartenance est quasi-concave, i.e., telle que $\forall x_1, x_2, \forall x_3 \in [x_1, x_2], \mu_A(x_3) \geq \min(\mu_A(x_1), \mu_A(x_2))$, est appelée intervalle flou (Dubois & Prade, 1988).

3.3.2 Opérations élémentaires avec les sous-ensembles flous

Après avoir décrit l'information disponible sous la forme des sous-ensembles flous, il est important de pouvoir combiner différentes informations, afin de pouvoir en dériver des informations nouvelles. Ceci nécessite un cadre mathématique permettant de traiter des informations sous la forme de sous-ensembles flous. La plupart des manipulations disponibles sont basées sur les opérations ensemblistes déjà existantes dans le cadre des sous-ensembles classiques. Elles sont brièvement résumées dans le Tableau 3.1 en considérant les sous-ensembles flous A , B et C dans le référentiel X . Les propriétés logiques qui peuvent être dérivées des opérations logiques sont présentées dans le Tableau 3.2.

Opération	Fonctions d'appartenance	Nom
$A = B$	$\forall x \in X \quad \mu_A(x) = \mu_B(x)$	Egalité
$A \subseteq B$	$\forall x \in X \quad \mu_A(x) \leq \mu_B(x)$	Inclusion
$C = A \cap B$	$\forall x \in X \quad \mu_C(x) = \min(\mu_A(x), \mu_B(x))$	Intersection
$C = A \cup B$	$\forall x \in X \quad \mu_C(x) = \max(\mu_A(x), \mu_B(x))$	Union
A^c	$\forall x \in X \quad \mu_{A^c}(x) = 1 - \mu_A(x)$	Complémentarité

Tableau 3.1: Les opérations ensemblistes principales.

$A \cup B = B \cup A; A \cap B = B \cap A$	Commutativité
$(A \cup B) \cup C = A \cup (B \cup C); (A \cap B) \cap C = A \cap (B \cap C)$	Associativité
$A \cap X = A; A \cap \emptyset = \emptyset$	
$A \cup X = X; A \cup \emptyset = A$	
$A \cup B \supseteq A \supseteq A \cap B$	

Tableau 3.2: Propriétés principales des opérations ensemblistes avec des sous-ensembles flous.

Les opérations présentées dans le Tableau 3.1 correspondent à ce qui s'appelle la théorie "max-min" des ensembles flous. Par rapport à l'algèbre booléenne qui sous-tend la théorie des ensembles classiques, deux propriétés ont été abandonnées: la loi du tiers-exclu (soit une proposition est vraie, soit elle est fausse) et la loi de non-contradiction (une proposition ne peut être à la fois vraie et fausse).

3.3.3 Relations floues

De même que pour les ensembles classiques, il est possible de définir des relations entre éléments d'ensembles flous. Une relation floue R est un ensemble flou (et est donc aussi définie par une fonction d'appartenance) sur un produit cartésien de référentiels. Considérons deux référentiels X et Y et soit $\mu_R : X \times Y \rightarrow [0,1]$ la fonction d'appartenance de la relation binaire de R , $\mu_R(x,y)$ mesure à quel point x est lié à y par R . Si R est une relation classique, alors soit x et y sont en relation par R , auquel cas $\mu_R(x,y) = 1$ soit il ne le sont pas et dans ce cas $\mu_R(x,y) = 0$.

Considérons, par exemple, un sol contenant deux horizons et dont la profondeur totale est définie comme "sol profond". La profondeur du sol est, par définition, la somme des épaisseurs des horizons. Ayant représenté "sol profond" par le sous-ensemble flou de la Figure 3.1, la somme des épaisseurs des horizons devient une relation floue : $\mu_R(x,y) = \mu_{\text{"sol profond"}}(x + y)$.

Un cas particulier de relation floue est le produit cartésien flou $R = A \times B$ de l'ensemble flou A sur X et B sur Y qui est défini par:

$$\forall x, \forall y, \mu_R(x,y) = \min(\mu_A(x), \mu_B(y)). \quad (3.8)$$

On peut ensuite projeter R sur X ou Y . Par exemple, la projection sur X notée $R \downarrow X$ se définit par:

$$\forall x, \mu_{R \downarrow X}(x) = \sup_{y \in Y} \mu_R(x, y). \quad (3.9)$$

Les variables V_1 et V_2 restreintes par les ensembles A et B sont dites non-interactives lorsque le domaine du couple (V_1, V_2) est restreint exactement par le produit cartésien de A et B , i.e.,

$$\mu_{V_1 V_2}(x, y) = \min(\mu_A(x), \mu_B(y)). \quad (3.10)$$

La non-interactivité de V_1 et V_2 signifie que les valeurs que peut prendre V_1 ne dépendent pas de la valeur de V_2 et inversement. Si A et B sont les seules contraintes connues de V_1 et V_2 , la fonction d'appartenance $\min(\mu_A, \mu_B)$ sur $A \times B$ est la façon la plus licite de définir la relation entre V_1 et V_2 ; c'est la plus grande au sens de l'inclusion. Si V_1 et V_2 avaient été interactives, on aurait eu $\mu_R \leq \min(\mu_A, \mu_B)$ ce qui signifie que l'introduction d'interactions entre les variables induit des relations floues plus spécifiques (précises).

3.3.4 La théorie de possibilité

La théorie de possibilité permet de formaliser des incertitudes de nature non probabiliste sur des événements. Elle donne donc un moyen de dire dans quelle mesure la réalisation d'un événement est possible (plausible) et dans quelle mesure on en est certain (Bouchon-Meunier, 1993; Dubois & Prade, 1998). Ce paragraphe présente la théorie de possibilité sous sa forme générale en insistant sur la formalisation de l'incertitude sur les événements: la mesure de possibilité et sa duale, la mesure de nécessité.

La théorie des sous-ensembles flous permet de traduire une connaissance disponible sous la forme d'une fonction d'appartenance attribuant à chaque événement élémentaire de l'ensemble de référence X un coefficient, compris entre 0 et 1, évaluant à quel point cet événement est possible. La mesure de possibilité Π permet de définir ce coefficient, qui est une fonction définie sur l'ensemble $\mathbf{P}(X)$ des parties de X et qui prend ses valeurs dans $[0, 1]$:

$$\begin{aligned} \text{i)} \quad & \Pi(\emptyset) = 0, \Pi(X) = 1, \\ \text{ii)} \quad & \forall A_1 \in \mathbf{P}(X), A_2 \in \mathbf{P}(X) \dots \\ & \Pi(\bigcup_{i=1,2,\dots} A_i) = \sup_{i=1,2,\dots} \Pi(A_i) \end{aligned} \quad (3.11)$$

Un événement est donc tout à fait possible si la mesure de sa possibilité est égale à 1, et impossible si celle-ci est nulle. Une distribution de possibilité $(\pi(x))$ est une fonction d'appartenance définie sur X , attribuant à tout élément de X une valeur entre 0 et 1 satisfaisant la condition de normalisation:

$$\sup_{x \in X} \pi(x) = 1. \quad (3.12)$$

Ceci signifie que au moins un élément x est complètement possible. Dès qu'une information imprécise n'est pas auto-contradictoire, sa distribution de possibilité est normalisée. Lorsqu'une distribution de possibilité n'est pas normalisée, cela signifie une incohérence ou une contradiction des connaissances disponibles.

La mesure de possibilité est construite à partir de la distribution de possibilité par la définition suivante:

$$\forall A \in \mathbf{P}(X) \quad \Pi(A) = \sup_{x \in A} \pi(x). \quad (3.13)$$

Une deuxième mesure importante dans la théorie des possibilités est la mesure de nécessité. Cette mesure a été introduite pour pouvoir exprimer la certitude de la réalisation d'un événement. Ceci était nécessaire parce que la mesure de possibilité n'exprime que la plausibilité de la réalisation. La mesure de nécessité peut être dérivée de la mesure de possibilité par l'intermédiaire du complémentaire A^c de toute partie A de X :

$$\forall A \in \mathbf{P}(X) \quad N(A) = 1 - \Pi(A^c). \quad (3.14)$$

Il est possible de définir maintenant la mesure de nécessité à partir d'une distribution de possibilité π de la manière suivante:

$$\forall A \in \mathbf{P}(X) \quad N(A) = \inf_{x \notin A} (1 - \pi(x)). \quad (3.15)$$

a. $N(A) \leq \Pi(A)$	e. $N(A) + N(A^c) \leq 1$
b. $N(A) > 0 \rightarrow \Pi(A) = 1$	f. $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$
c. $\Pi(A) + \Pi(A^c) \geq 1$	g. $N(A \cap B) = \min(N(A), N(B))$

Tableau 3.3: Quelques propriétés simples des mesures de possibilité et de nécessité.

Les mesures de possibilité et de nécessité contiennent certaines propriétés qui sont résumées dans le Tableau 3.3. On considère ici que les distributions de possibilité sont normalisées (cf. Equation 3.12). Considérons alors deux événements A et A^c dont l'un est le complémentaire de l'autre, alors $\Pi(A) = 1$ ou $\Pi(A^c) = 1$. En particulier, on peut exprimer que l'on ne sait rien sur A en posant $\Pi(A) = \Pi(A^c) = 1$, c'est-à-dire A et A^c sont aussi possibles l'un que l'autre.

La propriété a. du Tableau 3.3 exprime que la réalisation d'un événement est toujours plus plausible que certaine. De plus, pour qu'un événement soit un peu certain ($N(A) > 0$), cet événement doit être possible ($\Pi(A) = 1$) et pour qu'un événement soit complètement certain ($N(A) = 1$) il faut et il suffit que son complémentaire soit impossible ($\Pi(A^c) = 0$).

La théorie de possibilité utilise donc deux mesures pour exprimer l'incertitude sur la réalisation d'un événement, tandis que la théorie de probabilité n'en utilise qu'une seule (i.e. la probabilité). Supposons que, à cause de l'absence de données, la probabilité de l'événement

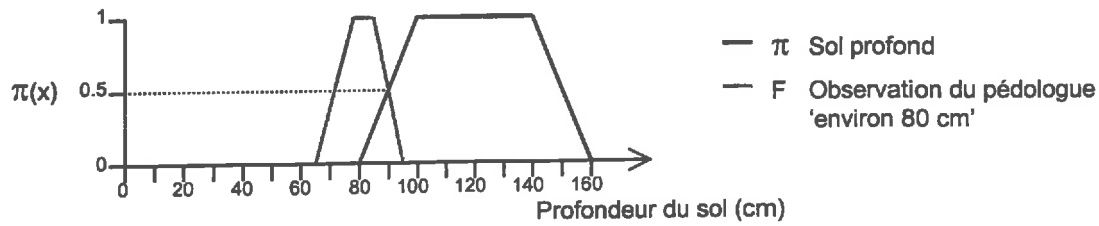


Figure 3.3 : L'évaluation de la compatibilité du sous-ensemble flou F, exprimant une observation d'un pédologue avec la distribution de possibilité π , exprimant le sous-ensemble flou de référence "Sol Profond".

ne soit pas déterminable ou accessible. Les mesures de possibilité et de nécessité peuvent ainsi être considérées comme l'encadrement de la mesure de probabilité (Dubois & Prade, 1998). La possibilité correspond alors à la probabilité supérieure et la nécessité à la probabilité inférieure.

La possibilité et la nécessité permettent, par extension des formules 3.13 et 3.15, d'évaluer la correspondance d'une description floue F avec une distribution de possibilité connue jouant le rôle de référence A. La possibilité exprime un degré d'intersection (appariement lâche) et elle est déterminée de la manière suivante:

$$\Pi(F) = \sup_{x \in X} \min(f_F(x), \pi_A(x)), \quad (3.16)$$

La nécessité exprime un degré d'inclusion (appariement strict) et elle est obtenue par $N(F) = 1 - \Pi(F^c)$, soit :

$$N(F) = \inf_{x \in X} \max(f_F(x), 1 - \pi_A(x)). \quad (3.17)$$

Une interprétation alternative est de considérer la possibilité comme une évaluation 'optimiste' et la nécessité comme une évaluation 'pessimiste' (Martin-Clouaire *et al.*, 1998). Pour clarifier le sens et l'intérêt de ces deux mesures d'incertitude considérons l'exemple de la Figure 3.3. Soit une connaissance préalable sur une propriété fournie par un sous-ensemble de référence A sur le référentiel d'une variable "profondeur du sol". Cet ensemble, décrit par la distribution de possibilité π , exprime la classe *Sol Profond* (voir la Figure 3.3). Un deuxième sous-ensemble flou exprime une observation de terrain du pédologue. L'information donnée par le pédologue est "environ 80 cm" qui peut être décrite sous la forme de la distribution de possibilité F. La possibilité que l'observation F corresponde à la classe *Sol Profond* est 0.5 et la nécessité est 0.0.

3.3.5 Propagation de l'imprécision à travers des relations floues

Une exigence imposée au cadre théorique choisi est la possibilité de propager l'imprécision au cours de la cartographie des propriétés hydriques. Dans le cadre de la théorie de possibilité, la propagation de l'incertitude s'effectue naturellement en généralisant les opérations arithmétiques au cas de valeurs floues représentées par des distributions de possibilité. Cette généralisation est une extension du calcul d'intervalles à des intervalles flous.

3.3.5.1 Calcul d'intervalles flous

La théorie des possibilités permet de généraliser le calcul d'intervalles à des intervalles flous. Comme nous l'avons déjà évoqué dans la sous section 3.3.3, un intervalle flou est une distribution de possibilité sur les réels. Le problème du calcul d'intervalle flou revient à celui de la propagation d'une distribution de possibilité par une fonction f réelle de X dans Y . Dans une opération arithmétique, on effectue ainsi le transport de la distribution par une fonction f dans Y ($Y = f(X)$).

Les opérations classiques de calcul d'intervalles (addition, soustraction, multiplication, division) permettent d'effectuer des calculs d'intervalles avec une série d'intervalles flous. Ces opérations sont brièvement résumées dans le tableau 3.4 où a , b , d et e indiquent respectivement les bornes inférieures et supérieures des intervalles flous A et B .

$$\begin{aligned}
 [a,b] + [d,e] &= [a+d, b+e], \\
 [a,b] - [d,e] &= [a-e, b-d], \\
 [a,b] * [d,e] &= [\min(ad,ae,bd,be), \max(ad,ae,bd,be)] \\
 \text{et pour } 0 \notin [d,e], \\
 [a,b] / [d,e] &= [a,b] * [1/e, 1/d] \\
 &= [\min(a/d,a/e,b/d,b/e), \max(a/d,a/e,b/d,b/e)],
 \end{aligned}$$

Tableau 3.4: Les quatre opérations arithmétiques sur des intervalles fermés.

$A + B = B + A$	}	Commutativité
$A * B = B * A$		
$(A + B) + C = A + (B + C)$	}	Associativité
$(A * B) * C = A * (B * C)$		
$A * (B + C) \subseteq A * B + A * C$		Distributivité faible

Tableau 3.5 : Les propriétés principales des opérations arithmétiques sur des intervalles flous.

En pratique, il est commode de faire le calcul en opérant sur les α -coupes pourvu que les distributions soient semi-continues (Klir & Yuan, 1995). En effet, la coupe de niveaux α

de $Y (=f(A,B))$ où A et B sont des intervalles flous, s'obtient en calculant $f(A_\alpha, B_\alpha)$ à l'aide du calcul d'intervalles classiques.

Les opérations arithmétiques sur des intervalles flous satisfont certaines propriétés. Les propriétés sont formulées pour les intervalles flous A , B et C , dans le Tableau 3.5.

3.3.5.2 Calcul d'intervalles flous avec des variables interactives

Il est possible de formuler une opération arithmétique d'une manière plus générale. Considérons un modèle numérique composé d'un ensemble de variables réelles $\{V_1, \dots, V_n\}$, d'une famille de relations $\{R_1, \dots, R_m\}$ liant ces variables, et d'une connaissance sur les domaines de variation $\{A_1, \dots, A_n\}$ de ces variables, tels que $V_i \in A_i$. La relation R_j concerne un sous-ensemble I_j de variables. Les ensembles A_i ainsi que les relations R_j peuvent être flous, et forment un ensemble de contraintes flexibles reliant les variables.

Supposons qu'on veuille connaître le domaine de variation d'une variable V_k , sous la forme d'une distribution de possibilité π_k . Zadeh (1978a) a montré que le calcul de cette distribution de possibilité X_k peut se faire en utilisant un principe général de raisonnement, dit de combinaison/projection, qui comporte les deux étapes suivantes.

a) combinaison conjonctive

Cette étape détermine la distribution de possibilité jointe π qui, à chaque vecteur de valeurs potentielles (u_1, \dots, u_n) des variables, associe un degré de possibilité $\pi(u_1, \dots, u_n) \in [0, 1]$ défini par:

$$\pi(u_1, \dots, u_n) = \min(\min_{i=1, n} \mu_{A_i}(u_i), \min_{j=1, m} \mu_{R_j}(u_{I_j})). \quad (3.18)$$

où u_{I_j} est un vecteur de valeurs des variables $V_i \in I_j$ et μ_{A_i} , μ_{R_j} sont les fonctions d'appartenance de A_i et R_j respectivement.

b) projection

Cette étape calcule la projection de π sur le domaine de V_k , c'est-à-dire,

$$\pi_{\pi \downarrow V_k}(u_k) = \sup_{u_i, i \neq k} \pi(u_1, \dots, u_n). \quad (3.19)$$

Cette formulation se ramène à un problème d'optimisation non-linéaire qui peut être coûteuse en temps de calcul si elle est implémentée naïvement.

Nous allons illustrer ce calcul d'intervalles avec des variables interactives en prenant un exemple. Supposons trois variables V_1 , V_2 , V_3 . La connaissance de V_1 , V_2 et V_3 est définie

par leur distribution de possibilité non floue ($A_1 = [0,1]$; $A_2 = [1,2]$; $A_3 = [0,2]$). La distribution de possibilité jointe (également non floue) est définie de la manière suivante:

$$\pi(u_1, u_2, u_3) = \begin{cases} 1 & \text{si } u_1 + u_2 = u_3 \\ 0 & \text{sinon} \end{cases}$$

On constate ensuite que l'Equation 3.18 rend certaines combinaisons impossibles (par ex. $\pi(u_1, u_2, u_3) = 0 \forall u_3 \in [0,1]$). Si on s'intéresse à savoir la valeur de V_3 , on effectue la projection définie par l'Equation 3.19 ce qui fournit le résultat suivant:

$$\pi_{\downarrow V_3}(u_3) = \begin{cases} 1 & \text{si } u_3 \in [1,2] \\ 0 & \text{sinon} \end{cases}$$

Cet exemple est relativement simple à résoudre; il s'agit d'un cas avec un nombre limité de variables et des contraintes simples et linéaires. La complexité du problème augmente si l'on ajoute des variables et des contraintes.

Une technique qui permet d'effectuer efficacement la combinaison/projection consiste à utiliser un solveur de Problèmes de Satisfaction de Contraintes (CSP : Constraint Satisfaction Problem) (Tsang, 1993; Schiex, 1997). Notre problème, reformulé dans ce cadre, se définit par:

- un ensemble de variables $V = \{V_1, \dots, V_n\}$;
- un domaine de définition U_i , $i = 1 \dots n$, pour chaque variable;
- un ensemble de contraintes floues $\{C_1, \dots, C_c\}$, chaque C_i contraignant un sous-ensemble $\{V_{j_1}, \dots, V_{j_k}\}$ de V (les contraintes désignent ici aussi bien les relations entre les propriétés de sol que les données floues sur chacune des variables, ces données correspondant simplement à des contraintes unaires);
- une requête portant sur la valeur (floue) d'une variable particulière, disons X_k .

L'outil utilisé dans cette étude est un outil développé pour la résolution des CSP flexibles (i.e. variables et contraintes floues) en appliquant des algorithmes spécifiques de résolution (Prosser, 1993). Ces algorithmes visent d'une part à réduire l'espace dans lequel il faut chercher des solutions (cette opération s'appelle le filtrage), et d'autre part à parcourir cet espace de manière arborescente pour y trouver les solutions. La combinaison de ces deux techniques revient à répéter l'application de la combinaison/projection sur des sous-ensembles de contraintes. Le filtrage dans le cas des contraintes numériques de notre application peut exploiter le calcul d'intervalles flous.

3.4 Interaction probabilité - possibilité

3.4.1 Comparaison des deux mesures

Les définitions et les propriétés énoncées plus haut laissent clairement apparaître des analogies entre les possibilités et les probabilités. En particulier, une distribution de possibilités joue un rôle comparable à une densité de probabilité comme base de définition de mesure d'incertitude: possibilité et nécessité dans un cas et probabilité dans l'autre. Une distribution de possibilité représente seulement l'ordre sur les possibilités des événements élémentaires. Cette représentation ordinale de l'incertitude contraste avec la nature numérique qui est attachée aux probabilités par l'hypothèse sous-jacente d'additivité.

La théorie des probabilités est particulièrement adaptée pour la modélisation du hasard, i.e., lorsqu'on dispose d'un modèle probabiliste. Mais, elle ne permet pas de représenter l'incertitude dans le cas d'une absence complète d'information. Par exemple, en utilisant la carte des pédopaysages, on veut déterminer l'UTS présente à un endroit précis. En considérant que la seule information dont on dispose est la localisation de l'UPP qui contient 3 UTS: UTS1, UTS2 et UTS3, les résultats s'exprimeraient en théorie de probabilités, par $P(\text{UTS1}) = P(\text{UTS2}) = P(\text{UTS3}) = 1/3$. Ceci entraîne, par exemple, qu'il est 2 fois plus probable que le point appartienne à UTS 1 ou UTS 2 qu'à UTS 3. Ceci est contradictoire avec le fait de déclarer que l'on est en situation d'ignorance totale. Donc, les probabilités ne permettent pas de représenter l'ignorance: supposer l'égalité de chance ne revient pas à exprimer l'ignorance.

La théorie des possibilités est mieux adaptée à la représentation des informations imprécises. Elle permet, en particulier, de coder de façon absolue l'ignorance totale, en considérant, pour les cas de notre exemple, qu'il est complètement normal que chacune des UTS soit trouvée. On obtient alors $\Pi(\text{UTS1}) = \Pi(\text{UTS2}) = \Pi(\text{UTS3}) = 1$.

Un autre avantage de la théorie des possibilités par rapport à la théorie de probabilités est que contrairement à cette dernière, elle limite la propagation d'éventuelles erreurs contenues dans les données initiales grâce à l'utilisation des opérateurs min et max à la place du produit et de la somme. Seule la notion d'ordre a réellement de l'importance.

3.4.2 Transformations de distributions de probabilité et de distributions de possibilité

Le paragraphe 3.1 a évoqué la possibilité d'utiliser un mélange de types de données dans la résolution d'un problème pratique. C'est le cas qui se présente dans notre étude. En effet, l'information pédologique principale à utiliser est composée d'expressions linguistiques données par le pédologue. Le traitement de ce type d'information s'effectue mieux dans le cadre de la théorie des possibilités. Cependant, il existe aussi des données quantitatives pour

lesquelles on est capable de déterminer la distribution de fréquence, et qu'on est donc capable de traiter de façon probabiliste (e.g. les données utilisées pour la construction des fonctions de pédotransfert).

L'utilisation de toute l'information disponible, relative à un problème donné, est souhaitable et ceci nécessite donc une transformation des mesures de possibilité en mesure de probabilité ou inversement (Dubois *et al.*, 1993; Klir & Parviz, 1992; Dubois & Prade, 1998). Pour la présente étude, nous nous intéressons seulement à la transformation probabilité \rightarrow possibilité car la transformation possibilité \rightarrow probabilité nécessite une hypothèse supplémentaire à introduire, une information arbitraire amenant un gain artificiel de précision. Une transformation d'une mesure de probabilité en mesure de possibilité se paye principalement par une perte de spécificité de l'information; cette perte n'ayant aucune conséquence sur la validité de ce qui est déduit.

La relation entre la probabilité et la possibilité est clairement décrite dans un principe proposé par Zadeh (1978b) (possibility/probability consistency principle). La connexion entre la possibilité et la probabilité exprimée dans ce principe est faible, mais importante pour les transformations probabilité \rightarrow possibilité. Ce principe dit que la probabilité d'un événement ne peut jamais dépasser la possibilité, c'est-à-dire, pour les cas extrêmes, qu'un événement impossible est aussi improbable et qu'un événement sûr (probabilité=1) est forcément possible. Ceci peut être traduit par:

$$\forall x \in X \quad P(x) \leq \Pi(x). \quad (3.20)$$

Dans la littérature, différentes transformations ont été présentées. Parmi elles, celle proposée par Dubois et Prade (voir Dubois *et al.*, 1993 ou Dubois & Prade, 1998), a été retenue. Cette transformation est basée sur un principe qui essaye de garder au mieux la richesse de la distribution de probabilité. Ceci aboutit à deux caractéristiques importantes de cette transformation. La première concerne la préservation du mode, c'est à dire, que dans cette transformation l'événement le plus probable devient l'événement le plus possible. La deuxième caractéristique concerne la préservation de l'ordre des événements à probabilité intermédiaire.

Considérons, $X = \{x_1, x_2, \dots, x_n\}$, $p_i = P(\{x_i\})$, $\pi_i = \Pi(\{x_i\})$ et l'ordre des valeurs de probabilité tel que: $p_1 \geq p_2 \geq \dots \geq p_n$. La solution pour obtenir la distribution de possibilité optimale (i.e. qui perd le moins d'information possible) proposée par Dubois & Prade est la suivante.

$$\forall i = 1, n, \quad \pi_i = \sum_{j=i, n} p_j \quad (3.21)$$

Par exemple, soit $X = \{x_1, x_2, x_3, x_4\}$ avec $p_1 = 0.4$, $p_2 = 0.3$, $p_3 = 0.2$ et $p_4 = 0.1$. En appliquant la formule 3.21, on obtient $\pi_1 = 1$, $\pi_2 = 0.6$, $\pi_3 = 0.3$ et $\pi_4 = 0.1$. L'annexe XIII

expose en détail la procédure suivie dans la présente étude pour la transformation d'une distribution de probabilité sous la forme d'un histogramme sur un domaine continu.

3.5 Application des sous-ensembles flous dans les domaines pédologiques et géographiques

Afin de compléter notre analyse de la théorie des possibilités, nous présentons ici une brève analyse bibliographique des applications des sous-ensembles flous et de la théorie des possibilités existant dans l'information pédologique et dans l'information géographique. Depuis la mise en place de leurs cadres théoriques, les ensembles flous et plus spécifiquement la théorie des possibilités ont trouvé des applications dans des domaines très divers comme, par exemple, la robotique, l'industrie ou l'aide à la décision en général (Dubois & Prade, 1988; Klir & Yuan, 1995). Dans les domaines de la pédologie et de l'information géographique, les applications des ensembles flous et de la théorie de possibilité ont également été nombreuses ces dernières années (McBratney & Odeh, 1997; Burrough & McDonnell, 1998).

La théorie des ensemble flous trouve son intérêt dans ces deux domaines pour deux raisons principales: (i) la connaissance du phénomène géographique dont on dispose (une propriété du sol, un paramètre d'un modèle de simulation ou une limite sur une carte) est généralement imprécise et (ii), beaucoup de phénomènes géographiques sont continus dans l'espace géographique ou dans l'espace multi-dimensionnel défini par les variables utilisées pour décrire ce phénomène. L'application des ensembles flous dans le domaine de la pédologie et de l'information géographique peut être résumée en différents thèmes de développement.

Premièrement, les sous-ensembles flous ont trouvé des applications dans l'analyse spatiale combinant plusieurs sources d'informations (Altman, 1994) et dans le domaine de l'évaluation des terres (Droesen & Geelen, 1993; Burrough, 1989; Tang & Van Ranst, 1992). De tels types d'analyse ont été réalisés en intégrant un système d'expert avec un SIG (Leung & Leung, 1993a,b; Saint-Joan & Desachy, 1995; Zhu *et al.*, 1996). La logique floue permet de prendre en compte la description imprécise des paramètres des paysages ou d'autres phénomènes géographiques utilisés pour la résolution d'un problème donné et l'incertitude associée, par exemple, aux erreurs d'observation. Les sous-ensembles flous permettent ainsi de traiter des requêtes mal définies comme celle-ci (Altman, 1994): *Quelles sont les surfaces proches de la ville, au sud de la rivière et aptes à l'activité agricole?*

La résolution du problème posé utilise une série de cartes floues contenant l'information sous la forme de degrés d'appartenance pour chaque variable nécessaire à l'analyse. Ces cartes floues sont ensuite utilisées dans des opérations logiques pour déduire l'information cherchée. Les résultats montrent généralement davantage de détail dans les

résultats; ce qui peut rendre service aux décideurs (Burrough, 1989). Ce type d'analyse est donc capable de traiter l'information imprécise, de répondre à des questions mal définies et de propager l'imprécision dans l'analyse. De plus, Heuvelink & Burrough (1993) démontrent un risque d'erreur moindre dans une analyse basée sur des ensembles flous que dans une analyse basée sur des ensembles booléens.

Un deuxième thème dans lequel les sous-ensembles flous ont été appliqués concerne la modélisation. Schulz & Huwe (1997) et Dou *et al.* (1995) ont utilisé les sous-ensembles flous pour pouvoir prendre en compte l'incertitude sur certains paramètres dans la modélisation hydrologique du sol. Dans le domaine de la géostatistique, les sous-ensembles flous ont été utilisés par Bardossy *et al.* (1990a,b) pour la représentation des paramètres incertains du variogramme.

En ce qui concerne la propagation d'incertitude, on trouve peu de références traitant de la comparaison entre des approches stochastiques et des approches basées sur des sous-ensembles flous (Dou *et al.*, 1995; Schulze & Huwe, 1997). Ces auteurs n'ont pas utilisé une approche CSP, mais d'autres méthodes d'optimisation pour chercher les bornes des intervalles flous. Ils concluent néanmoins que l'application des sous-ensembles flous pour traiter l'information imprécise s'avère plus flexible qu'une approche stochastique en particulier pour la modélisation hydrologique. On peut cependant déplorer que les auteurs n'aient pas validé leurs résultats en les comparant avec des valeurs réellement observées. Ceci aurait permis de déterminer avec plus de fondement l'intérêt de chaque approche. C'est ce que nous avons tenté de faire dans la présente étude.

Un dernier thème concerne l'utilisation des sous-ensembles flous dans la représentation des caractéristiques continues du sol et autres phénomènes géographiques. Cette continuité ou gradualité affecte deux aspects: (i) la gradualité des limites géographiques et (ii) la gradualité dans l'espace multivariable. L'application des sous-ensembles flous pour la représentation de la gradualité des limites géographiques a été étudiée par Leung (1987), Lagacherie *et al.* (1996) et Wang & Hall (1996).

En ce qui concerne la gradualité dans l'espace multivariable, la méthode de classification numérique, fuzzy c-means, est une méthode de classification basée sur la logique floue qui prend en compte cette gradualité naturelle entre les classes multivariées (Bezdek, 1981). Cette méthode effectue un groupement en déterminant pour chaque point d'observation un degré d'appartenance à chaque "cluster" prédéfini. Cette méthode de classification a été utilisée par de nombreux auteurs dans le domaine de la classification des données pédologiques (Burrough *et al.*, 1997; McBratney & De Gruijter, 1992; Odeh *et al.*, 1992; Lagacherie *et al.*, 1997; De Bruin & Stein, 1998).

3.6 Conclusion

Ce chapitre a exposé deux cadres théoriques différents permettant de traiter différents types d'information: la théorie des probabilités et la théorie des possibilités. La théorie de probabilité est un outil idéal pour la formalisation de l'incertitude dans les cas où la connaissance est basée sur la distribution de fréquences. En revanche, la théorie des possibilités est idéale pour la formalisation de l'information incomplète exprimée en terme de propositions imprécises et vagues.

On a vu que l'information disponible pour la cartographie des propriétés hydriques à l'échelle régionale est généralement imprécise et qu'elle peut être traduite sous forme d'intervalle de valeurs. De plus, dans la plupart des cas, il existe une incertitude sur les bornes inférieures et supérieures de ces intervalles. En tenant compte de ces caractéristiques des données à traiter, la théorie de possibilité est bien la plus appropriée pour répondre aux besoins de l'application visée. L'information nécessaire pour l'estimation des propriétés hydriques peut ainsi être représentée sous la forme d'une distribution de possibilité. Cette représentation est considérée comme plus fidèle à l'information réellement disponible qu'une fonction de densité de probabilité uniforme ou suivant la loi normale.

Les deux cadres théoriques possèdent des techniques permettant de propager l'imprécision et l'incertitude dans le cas où une série d'opérations est effectuée sur un ensemble de variables. La simulation de Monte Carlo est fréquemment appliquée pour ceci dans le cas où les variables sont décrites en distribution de probabilité. Il est à remarquer qu'une simulation de Monte Carlo peut tenir compte des interactions entre variables en effectuant un tirage aléatoire conditionnel.

Pour un système (i.e estimation de la réserve utile pour une UTS) dans lequel les variables sont représentées sous la forme de distribution de possibilité, il a fallu trouver un moyen d'effectuer le calcul d'intervalles avec des variables interactives. La technique proposée, basée sur la satisfaction de contraintes floues, est tout à fait adaptée à ce problème.

Ayant proposé le cadre théorique et les techniques de traitement nécessaires, il est maintenant possible de mettre en œuvre la cartographie des propriétés hydriques à l'échelle régionale. Cette mise en œuvre sera illustrée à l'aide d'une étude de cas développée dans les chapitres suivants.

4 La région d'étude et les données utilisées

Les trois premiers chapitres ont été consacrés à l'exposition de la problématique et des objectifs de la présente étude ainsi qu'au cadre théorique permettant la cartographie des propriétés hydriques à l'échelle régionale. Il est maintenant possible de mettre en pratique la cartographie des propriétés hydriques du sol dans le cadre de la théorie des possibilités.

Cette mise en œuvre de la cartographie des propriétés hydriques, à l'échelle régionale, nécessite de choisir une zone d'étude et de déterminer les données dont nous avons besoin concernant la zone d'étude pour pouvoir effectuer cette analyse.

Ce chapitre présente l'étude de cas permettant d'effectuer l'analyse cartographique proposée. Nous présentons, tout d'abord, le milieu naturel de la zone d'étude, pour ensuite, aborder, en détail, les informations nécessaires à cette analyse et surtout disponibles sur la zone d'étude VHOL. Ces données constituent donc l'ensemble des matériaux utilisés dans la suite de ce travail.

4.1 Présentation de la zone d'étude

Afin de pouvoir tester l'analyse cartographique proposée dans les chapitres précédents, il est nécessaire de choisir un terrain d'étude concret. Le choix de la zone d'étude s'est porté sur la zone définie par les moyennes et basses vallées de l'Hérault, de l'Orb et du Libron situées dans le département de l'Hérault (34). Cette région constitue une partie de la zone côtière languedocienne (voir Figure 4.1) dans laquelle de nombreuses études ont été effectuées par le laboratoire Science du Sol de l'INRA, Montpellier. Les études menées par ce laboratoire sont très diverses aussi bien au niveau du thème scientifique traité qu'au niveau de l'échelle géographique. Les sujets d'études varient depuis la cartographie générale du sol (Bonfils, 1993; Bornand *et al.*, 1992) jusqu'à la cartographie des propriétés hydriques du sol (Leenhardt, 1991), la cartographie automatisée (Lagacherie, 1992), le fonctionnement hydrologique d'un bassin versant (Voltz & Andrieux, 1994) ou la dynamique des nappes phréatiques (Tassinari, 1998). Quant à l'échelle à laquelle les études ont été effectuées, elle varie d'un bassin versant d'une centaine d'hectares (Voltz & Andrieux, 1994) jusqu'à la zone d'étude entière.

Il existe une forte correspondance entre la zone d'étude choisie et la petite région naturelle Moyenne Vallée de l'Hérault, étudiée par Lagacherie (1992). La zone des vallées de l'Hérault, de l'Orb et du Libron, abrégée VHOL, est une extension au sud et à l'ouest de la Moyenne Vallée de l'Hérault. Les caractéristiques naturelles de ces zones sont presque identiques.

Au sein de la zone d'étude VHOL, une zone de test a été définie. Cette zone de test à été choisie pour effectuer les analyses cartographiques préliminaires. Cette zone de test a déjà été utilisée dans d'autres études (Leenhardt, 1991; Lagacherie, 1992).

Pratiquement, on décrit successivement le milieu naturel de la zone des vallées de l'Hérault, de l'Orb et du Libron (§ 4.1.1), puis les caractéristiques naturelles de la zone de test (§ 4.1.2) et on explicite les raisons de son choix comme zone de test.

4.1.1 Les vallées de l'Hérault, de l'Orb et du Libron

4.1.1.1 Localisation et généralités géographiques

La zone d'étude est située dans la plaine côtière du Languedoc-Roussillon entre Béziers et Montpellier (voir Figure 4.1). Elle est constituée par les moyennes et basses parties des bassins versants de l'Hérault, de l'Orb et du Libron. L'activité fluviale de ces rivières a produit un relief peu accidenté et l'altitude varie entre 0 mètre au bord de la Méditerranée et 120 mètres pour les parties les plus à l'intérieur des terres.

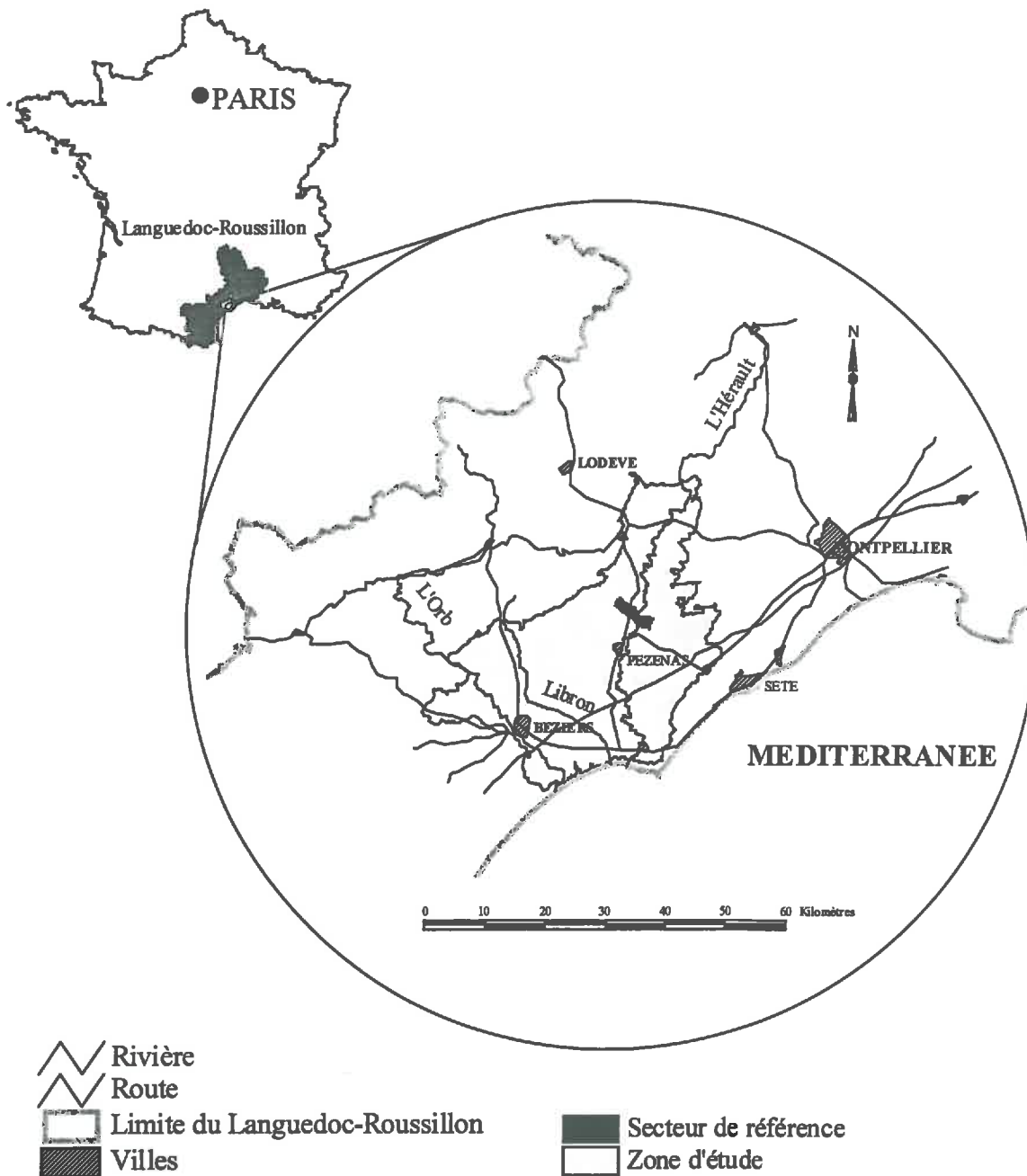


Figure 4.1: Localisation de la zone d'étude et du secteur de référence

Cette zone est entourée de secteurs plus montagneux: au nord-ouest, la Montagne Noire et au nord et à l'est, les plateaux calcaires comme le Causse d'Aumelas et le Causse du Larzac. Il existe une discontinuité remarquable entre les massifs montagneux et la zone d'étude en ce qui concerne le paysage et l'occupation du sol. Pour la zone d'étude, le paysage peut-être caractérisé avec un nombre réduit d'entités géomorphologiques: pentes quasi nulles sur les épandages d'alluvions étagées en terrasses successives, long glacis entrecoupé de talus

sur matériaux sableux marins tertiaires et pentes fortes, voir petites falaises, localisées au niveau des versants ou sommets des collines.

L'occupation végétale du sol de la zone est agricole avec une forte majorité de vignes. Le pourcentage de vignes a diminué récemment, mais il reste encore très majoritaire. Une reconversion vers d'autres cultures (blé dur, maïs, etc.) a été entreprise principalement pour les secteurs à sols profonds avec grandes parcelles facilement accessibles. Une partie des parcelles est laissée à l'abandon (friches).

Le climat de la zone d'étude est de type méditerranéen, sub-humide, à saison sèche prolongée (Tassinari, 1998). Les saisons sont contrastées. Les automnes et les hivers sont très pluvieux. Les étés sont chauds et secs avec des averses de forte intensité. Les précipitations s'étagent selon un gradient globalement Sud-Nord de 600 à 850 mm par an (Météo France, 1992). Cependant, la variabilité interannuelle de la pluviométrie est importante, illustrée par le Tableau 4.1.

Année	Pluie (mm)
1993	984
1994	765
1995	605
1996	1512
1997	964

Tableau 4.1: Variabilité interannuelle de la pluviométrie mesurée au site expérimental du Bassin Versant de Roujan (d'après Tassinari, 1998).

La variation entre la température minimale mensuelle (moyenne des températures minimales journalières) et la température maximale mensuelle (moyenne des températures maximales journalières) est comprise entre environ 3°C pour le mois de janvier et 30°C pour le mois de juillet (Météo France, 1992).

Les sols subissent généralement un excès d'eau en automne et en hiver, mais il existe un déficit hydrique sévère pendant l'été. Le bilan annuel (précipitation - évapotranspiration potentielle) est négative pour la plupart des années (Tassinari, 1998).

4.1.1.2 Principales entités géologiques

La géologie est un facteur important intervenant dans la pédogénèse et donc dans la distribution des types de sols. Ceci est particulièrement valable pour le milieu méditerranéen. C'est pourquoi il est nécessaire, dans le cadre de cette étude, de présenter brièvement la formation historique du substratum géologique de la zone d'étude et les formations géologiques affleurant, à présent, dans la zone d'étude.

La zone d'étude est couverte par six cartes géologiques au 1/50.000 du BRGM: les feuilles d'Agde (n°1040, 1978), de Béziers (n°1039, 1982), de St. Chinian (n°1014, 1982), de

Lodève (n°989, 1982), de Pézenas (n°1015, 1981) et de Sète (n°1016, 1967). Le substratum le plus ancien affleurant dans la zone date du milieu du Tertiaire (le Miocène). La zone d'étude est, à cette période, sous l'influence d'une transgression datée de l'Helvétien (env. 10 millions d'années). Les sédiments marins déposés à cette époque forment le substratum de base de la zone d'étude. Ce substratum correspond à l'unité m2a de la carte géologique et comprend des dépôts fins de marnes bleues et des dépôts sableux plus grossiers. Dans ce substratum épais, il existe des bancs de calcaires coquillers. La fin de cette période de transgression est marquée par des facies de calcaires durs et des molasses à dragées de quartz (unité m2b). Au Pliocène, la zone d'étude est affectée par des influences continentales. Ces dépôts caillouteux à galets affleurent principalement dans la partie ouest de la zone d'étude. Ensuite, le Quaternaire est dominé par l'influence fluviatile qui a façonné la géomorphologie actuelle. Les rivières de l'Hérault et de l'Orb ont creusé leurs lits dans le substratum en déposant des alluvions plus ou moins caillouteuses. De nombreux niveaux de terrasses fluviatiles ont été identifiés (Barrière, 1971a,b), mais ils ne sont pas toujours représentables sur les cartes géologiques. Les cartes géologiques distinguent trois niveaux principaux de terrasses: (i) les terrasses villafranchiennes (unité Fv), (ii) les terrasses du Pléistocène moyen, subdivisées en 2 sous-niveaux (unités Fya et Fyb) et (iii) les alluvions récentes (Fz).

Une activité volcanique pendant le Quaternaire a provoqué des épandages basaltiques traversant la zone d'étude de Clermont l'Hérault jusqu'à Agde. Ces coulées de basalte ont fortement marqué le paysage en certains endroits comme le plateau de Nizas et le Mont St. Loup près d'Agde. Enfin, la période actuelle est marquée par des activités géomorphologiques comme le colluvionnement modifiant légèrement la structure géologique.

4.1.1.3 Principaux types de sol

La couverture pédologique de la zone d'étude est fortement liée à la structure géologique. Ceci a été le sujet de nombreuses études pédologiques à différentes échelles. Deux cartes pédologiques couvrant la zone d'étude peuvent être citées ici:

- carte des sols au 1/100.000 feuille de Lodève (Bonfils, 1993)
- carte des pédopaysages au 1/250.000 du Languedoc-Roussillon (Bornand *et al.*, 1992)

Dans la suite de ce paragraphe, la caractérisation des différents types de sols présents dans la zone d'étude est basée sur ces deux sources d'information. L'Annexe I donne la partie de la carte des pédopaysages au 1/250.000 du Languedoc-Roussillon (Bornand *et al.*, 1992) correspondant à la zone d'étude. Cette carte montre la domination (environ 55%) de cinq unités cartographiques de sols: 173A, 309A, 309V, 372P et 552T. Seules ces cinq unités de pédopaysages principales de la zone d'étude seront présentées en détail. Les sols moins répandus dans la zone d'étude seront brièvement cités.

a. *Les sols développés sur 'molasse' de l'Helvétien-Miocène (UPP 552T).*

Les affleurements du substratum Helvétien-Miocène occupent une grande partie de la zone d'étude. Le terme molasse, utilisé ici, recouvre en fait des roches mères différentes et responsables de la différenciation des unités typologiques de sols. Les molasses sont de caractères très hétérogènes, ce qui a produit un grand nombre d'unités typologiques de sols au sein de cette unité de pédopaysage. Bornand *et al.* (1992) en distinguent douze. Seuls les sols les plus répandus sont présentés ici.

Un premier groupe de sols concerne les sols développés sur des faciès molassiques de calcaire gréseux dur ou de conglomérats caillouteux. Ces sols sont généralement non cultivés, peu profonds, très calcaires et classifiés en lithosols et rendzines (unité 60a: Bonfils, 1993; unité 514: Bornand *et al.*, 1992). Un deuxième groupe de sols concerne les sols développés sur les sédiments sableux. Ce sont les unités 64 et 65 de la carte de Bonfils (1993) et les unités 974 et 506 de la carte de Bornand *et al.* (1992). Ces sols bruns calcaires présentent une texture équilibrée (limono-sabo-argileuse à limono-argilo-sableuse), une faible charge en éléments grossiers. La profondeur est le facteur distinguant les deux types de sols (Bonfils, 1993). La profondeur du sol moyenne varie entre 50 et 100 cm pour l'unité 64 (unité 974: Bornand *et al.*, 1992) et elle est supérieure à 100 cm pour l'unité 65 (unité 506 : Bornand *et al.*, 1992).

b. *Les sols sur matériaux pliocènes et villafranchiens (UPP 372P et 309V).*

Bonfils (1993) distingue un groupe de quatre types de sols couvrant les glacis d'épandage pliocènes. Ce sont les sols 71, 72, 73, et 74 sur la carte de Bonfils (1993) qui correspondent respectivement aux unités 321, 341, 352, 351 sur la carte de Bornand *et al.* (1992). Ces sols se situent principalement dans la partie centrale de la zone d'étude (rive droite de l'Hérault). Les différences entre ces unités typologiques de sols sont principalement liées à la texture du sol et à la présence d'un encroûtement par exemple. En général, ces sols sont calciques ou calcaires, graveleux et de couleur rougeâtre.

Les sols sur les plateaux couverts par les dépôts du Villafranchien sont marqués par l'évolution fersiallitique qu'ils ont subie (unité 75 : Bonfils, 1993; unité 345 : Bornand *et al.*, 1992). Ces sols sont moyennement profonds et présentent une différenciation en horizons importante. L'horizon A, sableux, très faiblement argileux, contient 30 à 40 % de graviers quartzueux. L'horizon appauvri E, quand il est présent, contient des graviers quartzueux et du sable délavé. Les horizons rouges Bt sont à forte teneur en argile et à revêtements argileux. Ils ont un pH acide (5.0) et une charge caillouteuse composée de galets quartzueux.

c. *Les sols sur les terrasses du Pléistocène moyen (UPP 309A).*

Les restes des terrasses du Pléistocène moyen se retrouvent généralement, dans la morphologie régionale, entre les terrasses alluviales récentes et les formations miocènes. Les types de sols qui se sont formés sur les terrasses fluviales du Pléistocène moyen sont très divers, du fait des processus de colluvionnement issus des niveaux miocènes voisins. En l'absence de colluvionnement (unités 78 et 79 : Bonfils, 1993; unité 303 : Bornand *et al.*, 1992), les sols se développent directement sur les alluvions anciennes. Ces sols sont profonds, caillouteux dès la surface, de texture sableuse et toujours calciques.

Les zones colluvionnées sont localisées au contact des versants et des talus Miocènes (unité 305 : Bornand *et al.*, 1992). L'épaisseur des colluvions et la composition des matériaux sont très variables. Généralement, ces sols présentent un horizon d'épaisseur variable qui surmonte les alluvions anciennes. Cet horizon est limoneux, limoneux-argileux ou argilo-limoneux, calcaire et pauvre en éléments grossiers.

d. *Les sols de la plaine alluviale Holocène (UPP 173A)*

Les sols de la plaine alluviale sont récents et donc peu évolués (Fluvisols). Ils sont profonds, calcaires, généralement brun ou brun-jaunâtre (10YR) et peu chargés en éléments grossiers. Cette unité de pédopaysage est constituée de deux principales unités typologiques de sols: 82a et 82b (Bonfils, 1993) ou 101 et 102 (Bornand *et al.*, 1992). Selon la logique des processus de sédimentation fluviale, la texture varie progressivement avec la distance au fleuve. Les unités 82a et 102 des cartes respectives correspondent aux sols proches de la rivière. Ces sols sont plus sableux que les sols situés au contact des terrasses du Pléistocène moyen et plus éloignés de la rivière (unités 82b et 101 respectivement). Ces derniers sont de texture argilo-limoneuse.

e. *Les sols peu représentés dans la zone d'étude.*

- Les sols bruns andiques, formés sur les coulées de basaltes ou sur les débris volcaniques des plateaux basaltiques: sol moyennement profond (40-60 cm), pierreux et caillouteux, limono-sablo-argileux.
- Les sols hydromorphes dans la partie sud-est de la zone d'étude concernant les vallées des ruisseaux du bassin de Thau.

En résumé, l'organisation spatiale du sol est visible dans le paysage grâce à la forte relation entre le sol et les éléments paysagers comme la lithologie et le modelé du relief. Les cartes pédologiques au 1/100.000 (Bonfils, 1993) et au 1/250.000 (Bornand *et al.*, 1992)

représentent donc une source précieuse de connaissances sur la couverture pédologique de la zone d'étude. Ces informations s'avéreront très utiles dans la suite de ce travail.

4.1.2 Le secteur de référence Adissan-Cazouls-Montagnac

Au sein de la zone d'étude VHOL, une zone de test a été définie pour effectuer les analyses cartographiques préliminaires. Cette zone de test, le secteur Adissan-Cazouls-Montagnac (abrégié ACM), a déjà été utilisée dans d'autres études (Leenhardt, 1991; Lagacherie, 1992). Ce secteur de 940 ha est localisé dans la moyenne vallée de l'Hérault et il s'étend sur les 3 communes citées (voir Figure 4.1). Il a également été considéré comme secteur de référence dans la méthode de cartographie du sol à grande échelle proposée par Favrot (1989). Dans ce cadre méthodologique, une zone limitée est choisie de manière à ce qu'elle soit la plus représentative possible d'une zone environnante beaucoup plus vaste. Le positionnement du secteur a été choisi de manière à ce qu'il renferme le maximum d'unités de pédopaysages présentes dans la région dont il est sensé être représentatif (la petite région naturelle). Ceci explique sa forme globalement allongée selon un axe perpendiculaire à celui de la vallée, recoupant les grands ensembles morphologiques présents.

Le périmètre choisi a donc déjà été le sujet de nombreuses études effectuées au laboratoire. Premièrement, ce secteur a fait l'objet d'une cartographie à grande échelle réalisée par Clochard *et al.* (1991) (voir Annexe II). Cette carte a été réalisée selon une démarche classique de terrain, c'est-à-dire avec 855 sondages au total (1 sondage pour 0.9 ha). Leenhardt (1991) a ensuite caractérisé, avec plus de détail, les unités typologiques de sols inventoriées à l'aide d'une série de fosses pédologiques (description morphologique des horizons, caractérisations analytiques). Cette information a ensuite été utilisée pour la cartographie des propriétés hydriques du sol. Lagacherie (1992) utilise l'information pédologique disponible sur le secteur de référence pour l'automatisation de la cartographie des sols d'une petite région naturelle (Favrot, 1989).

Le choix de l'utilisation de ce secteur dans la présente étude est basé sur deux raisons principales:

- a) la représentativité du secteur pour la zone d'étude est importante. Quatre unités de pédopaysages majeures sont présentes dans le secteur. Seule l'unité 372P, présentée dans le paragraphe précédent, est absente du secteur.
- b) une connaissance importante de la couverture pédologique est déjà disponible. Ceci permettra d'évaluer les résultats obtenus au cours des différentes étapes de l'analyse cartographique proposée.

4.2 Présentation des données utilisées

L'approche cartographique proposée au cours des trois premiers chapitres nécessite un certain nombre d'informations différentes. Ces informations nécessaires peuvent être divisées en cinq types différents. Il s'agit, premièrement, de l'information pédologique issue de la carte pédologique régionale: elle fournit la localisation des unités cartographiques de sols et les descriptions des unités typologiques de sol. Deuxièmement, l'estimation de la réserve utile nécessite des fonctions de pédotransfert permettant l'estimation de θ_{cr} et θ_{pf} (cf. § 2.4.1). Troisièmement, nous avons besoin d'information exprimant les interactions entre les propriétés du sol intervenant dans l'estimation de la réserve utile. Le quatrième type d'information est l'information géographique nécessaire pour la cartographie des UTS au sein des UCS. Enfin, le dernier type de données dont nous avons besoin sont les données qui permettent d'évaluer l'estimation obtenue avec des données mesurées sur le terrain. Les paragraphes suivants exposent successivement les quatre types de données qui ont été utilisés dans la présente étude.

4.2.1 L'information pédologique régionale

L'information pédologique utilisée dans la présente étude est la base de données géographiques de sols du Languedoc-Roussillon au 1:250.000 (Bornand *et al.*, 1992). La cartographie du sol au 1:250.000 du Languedoc-Roussillon s'inscrit dans le cadre de l'opération Inventaire, Gestion et Conservation des sols du Ministère de l'agriculture (IGCS) (Bornand *et al.*, 1989; Favrot *et al.*, 1994; Jamagne *et al.*, 1995). La base de données géographiques de sols est composée de la carte en format papier et numérique et d'une base sémantique informatisée. L'Annexe I montre la partie de la carte qui couvre la zone d'étude VHOL.

Le Tableau 4.2 donne le nombre d'unités de pédopaysages et le nombre d'unités typologiques de sols pour l'ensemble de la zone d'étude et pour la zone de test.

La zone VHOL		Secteur ACM	
Nb UPP	Nb UTS	Nb UPP	Nb UTS
36	96	4	26

Tableau 4.2: Le nombre d'UPP et d'UTS par zones d'étude.

Nous utilisons cette base de données pour dériver deux types d'informations. Premièrement, nous avons besoin de l'information sur les propriétés du sol intervenant dans l'estimation de la réserve utile (voir Equation 2.1 et 2.2) (i.e. la texture, le taux de matière organique, la densité apparente, la profondeur). La base de données nous fournit la plupart de ces informations (toutes sauf la densité apparente), par l'intermédiaire d'une description tex-

tuelle ou d'une description à l'aide d'un profil de référence. Les caractéristiques de ces deux types de descriptions ont été exposées dans le paragraphe 2.3.3. Deuxièmement, l'information de la base de données est également utilisée dans la localisation des UTS au sein des UPP. Il s'agit ici de (i) l'information géographique (i.e. la localisation des UPP) et (ii) l'information descriptive des UTS concernant les variables environnementales.

4.2.2 Les fonctions de pédotransfert

Comme indiqué auparavant, l'estimation de la réserve utile nécessite des fonctions de pédotransfert permettant d'estimer θ_{pf} et θ_{cr} . Les caractéristiques hydrodynamiques des sols de la zone d'étude ont été étudiées par Leenhardt (1991) et Moulènes (1993). Au cours de ces études, des travaux expérimentaux ont été effectués pour obtenir les courbes de rétention en eau sur une série d'échantillons de terre venant de différents types de sol présents dans la zone d'étude VHOL.

Les analyses effectuées sur ces échantillons consistaient d'abord en la mesure des propriétés dites classiques comme la composition granulométrique, la densité apparente et la matière organique. Ces mêmes échantillons ont ensuite été analysés pour connaître leur comportement hydrodynamique en déterminant la courbe de rétention en eau. La courbe de rétention a été déterminée en mesurant la teneur en eau pour des points particuliers de potentiels matriciels, en l'occurrence aux pressions -3, -10, -30, -100 -300 et -1500 kPa. Pour cela, on utilise des presses à plaques appelées presses de Richards. Chaque presse est hermétique et doit pouvoir maintenir dans son enceinte la pression appliquée par air comprimé. Au total, 370 échantillons ont été analysés.

Cette série de données a permis d'établir les fonctions de pédotransfert déterminant θ_{cr} et θ_{pf} , ce qui correspond aux potentiels matriciels de -10 et -1500 kPa. Pour cela, on utilise la corrélation entre les variables hydrodynamiques et les propriétés du sol classiques. Or, la relation entre les variables en question est différente pour des types de sols différents. Il est alors nécessaire de classer les données en groupes homogènes pour lesquels des fonctions de pédotransfert seront établies séparément. La classification des données pour lesquelles les fonctions de pédotransfert ont été établies est basée sur l'origine des matériaux pédologiques. Moulènes (1993) propose le regroupement des observations en 6 classes de matériaux: alluvions, terrasses, terrasses villafranchiennes, molasses à argile de type 1-1, molasse à argile de type 1-2 et les calcaires marneux. Dans la suite de ce travail, une simplification de cette classification a été utilisée en regroupant les molasses et les calcaires marneux dans une seule classe.

Les fonctions de pédotransfert établies pour les quatre classes de matériaux ont la forme générale suivante:

	β_0	β_1 (da)	β_2 (arg)	β_3 (lim)	β_4 (sab)	β_5 (mo)	R^2
<i>Alluvions récentes (N = 70)</i>							
θ_{cr}	24.19	-11.80	0.14	0.28	0.00	0.00	0.79
θ_{pf}	11.91	-2.60	0.28	0.00	-0.07	0.14	0.90
<i>Terrasses moyennes (N = 52)</i>							
θ_{cr}	43.98	-11.12	0.00	0.00	0.22	0.00	0.58
θ_{pf}	39.90	-12.31	0.00	-0.09	-0.20	0.00	0.72
<i>Terrasses villafranchiennes (N = 59)</i>							
θ_{cr}	44.61	-13.01	0.05	0.00	-0.19	-0.86	0.81
θ_{pf}	29.12	-13.08	0.17	0.00	-0.09	0.00	0.76
<i>Molasses (N = 189)</i>							
θ_{cr}	35.24	-10.54	0.00	0.04	0.08	0.00	0.42
θ_{pf}	16.20	-6.11	0.20	0.00	-0.04	0.00	0.71

Tableau 4.3: Les fonctions de pédotransfert utilisées pour l'estimation de θ_{cr} et θ_{pf} .

$$\theta_X = \beta_0 + \beta_1 * da + \beta_2 * arg + \beta_3 * lim + \beta_4 * sab + \beta_5 * mo \quad (4.1)$$

où $\beta_0 \dots \beta_5$ constituent les coefficients de la régression multiple. Les fonctions de pédotransfert présentées par Moulènes (1993) utilisent les variables limon fin, limon grossier, sable fin et sable grossier en variables indépendantes. Ces variables ont été regroupées en limon et sable pour obtenir une seule variable exprimant le taux de sable et le taux de limon. Les fonctions de pédotransfert utilisées dans la présente étude sont présentées dans le Tableau 4.3. Ces fonctions ont été calculées en utilisant le module Régression Multiple du logiciel statistique Statistica™ (Statistica, 1995). Les régressions multiples sont calculées en ajoutant ou supprimant les variables individuellement du modèle, à chaque étape de la régression, jusqu'à ce que le "meilleur" modèle de régression soit obtenu. L'ajout ou la suppression d'une variable est fonction de la valeur F du modèle. Une valeur de $F = 1$ a été appliquée comme valeur seuil pour l'ajout de la variable au modèle. Un coefficient de 0 dans le tableau signifie l'absence de la variable dans le modèle. La valeur de R^2 indique la variabilité expliquée par la régression par rapport à la variabilité totale. Elle est alors une mesure de qualité de la régression. La majorité des régressions montre une qualité élevée (R^2 supérieur à 0.7). Seules les fonctions de pédotransfert pour θ_{cr} des Molasses et des Terrasses Moyennes montrent une qualité moindre.

4.2.3 Information disponible exprimant d'autres relations entre les propriétés du sol

Les fonctions de pédotransfert pour estimer θ_{pf} et θ_{cr} sont des relations indispensables pour pouvoir estimer la réserve utile. Cependant, nous avons vu que les variables intervenant dans cette estimation sont interactives. Il est donc nécessaire d'exprimer ces interactions de

manière mathématique (principalement sous forme d'égalités et d'inégalité). Pour ceci, on distingue deux sources d'information. Premièrement, ces relations mathématiques peuvent être issues d'une connaissance générale dans le sol (relations entre propriétés du sol au sein d'un horizon ou entre propriétés du sol issues d'horizons différents). Deuxièmement, nous pouvons dériver des relations empiriques entre les propriétés de sols à partir d'un ensemble de descriptions de profils de sol. Dans la présente étude, nous utilisons la base de données des profils de sols STIPA disponible au laboratoire Science du Sol de l'INRA Montpellier. Cette base de données contient un ensemble d'environ 2000 profils (4500 horizons environ), provenant principalement de la région Languedoc-Roussillon et dont les principales propriétés de sols ont été analysées. Ces relations concernent indifféremment des relations entre des propriétés du sol au sein d'un horizon ou entre des propriétés du sol issues d'horizons différents

4.2.4 Les données géographiques supplémentaires utilisées

L'analyse cartographique présentée dans les premiers chapitres propose de déterminer les unités typologiques de sol à l'intérieur des unités de pédopaysages à l'aide d'informations géographiques supplémentaires. Le paragraphe 2.3.2 a exposé quel type de données pouvait être utilisé pour effectuer une telle analyse. La disponibilité de l'information à l'échelle régionale est un critère important du choix des données utilisables dans la délimitation des unités typologiques. Les données utilisées dans la présente étude sont des données facilement accessibles et couvrant toute la zone d'étude VHOL.

L'analyse proposée est basée sur des opérations ensemblistes des différentes informations géographiques qui peuvent être effectuées à l'aide d'un système d'information géographique. Une analyse en mode raster est la mieux adaptée à notre problématique. Ceci nécessite la préparation de toutes les données dans le même format (rastérisation des couvertures en mode vecteur) et une décision concernant la résolution à utiliser. Une résolution de 50 mètres a été choisie pour l'analyse préliminaire sur la zone ACM. Une résolution plus grossière sera choisie pour l'analyse sur la zone VHOL.

4.2.4.1 *Les cartes géologiques*

Les cartes géologiques au 1/50.000 de BRGM (feuilles de Agde, Béziers, St Chinian, Lodève, Pézenas et Sète) (BRGM, 1981) ont été numérisées au format ARC/INFO™. Un regroupement de ces cartes a été effectué. L'unification de ces couvertures pose un problème pratique. Réalisées par des géologues différents, les cartes n'ont pas toutes le même niveau de détail et il n'y a pas toujours de continuité des unités entre deux cartes. Dans les cas où les

discontinuités étaient trop importantes, les limites des plages cartographiques ont été corrigées.

4.2.4.2 *L'information topographique*

Une grande quantité d'information concernant la topographie est stockée sous forme numérique dans la base de données topographiques (BDTopo[®]) éditée par l'IGN. La BDTopo[®] est constituée de plusieurs couvertures vectorielles contenant l'information des cartes topographiques au 1/25.000: les courbes de niveau, les routes, l'occupation de sol, les villes, etc.. Trois thèmes de la BDTopo[®] ont été utilisés dans la présente étude: les courbes de niveau, les points cotés pour la construction du MNA (voir le paragraphe suivant) et l'occupation du sol.

La localisation des plages cartographiques concernant l'occupation du sol est relativement précise. Cependant, dans la BDTopo[®], il existe des plages cartographiques non renseignées. Ces plages ont été corrigées à l'aide de l'occupation des sols trouvée sur la carte topographique. Les plages non renseignées correspondaient principalement à des zones viticoles.

4.2.4.3 *Le modèle numérique d'altitude (MNA)*

Le modèle numérique d'altitude a été construit à partir des informations extraites de la BDTopo[®]: couverture des courbes de niveau avec un intervalle de 5 m et couverture des points cotés. Ces couches d'information ont été utilisées dans la procédure de triangulation d'ARC/INFO[™]. La résolution du MNA issu de cette procédure était fixée à 50 m.

Le modèle numérique de terrain constitue l'information de base pour le calcul de la couverture de pente. Il existe plusieurs procédures pour la détermination de la pente (Skidmore, 1989). La procédure proposée par Horn (1981) a été appliquée dans la présente étude.

4.2.5 **Les données disponibles pour la validation**

Les paragraphes précédents ont exposé les données nécessaires à la cartographie de la réserve utile dans la zone d'étude VHOL. Cependant, il est important de savoir si les estimations produites seront compatibles avec la réalité. Ceci soulève donc le problème de la validation des résultats et la nécessité de trouver des données de terrain permettant une comparaison adéquate avec les résultats des modèles. Les données disponibles permettant d'effectuer une évaluation pour chacune des trois étapes principales de la démarche générale de la présente étude (cf. Figure 1.2.) sont explicitées ci-après.

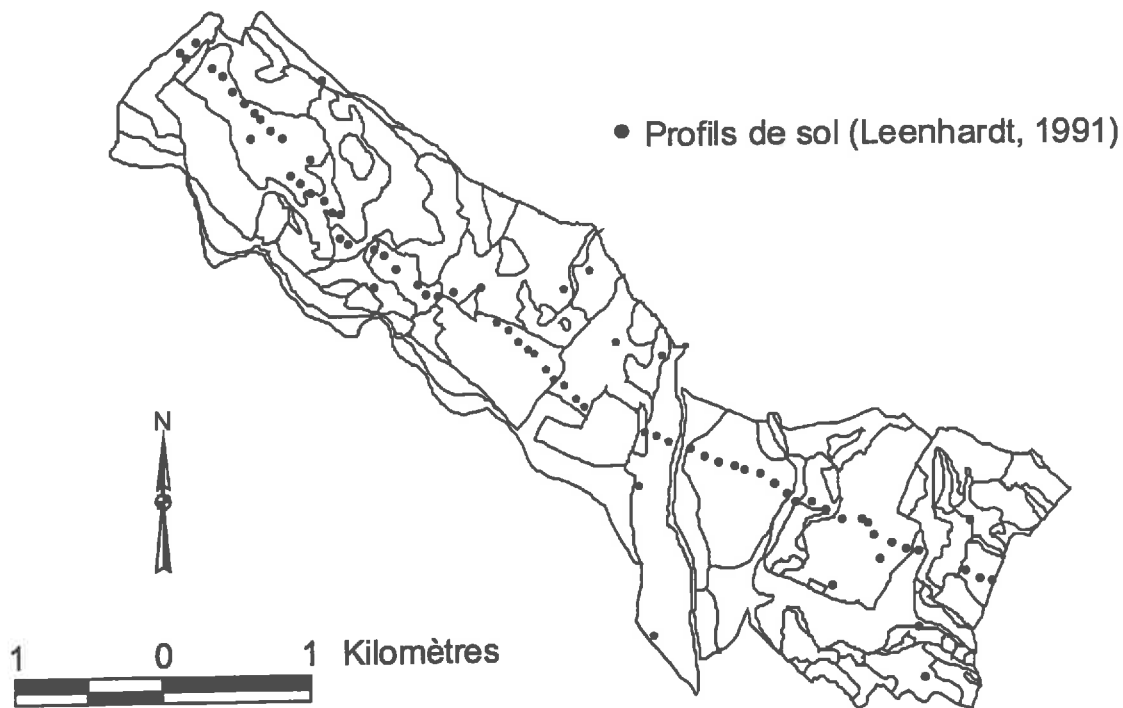


Figure 4.2: Localisation des profils pour lesquels la réserve utile est déterminée. L'arrière plan de cette carte est la carte de sols du secteur au 10.000^{ème} (voir Annexe II).

4.2.5.1 Observations ponctuelles des propriétés hydriques du sol

La présente étude se concentre sur la cartographie de la réserve utile du sol. Des observations ponctuelles de la réserve utile sur le terrain sont donc nécessaires pour pouvoir comparer les résultats obtenus par les modèles avec la réserve utile réelle observée sur le terrain. L'observation de la réserve utile réelle nécessite un travail de terrain extrêmement lourd qui n'est pas envisageable lorsque une grande série de mesures est nécessaire. Il est néanmoins possible d'estimer précisément la réserve utile en déterminant au laboratoire θ_{cr} et θ_{pf} pour chaque horizon distingué et prélevé sans perturbation. C'est ainsi que 66 profils de sol ont été utilisés pour déterminer la réserve utile totale du sol. Ces profils ont été décrits et analysés, pour les propriétés hydriques, dans le cadre des travaux de Leenhardt (1991). La figure 4.2 montre la localisation de ces profils au sein de la zone ACM. Le nombre d'observations par unité de sol reste néanmoins réduit ce qui limite une éventuelle évaluation statistique.

Une deuxième série de données de validation concerne les observations ponctuelles de la teneur en eau au point de flétrissement (θ_{pf}). Les observations sont néanmoins plus faciles à obtenir pour cette variable, si on s'intéresse à sa valeur à une profondeur donnée. La profondeur choisie pour le prélèvement des échantillons, dans la présente étude, était fixée à 40 cm.

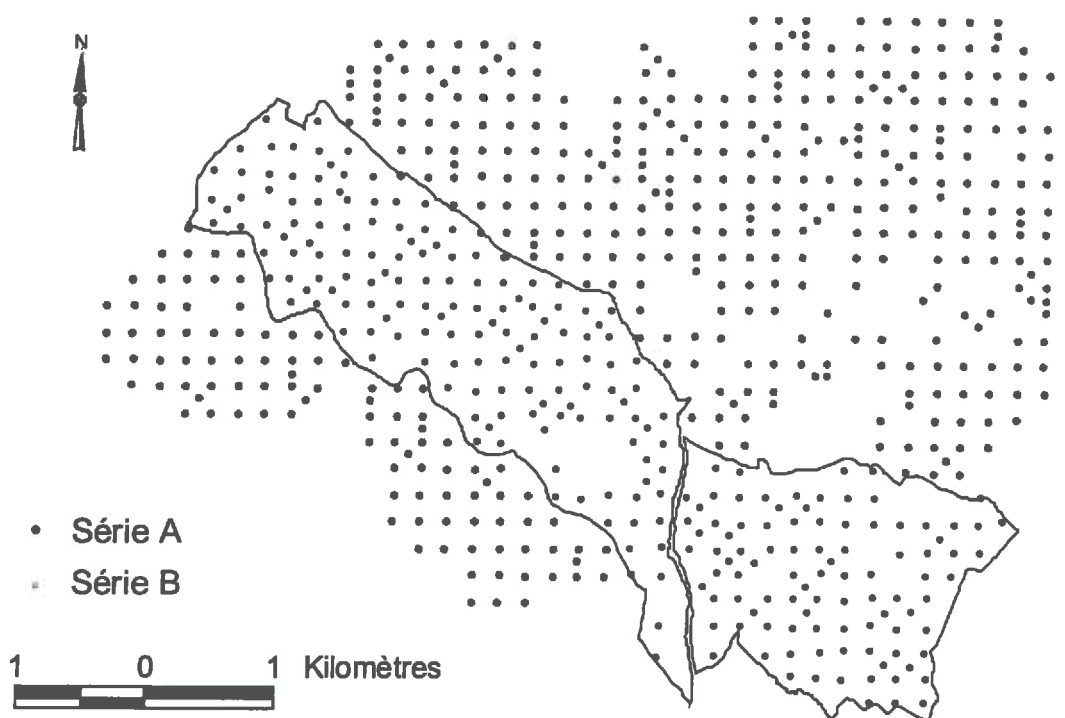


Figure 4.3: Localisation des observations de la teneur en eau au point de flétrissement.

Ceci correspond, dans la plupart des cas et pour la zone d'étude, à l'horizon perturbé par les opérations culturales.

Les échantillons ont été prélevés selon une grille régulière (voir la figure 4.3). Chaque point d'observation a été affecté à une UTS de la carte des pédopaysages, pendant les travaux de terrain. Les analyses ont été effectuées sur des échantillons remaniés et tamisés. La teneur en eau au point de flétrissement (sous pression de -1500 kPa) a été déterminée à l'aide des presses à plaques (presses de Richards). Au total, 673 échantillons ont été analysés.

Les observations sur θ_{pf} ont été effectuées à deux moments différents. Une première série de 247 observations a été effectuée au cours de la présente étude à l'intérieur de la zone ACM (série A). Une deuxième série de 426 observations a été menée à l'extérieur de la zone ACM (série B). Les données récoltées ont déjà été utilisées pour la cartographie à grande échelle (Voltz *et al.*, 1997).

4.2.5.3 Carte contenant la localisation exacte des UTS

Une partie de l'analyse cartographique adoptée dans la présente étude concerne la localisation des UTS au sein des UPP. Une évaluation de cette localisation est possible lorsque la localisation réelle des UTS est connue. Ceci revient ainsi à comparer la localisation de

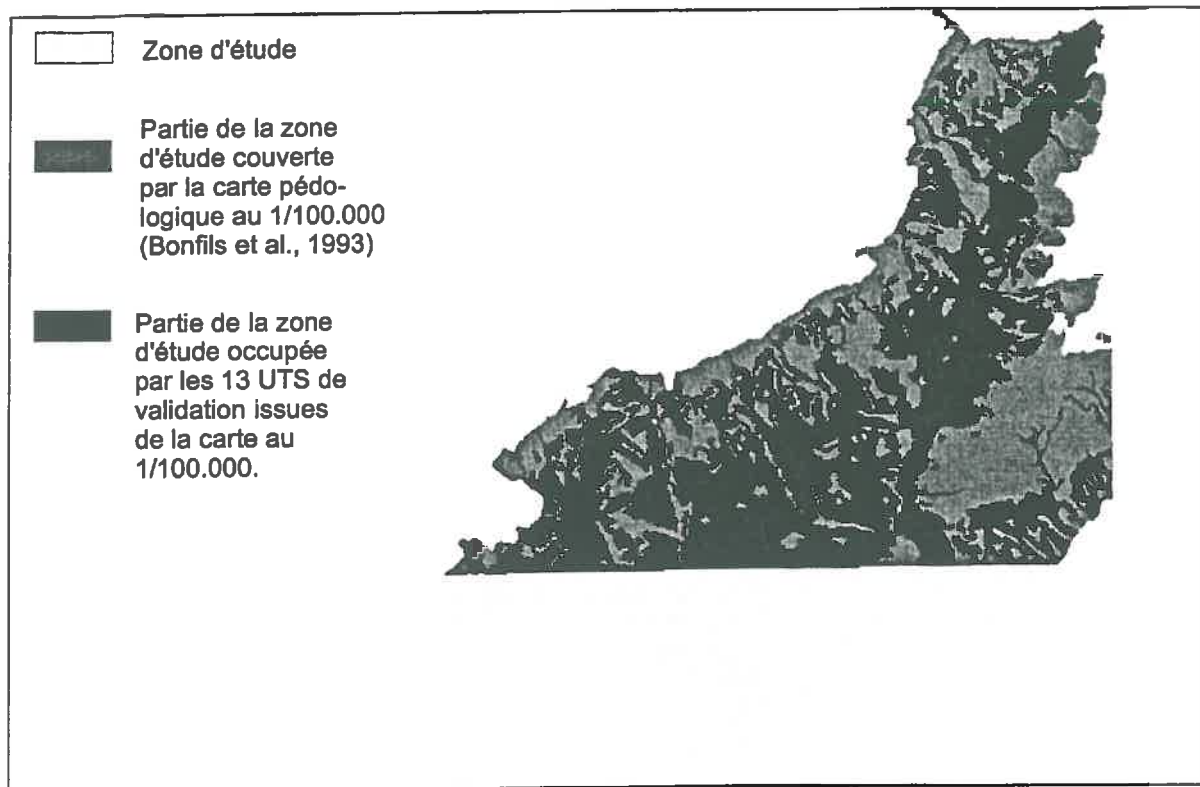


Figure 4.4: Partie de la zone d'étude utilisée pour l'évaluation de la cartographie des UTS au sein des UCS.

l'UTS obtenue par la modélisation avec la localisation de l'UTS représentée sur une carte plus précise.

Il est donc nécessaire de choisir une carte pédologique adaptée à ce type d'évaluation. Il faut, par conséquent, que cette carte représente des unités cartographiques de sols simples et que la définition de ces unités soit le plus proche possible des descriptions des UTS de la carte des pédopaysages. Une carte à grande échelle (par ex. 1:10.000) n'est pas adaptée à ce type d'évaluation, car elle donne une représentation des sols beaucoup trop précise, avec des descriptions de sols trop différentes de celles de la carte des pédopaysages. La carte pédologique au 100.000^{ème} (feuille de Lodève, (Bonfils, 1993)) est considérée comme la plus adaptée à l'évaluation proposée. Cette carte représente le sol à l'aide d'UCS simples (i.e. chaque UCS représente une seule UTS) et la description des UTS correspond bien aux descriptions des UTS de la carte des pédopaysages de Bornand *et al.* (1992). Un ensemble de 13 UTS, pour lesquelles la correspondance a pu être facilement effectuée, a pu être distingué. Ce rapprochement des UTS des deux cartes a déjà été présenté dans le paragraphe 4.1.1.3. La Figure 4.4 montre la répartition géographique des 13 UTS dans la zone d'étude VHOL. L'ensemble des 13 UTS couvre une surface d'environ $49 \cdot 10^3$ ha; ce qui correspond à 40% de la zone d'étude.

5 Représentation de la connaissance

Le chapitre précédent a présenté les données disponibles pour la cartographie de la réserve utile en ce qui concerne la zone d'étude des Vallées de l'Hérault, de l'Orb et du Libron. Dans le troisième chapitre, nous avons également vu que le cadre de la théorie des possibilités semblait le mieux adapté à notre situation, en tenant compte des caractères imprécis et incertains des données à manipuler.

Ce cadre théorique étant choisi, il est alors nécessaire de coder toute information en terme de distributions de possibilité permettant ensuite la manipulation de cette information. Nous sommes donc confrontés, ici, à un problème de représentation de la connaissance. Il est important de noter que cette connaissance est constituée de divers types d'information, chaque type d'information demandant une interprétation et une représentation spécifique. C'est pour cette raison qu'un chapitre entier est consacré à ce problème.

Ce chapitre a pour objectif d'exposer la représentation de la totalité de la connaissance en terme de distributions de possibilité et de contraintes, afin de répondre à la première question principale posée à la fin du deuxième chapitre.

Dans un premier temps, nous présentons la représentation de la connaissance intervenant dans l'estimation de la réserve utile pour chaque UTS. Il s'agit ici de l'interprétation et de la représentation des données issues de la description textuelle ou du profil de référence et de la représentation d'un maximum de connaissance en ce qui concerne les interactions entre variables. Dans un deuxième temps, nous exposons la représentation de l'information et de son incertitude intervenant dans la cartographie des unités typologique des sol.

5.1 Représentation de la connaissance intervenant dans l'estimation de la réserve utile

La première étape de la démarche générale de cette étude consiste à estimer la réserve utile pour chaque UTS présente dans la zone d'étude (voir Figure 1.2). Cette estimation s'effectue par une série d'opérations arithmétiques sur des variables imprécises et interactives. Nous avons vu que l'arithmétique définie par la théorie des possibilités permet d'effectuer un calcul d'intervalles avec des variables interactives (cf. chapitre 3). Ceci implique que la connaissance nécessaire à la détermination des réserve utiles des UTS soit préalablement exprimée sous une forme compatible avec les outils.

La totalité de la connaissance intervenant dans l'estimation de la réserve utile peut être divisée en deux parties. La première concerne l'information issue de la description des UTS (description textuelle ou description par profil de référence). Cette information fournit les données de base à l'estimation de la réserve utile. Il est alors nécessaire de traduire cette information généralement imprécise (qualitative ou par intervalle de valeurs) en distribution de possibilité. La deuxième partie de la connaissance à représenter est l'ensemble des interactions entre variables (i.e. les contraintes). Ces contraintes peuvent exprimer des interactions entre variables issues d'un même horizon ou des interactions liant les variables issues de deux horizons différents.

Ce paragraphe présente successivement l'interprétation et la représentation de l'information issue de la description des UTS ainsi que la définition des contraintes. Pour ces deux types de connaissance, on cherche également à quantifier et représenter leur incertitude à l'aide d'informations supplémentaires. Ces informations permettront, par exemple, de définir la largeur de l'intervalle du support des distributions de possibilité ou de quantifier l'incertitude sur les fonctions de pédotransfert.

5.1.1 Interprétation et représentation de la description textuelle

L'estimation de la réserve utile pour une unité typologique de sol, telle qu'elle est proposée dans le paragraphe 2.4.1, nécessite une information spécifique sur certaines propriétés de sol: composition granulométrique, profondeur du sol, pierrosité du sol, taux de matière organique et densité apparente. La description textuelle d'une unité typologique contient un ensemble d'expressions qualitatives exprimant de manière imprécise la valeur de ces propriétés de sol (toutes sauf la densité apparente: voir § 5.1.1.5). L'interprétation de cette information qualitative est basée sur les codifications de description des sols. Les guides de description de sol (Baize & Jabiol, 1995; Soil Survey Division Staff, 1993) donnent des intervalles de valeurs correspondant au code donné dans la description textuelle. La définition

d'une distribution de possibilité est donc possible en utilisant ces intervalles. Néanmoins, les bornes de ces intervalles doivent être interprétées avec précaution et il est important d'introduire une flexibilité sur leurs valeurs.

Une distribution de possibilité sous forme de trapèze pour une variable réelle est considérée comme la façon la plus naturelle de représenter une information qualitative (Klir & Yuan, 1995) lorsque l'on a seulement une connaissance sur le noyau et le support de la distribution; ce qui est souvent le cas dans la pratique. La transition entre le noyau et le support est alors supposée linéaire. Hootsmans (1996) a analysé l'influence de la forme de la transition (linéaire, courbe ou cloche) entre le noyau et le support sur la représentation floue des données spatiales. La transition linéaire a été considérée comme la mieux adaptée aux applications spatiales.

En fait, c'est une représentation simple (donc facile à coder) et qui, compte tenu des opérations min et max utilisées dans le traitement, n'amène pas de différence majeure avec un autre codage conservant le même ordre sur les valeurs possibles. Pour toute distribution de possibilités concernant une variable réelle X , le noyau et le support doivent être interprétés de la façon suivante:

- compte tenu de l'information disponible, rien n'empêche la variable X de prendre sa valeur dans le noyau (i.e. le noyau est l'ensemble des valeurs complètement possibles)
- il est sûr que la valeur de la variable appartient au sous-ensemble défini par le support (i.e. toute valeur hors du support est impossible)
- une valeur du support, non incluse dans le noyau, est d'autant plus surprenante (d'autant moins possible) qu'elle est éloignée du noyau

Pour l'interprétation de la description textuelle, nous pouvons établir maintenant deux règles générales pour la détermination du noyau et du support.

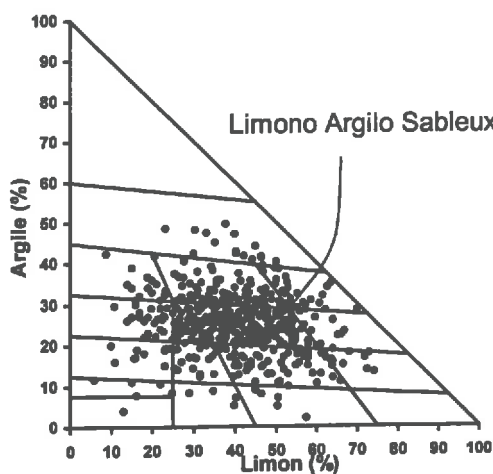
- Le noyau est défini par l'intervalle qui correspond au code donné par le guide de description des sols
- Le support est défini en tenant compte de l'erreur maximale tolérée sur la détermination de la propriété concernée, exclusion faite des situations exceptionnelles (grosses erreurs de transcription, pédologue peu expérimenté, moyens d'étude anormalement limités). Les bornes inférieures et supérieures du support sont alors définies respectivement en retranchant l'erreur maximale tolérée de la borne inférieure du noyau et en ajoutant l'erreur maximale tolérée à la borne supérieure du noyau.

La détermination du support des distributions de possibilité reste généralement délicate et elle est souvent définie de manière un peu plus arbitraire que le noyau. La suite de ce

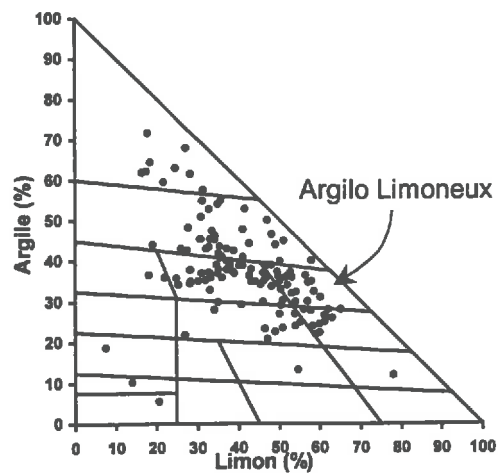
paragraphe expose, en détail pour chaque variable intervenant dans l'estimation de la réserve utile, la détermination des distributions de possibilités avec plus de détail.

5.1.1.1 Texture

La codification utilisée pour décrire la texture du sol est définie par un triangle de texture. Le triangle utilisé pour l'étude des pédopaysages du Languedoc-Roussillon est le triangle GEPPA (Baize, 1988). L'annexe III montre ce triangle avec les intervalles correspondants pour le taux d'argile, de limon et de sable. Ces intervalles ont été obtenus par projection des classes de texture sur les axes du triangle. Le noyau des distributions de possibilités est ainsi défini par ces intervalles. La détermination du support des distributions de possibilités est plus délicate. Elle nécessite une connaissance de l'incertitude sur l'appréciation de la texture par le pédologue sur le terrain (texture-terrain, Legros, 1996). Une méthode pour obtenir cette connaissance est de comparer la texture-terrain avec les résultats des analyses de la composition granulométrique effectuée au laboratoire. Les données disponibles dans la base de données STIPA, qui contient une grande quantité de descriptions de profils au sein du Languedoc-Roussillon, permettent d'effectuer cette comparaison. Les profils ont été décrits par un grand nombre de pédologues différents.



(a) N= 565



(b) N= 210

Figure 5.1: Comparaison entre l'appréciation de la texture par le pédologue sur le terrain et la composition granulométrique mesurée. Exemples de textures (a) Limono-Argilo-Sableuses et (b) Argilo-Limoneuses.

L'affichage du nuage de points obtenus par les analyses dans le triangle de texture révèle ensuite la qualité de l'appréciation de la texture. La Figure 5.1 présente deux exemples d'une telle comparaison. Ces deux exemples montrent une qualité d'appréciation différente pour les deux classes de texture. La même analyse pour les autres classes de texture montre

également une variation très importante de la qualité de l'appréciation. Le support des distributions de possibilités pour chaque composition granulométrique est, dans un premier temps, déterminé par une marge d'erreur de 10% de chaque côté du noyau. Dans la plupart des cas, cette marge d'erreur de 10 % couvre en grande partie la dispersion du nuage de points extérieurs à la classe de texture. Le pourcentage des points à l'extérieur de cette marge de 10 % est présenté dans le Tableau 5.1 pour les trois variables respectives.

On constate que le pourcentage des points qui n'est pas inclus dans les 10% de marge d'erreur de chaque côté du noyau est inférieure à 15% dans 34 cas sur 58. Pour certaines classes de texture, nous observons des erreurs très élevées (AL,LA, L et LS par ex.). Ces dernières classes se situent toutes au bord du triangle.

Classe	N	Argile	Limon	Sable
AA	12	8.3	0.0	0.0
A	274	10.2	0.0	0.0
As/AS	114	1.8	7.0	5.3
ALS	144	4.9	0.7	1.4
AL	210	13.3	32.9	25.2
LAS	565	5.0	3.2	2.8
LA	354	8.5	39.0	32.2
SA	138	13.0	10.1	21.0
SAL	81	6.2	3.7	7.4
LSA	539	18.6	11.9	7.8
L	236	8.9	56.4	47.5
S	348	6.6	7.2	11.5
SS	12	25.0	16.7	25.0
SL	600	13.0	9.7	13.8
LS	707	27.7	32.8	12.3
LL	1	100.0	0.0	0.0

Tableau 5.1: Pourcentage des observations, issues de la base de profil STIPA, dont le taux d'Argile, Limon ou Sable (mesuré au laboratoire) n'est pas inclus dans la marge d'erreur de 10% à chaque côté du noyau (le noyau étant défini par la projection des bornes des classes sur les axes du triangle de texture GEPPA).

Même si le Tableau 5.1 montre des pourcentages importants des observations non inclus dans la marge d'erreur de 10% à chaque côté du noyau, on considère ce marge de 10% acceptable pour exprimer le support pour deux raisons principales.

- Les concepts de texture-terrain et de granulométrie ne sont pas strictement équivalents (Legros, 1996). La granulométrie, nous l'avons, dit, donne la proportion de sables, limons et d'argile. La texture indique, ce qui est un peu différent, un comportement de sol sableux, de sol limoneux ou de sol argileux. Or des particules très fines, de la taille des argiles, peuvent être reliées si fortement et en assemblages si compacts qu'elles forment des petits agrégats de la taille de sables et se comportent comme sables (terres

légères et poreuses). Par ailleurs, le calcaire présent dans les sols sous forme de fine particules conduit à surestimer la proportion de limons par rapport à l'analyse granulométrique effectuée après destruction des calcaires (Baize, 1988).

- Lorsque le pédologue décrit l'UTS (i.e. la création de la description textuelle), il dispose généralement de toutes les informations recueillies sur le terrain (les analyses au laboratoire incluses). Ceci lui permet de fournir une information générale concernant l'UTS et de corriger ces observations de terrain erronées.

5.1.1.2 Profondeur du sol

L'interprétation de l'information concernant la profondeur du sol est difficile pour plusieurs raisons. D'abord, la profondeur peut être définie de deux façons différentes. Une façon de définir la profondeur du sol est la profondeur jusqu'à laquelle les racines ont été observées (Baize & Jabiol, 1995; Soil Survey Division Staff, 1993). Une deuxième définition de profondeur est la profondeur d'apparition du substrat géologique générant le sol. En interprétant la profondeur du sol dans la description textuelle, il faut donc savoir quelle définition le pédologue a appliqué pour la description de la profondeur totale du sol. Un deuxième problème lié à l'interprétation de la profondeur est l'absence d'une codification appliquée de façon générale. Seul le Soil Survey Division Staff (1993) donne une classification de la profondeur de l'enracinement (cf. le Tableau 5.2).

Classe	Nom	Traduction	Intervalle (cm)
1	very shallow	très peu profond	<25
2	shallow	peu profond	25-50
3	moderately deep	moyennement profond	50-100
4	deep	profond	100-150
5	very deep	très profond	>=150

Tableau 5.2: Classes de profondeur d'enracinement (d'après Soil Survey Division Staff, 1993).

Cette classification a été appliquée pour l'interprétation de la profondeur à partir des données textuelles. L'intervalle donné dans le Tableau 5.2 détermine alors le noyau des distributions de possibilités trapézoïdales dérivées de cette information. Le support est défini par une marge de 10 cm de chaque côté de l'intervalle. Cette marge de 10 cm de chaque côté représente l'erreur maximale tolérée après consultation de pédologues de terrain expérimentés, en l'absence d'autres données disponibles.

5.1.1.3 Pierrosité

La pierrosité est la proportion des éléments grossiers ramenée au volume de la masse de l'horizon ou du profil entier du sol. Les guides de description de sols proposent, comme pour la texture, différentes codifications pour la description de la pierrosité. Baize & Jabiol (1995) proposent une codification fortement liée à la possibilité de labourer le sol mécaniquement avec des machines agricoles. La codification proposée par le Soil Survey Division Staff (1993) (cf. le Tableau 5.3) est plus adaptée à la présente étude parce qu'elle propose un découpage en intervalles de valeurs qui correspond mieux aux différents types de sols présents dans la zone d'étude. Les intervalles proposés par cette codification définissent alors les noyaux des distributions de possibilités dérivées comme pour la variable précédente. Le support des distributions de possibilité est défini à partir de l'expérience de pédologues: une erreur maximale autorisée de 10% est considérée de chaque côté de l'intervalle du noyau.

Classes	Nom	Traduction	Intervalle (%)
1	no or slightly gravelly	légèrement caillouteux	0-15
2	gravelly	caillouteux	15-35
3	very gravelly	très caillouteux	35-60
4	extremely gravelly	extrêmement caillouteux	60-90
5	gravel	cailloutis	90-100

Tableau 5.3: Classes de pierrosité (d'après Soil Survey Division Staff, 1993).

Classe	Nom	Minimum	Minimum	Maximum	Maximum
		$\alpha=0.0$	$\alpha=1.0$	$\alpha=1.0$	$\alpha=0.0$
1	Absente	0	0	0	0
2	Faible	0	0	1	1.5
3	Moyenne	0	1	4	5
4	Assez forte	2.5	4	10	12.5
5	Forte	5	10	20	25
6	Très forte	15	20	30	35
7	Extrêmement forte	25	30	∞	∞

Tableau 5.4: Distributions de possibilité exprimant les différentes classes de teneur en matière organique (exprimées en %) (d'après Legros, 1996).

5.1.1.4 Matière Organique

L'appréciation de la teneur en matière organique est principalement basée sur la couleur du sol. A part cette relation forte entre la couleur et la teneur en matière organique du sol, le pédologue tient aussi compte de son expérience du milieu et de sa connaissance des sols. Legros (1996) propose une codification de teneurs approximatives en matière organique. Cette échelle d'appréciation a été adoptée dans la présente étude pour l'interprétation de

l'information textuelle concernant la teneur en matière organique. Le Tableau 5.4 donne les intervalles qui correspondent aux noyaux ($\alpha = 1.0$) et au support ($\alpha = 0.0$).

5.1.2 Interprétation des descriptions par profils de référence

L'information disponible dans les descriptions des profils de référence contient un intervalle de valeurs par variable d'intérêt. L'intervalle de valeurs donné dans le profil de référence définit le noyau des distributions de possibilité. Il reste maintenant à définir le support de ces distributions de possibilité.

Comme pour la définition du support des distribution de possibilité issues de la description textuelle, la support représente l'erreur maximale tolérée sur l'intervalle donné. Les bornes inférieures et supérieures du support sont alors définies respectivement en retranchant l'erreur maximale tolérée de la borne inférieure du noyau et en ajoutant l'erreur maximale tolérée à la borne supérieure du noyau. Il existe cependant peu d'information permettant la quantification de cette erreur maximale tolérée. La plupart des marges ont donc été choisies de manière arbitraire en restant cohérent avec les marges prises pour l'interprétation de la description textuelle. Ceci a conduit à la définition d'une marge de 10% pour les variables de texture et de pierrosité. Ensuite, une marge de 10 cm a été prise pour les distributions de possibilité sur la profondeur du sol.

L'utilisation des descriptions par profil de référence nécessite l'introduction d'une variable supplémentaire par rapport à l'ensemble des variables dérivées de la description textuelle. Il s'agit de l'épaisseur des horizons (voir §2.3.3.3). Une connaissance est également disponible pour cette variable sous la forme d'un intervalle de valeurs. Nous avons, comme pour la profondeur du sol, appliqué une marge de 10 cm de chaque côté de l'intervalle du noyau.

5.1.3 Densité apparente

A la différence des variables traitées précédemment, la densité apparente n'est pas usuellement fourni dans les base de données sol alors que c'est une variable importante dont il faut tenir compte dans cette étude. Ceci justifie donc un paragraphe particulier consacré à cette variable.

La densité apparente intervient dans l'estimation des propriétés hydriques pour chaque unité typologique de sols. Elle intervient dans chaque fonction de pédotransfert et elle explique généralement une grande partie de la variabilité de la régression multiple. De plus, la densité apparente est une variable intervenant dans l'estimation de la réserve utile. Cependant,

la disponibilité de l'information concernant la densité apparente par UTS est très réduite du fait de la difficulté à mesurer et apprécier sur le terrain cette variable.

Au cours de ses observations de terrain, le pédologue peut observer des différences de compacité du sol dans un profil, qui constituent des indications de différences en terme de masses volumiques (i.e. densité apparente). Cependant, les descriptions de sol données par le pédologue expriment souvent un changement relatif de la densité apparente et non pas une valeur exacte ou un intervalle de valeurs possibles pour cette densité comme c'est pourtant le cas pour les autres propriétés du sol mentionnées ci-dessus. Une appréciation précise de la densité apparente sur le terrain n'est pas possible et elle est généralement déterminée au laboratoire par des mesures sur des échantillons non remaniés, prélevés in situ. L'exécution de ces prélèvements et de ces analyses est relativement lourde et c'est pourquoi l'information sur la densité apparente des unités typologiques de sol est absente des descriptions textuelles et des descriptions des profils de référence. Il est donc important de trouver d'autres sources d'information pour la densité apparente.

Matériau	N	Moyenne	Minimum	Maximum	Ec-type
All. Réc.	70	1.42	1.16	1.90	0.13
Ter. Moy.	52	1.59	1.25	1.95	0.18
Ter. Villafr.	59	1.55	1.21	1.93	0.20
Molasse	189	1.48	1.03	1.98	0.15
Total	370	1.50	1.03	1.98	0.17

Tableau 5.5: Indicateurs statistiques exprimant les distributions de densité apparente par groupe de matériaux.

α -coupe	minimum	maximum	α -coupe	minimum	maximum
0.1	1.15	1.85	0.6	1.25	1.65
0.2	1.15	1.80	0.7	1.35	1.60
0.3	1.20	1.65	0.8	1.35	1.60
0.4	1.25	1.65	0.9	1.35	1.60
0.5	1.25	1.65	1.0	1.35	1.60

Tableau 5.6: Distribution de possibilité pour la densité apparente appliquée sur toute la zone d'étude.

Les données utilisées pour l'établissement des fonctions de pédotransfert pour les quatre groupes de matériaux lithologiques (voir le Tableau 4.2) sont les seules données considérées comme suffisamment représentatives de la zone d'étude. L'utilisation de ces données demande alors une transformation probabilité \rightarrow possibilité des histogrammes correspondants à chaque groupe de matériau, suivant la procédure décrite dans l'Annexe IV. Le Tableau 5.5 présente les indicateurs statistiques pour chaque groupe de matériaux et pour l'ensemble des observations de densité apparente. Ces indicateurs statistiques montrent une faible différence entre les différentes distributions de la densité apparente au sein de chaque groupe de maté-

riaux. Cette faible différence entre les groupes de matériaux permet d'établir une seule distribution de possibilité pour la zone d'étude entière. Le Tableau 5.6 montre les résultats obtenus après la transformation probabilité \rightarrow possibilité par α - coupe.

5.1.4 Contraintes exprimant les interactions entre variables dans l'estimation des propriétés hydriques

L'estimation de la réserve utile pour chaque UTS peut être considéré comme un problème de satisfaction de contraintes (voir §3.3.5.2). Dans la formulation d'un problème de satisfaction de contraintes, on cherche généralement à exprimer le maximum de connaissance pour orienter le plus efficacement la recherche des solutions. Ceci implique, pour l'estimation de la réserve utile que l'on cherche à représenter le maximum d'interactions entre les variables au moyen de contraintes. Ces contraintes peuvent être issues d'une connaissance générale sur le sol ou d'une connaissance empirique (provenant d'une série de données ponctuelles).

On distingue deux types de contraintes dans la présente étude: (i) les contraintes exprimant les interactions entre variables issues d'un même horizon et (ii) les contraintes liant les variables issues de deux horizons différents uniquement utilisée lorsque l'UTS est décrite sous forme de profil de référence. L'ensemble des contraintes utilisées dans la présente étude est exposé dans la suite pour les deux types cités précédemment.

5.1.4.1 *Interactions entre propriétés du sol appartenant à un même horizon*

Tout d'abord, la première série de contraintes à prendre en compte est l'ensemble des fonctions de pédotransfert déjà exposées dans le paragraphe 4.2.2. Ensuite, une deuxième contrainte utilisée dans la présente étude concerne l'équation mathématique linéaire qui détermine la réserve utile. La formulation de la RU donnée par l'Equation 5.1 (Baize & Jabiol, 1995) est légèrement différente de celle donnée par l'Equation 2.1 et contient deux variables de plus (voir § 2.4.1).

$$RU = (\theta_{cr} - \theta_{pf}) \times E \times da \times ((100 - pier)/100) \quad (5.1)$$

La réserve utile est calculée pour une profondeur E à partir de deux niveaux d'humidité: l'humidité à la capacité au champ, qui prend classiquement le nom de capacité de rétention (θ_{cr}), et l'humidité au point de flétrissement (θ_{pf}). Ces deux niveaux d'humidité correspondent respectivement à l'humidité volumique de pF2 (-10kPa) et pF4.2 (-1500kPa) et sont exprimés en pourcentages pondéraux (g/100g). La réserve utile doit être corrigée par la masse volumique du sol (densité apparente = da) et la pierrosité ($pier$) du sol.

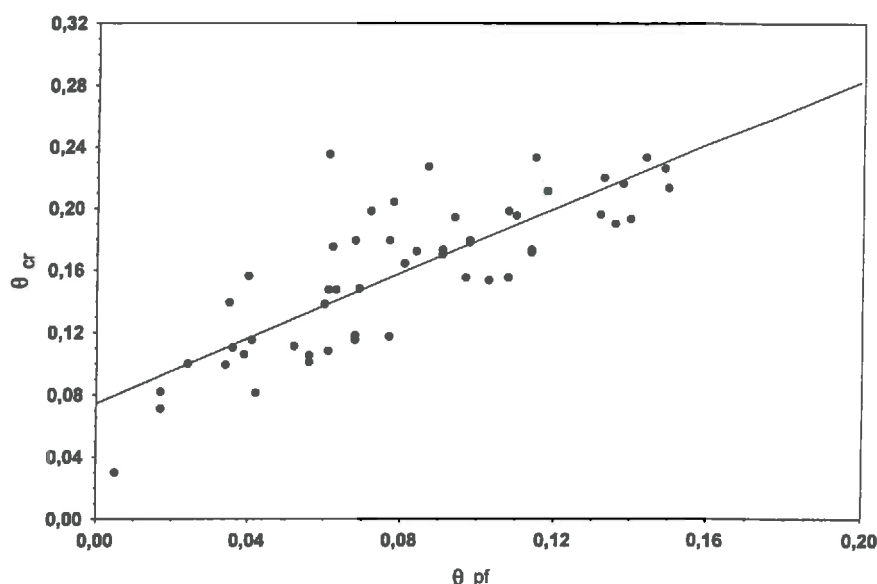


Figure 5.2: Nuage de points exprimant la relation entre le θ_{cr} et θ_{pf}

Les contraintes exprimées précédemment définissent les opérations arithmétiques qui déterminent les propriétés hydriques recherchées à partir de l'information disponible. Or, les variables impliquées dans ces opérations ne sont pas indépendantes. Il faut donc exprimer ces dépendances sous forme de contraintes.

Une première relation évidente concerne la relation entre les fractions granulométriques: argile, limon et sable. La composition granulométrique d'un sol, qui concerne seulement les particules fines du sol (< 2mm), est définie par les proportions des fractions granulométriques (Baize & Jabiol, 1995). Cette composition exprime donc les quantités relatives des fractions granulométriques et elle est normalisée sur 100% :

$$\%arg + \%lim + \%sab = 100 \quad (5.2)$$

Le rôle contraignant de l'Equation 5.2 peut être illustré par un exemple. Considérons les intervalles suivants pour les fractions granulométriques d'un sol argilo-sableux: Argile $\in [20, 50]$; Limon $\in [10, 30]$; Sable $\in [50, 80]$. La combinaison des fractions (Argile: 20, Limon:10, Sable:50) n'est alors pas possible. Lorsque deux fractions sont déterminées, la troisième est automatiquement connue.

Les fonctions de pédotransfert du Tableau 4.3 permettent alors d'estimer indépendamment les propriétés hydrodynamiques θ_{cr} et θ_{pf} avec une qualité statistique suffisante. Pourtant, cette indépendance entre θ_{cr} et θ_{pf} n'est pas réaliste, si l'on sait que ces deux propriétés sont fortement corrélées. Il est alors important d'introduire cette relation dans le système. La première relation est l'inégalité entre ces deux variables. θ_{cr} est toujours supérieur à θ_{pf} :

Matériau	Régression	R ²
Alluvions Récentes	$\theta_{cr} = 8.33 + 1.16 * \theta_{pf}$	0.51
Terrasses moyennes	$\theta_{cr} = 7.45 + 1.04 * \theta_{pf}$	0.65
Terrasses Villafranchiennes	$\theta_{cr} = 8.15 + 0.94 * \theta_{pf}$	0.68
Molasses	$\theta_{cr} = 10.46 + 0.772 * \theta_{pf}$	0.42

Tableau 5.7: Les régressions linéaires entre θ_{cr} et θ_{pf}

$$\theta_{cr} > \theta_{pf} \quad (5.3)$$

Cette inégalité exprime une relation incomplète puisqu'une corrélation forte entre ces deux variables est souvent constatée. Cette corrélation peut être visualisée par un nuage de points comme dans la Figure 5.2. Il est ainsi possible d'introduire une relation statistique entre ces deux variables. Les données utilisées pour l'établissement de fonctions de pédo-transfert permettent également de calculer ces fonctions de régression simple entre θ_{cr} et θ_{pf} . Le Tableau 5.7 montre les régressions pour les quatre classes de matériaux.

5.1.4.2 Interactions entre propriétés appartenant à deux horizons voisins

Les contraintes exposées dans le paragraphe précédent concernent des relations entre propriétés de sol appartenant au même horizon. Ces relations sont appelées des contraintes 'intra-horizon'. L'évaluation de la réserve utile est alors possible pour une UTS contenant un seul horizon. Si plusieurs horizons ont été définis, l'évaluation de la réserve utile totale d'un sol devient la simple addition des réserves utiles calculées pour chaque horizon (équation 5.4):

$$RU_{tot} = \sum_{i=1}^n RU_i \quad (5.4)$$

Une contrainte mettant en relation la profondeur du sol totale et l'épaisseur de chaque horizon ($Epais_i$) doit être ajoutée à celle-ci. Cette contrainte exprime que la profondeur du sol est, par définition, la somme des épaisseurs des horizons. L'expression mathématique de cette contrainte, qui a déjà été reconnue étant une relation floue dans le paragraphe 3.3.3, est la suivante:

$$Pr \text{ ofondeur_totale} = \sum_{i=1}^n Epais_i \quad (5.5)$$

Les interactions fonctionnelles existant entre les différents horizons d'un même profil de sol se traduisent par des relations de dépendances entre les propriétés de chaque horizon. Ces relations peuvent être prises en compte dans le calcul de la RU sous forme de contraintes 'inter-horizons'. Ces contraintes sont basées sur la proximité des valeurs des propriétés du sol des horizons voisins, en admettant ainsi peu de changements abrupts d'un horizon à l'autre.

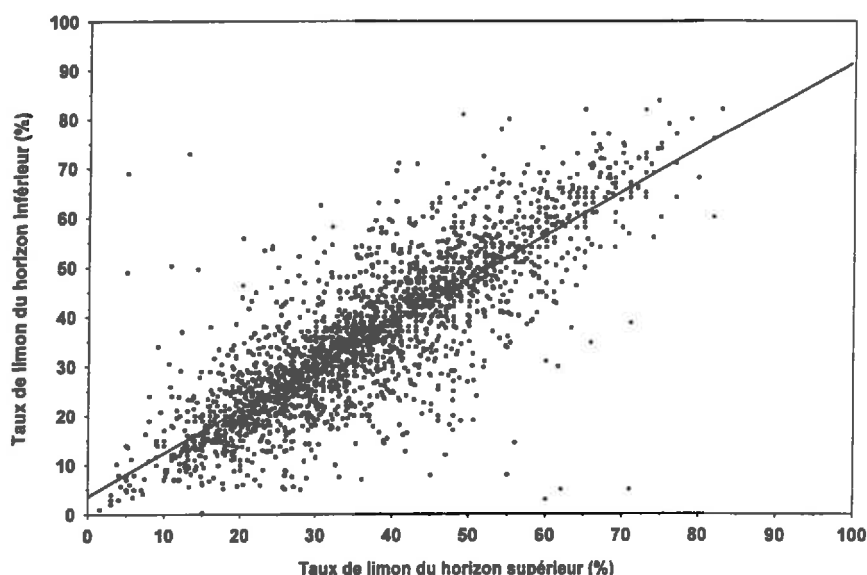


Figure 5.3: Nuage de points exprimant la relation entre le taux de limon de l'horizon inférieur et le taux de limon de l'horizon supérieur.

Variable	N	Régression	R ²
Argile	2721	$ARG_{inf} = 4.35 + 0.85 * ARG_{sup}$	0.55
Limon	2721	$LIM_{inf} = 3.65 + 0.88 * LIM_{sup}$	0.69
Sable	2721	$SAB_{inf} = 4.88 + 0.88 * SAB_{sup}$	0.67

Tableau 5.8: Les contraintes inter-horizons pour les composants granulométriques

Cette hypothèse est acceptable pour la majorité des sols présents dans la zone d'étude VHOL, mais elle est plus difficilement applicable aux sols à forte différenciation des horizons (par ex. les sols des niveaux Villafranchiens). Pour exprimer les relations trouvées, des relations mathématiques ont été construites entre une propriété du sol d'un horizon et la même propriété du sol de l'horizon sus-jacent.

Une première relation concerne la densité apparente. Dans un profil de sol, on constate généralement une augmentation de la compacité avec la profondeur (Baize & Jabiol, 1995). Ceci peut-être traduit par une inégalité entre la masse volumique (densité apparente) de l'horizon inférieur (DA_{inf}) et celui sus-jacent (DA_{sup}).

$$DA_{inf} \geq DA_{sup} \quad (5.6)$$

Un autre groupe de contraintes est représenté par les faibles variations des fractions granulométriques d'un horizon à l'autre. Les relations mathématiques pour exprimer ces faibles variations ont été dérivées de la base de données des profils de sols STIPA (voir § 4.2.3) disponible au laboratoire Science du Sol de l'INRA Montpellier.

Distributions exprimant l'erreur d'estimation des régressions

Exemple de la fonction de pédotransfert θ_{F} Alluvions récentes

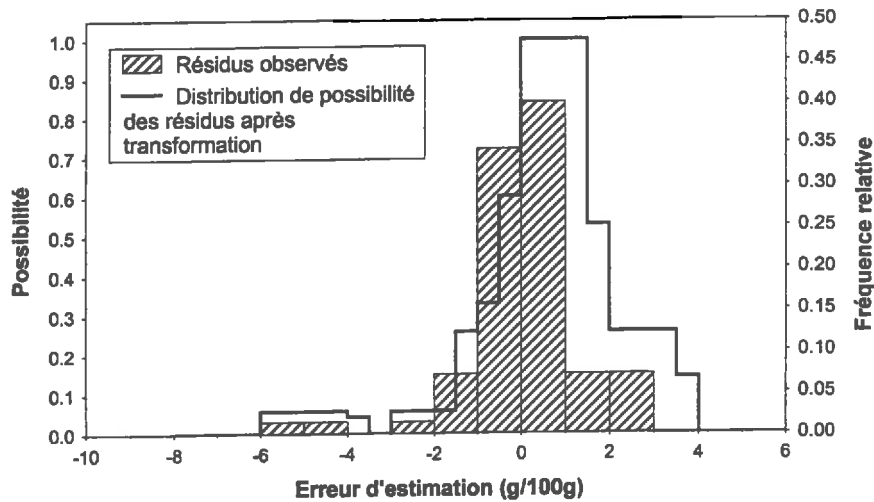


Figure 5.4: Détermination de la distribution de possibilité des résidus de la régression multiple.

Cet ensemble de données a permis d'établir une série de couples horizon inférieur / horizon supérieur pour le taux d'argile, de limon et de sable. A titre d'exemple, la Figure 5.3 explicite le nuage de points reliant le taux de limon de l'horizon supérieur et le taux de limon de l'horizon inférieur. Une corrélation forte a été constatée pour les trois variables et la relation peut alors être exprimée sous la forme d'une régression simple linéaire. Le Tableau 5.8 explicite ces régressions pour les trois variables.

5.1.5 Représentation de l'incertitude liée aux régressions

Une partie des relations utilisées dans l'estimation des propriétés hydriques est basée sur des régressions linéaires (multiples ou simples). Telles qu'elles sont présentées dans les paragraphes précédents, elle ne contiennent pas d'incertitude. Cependant, les équations linéaires dérivées d'une analyse de régression contiennent une incertitude sur l'estimation de la variable dépendante. Cette incertitude peut être exprimée, avec un terme d'erreur d'estimation E , pour une fonction de pédotransfert, de la manière suivante:

$$\theta_x = \beta_0 + \beta_1 * da + \beta_2 * arg + \beta_3 * lim + \beta_4 * sab + \beta_5 * mo + E \quad (5.7)$$

L'erreur d'estimation d'une régression est déterminée à partir de la distribution des résidus de régression (valeur observée - valeur estimée par la régression). Dans le cas où il n'existe pas de biais, cette distribution montre une forme suivant la loi normale avec une

moyenne de 0. Dans un cadre statistique, l'erreur d'estimation peut ainsi être exprimée sous la forme d'un intervalle de valeurs égal, par exemple, à deux fois l'écart type.

Dans le cadre de la présente étude, il est nécessaire d'exprimer l'information concernant la distribution des résidus de régression sous la forme d'une distribution de possibilité. Une distribution de possibilité donne ensuite par α -coupe l'intervalle de valeurs exprimant l'erreur d'estimation. L'introduction d'une distribution de possibilité transforme donc la relation booléenne en relation floue.

La détermination de la distribution de possibilité exprimant l'erreur d'estimation est effectuée grâce à la procédure de transformation probabilité \rightarrow possibilité décrite dans l'Annexe IV. Les résidus de régressions sont ainsi des données que l'on peut exprimer sous la forme d'un histogramme. La Figure 5.4 montre un histogramme des résidus de la fonction de pédotransfert déterminant θ_{pf} pour les alluvions récentes. La distribution de possibilité obtenue après transformation donne ensuite pour chaque α -coupe, l'intervalle d'erreur à appliquer.

5.2 Représentation de la connaissance dans la cartographie des UTS

La cartographie des UTS au sein des UCS constitue la deuxième étape de la démarche générale. Dans le cadre de la théorie des possibilités, cette étape revient à l'identification, en un point de l'espace, des UTS possibles sous l'hypothèse que le point appartient à une UCS donnée. La mathématique de la théorie de possibilité (calcul de compatibilité et opérations ensemblistes, voir § 3.3.4) va nous permettre d'effectuer cette identification.

Il est au préalable nécessaire de représenter en terme de possibilité la totalité de la connaissance intervenant dans l'analyse proposée en terme de possibilité. L'information nécessaire est composée de deux types: (i) l'information descriptive concernant les variables d'environnement (géologie, altitude, pente et occupation du sol) pour chaque UTS et issue de la base de données des sols régionale et (ii) les couches d'information géographique correspondantes.

Ce paragraphe expose successivement comment sont représentées les descriptions de l'environnement des UTS et les données correspondant aux couches d'information géographiques nécessaires à la cartographie des UTS. Pour ces deux types d'information, on cherche également à quantifier et à représenter l'incertitude à l'aide des informations supplémentaires.

5.2.1 Interprétation et représentation de l'information issue de la base de données sols

Trois variables géographiques supplémentaires ont pu être distinguées pour la localisation des UTS au sein des UCS: géologie, occupation du sol et pente (voir § 2.3.2.2). La variable altitude n'a pas été retenue car dans la base de données sol utilisée, cette variable intervient uniquement dans la description globale de l'UCS. Elle n'est donc pas discriminante vis à vis de la localisation des UTS. La base de données sol régionale contient, principalement dans la description textuelle des UTS, une information plus ou moins complète pour ces variables. Ce paragraphe présente l'interprétation et la représentation de l'information disponible pour chacune des variables mentionnées.

En ce qui concerne la variable pente, une information est disponible sous la forme d'une expression qualitative liée à une codification publiée dans les guides de description de sols. La classification des pentes proposée par le système français GEPPA (Baize & Jabiol, 1995) est détaillée en 10 classes. Cette classification a été simplifiée en 5 classes dont les bornes inférieures et supérieures (les noyaux des distributions de possibilités) sont indiquées dans le Tableau 5.9. Ce tableau montre également les marges floues qui ont été choisies pour définir le support des distributions de possibilités (cf. les minima et maxima sur $\alpha = 0.0$). Ces marges ont été choisies pour exprimer l'erreur maximale tolérée telle que définie en paragraphe 5.1.1.

Quant aux variables "géologie et occupation de sol", il est nécessaire de traduire l'information de la description textuelle en unité de légende de la carte correspondante. Cette traduction permet ensuite de déterminer les plages cartographiques correspondantes et un codage dans un SIG. On affecte ensuite une valeur de possibilité de 1 à la surface couverte par ces plages cartographiques. La façon dont on introduit l'incertitude sur la localisation des limites sera exposée dans le paragraphe suivant.

L'interprétation et le codage de la connaissance en ce qui concerne les variables géographiques permettant de traiter cette information est illustrée par le Tableau 5.10. Cette interprétation est alors effectuée pour chaque UTS présente dans la zone d'étude.

Classe	Nom	Minimum	Minimum	Maximum	Maximum
		$\alpha=0.0$	$\alpha=1.0$	$\alpha=1.0$	$\alpha=0.0$
1	Nulle à très faible	0	0	2	4
2	Faible	0	2	5	7.5
3	Modérée	2.5	5	15	17.5
4	Forte	10	15	30	35
5	Très forte	20	30	∞	∞

Tableau 5.9: Distributions de possibilité pour les différentes classes de pente.

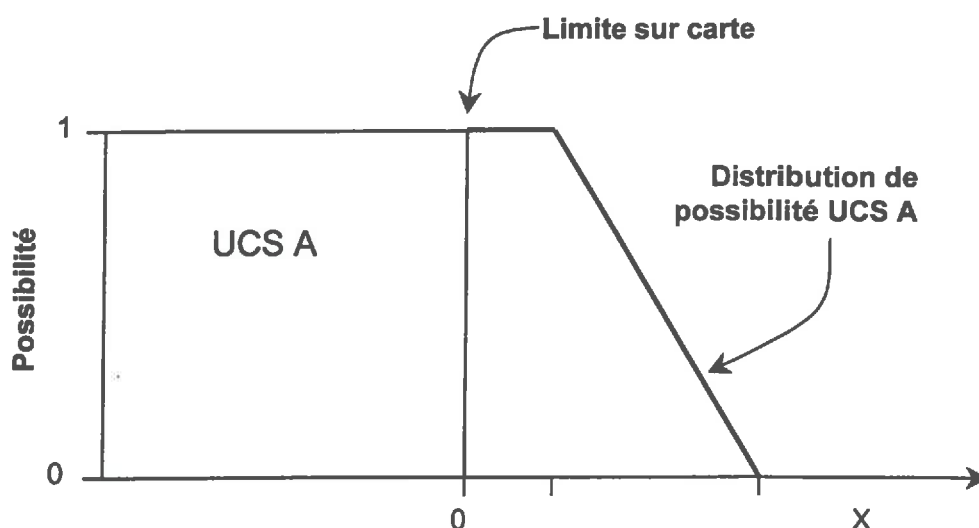


Figure 5.5: Représentation des limites floues dans le cadre de la théorie des possibilités.

Description textuelle pour UTS 102	Variable géographique	Codage dans SIG
Appartient à UCS 17301	→ Unité(s) de pédopaysage	Code 17301
Terrasses alluviales récentes	→ Unité(s) géologique	Code 2
Vignes et vergers	→ Occupation du sol	Code 3 ou code 4
Faible pente	→ Pente	[0,2,5,7.5]

Tableau 5.10: Illustration de la traduction de la description textuelle d'une UTS pour les quatre variables géographiques.

5.2.2 Détermination des limites floues des cartes choroplèthes utilisés

Le paragraphe 2.3.2 a présenté les sources d'incertitude dans l'utilisation des données géographiques présentes dans des cartes numérisées. L'imprécision sur le positionnement des limites est une source d'incertitude considérable pour les cartes à petite échelle. Pour la présente étude, on cherche à représenter l'imprécision des limites issues de la carte des pédopaysages et de la carte géologique. Le positionnement des limites des unités d'occupation du sol, issue de la BDTopo, a été jugé comme suffisamment précis pour l'échelle considérée.

Dans notre étude nous proposons de représenter les limites imprécises sous la forme de distributions spatiale de possibilité. Les règles de construction de ces distributions de possibilité sont les suivantes:

- Il est complètement possible (rien n'empêche) que l'UCS A soit présente sur le site x:
 - ✓ si x est localisé à l'intérieur de la plage cartographique définissant l'UCS A
 - ✓ ou si x, tout en étant localisé dans un autre UCS que X, se trouve à une distance d'une limite de l'UCS A tellement proche que rien ne peut exclure que x appartienne quand même à A, compte tenu de l'imprécision des limites de l'UCS A.

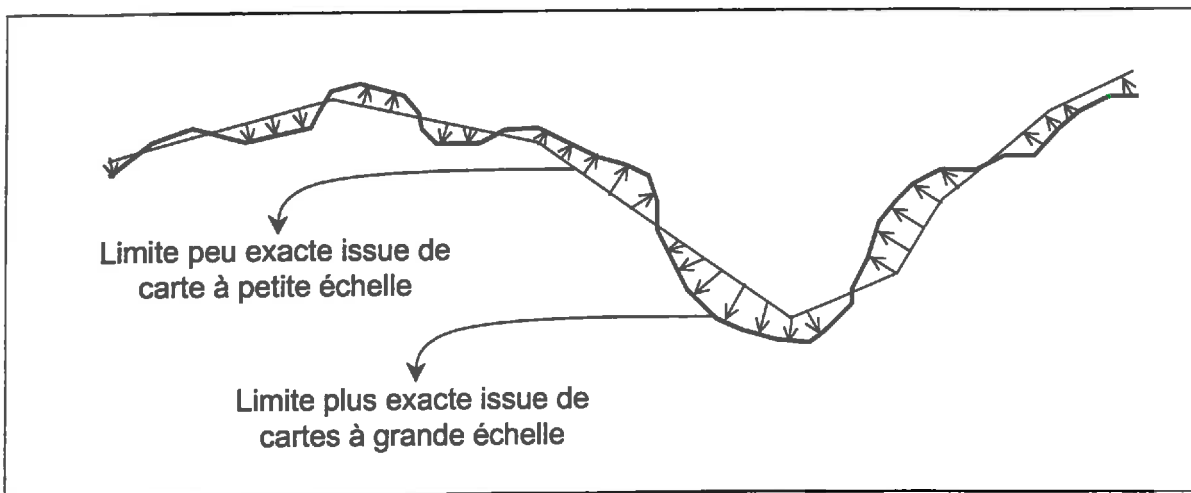


Figure 5.6: Deux limites géographiques exprimant la même limite entre deux unités cartographiques.

- Il est complètement impossible que l'UCS A soit présente sur le site x si celui-ci est localisée à une distance d'une limite de A qui est supérieure à l'erreur maximale tolérée sur les limites de l'UCS A.
- Dans le cas où l'appartenance de x à l'UCS A n'est ni complètement possible, ni complètement impossible, il est d'autant plus surprenant que x appartienne à A qu'il est éloigné d'une limite de A

La Figure 5.5 représente la traduction des règles précédentes sous forme d'une distribution de possibilité autour de la limite d'une UCS A. La première règle permet de définir le noyau de cette distribution (i.e. $\Pi(A) = 1$), la seconde le support (i.e. $\Pi(A) > 0$) et la troisième justifie la décroissance des possibilités entre les bornes du noyau et du support.

Cette représentation de la limite des unités cartographiques apparaît différentes des représentations proposées antérieurement qui utilisent aussi une approche floue (Burrough, 1989; McBratney & De Grujter, 1992; Lagacherie *et al.*, 1996; Brown, 1998). La principale différence réside dans le choix de normalisation retenu: les approches citées impose que les degrés d'appartenance des UC somment à 1. Dans le cadre théorique choisi, cette contrainte étant en contradiction avec la nécessité de condition de normalisation des distributions de possibilité (cf. Equation 3.12) qui, dans notre cas se traduit par la nécessité qu'au moins une UCS soit complètement possible au lieu donnée. Par contre, conformément au cadre théorique possibiliste, il accepte dans la représentation proposée que plusieurs UCS soient complètement possibles ($\Pi(A) = \Pi(A^c) = 1$ (cf. §3.3.4).

La représentation théorique de la limite d'une UCS étant choisie, il faut maintenant préciser la forme de la distribution de possibilité associée, en particulier pour la partie située à l'extérieur de l'UCS étudiée. Ceci est réalisé à partir d'une la comparaison entre la limite pour laquelle on cherche à déterminer l'imprécision et une limite géographique dont on connaît

précisément la localisation, mais elle représente exactement le même objet géographique (Dunn *et al.*, 1990; Bolstad *et al.*, 1990). La Figure 5.6 illustre graphiquement le principe de cette comparaison. La distance entre les deux limites est alors considérée comme mesure d'incertitude sur le positionnement de la limite peu exacte.

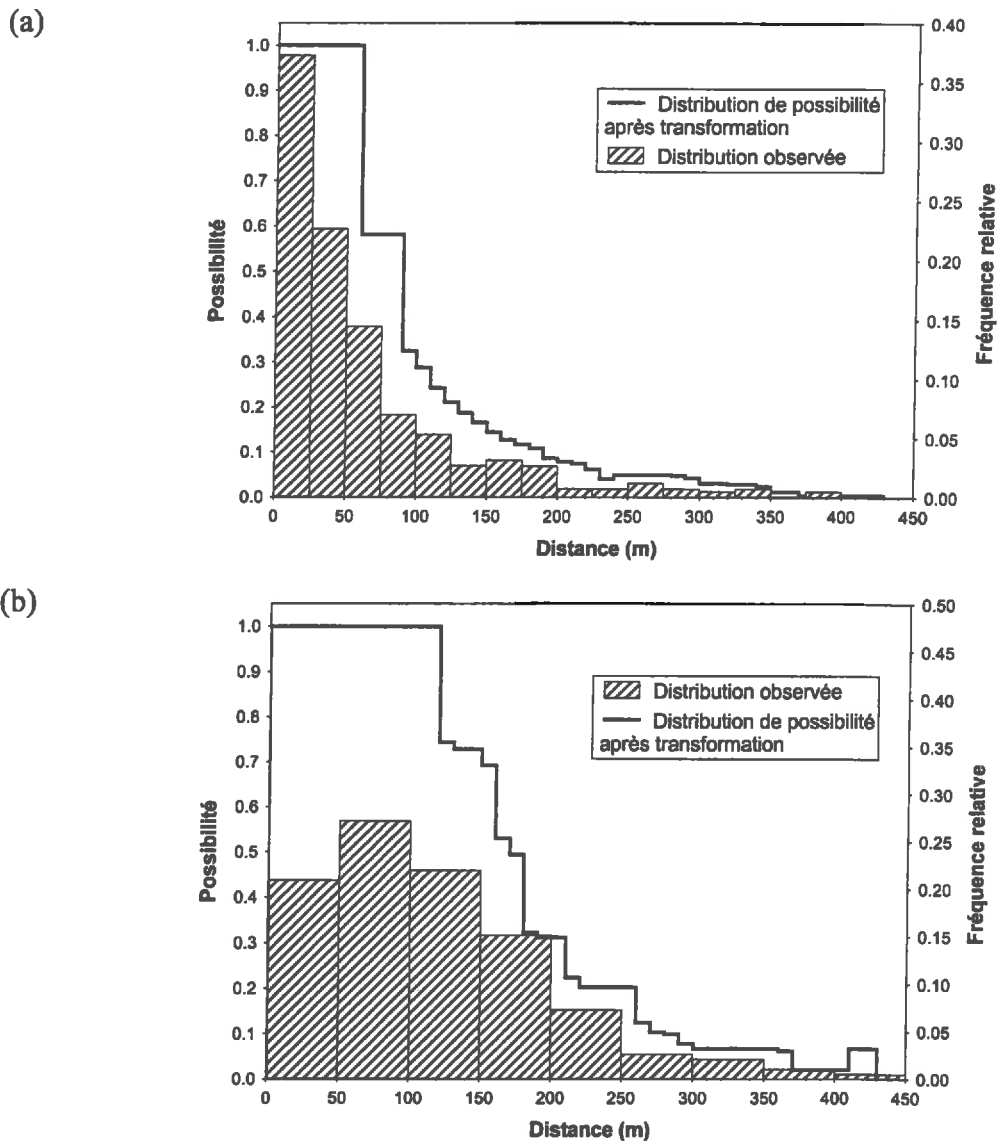


Figure 5.7: Détermination de la distribution de possibilité d'apparition pour les unités cartographiques de la carte géologique (a) et de la carte des pédopaysages (b). Les distributions de possibilité définissent la possibilité d'apparition d'une unité cartographique à l'extérieur de ses limites.

En se basant sur les cartes pédologiques à grande échelle disponibles dans la zone d'étude, une série de situations pour laquelle cette analyse pouvait être effectuée a pu être déterminée. Certaines limites sur ces cartes correspondent aux limites des unités géologiques représentées sur la carte géologique ou aux limites des unités de pédopaysages représentées

sur la carte de pédopaysages. Le calcul de la distance entre deux limites s'effectue pour une série de points définis sur la limite. La Figure 5.7 montre les résultats de ces calculs de distance. Cette figure montre la distribution de probabilité de la distance sous la forme d'un histogramme pour les limites issues de la carte géologique (a) et d'un autre histogramme pour celles de la carte des pédopaysages (b). Ces distributions de probabilité respectives ont été transformées en distributions de possibilité selon la procédure proposée dans l'Annexe IV.

Les distributions de possibilité ainsi construites, définissent les possibilités d'apparition d'une UCS à l'extérieur des ses limites, conformément au cadre théorique visualisée dans la Figure 5.5. L'utilisation de comparaison avec des cartes plus précises a permis, dans ces cas, de préciser la distribution trapézoïdale initialement définie.

5.2.3 Quantification de l'incertitude sur la pente

Il est important de prendre en compte l'incertitude liée au MNA. La quantification de l'incertitude sur les valeurs du MNA est basée sur l'erreur d'estimation, ce qui est comparable à l'approche utilisée pour la détermination de l'incertitude liée aux régressions.

Deux modèles numériques d'altitude précis ont été utilisés pour la détermination l'incertitude sur les valeurs d'altitude et de pente. Il s'agit du MNA précis du secteur ACM et d'un deuxième MNA disponible couvrant une partie de la commune de Roujan (Moussa *et al.*, 1998). Ces deux modèles numériques d'altitude permettent de calculer l'erreur d'estimation de MNA utilisé pour la zone VHOL. La figure 5.8 montre la distribution de fréquence relative de l'erreur d'estimation sur la pente. Les histogrammes ont ensuite été transformés en distribution de possibilité.

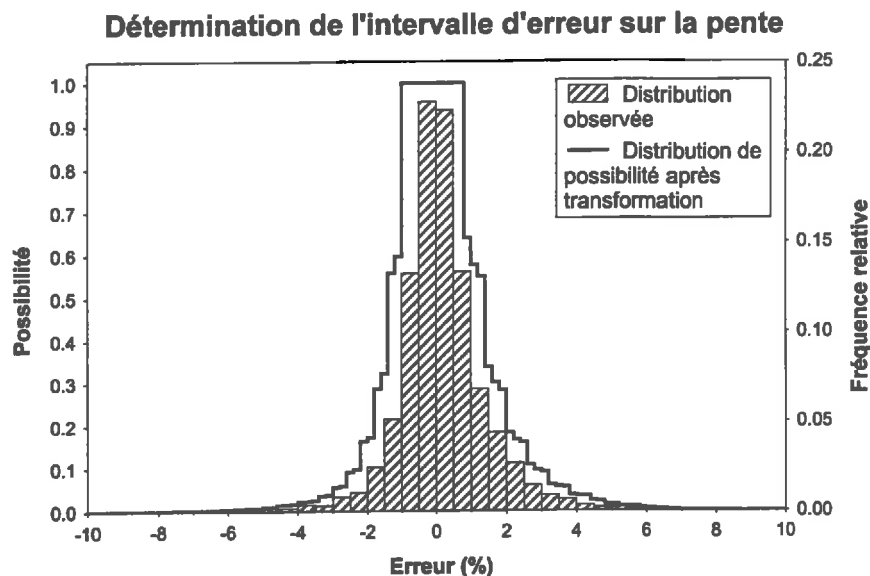


Figure 5.8: Détermination des distributions de possibilité exprimant l'incertitude sur la pente.

Nous pouvons illustrer l'intérêt de tenir compte de cette incertitude par un exemple. Considérons qu'on cherche à déterminer la cohérence (i.e. exprimé par la mesure de possibilité) d'une valeur donnée par le MNA pour la pente avec la classe "pente faible" (issue d'une description textuelle d'une UTS). La distribution de possibilité trapézoïdale de "pente faible" est ainsi définie par [0.0, 2.0, 5.0, 7.5]. Si le MNA donne une valeur de 6% pour la pente sur un point d'intérêt, la compatibilité des deux informations exprimée en mesure de possibilité est égale à 0.6. Cependant, la valeur de la pente de 6% doit être interprétée comme [5.0, 7.5] pour l' α -coupe 1.0, tenant compte de l'intervalle d'erreur correspondant à l' α -coupe 1.0 (i.e. [-1.0, 1.5]). Comme l'intervalle [5.0, 7.5] a une intersection non vide avec l'intervalle [2.0, 5.0] (i.e. l' α -coupe 1 de l'intervalle flou exprimant "faible pente") la compatibilité entre la donnée et la référence est 1.

5.3 Conclusion

Dans ce chapitre, des données nécessaires au calcul de la réserve utile des UTS et à la cartographie des UTS au sein des UCS ont été représentées dans le cadre possibiliste choisi au chapitre précédent. Deux grand voies ont été choisies en fonction de la nature des données permettant de caractériser l'incertitude.

- Lorsque des données expérimentales pouvaient être raisonnablement utilisées, nous avons, dans un premier temps traduit l'incertitude par des histogrammes (résidu fonction de pédotransfert, imprécision sur les limites de cartes, erreur du MNA) Puis une transformation de ce histogramme en distribution de possibilité a été appliquée selon une méthode que nous avons mise au point (cf. Annexe IV).
- Lorsque ces données expérimentales ne pouvaient être rassemblées (cas des descriptions textuelles et des profils de référence), une représentation en distribution de possibilité trapézoïdale a été choisie comme la plus simple à paramétrer en fonction des données disponibles. Après avoir défini des règles générales permettant d'interpréter la signification du noyau et du support de ces distribution, nous avons utilisé l'information disponible (guide de description pédologique, connaissance de terrain) pour paramétrer au mieux ces distributions. Cependant, la détermination du support reste généralement sujette aux choix arbitraires à cause d'une manque de données et de méthodes efficaces pour sa détermination. Une recherche supplémentaire est nécessaire pour mieux (i.e. moins arbitrairement) déterminer le support des distributions de possibilité issues de la connaissance pédologique.

A l'issue de cette étape, nous disposons d'une formalisation de la connaissance nécessaire sous forme d'un cadre théorique unifié. Ceci nous permet maintenant d'envisager les 2 problèmes suivant tels qu'ils ont été exposés à l'issue du chapitre 2:

- a) **Comment estimer, par des opérations arithmétiques, la réserve utile pour chaque UTS en propageant l'incertitude sur chaque variable intervenant dans l'estimation et en tenant compte des interactions entre ces variables?**

- b) **Comment localiser, par des combinaisons logiques, les UTS au sein des UPP en prenant en compte l'imprécision de la description des UTS et l'incertitude sur l'information géographique supplémentaire utilisée?**

6 Estimation des propriétés hydriques par UTS

Dans ce chapitre nous étudions plus particulièrement le problème de calcul des propriétés hydriques des unités typologique des sol à partir de leurs descriptions imprécises et d'informations supplémentaires qui ont été détaillées au chapitre précédent. Nous avons reconnu qu'il s'agissait d'un problème de propagation d'incertitude dans une opération arithmétique utilisant des variables interactives. Deux questions principales se posent:

a) Quelle approche faut-il utiliser pour traiter le problème posé?

Au vu des développements théoriques évoqués dans le troisième chapitre, une méthode a été identifiée à priori. Cette méthode utilise l'information sol formulée dans un cadre possibiliste, la propagation de l'incertitude étant réalisée grâce aux techniques de CSP. Il convient de la comparer à d'autres approches plus conventionnelles. Deux méthodes seront ainsi considérées:

- calcul d'intervalles flous sans prise en compte des dépendances induites sur certaines variables, du fait de leur appartenance à plusieurs relations.
- calcul en considérant que les variables sont des variables aléatoires, ce calcul étant réalisé par simulation de Monte Carlo (cf. §3.2.2) tout en tenant compte également des interactions entre variables.

b) Comment varient les estimations des propriétés hydriques en fonction de l'information sol utilisée?

Nous avons vu dans le deuxième chapitre que les UTS pouvaient être décrites avec des précisions différentes: une description "textuelle" et une description par profil de référence. Nous comparons les estimations obtenues avec ces deux types de forme et ce, pour deux types d'UTS différentes: des UTS issues d'une cartographie à petite échelle et des UTS issues d'une carte à grande échelle.

Le premier paragraphe sera consacré à la mise en œuvre des modèles permettant d'effectuer les différents calculs nécessaires. Ensuite, nous présentons en détail les protocoles d'expérimentation puis les résultats obtenus pour les différentes expérimentations. Enfin, les conclusions sont dérivées de l'ensemble des résultats obtenus.

6.1 Mise en œuvre des modèles

Dans ce paragraphe, nous présentons successivement la mise en œuvre de l'estimation de propriétés hydriques par la résolution des problèmes de satisfaction de contraintes (CSP), l'outil utilisé pour les calculs d'intervalles flous et la mise en place d'un modèle permettant d'effectuer la simulation de Monte Carlo.

6.1.1 Résolution de CSP avec CON'FLEX

Le résolveur de contraintes CON'FLEX¹ (Rellier & Vardon, 1998) a été utilisé pour résoudre le problème de la propagation de l'incertitude dans une série d'opérations arithmétiques avec des variables interactives (cf. § 3.3.5.2). Il propose un langage d'expression du problème en termes de variables, de contraintes et de requêtes, et fournit un ensemble d'algorithmes de résolution (le filtrage et la recherche arborescente des solutions). Les Annexes V et VI montrent deux exemples de ce type d'expression d'un problème dans le langage proposé par CON'FLEX.

Nous présentons successivement les variables et contraintes nécessaires à la résolution de notre problème, les requêtes exprimant l'information cherchée et la façon technique dont les solutions ont été recherchées. Il est également nécessaire de consacrer un paragraphe aux problèmes de temps de calculs rencontrés et à la façon dont ces problèmes ont été traités.

6.1.1.1 Données d'entrées du problème

L'information élémentaire d'un problème de satisfaction de contraintes est constituée d'un ensemble de variables et d'un ensemble de contraintes exprimant les relations entre ces variables (cf. 3.3.5.2). Il est donc nécessaire de spécifier précisément les variables et contraintes utilisées pour la détermination de la réserve utile et de la teneur en eau au point de flétrissement.

a Variables

Le chapitre 5 a présenté l'interprétation et la représentation des données nécessaires à l'estimation de la réserve utile par UTS. Cette analyse nous a permis de définir un jeu de

¹ CON'FLEX: Unité de Biométrie et Intelligence Artificielle, INRA, Toulouse, 1996. L'exécutable (sous UNIX Solaris 5 et Linux) et le manuel utilisateur sont disponible sur internet à <http://www-bia.inra.fr/>

Nom Variable	Explication des variables
Arg	Taux d'argile (%)
Lim	Taux de limon (%)
Sab	Taux de sable (%)
Pier	Pierrosité, taux d'éléments grossiers (%)
MO	Matière organique du sol (%)
Da	Densité apparente
Prof	Profondeur du sol (dm)
θ_{cr}	Teneur en eau au capacité de rétention = θ_{cr} (g/100g)
θ_{pf}	Teneur en eau au point de flétrissement = θ_{pf} (g/100g)
RU	Réserve utile du sol (mm)

Tableau 6.1: Le jeu des variables d'entrées dans la définition de CSP afin d'estimer la RU et θ_{pf} .

variables d'entrée nécessaires à cette estimation. Le jeu de variables à utiliser est différent selon le type d'information utilisée (description textuelle ou description par profil de référence). La description textuelle ne contient généralement pas de subdivision en horizon; ce qui permet de définir un jeu de variables pour un seul horizon, c'est-à-dire, pour la couche qui correspond à la profondeur totale du sol. Les variables introduites dans les CSP sont présentées dans le Tableau 6.1.

Lorsque l'on utilise la description par profil de référence, on définit un jeu de variables, tel qu'il est présenté dans le Tableau 6.1, par horizon décrit. On ajoute à ces variables la variable 'épais' exprimant l'épaisseur de chaque horizon (en dm). La totalité des variables pour une description issue d'un profil de référence contenant deux horizons est donc approximativement deux fois le nombre de variables utilisées dans une description textuelle exprimant le sol en un seul horizon (voir les exemples des Annexes V et VI). Le nombre de variables dans un CSP représentant un sol en trois horizons est ainsi encore plus important.

Il est ensuite nécessaire de représenter ou déclarer la connaissance sur chaque variable sous la forme d'une distribution de possibilité. Deux cas sont à distinguer: les variables issues de la description des UTS (Arg, Lim, Sab, Pier, MO et Prof) pour lesquelles pré-existe une représentation sous forme de distribution de possibilité (cf. chapitre 5) et les variables intermédiaires ou variables résultats (θ_{pf} , θ_{cr} et RU), dont il faut fixer les domaines de variation à priori. Les premières sont représentées dans le CSP par leur distribution de possibilité (i.e. distribution trapézoïdale du type [a, b, c, d], où a et d définissent le support et b et c le noyau). Pour ce qui concerne les secondes, nous avons fixé les domaines de variation par des intervalles non flous. Pour la réserve utile, nous avons choisi un intervalle large de 0 à 1000mm. Quant aux variables θ_{cr} et θ_{pf} , les bornes de l'intervalle correspondent aux valeurs maximales et minimales des données utilisées pour l'établissement des fonctions de pédotransfert respectives (cf. § 2.3.4.2).

b Contraintes

Nous avons présenté, dans le chapitre 5, l'ensemble des contraintes à introduire dans l'estimation de la réserve utile et de la teneur en eau au point de flétrissement. L'introduction de cette connaissance est facile à l'aide du langage proposé par CON'FLEX. Les contraintes peuvent être exprimées soit en *extension*, en décrivant les combinaisons possibles de valeurs pour les variables entrant dans la contrainte, soit en *intension*, par une équation liant les variables. Les contraintes utilisées pour l'estimation de la réserve utile et de θ_{pf} sont toutes des contraintes exprimées en *intension*. La déclaration des contraintes est illustrée dans l'Annexe V pour le cas d'une description textuelle (sol composé d'un seul horizon) et dans l'Annexe VI pour le cas d'une description par profil de référence (sol composé de deux horizons).

Le nombre de contraintes dans le CSP pour un sol composé d'un horizon est beaucoup moins important que dans le cas d'un sol composé de différents horizons. Pour un sol composé d'un horizon, il suffit de déclarer une fois l'ensemble des contraintes 'intra-horizon' (cf. § 5.1.3.1). Pour un sol composé de deux horizons, il est nécessaire de déclarer ces contraintes deux fois. De plus sont introduites les contraintes 'inter-horizon' (cf. § 5.1.3.2).

6.1.1.2 Formulation des requêtes

Pour la présente étude nous cherchons à estimer par UTS deux propriétés hydriques: la réserve utile et la teneur en eau au point de flétrissement. A présent, il est nécessaire de formuler plus précisément l'information cherchée sous la forme d'une requête.

Le premier paramètre à fixer est la profondeur sur laquelle sera estimée la réserve utile. En effet, nous considérons que les estimations de profondeur de sol effectuées au delà de 120 cm comportent un risque d'erreur trop élevé, compte tenu des profondeurs maximales auxquelles sont effectuées les observations servant à définir les UTS (profils et sondages). Les erreurs auraient pu fausser les comparaisons d'estimation de RU que nous envisageons. C'est pourquoi nous proposons d'estimer la réserve utile sur une profondeur de sol limitée à 120 cm. Cette requête peut être définie dans le CSP en modifiant la distribution de possibilité pour la variable profondeur. Par exemple, l'expression "sol profond" fixée initialement à la distribution de possibilité [90, 100, 150, 160] est ramenée à [90, 100, 120, 120].

En ce qui concerne la teneur en eau au point de flétrissement, nous cherchons à l'estimer à la profondeur de 40 cm pour pouvoir comparer les résultats avec les données de validation mesurées également à cette profondeur (cf §4.2.4.1). Ce type de requête introduit une difficulté supplémentaire. Considérons, par exemple, le CSP donné dans l'Annexe VI. La profondeur de 40 cm peut se situer aussi bien dans le premier horizon que dans le deuxième horizon. Lorsque l'épaisseur du premier horizon est supérieure à 40 cm, la variable cherchée (i.e. θ_{pf} à 40 cm) est la θ_{pf} du premier horizon. En revanche, dans le cas où l'épaisseur du

premier horizon est inférieure à 40 cm, la θ_{pf} à 40 cm est la θ_{pf} du deuxième horizon. CON'FLEX permet de résoudre ce problème en introduisant une contrainte en *intension* définie par morceaux (voir Rellier & Vardon, 1998). Une contrainte par morceaux permet d'introduire une discontinuité dans la représentation d'une relation (voir l'exemple donné dans l'Annexe VI).

Comme nous l'avons indiqué dans le paragraphe 3.5.1.2, la résolution de CSP permet d'effectuer la combinaison d'un ensemble de contraintes pour déterminer efficacement la distribution de possibilité jointe. Le nombre de solutions trouvées pour cette distribution de possibilité jointe est généralement très important pour les types de CSP présentés en annexes V et VI. Il en résulte parfois des temps de calcul très élevés et la génération de résultats inutilisables. En fait, pour notre étude, nous nous intéressons généralement seulement aux solutions trouvées pour une variable spécifique (RU sur une profondeur limitée à 120 cm et θ_{pf} à 40 cm). Il est alors nécessaire d'effectuer la projection de la distribution de possibilité jointe sur le domaine de la variable cherchée. Ceci revient donc à déterminer le noyau et le support de la distribution de possibilité de ces variables.

Le noyau correspond à l'intervalle des solutions trouvées pour l' α -coupe = 1 et il est donc facilement déterminable. Le support, quant à lui, correspond à l'intervalle des solutions trouvées lorsque l' α -coupe tend vers 0 (l'intervalle des solutions trouvées pour l' α -coupe = 0.0 est égal à l'univers). Seule une approximation a donc pu être faite, en identifiant le support à l'intervalle des solutions trouvées pour l' α -coupe = 0.1.

Dans la mesure où seules les bornes inférieures et supérieures du noyau et du support nous intéressaient dans cette application, nous avons pu adapter la formulation de la requête dans CON'FLEX de façon à limiter un maximum le temps de calcul. Plutôt que de déterminer ces bornes à la suite d'une recherche de l'ensemble des solutions satisfaisant aux contraintes du problème posé qui aurait été trop coûteuse en temps de calcul, nous avons décomposé le problème en une série de sous problèmes pour lesquels il était possible d'arrêter les recherches dès la première solution trouvée. Les sous problèmes étaient définis en ajoutant au CSP initial un contrainte supplémentaire portant sur la variable recherchée (par ex. $RU \geq X$). Pour chacun des CSP ainsi défini, on cherche alors à savoir uniquement s'il y a une solution, ce qui est peu coûteux en temps de calcul. En faisant varier X selon un pas de discrétisation prédéfini (0.5 g/100g pour θ_{pf} et 5 mm pour la RU), on peut ainsi définir les bornes inférieures et supérieures du domaine de X pour lequel il y a une solution aux sous problèmes défini plus haut.

6.1.1.3 Paramétrage de la recherche des solutions avec CON'FLEX

Le lecteur trouvera la description complète du paramétrage dans le manuel de l'utilisateur de CON'FLEX (Rellier & Vardon, 1998). Seuls certains aspects importants du paramétrage sont exposés ici.

La recherche de solutions pour un CSP s'effectue pour une ou plusieurs α -coupes. CON'FLEX demande alors aux utilisateurs de définir les α -coupes, c'est-à-dire, le pas de discrétisation entre la possibilité 0 et 1. Il est également possible de définir le niveau minimal de α à partir duquel des solutions devront être cherchées, si on ne s'intéresse, par exemple, qu'aux solutions supérieures à un certain α -coupe.

Avant d'effectuer la recherche arborescente des solutions, CON'FLEX effectue un filtrage, éventuellement suivi par un ultra-filtrage (voir Rellier & Vardon. 1998). Pour la présente étude, il est préférable d'effectuer un ultra-filtrage. Il permet en effet un filtrage renforcé en découpant les domaines des variables réelles (i.e. l'unique type de variable utilisé dans la présente étude). En considérant les domaines des variables réelles comme un ensemble de sous intervalles, le processus d'ultra-filtrage permet d'éliminer plus précisément les parties des domaines de variables qui sont a priori exclues des solutions. La recherche des solutions sera ensuite plus rapide. Il est également possible de jouer sur plusieurs paramètres qui permettent d'optimiser la stratégie de la recherche arborescente. Enfin, CON'FLEX permet d'afficher les résultats selon plusieurs types de présentations.

La recherche des solutions par ordre arborescent nécessite un découpage ou discrétisation du domaine continu en un nombre limité de valeurs réelles. C'est cette raison qu'un pas de discrétisation est demandé pour chaque variable continue. La détermination des pas de discrétisation doit respecter un certain nombre de règles:

- un pas trop grand risque d'être trop imprécis.
- un pas trop petit risque d'engendrer un trop grand nombre de solutions et un temps de calcul trop élevé.
- enfin, il convient de respecter une cohérence entre les pas des différentes variables. Il est inutile de discrétiser une variable avec un pas très petit si celle-ci est liée à une autre variable discrétisée avec un pas beaucoup plus grand.

Les pas de discrétisation retenus pour chaque variable sont résumés dans le Tableau 6.2.

Variable	Pas de discrétisation	Variable	Pas de discrétisation
Arg	5 %	Epais	0.5 dm
Lim	5 %	Prof	0.5 dm
Sab	5 %	θ_{cr}	0.5 g/100g
Pier	5 %	θ_{pf}	0.5 g/100g
Mo	5 %	RU	5 mm
Da	0.1		

Tableau 6.2: Les pas de discrétisation utilisés pour les variables utilisées dans la présente étude.

6.1.1.4 Résolution des CSP et temps de calcul

Lorsque le sol est décrit en plusieurs horizons, la recherche des bornes supérieures et inférieures risque de provoquer des temps de calculs très élevés². Ceci est surtout le cas pour la recherche des bornes du support pour la réserve utile et en particulier pour la recherche de la borne supérieure. En effet, il n'existe qu'un petit nombre de solutions aux valeurs élevées de RU. Ainsi, en s'approchant de la borne supérieure, le filtrage et l'ultra-filtrage effectués par CON'FLEX n'éliminent qu'une partie de l'espace des solutions potentielles (Rellier & Vardon, 1998). Les algorithmes de recherche arborescente ne s'avèrent ensuite pas suffisamment efficaces pour trouver rapidement la première solution. La recherche arborescente de la première solution devient alors une recherche par instanciation systématique de chaque variable déclarée dans le CSP, en vérifiant chaque fois les contraintes imposées. Le nombre de variables déclarées dans le CSP influence ainsi le temps de calcul, ainsi que le pas de discrétisation défini pour chaque variable. Le temps de calcul pour la détermination de l'enveloppe des variables cherchées risque d'être plus élevé pour un CSP de deux horizons. Le CSP utilisé pour la description par profil de référence pour l'UTS 974 contient même trois horizons. L'estimation de la RU pour cette UTS a pris plus de temps de calcul que les CSP avec deux horizons.

Le temps de calcul est également affecté par les contraintes déclarées. Certaines relations entre variables sont, par exemple, plus contraignantes que d'autres. Si le CSP contient des contraintes fortes, le filtrage et la recherche arborescente s'avèrent très efficaces et la première solution est trouvée plus rapidement. L'influence de ce facteur a été observée pour les différences de temps de calcul entre l'estimation pour l' α -coupe 1 et l' α -coupe 0.1. Les relations entre variables sont en effet moins contraignantes pour l' α -coupe 0.1 du fait des différences entre les intervalles d'erreur des régressions (cf. § 5.1.4). Les temps de calcul pour l' α -coupe 0.1 sont donc plus importants que pour l' α -coupe 1.0.

La recherche approximative du noyau et du support utilise un pas de discrétisation minimal de 0.5 pour θ_{pf} et de 5 pour la RU (cf. § 6.1.1.2). Cependant, lorsque les temps de calculs étaient très élevés, ce pas minimal n'a pas pu être respecté. Ceci n'a concerné qu'un nombre restreint d'estimations de la borne supérieure de la RU. Pour ces cas particuliers, au lieu d'identifier cette borne par une valeur précise, l'utilisation de CSP conduit à définir un intervalle de valeurs.

Les problèmes de temps de calcul rencontrés dans la présente étude sont des problèmes fréquents lors de l'optimisation des paramètres dans la modélisation en général. Lors d'une

² La résolution des CSP avec CON'FLEX a eu lieu sous système Red Hat Linux sur un PC équipé avec un processeur de 200MHz. Le temps de calcul de la résolution peut varier de plusieurs minutes à quelques jours et même des semaines.

optimisation on cherche à déterminer la valeur minimum ou la valeur maximum d'un système mathématique. La recherche des bornes inférieure et supérieure pour chaque α -coupe pourra être considérée comme un problème d'optimisation. Dou *et al.* (1995) proposent une telle approche pour la modélisation hydrologique dans le cadre des sous-ensembles flous où l'optimisation s'effectue par des algorithmes de programmation non-linéaire.

Un outil d'optimisation sous contrainte a été développé par Van Hentenryck *et al.* (1997) (logiciel Numerica gratuit pour une période d'essai). Quelques essais ont été effectués avec cet outil pour comparer les temps de calculs avec ceux obtenus par la résolution de contraintes proposée (CSP) dans la présente étude. Ces essais n'ont pas démontré une réduction significative des temps de calculs.

6.1.2 Calcul d'intervalles flous

Nous avons vu que le calcul d'intervalles flous (cf § 3.3.5.1) permet de combiner des distributions de possibilité, ce qui permet donc de propager l'imprécision dans une opération arithmétique. A la différence de la méthode utilisant un résolveur de contraintes, les relations entre variables impliquées dans l'opération arithmétique ne sont pas prise en compte, ce qui devrait amener à surestimer les noyaux et support des distributions de possibilité sur les propriétés calculées. Pour vérifier ceci, nous proposons de déterminer la réserve utile et la teneur en eau au point de flétrissement en appliquant les opérations présentées dans le Tableau 3.4 pour les mêmes α -coupes proposées précédemment pour la résolution des CSP (α -coupes 0.1 et 1.0). Pour ces estimations, nous avons utilisé le minimum de relations nécessaires, c'est à dire, l'Equation 5.1 pour le calcul de la RU et l'Equation 5.7 exprimant les fonctions de pédotransfert pour respectivement θ_{pf} et θ_{cr} . On ne tient pas compte des dépendances entre certaines variables impliquées dans les équations (par ex. " $\theta_{pf} < \theta_{cr}$ " et " $\%arg + \%lim + \%sab = 100\%$ ").

Nous avons effectué ces calculs d'intervalles en appliquant le résolveur de contraintes CON'FLEX. En effet, en éliminant les contraintes citées ci-dessus, cet outil effectue un calcul d'intervalle classique pour les différentes α -coupes définies.

6.1.3 Simulation Monte Carlo

L'estimation de la réserve utile à l'aide des fonctions de pédotransfert pour chaque unité typologique de sols constitue une étape où la simulation Monte Carlo pourra servir comme méthode de propagation d'incertitude. L'application de la simulation de Monte Carlo à cette partie de l'étude introduit deux problèmes principaux.

Le premier problème concerne la détermination de la distribution de probabilité pour chaque variable d'entrée. L'information disponible se présente sous la forme d'intervalles de valeurs et les paramètres "moyenne arithmétique" et "écart type", nécessaires pour la définition d'une fonction de densité de probabilité normale, ne sont donc pas disponibles. Une solution est proposée par l'établissement de distributions de densité de probabilité trapézoïdale; ce que nous considérons comme les distributions les plus proches des distributions de possibilités définies dans le paragraphe 5.1.1.

Le deuxième problème concerne la représentation de la dépendance entre les variables. Dans la présente étude, on ne dispose pas de la matrice de covariance ni de la matrice de corrélations des variables nécessaires pour l'estimation des propriétés hydriques. Il existe néanmoins des relations entre variables (i.e. les contraintes) sous la forme de régressions et d'inégalités. Il est donc important d'introduire ces relations dans l'échantillonnage aléatoire des valeurs.

Un modèle permettant d'effectuer une simulation Monte Carlo a été développé en collaboration avec Robert Faivre³ en utilisant le logiciel statistique Splus™ (Statistical Sciences, 1993). Ce modèle a été conçu pour pouvoir comparer l'approche CSP avec cette approche stochastique (simulation Monte Carlo). Ce modèle présente donc deux caractéristiques spécifiques. Premièrement, les variables d'entrées sont exprimées sous la forme de distributions de densité de probabilités trapézoïdales. Deuxièmement, le tirage aléatoire est conditionné par les relations entre variables. A ce titre, le modèle construit utilise les mêmes informations que celles introduites dans le CSP.

6.1.2.1 *Création de la fonction de répartition à partir d'une fonction de densité de probabilité pour le tirage aléatoire*

Comme on l'a vu dans le paragraphe 3.2.2, la simulation Monte Carlo nécessite le tirage aléatoire d'un jeu de valeurs pour les variables d'entrée. Le jeu de valeurs est en fait choisi à partir de la fonction de répartition de chaque variable. Le modèle transforme donc la fonction de densité trapézoïdale en fonction de répartition permettant un tirage aléatoire dans l'intervalle [0,1] et donc la création d'une distribution de probabilité. La Figure 6.1 montre la création de la fonction de répartition à partir d'une fonction de densité de probabilité sous forme trapézoïdale.

Les triangles (abb') et (c'cd) et le rectangle (b'c'cb) représentent respectivement les zones de surface X, Z et Y. Ces zones peuvent être calculées de la manière suivante:

³ chercheur au laboratoire de Biométrie et d'Intelligence Artificielle de l'INRA, Castanet-Tolosan (faivre@inra.toulouse.fr).

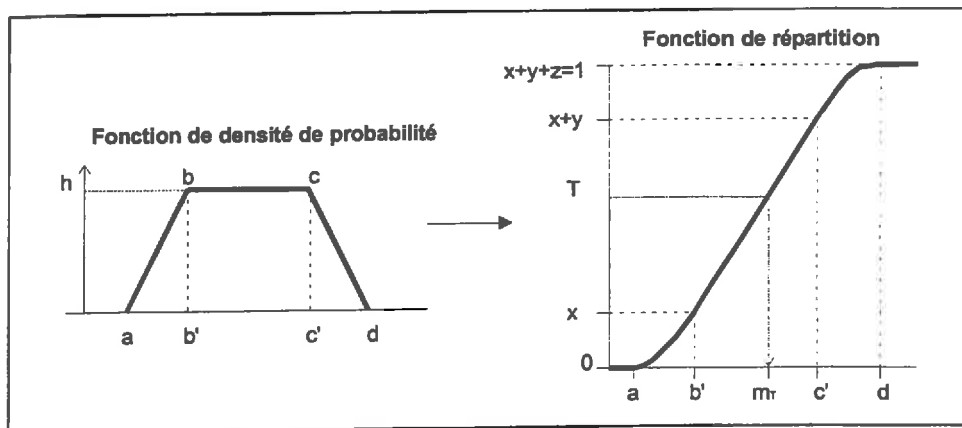


Figure 6.1: Création d'une fonction de répartition à partir d'une fonction de densité de probabilité trapézoïdale.

$$X = h \times (b' - a) / 2$$

$$Y = h \times (c' - b')$$

$$Z = h \times (d - c') / 2$$

Les valeurs obtenues pour X, Y et Z sont normalisées par $(X + Y + Z)$ pour donner x, y et z ($x = X/(X+Y+Z)$; $y = Y/(X+Y+Z)$; $z = Z/(X+Y+Z)$). On obtient ainsi les pourcentages relatifs de chaque zone. Une déformation quadratique est alors introduite au début et à la fin de cette fonction de répartition. Pour le nombre aléatoire T, on retrouve la valeur $m\pi$. La distribution de probabilité est ainsi obtenue par une succession de tirages aléatoires (10.000 pour la présente étude).

6.1.2.2 Prise en compte des contraintes

L'introduction des relations entre variables peut s'effectuer à deux niveaux: en entrée, au niveau du tirage aléatoire conditionné et en sortie, au niveau du tri des résultats de la simulation. Le tirage aléatoire conditionné s'applique, par exemple, à la contrainte $Arg + Lim + Sab = 100$. En tirant une valeur pour Arg et une valeur pour Lim, la valeur pour Sab est conditionnée ($Sab = 100 - (Arg + Lim)$). Pour cette contrainte, il est ensuite important de prendre en compte l'ordre du tirage aléatoire pour éviter que toujours la même variable soit conditionnée. Par exemple, il est possible de commencer avec la variable Sab, ensuite Lim, ce qui conditionne la valeur pour Arg. L'ordre du tirage aléatoire est pris en compte par un choix aléatoire de l'ordre. En sortie, les résultats peuvent être triés, de façon que la contrainte $\theta_{cr} > \theta_{pf}$ (voir Equation 5.3) soit toujours respectée.

Ainsi, nous avons pu élaborer un modèle de simulation de Monte Carlo pour le calcul de RU et θ_{pf} pour un sol contenant un seul horizon. Nous avons également tenté de construire un modèle pour un sol décrit en plusieurs horizons en tenant compte des contraintes "intra-horizon" pour chaque horizon et des contraintes "inter-horizon". Des grandes difficultés ont été rencontrées pour prendre en compte la totalité de contraintes à cause de la complexité du

système à représenter. C'est pour cette raison, que nous avons décidé d'appliquer la simulation de Monte Carlo qu'aux sols décrits en un seul horizon.

6.2 Protocoles d'évaluation

Les différentes méthodes et techniques étant présentées, il est maintenant nécessaire de mettre en place les différentes expérimentations afin de pouvoir répondre aux questions principales:

- Quel est la meilleure méthode pour propager l'imprécision au calcul de deux variables prises comme exemple (θ_{pf} à 40 cm et RU sur 120cm)?
- Quelle source d'information (textuelle ou profil de référence) faut-il utiliser pour caractériser ces deux propriétés?

Nous présentons l'ensemble des expérimentations effectuées et les méthodes d'évaluation utilisées permettant d'interpréter et d'analyser les résultats obtenus.

6.2.1 Expérimentations effectuées

Au total, huit unités typologiques de sols (UTS) ont été choisies pour l'ensemble des expérimentations. Pour ces 8 UTS, nous disposons d'une description textuelle et d'une description par profil de référence. Le Tableau 6.3 résume l'origine des descriptions de sol et le nombre d'horizons décrit. Les UTS choisies sont issues de deux cartes pédologiques différentes. Les UTS 102, 303, 345 et 974 sont issues de la carte au 250.000^{ème} du Languedoc-Roussillon. (Bornand *et al.*, 1992). Les descriptions textuelles et les descriptions par profil de référence ont été présentées dans les Tableaux 2.2 et 2.3. Les UTS 2, 8a, 14 et 16 sont issues de la carte pédologique au 10.000^{ème} de la zone ACM (Clochard *et al.*, 1991) (voir Annexe II). L'information textuelle pour les 4 dernières UTS a été dérivée de la légende de la carte. Les descriptions par profil de référence ont été créées à partir des descriptions de profils de Leenhardt (1991). Le Tableau 6.4 montre ensuite l'ensemble des calculs effectués pour chaque UTS.

Quatre UTS ont été choisies pour la comparaison entre les estimations obtenues par l'approche CSP, par le calcul d'intervalles flous et par la simulation de Monte Carlo (8a, 16, 303 et 974). Ces UTS représentent deux groupes de matériaux pour lesquels les différentes fonctions de pédotransfert ont été établies (cf. § 4.2.2). UTS 8a et UTS 303 représentent les sols des terrasses moyennes (UPP 309A) et les UTS 16 et 974 les sols des Molasses (UPP 552T)). Nous avons utilisé uniquement l'information issue de la description textuelle qui

UTS	Etude	NHDT	NHPR
2	*	2	1
8a	*	1	2
14	*	2	2
16	*	1	2
102	**	1	1
303	**	1	2
345	**	1	2
974	**	1	3

* Carte de sol de la zone ACM (Clochard *et al.*, 1990)

** Carte des pédopaysages du Languedoc-Roussillon (Bornand *et al.*, 1992)

NHDT Nombre d'horizons de la description textuelle

NHPR Nombre d'horizons dans la description par profil de référence

Tableau 6.3: Origines des descriptions des UTS et nombre d'horizons décrits dans les deux types de descriptions.

UTS	Type d'info	2		8a		14		16		102		303		345		974		
		T	PR	T	PR	T	PR	T	PR	T	PR	T	PR	T	PR	T	PR	
CIF	→			*				*				*					*	
CSP	→	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
SMC	→			*				*				*					*	
CIF	Calcul d'intervalles flous sans dépendance							T	Description textuelle									
CSP	Problèmes de Satisfaction de Contraintes							PR	Description par profil de référence									
SMC	Simulation de Monte Carlo																	

Tableau 6.4: Ensemble des expérimentations effectuées pour les différentes UTS choisies.

suppose un unique horizon pour ces sols. Il n'est ainsi pas nécessaire de tenir compte des contraintes entre variables des horizons voisins (les contraintes inter-horizons).

Ensuite, l'ensemble des 8 UTS a été choisi pour analyser l'influence des différents types d'information sur l'estimation de la RU et de θ_{pf} en appliquant la méthode de CSP. Les estimations de la réserve utile et de la teneur en eau au point de flétrissement, en utilisant les descriptions textuelles et les descriptions par profil de référence pour ces 8 UTS, permettront d'analyser (i) les différences entre les estimations issues de type d'information différents et (ii) les différences d'estimation en utilisant l'information issu de cartes pédologiques à différentes échelles.

6.2.2 Méthodes d'évaluation

Pour évaluer les résultats obtenus selon le protocole expérimental, deux aspects sont à considérer: (i) la comparaison des estimations entre elles et (ii) la comparaison des estima-

tions avec des données de terrain. Le premier aspect d'évaluation permettra de déterminer si les estimations sont cohérentes entre elles et d'observer les différences de précision d'estimation. Pour permettre une meilleure comparaison entre les résultats obtenus il convient de présenter tous les résultats sous la même forme. Il a donc fallu transformer la distribution de probabilité issue de la simulation Monte Carlo en distribution de possibilité (cf. Annexe XIII). La comparaison des distributions entre elles est principalement qualitative, en analysant les graphiques présentant les distributions.

Le deuxième aspect d'évaluation utilise les données de validation présentées dans le paragraphe 4.2.5. En ce qui concerne les données mesurées sur le terrain, il a fallu distinguer les données de validation disponible pour le θ_{pf} et les données de validation disponible pour la RU. Les données de θ_{pf} étaient présentes en quantité suffisante pour pouvoir subir une transformation de l'histogramme en distribution de possibilité. En revanche, le nombre d'observations très réduit de la réserve utile ne le permettait pas. Ces observations de RU ont donc été représentées non pas par une distribution de possibilité issue d'une transformation mais par des points en affectant une possibilité de 1 à chaque point.

L'objectif de la comparaison avec les données de terrain est de déterminer quelle distribution représente au mieux la distribution des mesures de terrain. Pour cela, on va se baser sur deux critères:

- 1) le premier critère choisi exprime l'imprécision de l'estimation. Pour ceci, on détermine la largeur de l'intervalle des estimations pour deux α -coupes, notamment, l' α -coupe 0.1 et 1.0.
- 2) le deuxième critère choisi exprime la qualité de l'estimation par rapport à la distribution des données de validation. Cette mesure de qualité est déterminée, pour les α -coupes 1.0 et 0.1, par le taux des observations tombant à l'intérieur de l'intervalle estimé. On appelle ce critère le taux de bonnes estimation.

6.3 Comparaison des différentes méthodes d'estimation des propriétés hydriques du sol

Nous présentons, dans ce paragraphe, l'ensemble de résultats qui vont permettre de comparer les estimations obtenus par l'approche CSP avec les deux alternatives proposées. Dans un premier temps nous présentons l'analyse graphique des différents résultats obtenus. Dans un deuxième temps nous analysons les différentes estimations à l'aide des deux critères mentionnés ci-dessus.

6.3.1 Présentation et interprétation des résultats bruts

Une première interprétation des résultats s'effectue à l'aide d'une représentation graphique des résultats (voir la Figure 6.2). Cette figure montre 8 graphiques, qui correspondent aux 4 UTS d'évaluation choisie (8a, 16, 303 et 974) et les deux propriétés estimées (la réserve utile en eau sur 120 cm (RU) et la teneur en eau au point de flétrissement (θ_{pf}). Chaque graphique montre les trois estimations différentes.

Tout d'abord on compare les différences entre les estimations issues de CSP et avec les estimations obtenues par le calcul d'intervalle flous (CIF). On constate que les différentes distributions sont cohérentes entre elles et, comme nous l'avons attendu, les estimations issues de calcul d'intervalle flous sont généralement plus imprécises. Pour un seul cas (l'estimation de θ_{pf} pour UTS 16) on observe une estimation identique: les distributions coïncident complètement.

Ensuite, les graphiques de la Figure 6.2 permettent de faire un certain nombre d'observations sur les différences obtenus entre les estimations par CSP et les estimations par la simulation Monte Carlo (SMC). Tout d'abord, les résultats obtenus sont cohérents entre eux. En effet, les distributions de possibilité issues de la SMC sont systématiquement incluses dans celles issues de la résolution du CSP, aussi bien pour la RU que pour θ_{pf} . Les distributions de possibilité obtenues par la simulation Monte Carlo sont plus "pointues". Ceci illustre les différences fondamentales entre l'approche CSP (principalement basée sur le calcul d'intervalles flous) et une approche stochastique. Une approche stochastique a généralement tendance à produire une concentration de simulations autour d'une valeur modale (Schulze & Huwe, 1997). Le calcul de la distribution de probabilité (i.e. fréquence relative pour une série de classes définie sur le domaine de la variable recherchée) fournit alors une probabilité importante, dans laquelle se trouve la valeur modale, ce qui est seulement lié au comptage de simulations se situant à l'intérieur de cette classe.

On peut également noter une différence de forme entre les distributions de la réserve utile et celles de la teneur en eau au point de flétrissement (ceci est observable aussi bien pour l'approche CSP que pour la simulation Monte Carlo). Les distributions pour la réserve utile présentent plutôt une forme asymétrique, tandis que les distributions pour la θ_{pf} présentent une distribution symétrique. En ce qui concerne les résultats de RU obtenus par la simulation de Monte Carlo, l'approche stochastique a tendance à privilégier les petites valeurs de la réserve utile; il est en effet plus difficile de trouver un jeu de valeurs des variables donnant une grande valeur de la RU qu'un jeu donnant une petite valeur. Cette observation a été confirmée par la recherche par l'approche CSP de l'ensemble des solutions à une α -coupe donnée: on trouve un nombre plus faible pour les grandes valeurs de RU. Le paragraphe 6.2.1.4 a présenté l'influence de ce phénomène sur le temps de calcul de la recherche de la borne supérieure de la réserve utile.

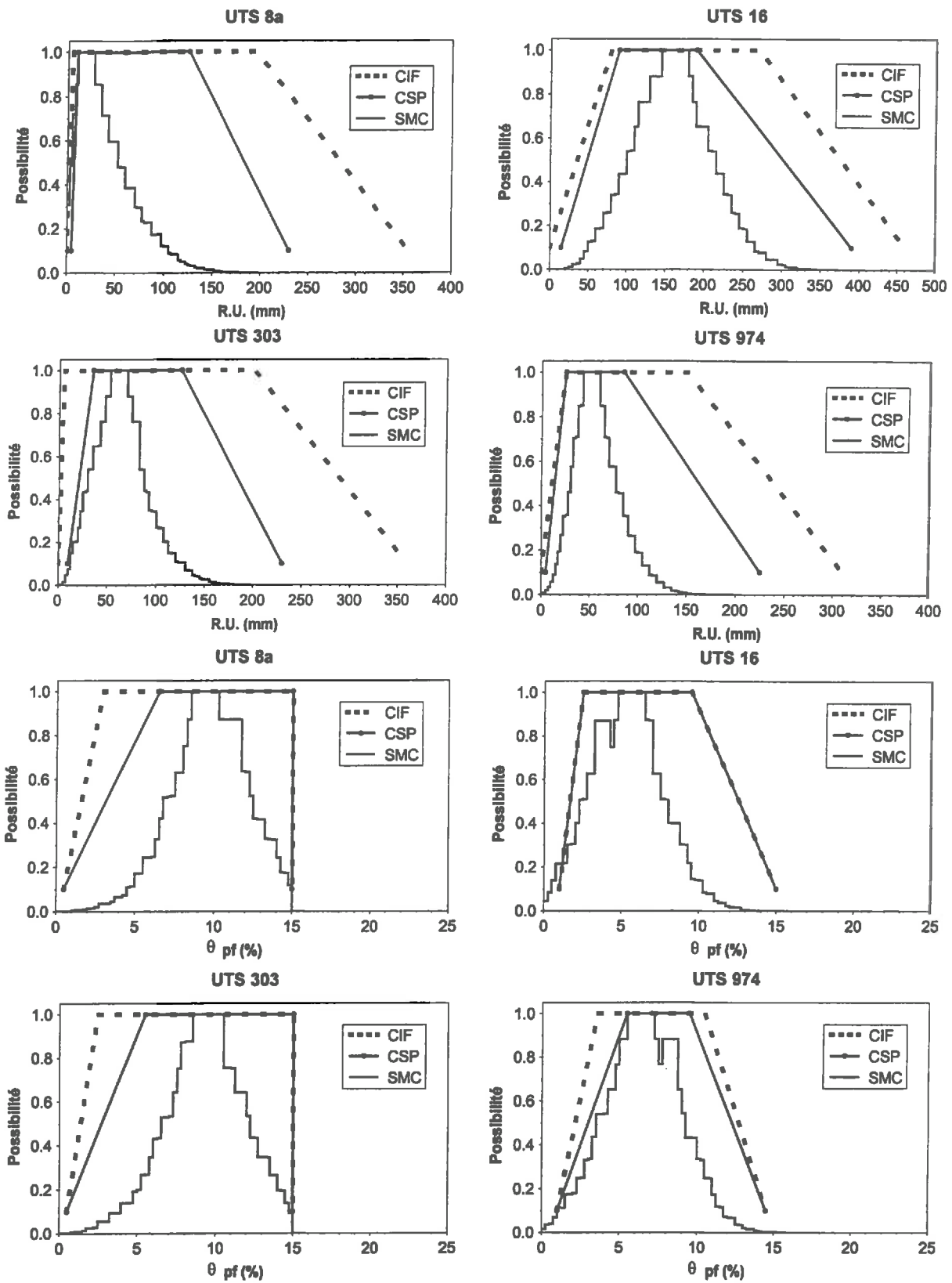


Figure 6.2: Représentation de l'ensemble des estimations issues des trois types de calcul différents pour les quatre UTS d'évaluation et les deux variables considérées.

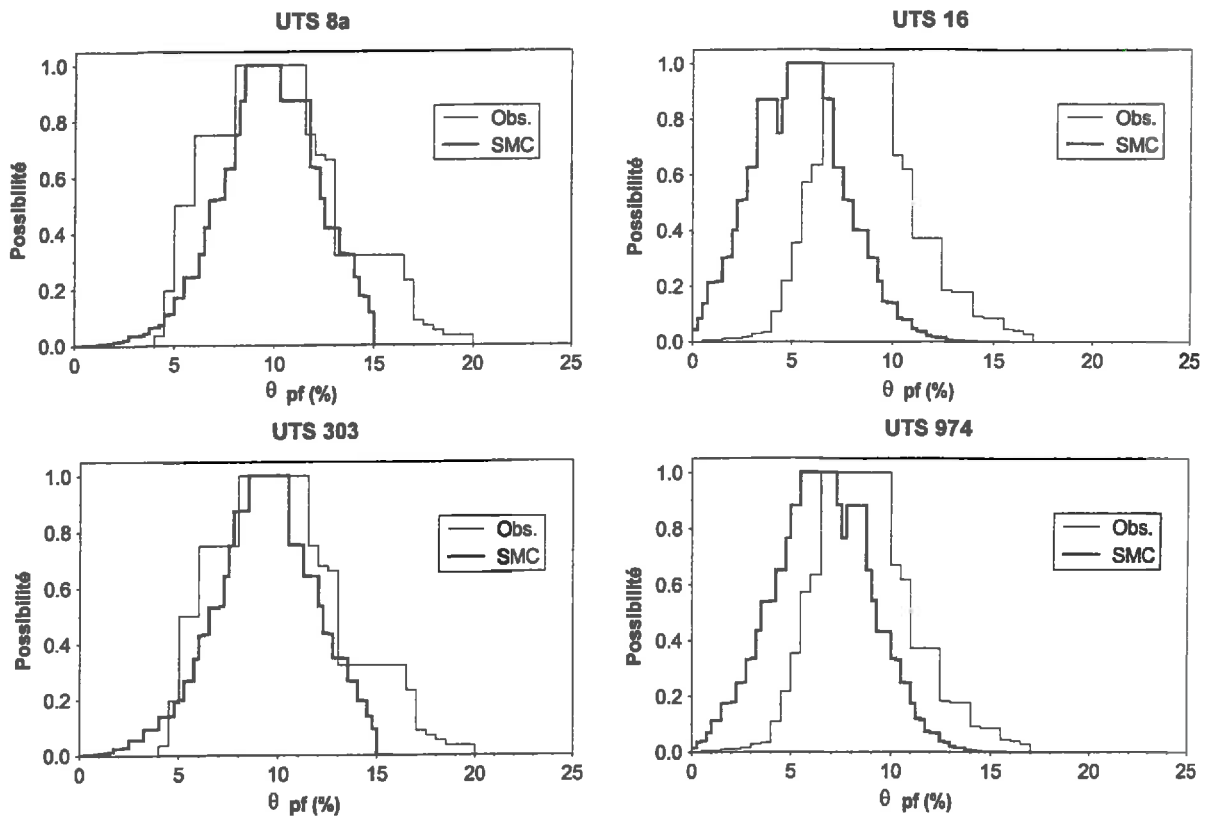


Figure 6.3: Comparaison graphique entre les estimations obtenues par la simulation de Monte Carlo et les mesure de terrain.

Pour les différentes méthodes de propagation, on observe une limitation à 15% des estimations de θ_{pf} pour les UTS 8a et 303. Cette limitation est dû au fait que les fonctions de pédotransfert ont été établies avec un nombre limité de mesures de terrain pour lesquelles θ_{pf} ne dépassait pas 15% (la valeur de θ_{pf} était alors contrainte à un maximum de 15% dans les estimations par les approches différentes). Les données mesurées, par la suite, sur le terrain ont, par contre, montré des valeurs de θ_{pf} supérieures à 15%.

6.3.2 Comparaison des estimations avec des mesures de terrain

Après avoir analysé les différences entre les résultats produits par les trois méthodes appliquées, il convient de comparer ces résultats avec les mesures de terrain.

Tout d'abord, les annexes VIIa et VIIb permettent aussi une comparaison graphique des résultats. On observe que les estimations issues de l'approche CSP et CIF englobent plus largement les valeurs mesurées sur le terrain que la simulation de Monte Carlo. Il est à noter également que l'interprétation graphique pour la réserve utile est rendue difficile par le faible nombre d'observations. Si l'on compare plus en détail seulement les différentes estimations de θ_{pf} avec les données mesurées (voir la Figure 6.3) on constate premièrement que les distri-

butions se présentent sous la même forme suivant le lois normal. La distribution statistique d'une propriété du sol à l'intérieur d'une unité de sol se présente généralement sous la forme d'une distribution normale (Webster & Oliver, 1990). Cependant, on observe un décalage de la totalité de la distribution pour l'UTS 16 et 974, tandis que les distributions estimées avec la simulation de Monte Carlo pour UTS 8a et 303 correspondent relativement bien avec la distribution exprimant les mesures de terrain. Pour les deux cas où les modes des distributions ne correspondent pas, on considère que l'estimation issue de la simulation de Monte Carlo est erronée.

Nous analysons les différences entre les méthodes d'estimation de façon plus approfondie à l'aide des deux critères mentionnés au paragraphe 6.2. Le Tableau 6.5 montre les largeurs des intervalles produites par les 3 méthodes testées pour estimer les variables RU et θ_{pf} . Ces largeurs, fournies pour les UTS différentes et deux α -coupes, sont une mesure de l'imprécision de l'estimation. En observant les valeurs moyennes de la largeur des intervalles, on constate que les estimations par calcul d'intervalles flous sont plus imprécises que celles obtenues par la simulation de Monte Carlo. Les estimations obtenues par CSP ont une précision intermédiaire. Ceci est valable pour les deux variables ainsi que pour les deux α -coupes calculées. On constate cependant que les différences de largeur d'intervalle sont généralement plus important entre l'approche CSP et la simulation Monte Carlo que entre l'approche CSP et le calcul d'intervalles flous classique.

En examinant le détail des résultats par UTS, on s'aperçoit que cette hiérarchie s'observe également systématiquement. Donc, conformément aux comparaisons graphiques précédemment évoquée, la simulation de Monte Carlo délivre des estimation sensiblement plus précise que l'approche CSP, lui même délivrant des estimation plus précises que le calcul d'intervalle flous classique.

Le Tableau 6.6 présente des taux de bonnes estimations pour les mêmes méthodes, unités et α -coupes que précédemment. Les moyennes obtenues toutes UTS confondues révèle un hiérarchie inverse par rapport aux résultats précédents: le calcul d'intervalle flou obtient les meilleurs taux d'estimations correctes alors que la simulation de Monte Carlo obtient des moins bons. Il faut cependant souligner la faible différence enregistrée en moyenne entre les résultats de l'approche CSP et le calcul d'intervalles flous.

Si on analyse les résultats plus en détail pour la teneur en eau au point de flétrissement, on peut considérer que les taux de bonnes estimations obtenues pour l'approche CSP et le calcul d'intervalle flous (entre 63.43 et 98.86 %) sont de très bonne qualité. De plus il n'existe que peu de différence entre les valeurs obtenues pour les α -coupes différentes. En ce qui concerne les estimations de θ_{pf} par la simulation Monte Carlo, on constate une forte différence entre les taux de bonnes estimations par α -coupes (moyenne de 75.23% pour α -coupe 0.1 et de 23.04% pour α -coupe 1).

UTS	ALPHA	θ_{pf} (%)			Ru (mm)		
		CSP	CIF	SMC	CSP	CIF	SMC
16	0.1	14.0	14.0	9.75	375	460	215
	1.0	7.0	7.0	1.75	100	190	35
8a	0.1	14.5	14.5	11.50	225	355	105
	1.0	8.5	12.0	1.75	115	190	17.5
303	0.1	14.5	14.5	11.00	220	360	117.5
	1.0	9.5	12.5	2.00	90	195	17.5
974	0.1	13.5	13.5	10.75	220	310	97.5
	1.0	4.0	7.0	1.75	60	125	17.5
moyenne	0.1	14.13	14.13	10.75	260	371	134
moyenne	1.0	7.25	9.64	1.81	91	175	21.9

Tableau 6.5: Largeur des intervalles estimés pour les trois approches évaluées et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).

UTS	ALPHA	θ_{pf} (%)			Ru (mm)		
		CSP	CIF	SMC	CSP	CIF	SMC
16	0.1	98.86	98.86	76.00	100	100	100
	1	66.29	66.29	12.57	100	100	17
8a	0.1	85.71	85.71	51.79	100	100	67
	1	78.57	85.71	28.57	100	100	0
303	0.1	85.71	85.71	85.71	100	100	100
	1	85.71	85.71	32.14	67	100	17
974	0.1	97.71	97.71	87.43	86	100	29
	1	63.43	76.00	18.86	14	43	0
moyenne	0.1	92.00	92.00	75.23	96.5	100	74
moyenne	1	73.50	78.43	23.04	70.25	85.75	8.5

Tableau 6.6: Taux des bonnes estimations pour les trois approches évaluées et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).

En ce qui concerne les estimations de la réserve utile, on observe des fortes taux de bonnes estimations pour l'approche CSP et l'approche par calcul d'intervalles flous (moyennes entre 70,25 et 100%). On considère ces taux comme satisfaisants. On constate un taux très faible (moyenne de 8.5%) pour les estimations par simulation de Monte Carlo pour l' α -coupe 1.0, mais satisfaisant pour α -coupe 0.1 (74%).

L'analyse plus précise indique que les résultats ne sont pas homogènes d'une UTS à l'autre. Particulièrement, l'UTS 974 enregistre de faibles taux de bonnes estimations de la réserve utile quel que soit les méthodes employée. Cette forte différence entre les estimations de RU et les données de terrain est dû à l'incohérence de l'information sur la profondeur dérivée de la description textuelle et la profondeur des données des profil de validation. L'interprétation de la description textuelle de l'unité 974 amène à définir un intervalle flou correspondant à "sol peu profond" alors que les profils de validation, décrits jusqu'à 120 cm ne révèlent pas d'horizon limitant l'enracinement, ce qui amène à définir leur profondeur à au

moins 120cm. Cette différence est vraisemblablement due à l'ambiguïté de la description de la profondeur déjà évoquée dans le paragraphe 5.1.1.2 (profondeur du sol ou profondeur de l'enracinement?).

Pour finir, nous pouvons, à présent, déterminer la méthode qui constitue le meilleur compromis entre l'imprécision et l'incertitude. On cherche ainsi la méthode qui fournit des estimations avec une faible largeur d'intervalle accompagnée d'une incertitude acceptable. Pour les moyennes observées pour α -coupe 1.0 on constate que ce compromis est constitué par l'approche CSP ayant des taux de bonnes estimations de 73.5 et 70.25 pour les variables respectives. Ces taux ont été considérés comme satisfaisants. En ce qui concerne les moyennes observés pour α -coupe 0.1 on constate que ce meilleur compromis est donné par la simulation de Monte Carlo (taux de bonnes estimations de 75.23 et 74).

6.3.3 Conclusion

Le calcul des propriétés hydriques par UTS a été reconnue comme un problème de propagation d'incertitude dans une opération utilisant des variables interactives. Nous avons proposé de résoudre ce problème en le formalisant dans un cadre possibiliste; et la propagation de l'incertitude étant réalisée grâce aux techniques CSP.

Ce paragraphe a présenté les résultats des expérimentations effectuées afin d'évaluer l'approche CSP confrontées à deux alternatives (calcul d'intervalles flous et simulation de Monte Carlo étant des méthodes conventionnelles dans la résolution de notre type de problème).

L'analyse graphique et quantitative des résultats présentés ci-dessus démontrent que:

- le calcul d'intervalles flous produit des estimations très imprécises, mais elles englobent généralement les observations de terrain. Ces estimations sont donc correctes, mais peu informatives.
- le calcul par simulation de Monte Carlo fournit des estimations plus précises et donc, a priori plus informatives. La comparaison avec des données de terrain nous apprend que ces estimations n'englobent généralement qu'une partie des mesures de terrain. Ces estimations s'avèrent donc précises mais fausses.
- le calcul d'intervalles flous avec prise en compte des dépendances induites sur certaines variables du fait de leur appartenance à plusieurs relations (les calculs étant effectués par des techniques de CSP) paraît une très bonne alternative aux méthodes conventionnelles. Les résultats montrent une précision supérieure à celle obtenue par le calcul d'intervalles flous classique et une précision inférieure à celle obtenue par la simulation Monte Carlo. En revanche,

l'amélioration du taux des estimations correctes l'approche CSP par rapport aux taux de bonnes estimations observé pour la simulation de Monte Carlo est très important, tandis qu'une faible perte de taux de bonnes estimations a été constatée par rapport au calcul d'intervalles flous classique.

6.4 Comparaison des estimations issues de différentes données

L'analyse présentée dans le paragraphe précédent a permis de justifier l'utilisation de l'approche CSP pour l'estimation de la réserve utile et de la teneur en eau point de flétrissement. C'est donc cette méthode qui va être utilisée pour la comparaison des estimations obtenues par les différents types d'information.

Les résultats obtenus par la résolution des CSP sont présentés dans les Figures 6.4 et 6.5. La Figure 6.4 montre les résultats de l'estimation de la teneur en eau au point de flétrissement et la Figure 6.5 celle de la réserve utile pour les 8 UTS d'évaluation. Chaque graphique montre trois distributions différentes: (i) la distribution de possibilité obtenue en utilisant l'information de la description textuelle, (ii) la distribution de possibilité obtenue en utilisant l'information de la description par profil de référence (en pointillé) et (iii) la distribution de possibilité exprimant les mesures de terrain.

L'interprétation des estimations obtenues grâce aux deux types d'information (description textuelle et description par profil de référence) s'effectue en deux étapes. Dans un premier temps, il est important de comparer les résultats issus des différents types d'information entre eux. Dans un deuxième temps, une comparaison avec les mesures de terrain sera effectuée à l'aide des deux critères d'évaluations proposés auparavant.

6.4.1 Analyse graphique des résultats bruts

Les graphiques présentés dans les Figures 6.4 et 6.5 permettent de faire un certain nombre d'observations. Tout d'abord, il existe quelques cas, où la borne supérieure pour la réserve utile n'a pas pu être estimée avec la précision souhaitée à cause des problèmes de temps de calcul évoqués auparavant. Il s'agit de l'estimation issue de la description textuelle pour l'UTS 14 et des estimations issues du profil de référence pour les UTS 8a, 303 et 974. Les graphiques correspondants montrent une double ligne, indiquant l'intervalle des valeurs de la borne supérieure de RU (cf. § 6.2.1.5).

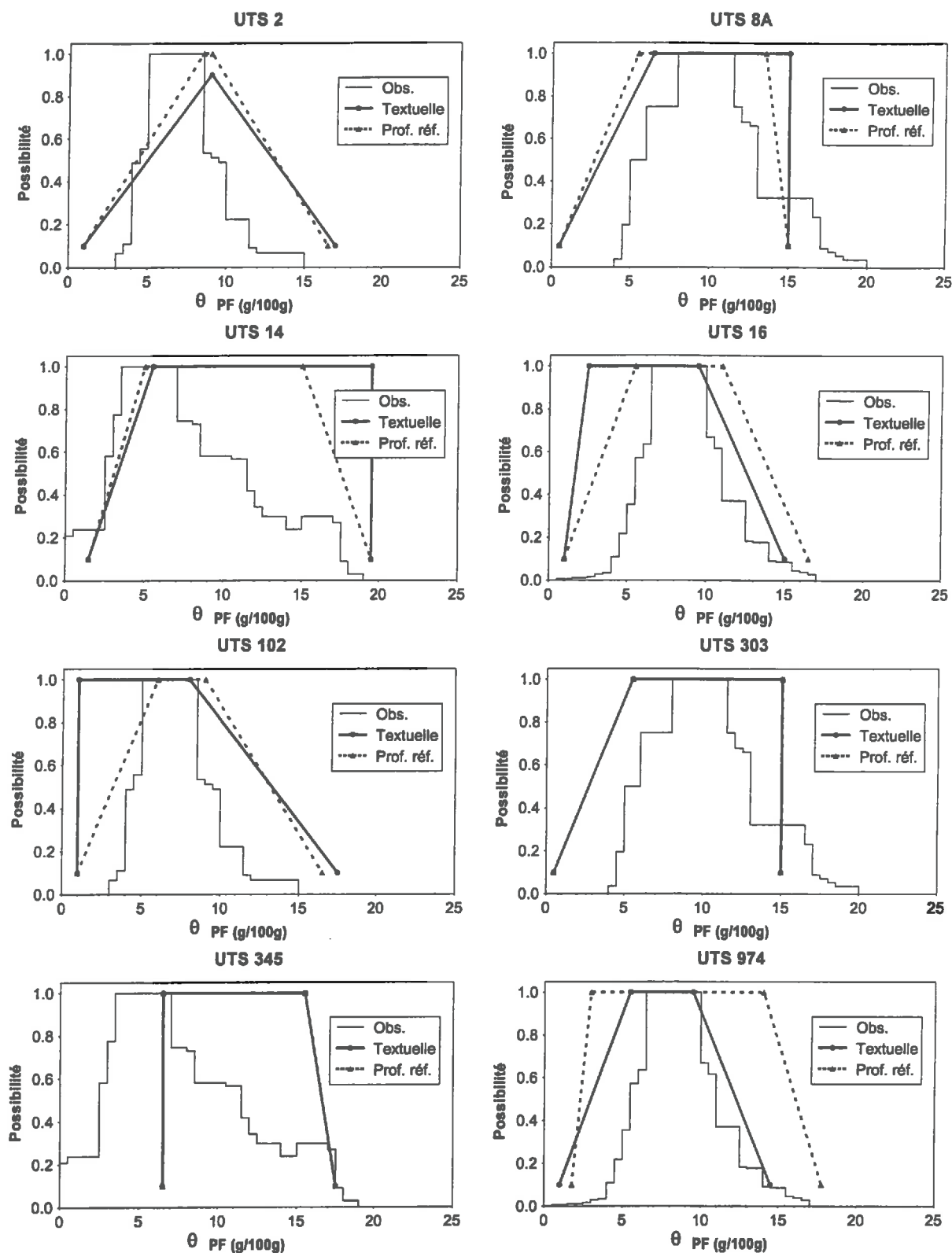


Figure 6.4: Estimations de θ_{pr} issues des deux types de description des UTS différents (profil de référence et description textuelle) avec la distribution des mesures de terrain.

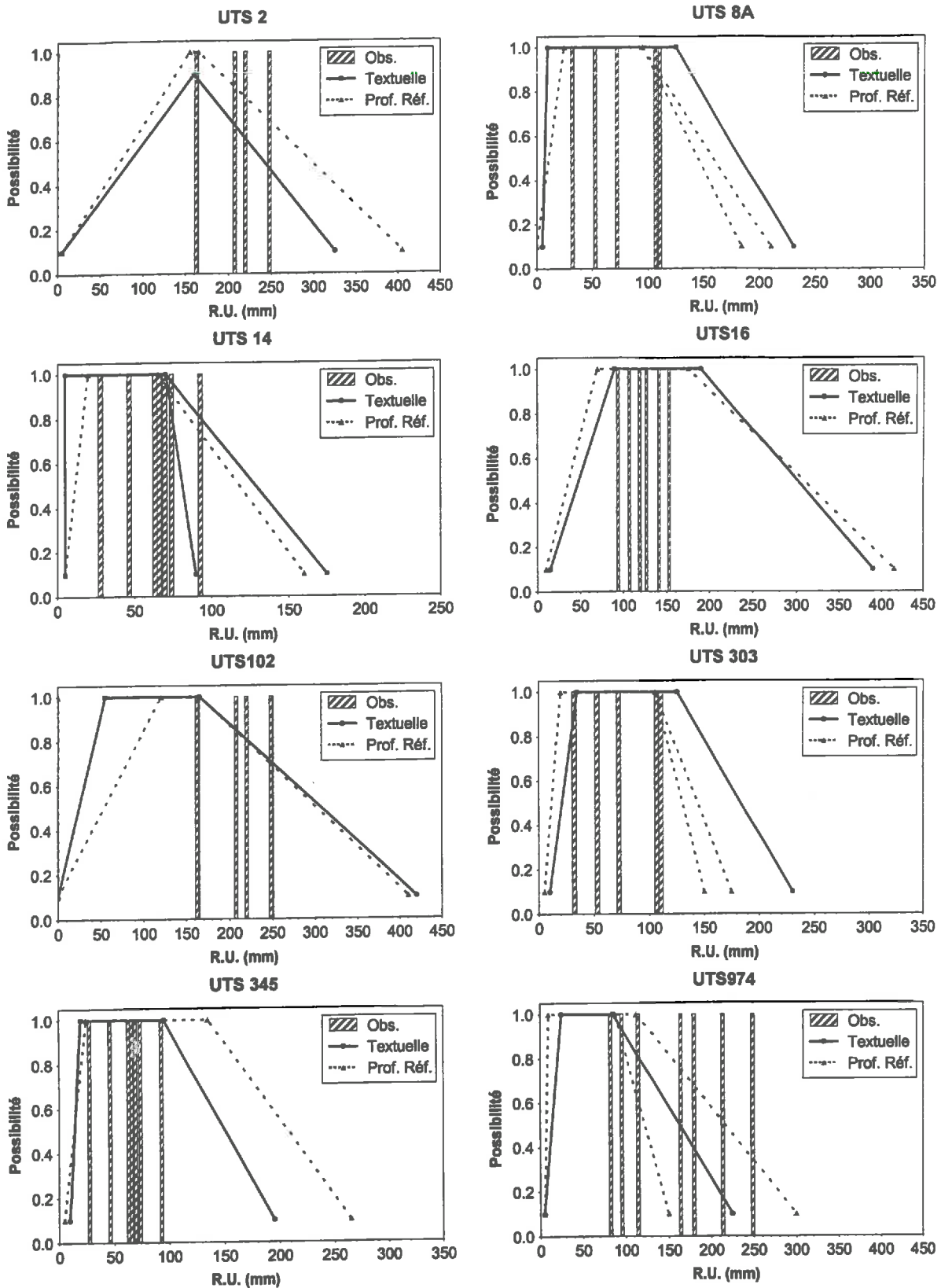


Figure 6.5: Estimations de la réserve utile issues des deux types de description différents (profil de référence et description textuelle) avec la distribution des mesures de terrain.

En ce qui concerne les deux variables estimées, on observe une possibilité maximale de 0.9 pour les estimations de RU et de θ_{pf} issues de la description textuelle de l'UTS 2. Pour ce cas précis, aucune solution n'a été trouvée pour le CSP pour l' α -coupe 1. Ceci signifie donc une incohérence des connaissances disponibles (voir la définition des distributions de possibilité normalisées en 3.3.4). Une analyse des données introduites pour cette UTS spécifique nous apprend que cette unité contient deux horizons qui sont très contrastés. Ceci provoque ainsi des incohérences avec les contraintes 'inter-horizon' qui supposent une proximité entre les propriétés du sol des horizon voisins. Il conviendrait donc de supprimer ces contraintes dans les descriptions d'UTS où une forte discontinuité serait signalée.

Ensuite, on peut constater que les résultats obtenus grâce aux deux types d'information correspondent relativement bien. A première vue, pour l'ensemble des graphiques présentés, on n'observe pas une augmentation nette de la précision des estimations de RU et θ_{pf} issues de la description par profil de référence par rapport aux estimations issues de la description textuelle. A priori on s'attendait à des estimations plus précises avec le profil de référence car les descriptions ne se font pas au moyen de classe pré-établie qui limite la précision. De plus, les descriptions textuelles ne distinguent généralement pas des horizons. Nous allons analyser ce résultat surprenant de manière plus approfondie dans le paragraphe suivant.

6.4.2 Analyse des résultats à l'aide des deux critères d'évaluation

L'objectif de la comparaison avec des mesures de terrain est de déterminer quel type d'information permet la meilleure estimation des propriétés hydriques du sol. Pour cela, on va se baser sur les critères définis dans le paragraphe 6.3.2. Dans le cas des UTS pour lesquelles les bornes du noyau et du support sont eux-mêmes représentés par un intervalle, le calcul de ces critères se fait en considérant la valeur qui définit le noyau ou le support le plus large. Le Tableau 6.7 montre les largeurs des intervalles produites en utilisant les 2 types d'informations testés pour les variables RU et θ_{pf} . En observant les valeurs moyennes de la largeur des intervalles, on constate qu'il existe peu de différences entre les deux types d'informations introduits; ce qui confirme l'observation faite dans le paragraphe précédent. Ceci est le cas pour les deux α -coupes considérées et les deux variables estimées. Les estimations issues du profil de références sont légèrement plus précises pour α -coupe 1.0 (7.4 vs. 7.1 et 77 vs. 72), tandis qu'elles sont moins précises pour l' α -coupe 0.1. (14.8 vs. 15 et 267 vs. 288). En examinant en détail les résultats par UTS (cf. également la Figure 6.4 et 6.5), on s'aperçoit que la différence de précision entre les deux types d'information varie fortement d'une UTS à l'autre. Nous allons analyser les résultats au cas par cas par UTS. Les déclara-

UTS	α -coupe	θ_{pf} (%)		RU (mm)	
		T	PR	T	PR
2	0.1	16.0	15.5	320	400
	1.0	0	0.5	0	10
8a	0.1	14.5	14.5	225	210
	1.0	8.5	8.0	115	70
14	0.1	18.0	18.0	170	155
	1.0	14.0	10.0	65	45
16	0.1	14.0	15.5	375	405
	1.0	7.0	5.5	100	105
102	0.1	16.5	15.5	420	410
	1.0	7.0	3.0	110	45
303	0.1	14.5	14.5	220	170
	1.0	9.5	9.5	90	85
345	0.1	11.0	11.0	185	260
	1.0	9.0	9.0	75	110
974	0.1	13.5	15.9	220	295
	1.0	4.0	11.0	60	102.5
moyenne	0.1	14.8	15.0	267	288
moyenne	1.0	7.4	7.1	77	72

Tableau 6.7: Largeur des intervalles estimés pour les deux types d'information évalués et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).

tions des variables dans les CSP respectives pour chaque UTS et par type d'information sont présentées dans l'annexe XII.

On constate une amélioration de précision pour les UTS 8a et 14. Cette amélioration est valable pour l'estimation de la réserve utile ainsi que l'estimation de θ_{pf} . Les descriptions par profil de référence contiennent des horizons contrastés bien différenciés. De plus, l'information disponible sur l'épaisseur des horizons est très précise. L'information est plus précise dans sa totalité; ce qui fournit des estimations plus précises pour les deux variables.

Pour deux UTS (303 et 345), on observe une précision équivalente de l'estimation de θ_{pf} . Pour ces deux UTS, l'imprécision des épaisseur des horizons du profil de référence est telle que, à 40 cm, il est également possible qu'on ait les propriétés du premier horizon ou du deuxième horizon. La différenciation en horizons ne joue donc pas car la combinaison disjonctive des intervalles définie par horizon ressemble beaucoup aux intervalles de la description textuelle; ce qui produit une estimation équivalente pour ces deux UTS. Dans le cas de l'estimation de θ_{pf} pour UTS 345 et 303, les éventuelles différences, qu'on pourrait trouver, ne s'observent même pas, compte tenu du pas de discrétisation appliqué (cf. Tableau 6.2).

Pour deux UTS (345 et 974), on observe une diminution de la précision de l'estimation de la réserve utile en utilisant la description par profil de référence. En ce qui concerne l'UTS 345, on constate que l'information pour l'horizon principal est plus imprécise

UTS	α -coupe	θ_{pf} (%)		RU (mm)	
		T	PR	T	PR
2	0.1	100.00	100.00	100	100
	1.0	0.00	4.44	0	33
8a	0.1	98.86	98.86	100	100
	1.0	66.29	95.43	100	67
14	0.1	85.71	85.71	100	100
	1.0	78.57	78.57	75	63
16	0.1	95.52	95.52	100	100
	1.0	65.67	62.69	100	100
102	0.1	100.00	100.00	100	100
	1.0	62.22	62.22	33	33
303	0.1	85.71	85.71	100	100
	1.0	85.71	85.71	67	67
345	0.1	53.73	53.73	100	100
	1.0	46.27	46.27	100	100
974	0.1	97.71	100.00	86	100
	1.0	63.43	96.00	14	28
moyenne	0.1	89.66	89.94	98	100
moyenne	1.0	58.52	66.42	61	61

Tableau 6.8: Taux des bonnes estimations pour les deux types d'informations évaluées et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).

que la description textuelle. Cette imprécision des données d'entrée est liée au choix des profils utilisés dans la construction du profil de référence. Pour la description de l'unité 345, le pédologue a utilisé un profil à charge en cailloux en profondeur exceptionnellement faible qui étant seul disponible à l'échelle considérée. Ce profil, apparemment représentait une variante de l'UTS, n'a par contre pas été pris en compte par le pédologue dans la description textuelle. En ce qui concerne l'UTS 974, on constate que les horizons sont contrastés, mais que leurs épaisseurs sont définies de manière très imprécise. Dans ce cas très particulier, les 3 horizons peuvent être présents à la profondeur de 40 cm, ce qui provoque une imprécision importante pour l'estimation de θ_{pf} comme indiqué ci-dessus. L'imprécision sur les épaisseurs d'horizon se répercute ainsi également pour l'estimation de la réserve utile. Il est à noter ici, que la résolution des CSP de l'UTS 974 a été très coûteuse en temps de calcul du fait de cette grande imprécision (cf. § 6.1.1.4).

Le Tableau 6.8 présente des taux des bonnes estimations pour les deux types d'information et les mêmes unités et α -coupes que précédemment. Les moyennes obtenues toutes UTS confondues montrent des taux de bonnes estimations très proches pour les deux types d'informations différents. La différence la plus importante entre les deux types d'informations s'observe pour les estimations de θ_{pf} pour α -coupe 1.0 (58.52 pour la description textuelle vs. 66.42 pour la description par profil de référence). Cette différence est prin-

	α -coupe	θ_{pf} (%)		RU (mm)	
		T	PR	T	PR
Carte pédologique au 10.000^{ème}					
Moyenne	0,1	15.6	15.9	273	293
Moyenne	1	7.4	6.0	70	58
Carte pédologique au 250.000^{ème}					
Moyenne	0,1	13.9	14.2	261	284
Moyenne	1	7.4	8.1	84	86

Tableau 6.9: Moyennes des largeurs d'intervalle pour les deux groupes d'unités typologiques de sol issus de cartes pédologiques à différente échelle pour les deux variables estimées, par α -coupe et par type d'information utilisé.

cipalement due aux différences obtenues pour UTS 16 et 974. Ces résultats confirment les faibles différences entre la description textuelle et la description par profil de référence en moyenne.

En résumé, les profils de références ne s'avèrent utiles que dans les cas où les horizons sont bien différenciés et leurs épaisseurs connues avec précision. Dans le cas où les horizons sont peu distincts, mais les épaisseurs sont bien connues, on obtient des estimations équivalentes à celles obtenues par l'utilisation de la description textuelle. Dans le cas où les horizons sont bien distincts, mais les épaisseurs respectives définies de manière imprécises, l'estimation par profil de référence s'avère moins performante.

6.4.3 Différences de précision des estimations issues des cartes pédologiques à différentes échelles

L'approche de l'estimation des propriétés hydriques à partir d'une information imprécise a été initialement développée pour l'information issue des études pédologiques à petite échelle. Dans l'analyse précédente, nous avons introduit des descriptions textuelles et des descriptions par profil de référence issues d'une carte pédologique à grande échelle. Les UTS 2, 8a, 14, et 16 (voir le Tableau 6.3) sont issues de la cartographie au 1/10.000 de la zone de test ACM, alors que les UTS 102, 303, 345 et 974 sont issues de la carte des pédopaysages qui couvre la zone d'étude VHOL (voir la Figure 4.1). Une comparaison des résultats pour les deux groupes d'unités typologiques de sols permet d'observer des éventuelles différences de précision entre les estimations des cartes pédologiques à échelles différentes et donc d'affiner le diagnostic précédent. A priori, on attend une estimation plus précise pour les UTS issues des cartes à grande échelles. Ce gain de précision est principalement attendu pour les descriptions par profil de référence car, à grande échelle, la connaissance de l'unité de sol

	α -coupe	θ_{pf} (%)		RU (mm)	
		T	PR	T	PR
Carte pédologique au 10.000^{ème}					
Moyenne	0,1	95	95	100	100
Moyenne	1	53	60	75	69
Carte pédologique au 250.000^{ème}					
Moyenne	0,1	85	85	97	100
Moyenne	1	64	73	54	57

Tableau 6.10: Moyennes des taux de bonnes estimations pour les deux groupes d'unités typologiques de sol issus de cartes pédologiques à différentes échelles pour les deux variables estimées, par α -coupe et par type d'information utilisé.

est plus complète; ce qui permet au pédologue de mieux choisir les profils de sol pour établir plus précisément la description par profil de référence.

Le Tableau 6.9 présente les moyennes des largeurs des intervalles pour les deux groupes d'unités typologiques de sol pour les deux variables estimées, par α -coupe et par type d'information utilisé. On observe une précision égale ou supérieure pour les estimations issues de la carte à grande échelle par rapport aux estimations issues de la carte à petite échelle pour α -coupe 1.0. On observe également que les différences sont les plus élevées pour les estimations par profil de référence; ce qui correspond aux attentes mentionnées ci-dessus.

En ce qui concerne l' α -coupe 0.1, on observe une précision légèrement supérieure pour les estimations issues de la carte à petite échelle que celles de la carte à grande échelle. Cependant, ces différences sont proportionnellement moins importantes que celles observées pour l' α -coupe 1.0. On considère alors que la précision est équivalente pour les deux groupes d'unité pour l' α -coupe 0.1. Ces faibles différences sont liées au mêmes marges floues utilisées pour les deux groupes d'unités de sols.

Le Tableau 6.10 montre ensuite les moyennes des taux de bonnes estimations pour les deux groupes d'UTS. On constate que les taux de bonnes estimations issues de la carte à grande échelle sont généralement supérieurs à celle de la carte à petite échelle. Cependant, pour l'estimation de θ_{pf} on observe des taux supérieures pour les estimations issues de la carte à grande échelle. Cette différence pour les deux type d'informations est fortement biaisée par l'incohérence rencontrée dans la résolution de CSP pour l'UTS 2 évoque précédemment (cf § 6.4.1).

Au vu des résultats présentés dans les Tableaux 6.9 et 6.10, nous pouvons conclure que les estimations pour les UTS issues de la carte à grande échelle sont généralement plus précises. Ce gain de précision n'est cependant pas accompagnée d'une augmentation d'incertitude.

6.5 Conclusion

Ce chapitre a présenté l'approche permettant d'estimer les propriétés hydriques des UTS en appliquant l'approche CSP proposée dans le troisième chapitre ainsi que l'évaluation des résultats obtenus. Nous avons essayé de répondre à deux questions différentes concernant l'estimation des propriétés hydriques pour chaque UTS.

La première question concerne la justification de l'utilisation de l'approche CSP. Pour ceci, nous avons comparé les résultats issus de l'approche CSP avec les résultats obtenus par deux approches plus conventionnelles: (i) le calcul d'intervalles flous sans prise en compte des dépendances induites sur certaines variables, du fait de leur appartenance à plusieurs relations et (ii) la simulation de Monte Carlo.

Cette comparaison nous apprend que les estimations par l'approche CSP et le calcul d'intervalles flous classique sont imprécises, mais couvrent généralement la variabilité observée. Une approche stochastique, quant à elle, fournit des résultats généralement plus précis, mais souvent erronés. L'approche CSP (i.e. calcul d'intervalles flous avec prise en compte des dépendances induites sur certaines variables du fait de leur appartenance à plusieurs relations) paraît une très bonne alternative aux méthodes conventionnelles. Les résultats montrent une précision supérieure à celle obtenue par le calcul d'intervalles flous classique et une précision inférieure à celle obtenue par la simulation Monte Carlo. En revanche, l'amélioration du taux des bonnes estimations par l'approche CSP par rapport aux taux de bonnes estimations pour la simulation de Monte Carlo est très importante, tandis qu'une faible perte de taux de bonnes estimations a été constatée par rapport au calcul d'intervalles flous classique.

La deuxième question traitée dans ce chapitre concerne l'utilisation de différentes sources de données. Nous avons analysé l'influence des différents types de données (description textuelle et description par profil de référence) dans l'estimation de la réserve utile et de la teneur en eau au point de flétrissement. Cette analyse a mis en évidence la proximité des résultats obtenus par les deux types d'informations. Il est intéressant d'appliquer les profils de référence seulement si toutes les propriétés du sol sont décrites de manières précises.

En ce qui concerne la comparaison de la précision des estimations issues des cartes pédologiques à différente échelle, nous pouvons conclure que la précision la plus élevée s'obtient en appliquant les profils de référence des cartes à grande échelle. Ces descriptions contiennent l'information la plus riche et la plus précise, ce qui permet alors d'estimer les propriétés hydriques avec beaucoup de précision. Cette avantage de précision n'est pas accompagnée d'une incertitude importante.

7 Cartographie des unités typologiques de sols

L'analyse présentée dans le chapitre précédent a permis de justifier l'utilisation de l'approche CSP et des données issues de la description textuelle de chaque UTS pour l'estimation des propriétés hydriques. A présent, il est possible d'estimer la réserve utile et la teneur en eau au point de flétrissement pour chaque UTS définie dans la zone d'étude, ce qui correspond à la première étape de la démarche générale.

La cartographie de la réserve utile et de la teneur en eau au point de flétrissement nécessite également que soient localisées les UTS pour lesquelles les RU ont été calculées. Or, la carte pédologique régionale utilisée dans la présente étude ne fournit pas la localisation exacte des UTS, car les unités cartographiques des sols sont des unités complexes contenant plusieurs unités typologiques de sols.

Dans ce chapitre, nous présentons et évaluons une méthode de localisation des UTS à l'aide de la description de leur environnement stockée dans une base de données sol et des informations sur les facteurs déterminant l'organisation des UTS (géologie, pente et occupation du sol). Cette méthode s'appuie sur le cadre théorique défini dans le troisième chapitre.

7.1 Le problème de la localisation des UTS au sein des UCS

Comme évoqué précédemment, la carte des sols au 1:250.000 ne fournit pas une localisation précise de chaque UTS. A partir de ce document, il est seulement possible d'identifier en chaque point de l'espace, une liste d'UTS "possibles" qui correspondent aux UTS de l'UCS dans laquelle se trouve le point. Cette imprécision a bien entendu des conséquences sur l'imprécision de l'estimation finale des propriétés hydriques des sols. Il est donc important d'essayer de la réduire.

Pour la présente étude, nous avons proposé d'améliorer la précision de la localisation des UTS au sein des unités de pédopaysages en exploitant à la fois la description environnementale des UTS et l'information géographique supplémentaire généralement disponible à l'échelle régionale. Le paragraphe 5.2 a exposé la façon dont les connaissances ont été représentées afin de permettre le traitement de l'information dans le cadre de la théorie des possibilités. Dans ce paragraphe, nous avons pu distinguer trois variables géographiques potentiellement utilisables pour la localisation des UTS (géologie, occupation du sol, pente). On suppose donc qu'en introduisant cette connaissance, la localisation des UTS sera plus précise que celle obtenue en utilisant seulement la carte des pédopaysages (i.e. qu'une partie de l'imprécision mentionnée ci-dessus sera éliminée). Pour pouvoir évaluer ce gain d'information, il est nécessaire de vérifier la localisation ainsi modélisée avec la localisation des UTS issue d'une carte pédologique à plus grande échelle (i.e. les données de validation). Ceci nécessite également l'introduction de certaines mesures d'évaluation. Nous en proposerons trois permettant de quantifier à la fois l'incertitude et l'imprécision de la localisation.

Ces mesures d'évaluation permettront d'analyser la qualité de la localisation des UTS à l'aide d'autres combinaisons de variables géographiques (par ex. géologie et occupation du sol ou pente, occupation du sol et altitude, etc.). Cette analyse permettra de déterminer la sensibilité de la localisation des UTS à certaines variables et éventuellement de déterminer une combinaison optimale de variables.

Dans les paragraphes suivants, nous présentons successivement la méthode retenue, le protocole d'évaluation et les résultats obtenus.

7.2 Méthodes et techniques appliquées pour la localisation des UTS

Il est nécessaire d'exposer en détail les méthodes et techniques appliquées pour améliorer la localisation des UTS. Premièrement, nous présentons les aspects théoriques des

opérations nécessaires et issues de la théorie de possibilité. Deuxièmement, nous explicitons la façon dont l'analyse a été implémentée dans un SIG.

7.2.1 Localisation par appariement partiel et opération ensembliste

L'identification, en un point de l'espace, des unités typologiques de sols possibles, sous l'hypothèse que le point appartienne à une unité de pédopaysage donné, correspond à une application de l'appariement partiel (*fuzzy pattern matching*) (Cayrol *et al.*, 1982, Dubois & Prade, 1987). Pour chaque unité typologique de sols de l'unité de pédopaysage, l'appariement partiel confronte la description de chaque propriété (géologie, occupation du sol et pente) exprimée par l'appartenance à une classe (cf. § 5.2.1) avec la valeur de cette propriété connue au point considéré. Soit μ_P la distribution de l'ensemble flou décrivant la classe de valeurs de la propriété P_i (e.g. la pente) pour l'unité typologique de sols considérée et soit π_{P_i} la distribution de possibilité de cette propriété au point considéré. La compatibilité de la propriété P_i de l'unité typologique de sol avec celle du point considéré est donnée par la mesure de possibilité (cf. § 3.3.4, p. 46):

$$\Pi(P_i) = \sup_s \min(\mu_{P_i}(s), \pi_{P_i}(s)) \quad (7.1)$$

L'agrégation des compatibilités sur chacune des propriétés est obtenue par l'équation 7.2 où w_i est un degré d'importance relative de la propriété P_i dans l'identification d'une unité de sol. w_i a une valeur dans $[0,1]$, 1 représentant l'importance maximale et 0 signifiant que P_i est sans importance.

$$\Pi_{UTS} = \min_{i=1,n} \max(1-w_i, \Pi(P_i)) \quad (7.2)$$

L'analyse proposée dans cette formulation est basée sur la mesure de possibilité. Il est envisageable ici d'effectuer l'identification des unités typologiques de sol en utilisant aussi la mesure de nécessité dans l'appariement partiel. On pourrait alors dériver deux évaluations de la propriété hydrique du sol recherchée en chaque point: l'évaluation 'optimiste' telle quelle est présentée ci-dessus et l'évaluation 'pessimiste' basée sur la mesure de nécessité. Dans l'état actuel de l'analyse, seule la mesure de possibilité est appliquée. Ainsi, l'exécution de cette combinaison sur chaque point de l'espace permet de cartographier la distribution de possibilité d'apparition de chaque UTS.

Sans introduction de l'information géographique supplémentaire, le nombre d'UTS candidates sur un point d'intérêt est théoriquement maximal et il correspond au nombre d'UTS défini pour l'UCS dans laquelle se situe le point considéré. Lorsque le point considéré se situe dans la zone frontalière entre deux UCS, deux UCS peuvent être complètement possibles (voir la représentation des limites dans le paragraphe 5.2.). Le nombre total d'UTS pos-

sibles sur ce point est alors la somme des UTS possibles pour les deux UCS voisines. Ceci introduit ainsi une imprécision sur la localisation considérable pour les zones frontalières des UCS.

7.2.2 L'utilisation d'un SIG en mode raster

Un SIG en mode raster¹ permet d'effectuer des opérations ensemblistes sur les différents couches d'information géographique. Ceci nécessite une transformation de l'information brute (les cartes vectorielles concernant la topographie, la géologie, etc.) en cartes sous format raster. Ensuite, l'information des cartes doit être exprimée sous formes de possibilité (cf. § 5.2) pour pouvoir effectuer les combinaisons proposées ci-dessus. La procédure développée pour la localisation des UTS comprend quatre étapes. Une série de procédures écrites en langage Arc-Info Macro Language (AML) a permis l'automatisation de la plupart des étapes.

a. *Création des couvertures de base*

Nous avons superposé à la zone VHOL une grille régulière avec un pas de 50 mètres déterminant l'ensemble des points sur lequel l'analyse sera effectuée. Les cartes d'origine en mode vecteur (carte des pédopaysages, carte géologique et carte d'occupation du sol) ont ensuite été rastérisées avec la même résolution. Le modèle numérique d'altitude (MNA) était déjà de même résolution.

b. *Calcul des limites floues*

Pour chaque unité de pédopaysage et pour chaque unité géologique considérée séparément, on a créé des cartes en introduisant les limites floues. Le degré d'appartenance de chaque point se situant à l'extérieur de l'unité considérée est fonction de la distance entre ce point et la limite de l'unité considérée (cf. paragraphe 5.2.2). Cette distance a été calculée pour chaque point situé à l'extérieur de l'unité considérée appliquant une fonction de calcul de la distance euclidienne. Cette distance a été ensuite traduite en degré de possibilité d'apparition en utilisant les fonctions d'appartenance présentées dans le paragraphe 5.2.2.

¹ Dans la présente étude on a utilisé le module GRID™ sous ARC/INFO™ version 7.1 de ESRI™.

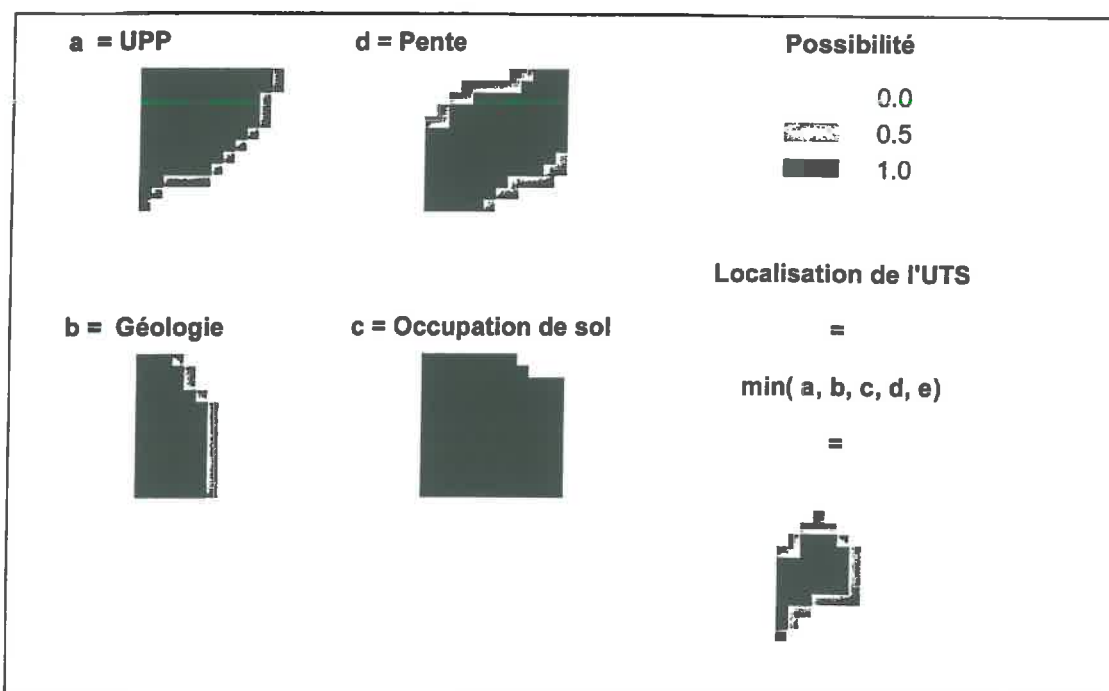


Figure 7.1: Combinaison logique des variables géographiques en vue de localiser les UTS.

c. L'appariement partiel pour chaque variable

L'appariement partiel sur l'ensemble des points défini par la grille régulière produit, pour chaque variable, une carte intermédiaire exprimant la partie de la zone d'étude compatible (en terme de possibilité) avec la classe issue de la connaissance de l'UTS. Une illustration des résultats de l'appariement partiel (i.e. les cartes intermédiaires) pour chaque variable est illustré par les cartes a, b, c, d dans la Figure 7.1.

Les cartes intermédiaires pour les variables géologie, et les unités de pédopaysage sont des cartes floues dues à l'introduction des limites floues comme indiqué ci-dessus. Ces cartes représentent donc la possibilité d'apparition de l'unité géologique et de l'unité de pédopaysage correspondant celles mentionnées dans la description de l'unité de sol. En ce qui concerne l'occupation du sol, cette information géographique est discrète, car les limites ont été considérées comme suffisamment précises à l'échelle considérée. Lorsque la description de l'unité typologique de sol fournit plusieurs unités géologiques ou types d'occupation du sol possibles, la carte intermédiaire est constituée par la combinaison disjonctive (union) des cartes respectives. Ceci est, par exemple, le cas pour la variable "occupation du sol" dans l'exemple du Tableau 5.9.

Pour les données se présentant initialement sous format raster (la variable pente), l'appariement partiel consiste à déterminer, en chaque point de la zone d'étude, la compatibilité (en terme de possibilité) entre la valeur observée en ce point et la distribution de possibi-

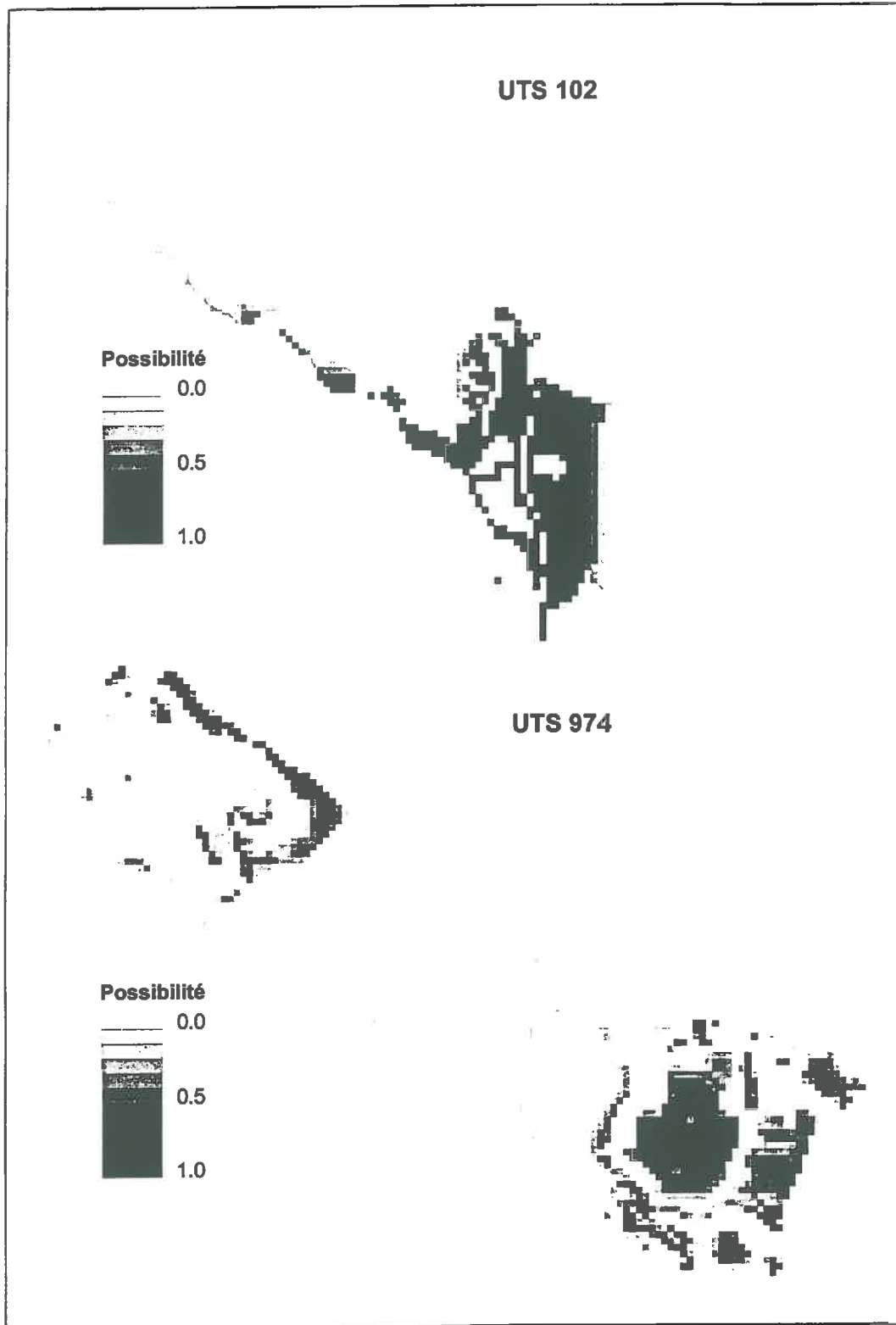


Figure 7.2: Cartes floues exprimant la localisation des UTS 102 et 974 pour le secteur de référence ACM.

lité (par ex. "pente faible") issue de la description de l'UTS. Ce calcul de compatibilité a été illustré par l'exemple donné dans le paragraphe 5.2.3.

d. Combinaison disjonctive des variables

Une fois l'appariement partiel effectué pour chaque variable, on dispose de quatre couches d'information avec lesquelles on détermine la localisation de l'UTS considérée. Dans la présente étude, on évalue la possibilité par la combinaison conjonctive (min) des différentes couches d'information (cf. Equation 7.2). La Figure 7.1 présente, de manière graphique, la combinaison logique des quatre variables géographiques aboutissant à la cartographie, en terme de possibilité d'apparition, d'une UTS. Le résultat est autrement dit une "carte floue". La Figure 7.2 montre deux exemples de cartes floues pour les UTS 102 et 974, dans la zone de teste utilisée dans la présente étude (i.e. le secteur de référence ACM). Pour ces deux exemples, nous avons affecté un degré d'importance (w_i) de 1 à chaque variable.

7.3 Protocole d'évaluation

Pour chaque UTS présente dans la zone d'étude, nous pouvons donc cartographier sa distribution de possibilité d'apparition (i.e. les deux cartes floues de la Figure 7.2). La localisation des UTS, telle qu'elle a été présentée dans le paragraphe précédent, utilise les trois variables géographiques supplémentaires (géologie, occupation de sol et pente) avec un degré d'importance égal pour chaque variable (voir Equation 7.2).

L'évaluation de la démarche proposée suscite deux questions principales auxquelles on va devoir s'efforcer de répondre.

- a) Quel est le gain de précision de localisation des UTS en utilisant toute la connaissance géographique supplémentaire par rapport à l'utilisation de la carte de pédopaysage seule?
- b) S'il existe un gain significatif, est-il ensuite nécessaire d'utiliser la totalité des connaissances pour obtenir une prévision acceptable de localisation?

Pour pouvoir répondre à ces deux questions, il est nécessaire de (i) déterminer un ensemble de données de validation, (ii) d'effectuer la localisation des UTS pour des combinaisons de variables différentes et (iii) d'évaluer les résultats obtenus avec les mesures quantitatives choisies.

Combinaison	Nom	Var1	Var2	Var3	Var4
a	gop	UCS	géologie	occup. sol	pen­te
b	gp	UCS	géologie	pen­te	
c	go	UCS	géologie	occup. sol	
d	op	UCS	occup. sol	pen­te	
e	g	UCS	géologie		
f	o	UCS	occup. sol		
g	p	UCS	pen­te		
h	rien	UCS			

Tableau 7.1: 8 combinaisons de variables géographiques différentes utilisées pour l'évaluation de la localisation des UTS

7.3.1 Données de validation disponibles

La détermination de la qualité de la localisation des UTS nécessite des informations permettant la comparaison entre la localisation estimée et la localisation réelle issue d'une carte de référence. La carte pédologique au 1/100.000 (feuille de Lodève) de Bonfils (1993) a été choisie comme carte de référence (cf. § 4.2.5.3). Cette carte, qui couvre seulement une partie de la zone VHOL (cf. Figure 4.4), contient la localisation exacte des UTS, c'est-à-dire, que les UCS sont constituées d'une seule UTS. On a pu distinguer 13 UTS pour lesquelles on connaît la correspondance entre les UTS de la carte de pédopaysages au 1/250.000 et celles de la carte au 1/100.000. Ces 13 UTS constituent l'ensemble des données de validation.

7.3.2 Série d'expérimentations à effectuer

La localisation des UTS telle qu'elle a été présentée dans le paragraphe précédent utilise les cinq variables géographiques avec un degré d'importance égal pour chaque variable (voir Equation 7.2). Il est cependant possible d'affecter des degrés d'importance différents à chaque variable. En appliquant simplement des degrés d'importance 0 et 1, on peut créer une série de combinaisons de variables intervenant dans la localisation des UTS. Le Tableau 7.2 résume ces 8 combinaisons possibles.

Nous avons donc effectué la localisation de toutes les UTS définies dans la zone VHOL (96 au total) pour ces 8 combinaisons de variables différentes (i.e. $8 * 96 = 768$ cartes floues au total).

7.3.3 Trois mesures quantitatives pour évaluer la qualité de la localisation des UTS

La méthode la plus courante d'évaluation de la qualité des informations géographiques en mode raster (cartes thématiques issues d'images satellites ou cartes pédologiques en mode raster) avec une information de référence est basée sur la détermination de la matrice d'erreur ou matrice de confusion (Lillesand & Kiefer, 1994; Legros, 1996). La matrice de confusion permet de calculer un certain nombre de mesures d'évaluation comme l'erreur par excès (Commission Error), l'erreur par défaut (Omission Error) et le coefficient de Kappa exprimant dans quelle mesure la distribution spatiale de l'information testée n'est pas uniquement due au hasard (Lillesand & Kiefer, 1994).

Ces méthodes traditionnelles d'évaluation ont une limitation principale qui veut que chaque surface sur la carte (pixel ou polygone) a été identifiée par une seule valeur (la somme des totaux des lignes de la matrice de confusion est ainsi égale à la somme des totaux des colonnes). Dans la présente étude, ceci n'est donc pas le cas, car plusieurs UTS peuvent être plus ou moins possibles sur un point d'intérêt. Une méthode d'évaluation alternative permettant de déterminer des mesures de qualité des cartes dans un cadre théorique des sous-ensembles flous a été présentée par Gopal & Woodcock (1994). Ils proposent une méthode basée sur l'évaluation de la qualité à l'aide d'une échelle qualitative. Cette échelle de 5 classes peut ensuite être utilisée par un expert pour déterminer l'appariement entre la valeur issue de la carte et la valeur de la référence. Ils utilisent ensuite des opérations floues pour déterminer des mesures de confusion et d'ambiguïté. Les inconvénients de cette méthode résident dans l'interprétation qualitative de la qualité de la carte par l'expert. De plus, la détermination de la qualité par l'expert sur un grand nombre de points (une couverture en mode raster) n'est pas envisageable.

Nous proposons d'évaluer la qualité de la localisation des UTS en se basant sur trois critères: un critère exprimant la proportion des pixels bien classés; ce qui correspond à la précision-produit (Legros, 1996), un critère exprimant l'imprécision de la localisation et un critère de compatibilité totale. Nous présentons ces critères en détail dans les paragraphes suivants.

7.3.3.1 *Proportion de pixels bien classés*

Pour définir le premier critère d'évaluation, nous nous intéressons aux situations où la localisation de l'UTS estimée correspond bien à la localisation donnée par la carte au 1:100.000. Considérons le cas illustré par la Figure 7.3. La surface Ao_i est la localisation de l'UTS i issue de la carte au 1:100.000. Les ellipses concentriques $Ae_{i0.1}$, $Ae_{i0.3}$, $Ae_{i0.6}$ et $Ae_{i1.0}$, correspondent à la localisation modélisée de la même UTS pour les différentes α -coupes.

L'intersection des surfaces observées avec les surfaces estimées donne maintenant toutes les situations où l'UTS a été correctement prédite. Les surfaces correspondant aux intersections ont été appelées $Ae'_{i0.1}$, $Ae'_{i0.3}$, $Ae'_{i0.6}$ et $Ae'_{i1.0}$.

On définit alors le critère d'évaluation suivant: quelle est la proportion de la surface de l'unité de référence pour laquelle on estime correctement l'UTS? Pour une α -coupe considérée, cette Proportion des Estimations Correctes (PEC) est calculée pour chaque UTS de la manière suivante:

$$PEC_{i\alpha} = A_{o_{i\alpha}} / A_{e_{i\alpha}} \quad (7.3)$$

où $A_{o_{i\alpha}}$ exprime la surface réelle observée pour l'UTS i et où $A_{e'_{i\alpha}}$ exprime la surface estimée au sein de l'UTS de référence (i.e l'intersection de $A_{o_{i\alpha}}$ et $A_{e_{i\alpha}}$). La moyenne arithmétique avec les valeurs de PEC_{α} des 13 UTS permet d'obtenir la valeur de PEC_{α} pour l'ensemble des 13 UTS de validation.

Remarque: il est possible de définir la surface en nombre de pixels en sachant que chaque point représente 2500m².

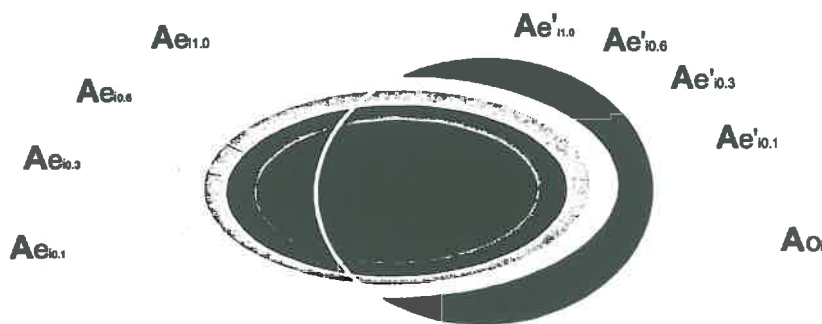


Figure 7.3: Comparaison entre la localisation estimée (par différentes α -coupes) et la localisation observée (information non-floue).

7.3.3.2 Imprécision de la localisation des UTS

La mesure précédente établit l'adéquation de la prédiction avec la réalité. Cependant ceci doit être complété par une mesure de précision de la prédiction. En effet, il serait absurde de considérer comme une prédiction efficace celle qui consisterait à prédire en tout point de l'espace toutes les unités possibles, et ce, bien que cette prédiction produise 0% d'erreur.

Nous envisageons donc comme mesure d'imprécision le nombre moyen d'UTS prédit. Ce nombre est calculé après avoir sélectionné sur chaque pixel les UTS "possibles" à une α -

coupe fixée (c'est-à-dire, dont la possibilité est supérieure à α). Le nombre d'UTS ainsi sélectionné par pixel est moyenné sur l'ensemble de la zone pour obtenir le critère recherché pour une α -coupe donnée. Le nombre d'UTS moyen est maximal lorsque aucune information géographique autre que la carte des UCS n'est introduite. Il correspond alors à la moyenne du nombre d'UTS prédite par UCS, pondéré par la surface des UCS.

7.3.3.3 *Cohérence de l'estimation*

Dans le cas où aucune UTS prédite n'atteint une possibilité égale à 1, la distribution de possibilité d'UTS n'est pas normalisée. On considère alors qu'il existe une incohérence dans la connaissance disponible. Puisque, dans ce cas, la combinaison des valeurs des variables géographiques caractérisant le pixel n'est compatible (au niveau de possibilité 1) avec aucune des UTS de l'UCS contenant ce pixel. En d'autres termes, l'ensemble des descriptions de l'environnement des UTS ne couvrent pas l'ensemble des situations géographiques qu'il est possible de rencontrer dans une UCS. Il s'agit donc d'une incohérence des données que l'approche utilisée permet de détecter. L'importance de ce problème est évaluée par le pourcentage de points pour lesquels au moins une UTS atteint la possibilité de 1 (= la pourcentage complètement compatible: PCC).

Pour les points où aucune UTS atteint la possibilité de 1, il ne sera donc pas possible d'estimer la réserve utile ou θ_{pf} . Dans la suite de cette étude, ces zones seront indiquées comme "situations non-renseignées" lorsque l'on présentera les résultats de la cartographie des propriétés hydriques (cf. chapitre 8).

7.4 Résultats obtenus pour les trois mesures d'évaluation

Les résultats des expérimentations permettant d'évaluer la localisation des UTS au sein des UCS sont présentés dans les Figures 7.4, 7.5, 7.6. Pour ces trois graphiques, nous avons représenté les différentes combinaisons de variables sur l'axe X et la mesure d'évaluation sur l'axe Y. Chaque graphique sera analysé indépendamment dans les sous-sections suivantes.

7.4.1 Proportion d'estimations correctes

La Figure 7.4 montre les résultats des Proportions des Estimations Correctes (PEC) pour les différentes combinaisons de variables géographiques. Ils ont été rangés selon un

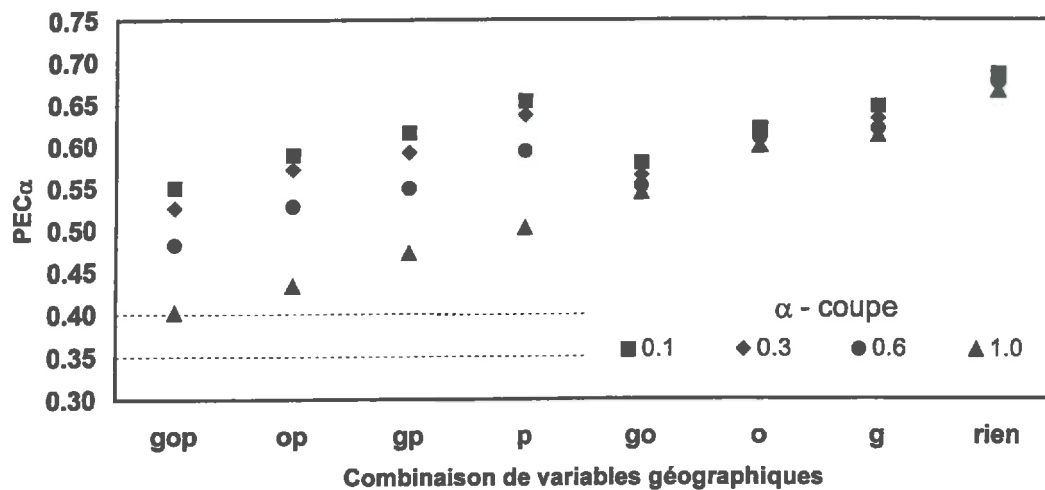


Figure 7.4: Proportion des points où l'UTS est correctement prédite pour les différentes combinaisons de variables et quatre α -coupes différentes.

ordre croissant de la valeur de PEC pour l' α -coupe 1. Cette présentation des résultats nous permet de faire un certain nombre d'observations:

- On constate en général une diminution de proportion d'estimation correcte avec le nombre de couches d'information ajouté. Le taux d'estimation le plus élevé est obtenu lorsque aucune variable supplémentaire n'a été introduite pour prédire les UTS (**rien** sur le graphique) alors que la situation où l'ensemble des trois variables supplémentaires a été introduit produit les plus faibles taux d'estimation. Ces résultats étaient prévisibles puisque le recours à une nouvelle couche d'information introduit une nouvelle erreur associée qui se cumule avec les précédentes suivant la loi de propagation dans les combinaisons logiques (Burrough & McDonnell, 1998). Il est à remarquer que l'utilisation seule de la variable "pente" produit des valeurs de PEC plus mauvaises pour α -coupe 1.0 que la combinaison de "géologie" et "occupation de sol"
- On observe également que les valeurs les plus faibles de taux d'estimation sont obtenues avec les combinaisons de variables utilisant la pente (partie gauche de la graphique) au niveau d' α -coupe 1.
- Uniquement lorsque la variable pente est impliquée dans les combinaisons des variables prédictives, les taux de bonnes estimations varient fortement avec le choix de l' α -coupe. Ceci est le résultat du choix de largeur des marges autour du noyau définissant cette variable. Cette largeur se traduit géographiquement par une zone floue beaucoup plus importante que pour les deux autres variables ce qui influence ainsi fortement les prédictions d'UTS.

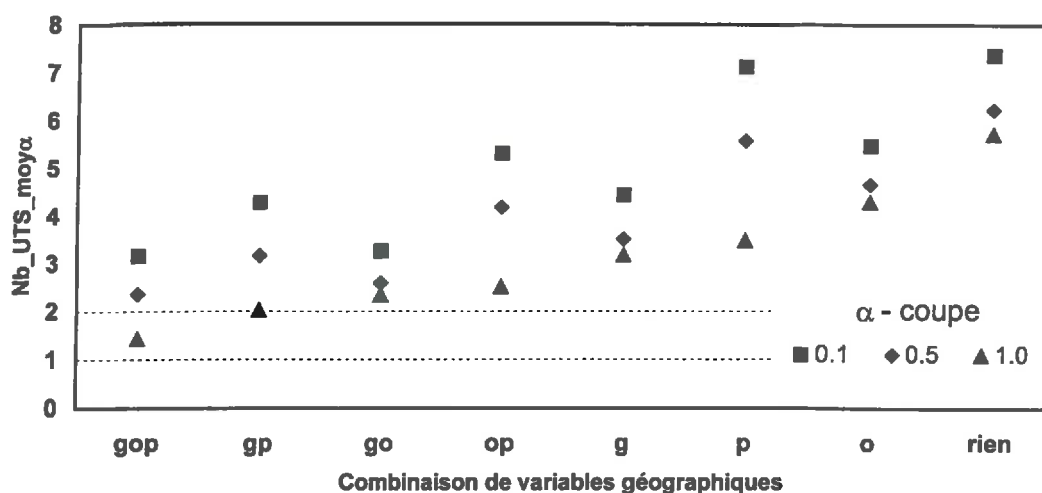


Figure 7.5: Nombre d'UTS moyen estimé sur l'ensemble des points de la zone de validation pour chaque combinaison de variables géographiques et pour trois α -coupes différentes.

7.4.2. Evaluation de l'imprécision sur la localisation

La Figure 7.5 montre le nombre d'UTS moyen issu des localisations par les différentes combinaisons pour trois α -coupes. Comme pour la Figure 7.4, les résultats ont été rangés par ordre croissant du nombre d'UTS pour α -coupe 1. Cette présentation nous permet de constater les choses suivantes:

- On constate en général une diminution du nombre d'UTS moyen avec le nombre de couches d'information ajouté. Les cas extrêmes étant la situations où aucune variable supplémentaire a été introduite (**rien**) et la situation où l'ensemble des trois variables supplémentaires ont été introduites (**gop**). Ce résultat correspond également à ce qu'on attendait: plus des variables sont ajoutées, plus la prédiction d'UTS est sélective et moins d'UTS sont prédites comme possibles.
- Lorsqu'on examine séparément les combinaisons de variables impliquant le même nombre de couches d'information (1 ou 2), il se dégage clairement une hiérarchie entre les variables. La variable "géologie" semble la plus discriminante alors que l'occupation du sol joue un rôle moins important, la variable pente ayant une position intermédiaire.
- La différence entre les nombres prédits pour les différentes α -coupes est plus importante lorsque la variable pente participe à la localisation des UTS; ce que nous avons constaté et justifié dans le paragraphe précédent. On constate également que cette différence diminue lorsque le nombre de variables augmente: par exemple pour les combinaisons **p**, **op**, **gp** et **gop**).

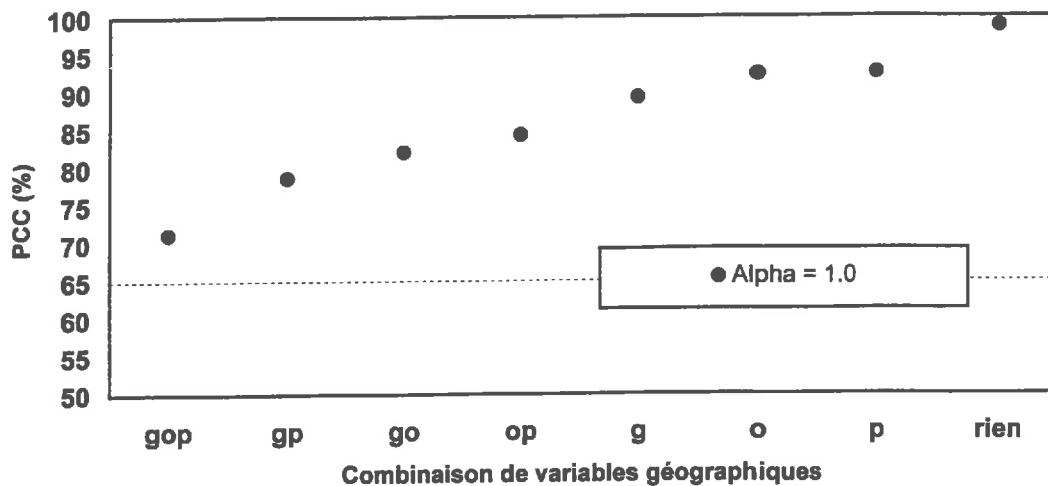


Figure 7.6: Proportion de la zone d'évaluation où au moins une UTS atteint la possibilité 1 (PCC = pourcentage complètement compatible).

7.4.3 L'évaluation de la cohérence de l'information pour chaque combinaison.

La Figure 7.6 expose le pourcentage de points pour lesquels au moins une UTS est trouvée complètement possible, pour chaque combinaison de variables (pourcentage complètement compatible: PCC). Les résultats sont rangés dans l'ordre croissant. L'ordre obtenu est presque identique à celui présenté dans la Figure 6.5. La seule différence concerne l'ordre des variables observé lorsqu'on examine séparément les combinaisons à une couche: la variable "occupation du sol" produit moins d'incohérences que la variable "pente".

Le pourcentage complètement compatible atteint 100% pour la combinaison p et 67% pour la combinaison a. Entre ces deux extrêmes on observe une répartition de valeurs entre 70 et 95 %. La tendance générale montre une augmentation du pourcentage de points où au moins une UTS est complètement possible en diminuant le nombre de variables dans la combinaison.

7.4.4 Evaluation globale en combinant les critères d'évaluation

Les paragraphes précédents ont présenté les résultats obtenus pour les trois critères d'évaluation. Nous avons pu constater que la combinaison de toutes les variables d'environnement (combinaison gop) produit la localisation des UTS la plus précise, mais elle est perdante en terme PEC et de PCC. En revanche, la localisation des UTS sans variables d'environnement est très imprécise (i.e. nombre d'UTS moyen élevé), mais elle est, comme nous l'attendions, plus performante en ce qui concerne la PEC et la PCC. En résumant, les

Comb.	Σ Rangs	Comb.	Σ Rangs
gop	9	og	7
op	11	o	10
gp	8	g	7
p	11	rien	9

Tableau 7.2: Somme des rangs pour deux critères ($PEC_{1,0}$ et $NB_UTS_moy_{1,0}$) pour les huit combinaisons de variables évaluées.

deux cas extrêmes (utiliser tout connaissance ou rien) se trouvent toujours aux extrémités des ordres des Figures 7.4, 7.5 et 7.6.

Les résultats des autres combinaisons possibles se trouvent ainsi entre ces deux cas extrêmes. Il est intéressant de déterminer s'il existe des combinaisons de variables qui sont performantes pour les trois critères. Nous proposons d'analyser les différences des combinaisons intermédiaires avec plus de détail.

Tout d'abord, nous avons constaté que les résultats pour le deuxième et troisième critère (Nb_UTS_moy et PCC) montrent une image approximativement identique pour la série d'expérimentations. C'est pour cette raison que nous avons décidé d'éliminer le critère PCC pour l'analyse plus approfondie permettant de mieux comparer la localisation obtenue par les différentes combinaisons de variables. Nous proposons une analyse basée sur l'affectation des rangs aux combinaisons de variables et l'addition des rangs.

On affecte à chaque combinaison de variables la valeur de son rang dans le classement par ordre décroissant de $PEC_{1,0}$, puis la valeur de son rang dans le classement par ordre croissant de $Nb_UTS_{1,0}$. On additionne ensuite les deux valeurs de rang obtenues. Ces valeurs ont été représentées dans le Tableau 7.2

Nous pouvons ensuite constater que les sommes de rangs sont les plus basses pour les combinaisons **og**, et **g**. Ces combinaisons ne contiennent pas la variable "pente" et elles contiennent toutes la variable "géologie". Les plus grandes sommes de rangs sont obtenues pour les combinaisons **op** et **p**. Ces combinaisons de variables ne contiennent pas la variable géologie et, en revanche, ces combinaisons contiennent toutes deux la variable "pente".

Ces observations amènent à proposer un certain nombre de conclusions. Premièrement, l'introduction de l'information géologique s'avère essentielle pour pouvoir localiser les UTS avec une qualité acceptable en terme de précision (NB_UTS_moy) et en terme de proportion estimée correctement (PEC). L'introduction de la variable "pente" provoque généralement une qualité moindre des localisations obtenues. On peut alors dire que la localisation des UTS est sensible à la combinaison de variable utilisée, dans laquelle la variable "géologie" joue un rôle positif et la variable "pente" un rôle négatif. La localisation des UTS s'avère moins sensible à l'absence ou à la présence de la variable "occupation du sol" dans la combi-

raison de variables. Cette variable est peu discriminante dans notre étude. Ceci s'explique par l'uniformité de l'occupation du sol dans la zone d'étude (i.e. vigne).

Les résultats obtenus pour les deux combinaisons "optimales" sont présentés dans le Tableau 7.3 pour α -coupe 0.1 et 1.

Combinaison	PEC0.1	PEC1.0	NBUTS0.1	NBUTS1.0	PCC1.0
g	0.646	0.611	4.44	3.18	89
og	0.579	0.544	3.26	2.36	82

Tableau 7.3: Les résultats obtenus pour les deux combinaisons "optimales" (cf. Tableau 7.2)

7.5 Conclusion et discussion sur la localisation des UTS

La démarche suivie pour la localisation des UTS au sein des UCS s'appuie sur la théorie des possibilités. L'utilisation de l'information géographique supplémentaire permet de déterminer plus précisément la localisation de l'UTS par rapport à celle fournie par la carte des pédopaysages (les UCS sont des unités complexes contenant un ensemble d'UTS possibles). L'intérêt principal de la démarche est donc la réduction du nombre d'unités typologiques du sol potentiellement possibles en un point d'intérêt donné, ce qui réduira également l'imprécision des estimations des propriétés hydriques sur la région.

La démarche proposée a été évaluée en introduisant chaque fois un jeu de variables différent dans le modèle. Nous avons pu comparer les différents résultats obtenus à l'aide de trois critères distincts: un critère exprimant la proportion estimée correctement, un critère d'imprécision exprimant le nombre d'UTS candidate moyen et le pourcentage de la zone où au moins une UTS a été trouvée complètement possible.

L'interprétation des résultats pour chaque critère fournit certains enseignements majeurs. De façon générale, en introduisant les variables géographiques nous gagnons en terme de précision, mais ce gain est accompagné (i) d'une diminution de la proportion des estimations correctes (PEC) et (ii) d'une diminution de la surface pour laquelle au moins une UTS est complètement possible. L'utilisation de toutes les variables dans le modèle fournit la localisation la plus précise (NB_UTS_moy le plus bas) et donc les valeurs pour PEC et NB_UTS_moy les plus basses.

Nous avons donc démontré l'intérêt d'introduire une connaissance géographique supplémentaire pour la localisation des UTS au sein des UCS, même si cette information ne permet pas de localiser avec certitude chaque UTS. Des marges de progrès sont à rechercher vers une description plus précise des UTS dans les bases de données en ce qui concerne leur variables d'environnement. Par ailleurs, une description des relations de voisinage existant entre les UTS d'une même UCS pourrait également être mise en œuvre. Les blocs diagrammes illustrant la distribution spatiale des UTS sont un premier essai en ce sens. Traduire un

bloc diagramme sous forme de relations spatiales telle celles décrites par les lois de voisinage (Lagacherie *et al.*, 1995) est à envisager pour l'avenir.

8 Cartographie des propriétés hydriques

Dans les chapitres précédents, nous avons présenté les approches permettant (i) d'estimer pour chaque unité typologique de sol ses propriétés hydriques, (ii) de localiser ses unités dans l'espace. Afin d'obtenir une cartographie des propriétés hydriques sur une région, il est nécessaire d'intégrer ces deux résultats. Cette intégration doit être conçue dans le cadre de la théorie des possibilités qui est le cadre théorique choisi dans cette étude.

Dans ce chapitre nous présentons la méthode d'intégration retenue, nous évaluons les estimations des propriétés hydriques et nous proposons des méthodes permettant de cartographier l'information complexe que constitue les distributions de possibilité.

8.1 Estimation spatiale des propriétés hydriques

Comme nous l'avons présenté auparavant, l'estimation spatiale des propriétés hydriques nécessite l'intégration des résultats issus de (i) l'estimation des propriétés hydriques par UTS (cf. chapitre 6) et (ii) la cartographie des UTS (cf. chapitre 7). Dans un premier temps, nous présentons brièvement, les aspects méthodologiques de la démarche suivie pour cette estimation spatiale. Dans un deuxième temps, nous proposons une évaluation succincte permettant de comparer les résultats obtenus par le modèle avec les mesures de terrain (ceci en raison de la disponibilité des données seulement pour θ_{pf}).

8.1.1 Intégration de la première et deuxième étape par combinaison disjonctive

Considérons, en un point, différentes UTS possibles. Pour chaque UTS, nous avons donc une distribution de possibilité de la propriété hydrique du sol. Ces différentes distributions de possibilité de la propriété hydrique du sol sont agrégées par une combinaison disjonctive définie par:

$$\forall u, \pi_{PHS}(u) = \max_{i=1,s} \min(\pi_{PHSi}(u), \Pi_{UTSi}) \quad (8.1)$$

où π_{PHSi} est la distribution de possibilité de la propriété hydrique du sol en faisant l'hypothèse de l'appartenance du point au type de sol $UTSi$, Π_{UTSi} est la possibilité que le point appartienne à $UTSi$ et s est le nombre d'unités typologiques de sols envisageables pour le point. Cette combinaison réalise l'union des intervalles flous liés à chaque type de sol possible en les tronquant selon les degrés de possibilité d'appartenance à l'unité typologique correspondante. Ceci est illustré par la Figure 8.1. Dans le cas où aucune UTS n'est complètement possible (cf. § 7.3.3.3), la distribution obtenue ne sera pas normalisée (cf. § 3.3.4).

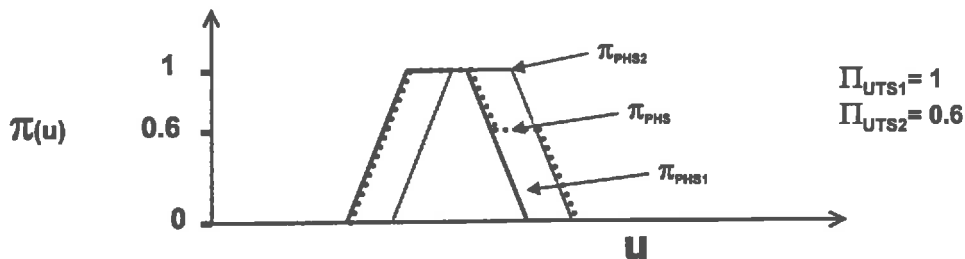


Figure 8.1 : Combinaison disjonctive de deux sources d'information.

Cette combinaison disjonctive a été appliquée sur chaque pixel de la zone d'étude afin d'obtenir une cartographie de la teneur en eau au point de flétrissement et la réserve utile.

8.1.2 Protocole d'évaluation des estimations spatiales

L'évaluation de la qualité des estimations des propriétés hydriques est réalisée en comparant les distributions de possibilité obtenus pour un ensemble de points avec les mesures de terrain. Un ensemble de données de terrain suffisamment grand est nécessaire pour pouvoir évaluer les résultats du modèle de manière satisfaisante. Les données de validation choisies pour cette partie sont les données de validation (série A) prises à l'intérieur du secteur de référence ACM (voir la Figure 4.3). Il s'agit donc des données de validation de θ_{pf} issues des échantillons pris à 40 cm de profondeur (voir l'Annexe X). Les mesures de terrain de RU ne sont pas suffisamment nombreuses pour l'analyse qu'on propose d'effectuer.

On dispose d'un ensemble de 246 points pour lesquels on possède d'un côté l'estimation sous la forme d'une distribution de possibilité et, d'un autre côté, de la mesure de terrain. Deux mesures d'évaluation ont ensuite été calculées. La première mesure exprime l'imprécision des estimations et elle est définie par la largeur moyenne de l'intervalle estimé par α -coupe. La deuxième mesure exprime l'erreur de l'estimation qui est définie par le pourcentage des mesures de terrain dont la valeur n'était pas incluse dans l'intervalle estimé (ceci également par α -coupe).

Les mesures d'évaluation ont été calculées pour chaque α -coupe sur l'ensemble de points où un intervalle était disponible, c'est-à-dire, qu'au moins une UTS doit être possible pour ce niveau de coupe α .

8.1.3 Résultats de l'évaluation des estimations spatiales

Les résultats obtenus pour les deux mesures d'évaluation sont présentés dans les Figures 8.2 et 8.3. En ce qui concerne le pourcentage des points de validation dont la valeur de θ_{pf} se situe à l'extérieur de l'intervalle estimé (Figure 8.2), on constate que ce pourcentage est très faible (1 à 2 %) pour les faibles niveaux de α . Il augmente graduellement en augmentant le niveau de α , jusqu'à un pourcentage d'environ 10 % pour α -coupe 1.0. Ces résultats montrent donc que l'intervalle estimé englobe pour 98 à 99 cas sur 100 la réalité pour α -coupe 0.1 et l'intervalle estimé englobe 9 cas sur 10 la réalité pour α -coupe 1.

Avant de pouvoir affirmer que ces résultats sont satisfaisants ou non, il est nécessaire d'analyser également l'imprécision des résultats. Ceci est présenté dans la Figure 8.3. Cette figure montre une largeur moyenne considérable de l'intervalle estimée pour θ_{pf} sur l'ensemble des points de validation (elle varie entre 17.5 g/100g pour α -coupe 0.1 et 12 g/100g pour α -coupe 1). On constate donc que les estimations sont très imprécises. Il n'est alors pas surprenant de trouver des faibles pourcentages de points de validation dont la valeur de θ_{pf} se situe à l'extérieur de l'intervalle estimé (Figure 8.2).

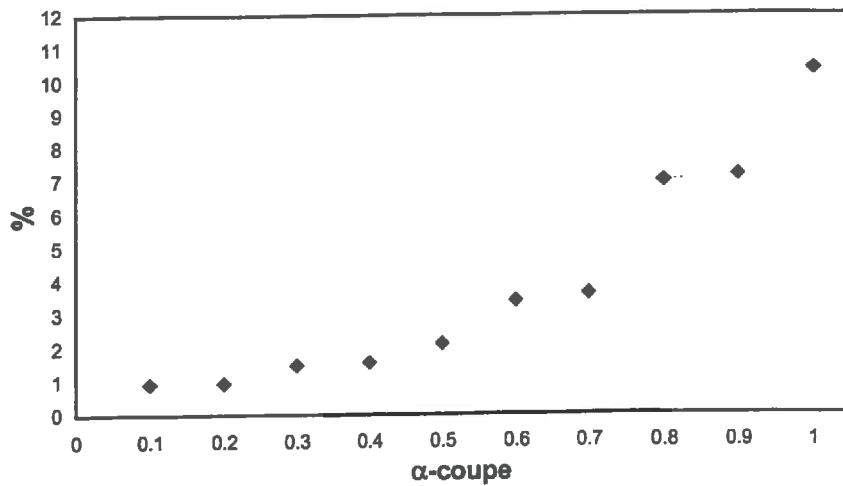


Figure 8.2: Pourcentage des points de validation dont la valeur de θ_{pf} se situe à l'extérieur de l'intervalle estimé.

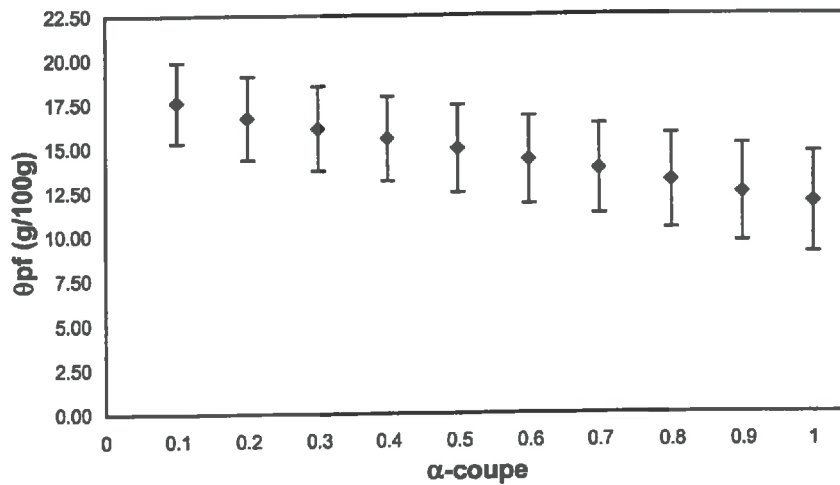


Figure 8.3: Largeur moyenne (plus ou moins l'écart type) de l'intervalle estimée pour θ_{pf} sur l'ensemble des points de validation.

En conclusion, il est possible de dire que l'estimation spatiale est imprécise mais relativement peu informative. Elle englobe, pour une grande part, les données issues de mesures de terrain. Cependant, il est à remarquer que, même si elle est imprécise, l'estimation spatiale n'englobe pas la totalité (100%) des données mesurées (ceci est surtout le cas pour les estimations d' α -coupe 1).

8.2 Représentation cartographique de l'estimation spatiale

Le résultat obtenu en chaque point se présente donc sous la forme d'une distribution de possibilité comme présenté dans la Figure 8.1. Nous cherchons ensuite à représenter cette information disponible en chaque point sous la forme d'une carte pour pouvoir communiquer les résultats aux utilisateurs ou aux décideurs. Il est alors essentiel de pouvoir communiquer (i) l'information primaire et (ii) l'information concernant son incertitude (Kraak *et al.*, 1995). Ceci implique pour notre étude de pouvoir représenter, au mieux possible, la totalité de l'information exprimée par la distribution de possibilité obtenue.

Hootsmans (1997) a analysé différentes méthodes pour la représentation de l'incertitude attachée à l'information primaire lorsque l'incertitude est traitée dans le cadre de la théorie des sous-ensembles flous. Il propose des représentations cartographiques en combinant plusieurs variables graphiques (couleur, saturation des couleurs, degré de grisé, etc.). Certaines combinaisons de variables comme, par exemple, la couleur et la luminosité s'avèrent très utiles pour la représentation du degré d'appartenance à une classe: une couleur pour la classe et la luminosité pour le degré d'appartenance à cette classe. De Gruijter *et al.* (1997) présentent également une représentation floue des résultats issus d'une classification floue des unités de sol en utilisant une technique basée sur la mélange des couleurs associées à chaque unité de sol. Ceci permet de visualiser la transition graduelle d'une unité de sol à une autre.

Hootsmans (1995) met également en évidence que la création d'une série de cartes peut être utile dans un certain nombre de cas. En utilisant des outils informatiques, il est possible de créer des cartes interactives. Ceci permet, par exemple, de visualiser rapidement différentes cartes pour différentes α -coupes.

Pour la présente étude, nous proposons de se baser sur deux types de représentation. Le premier utilise le principe de la détermination de l'estimation par rapport à une valeur seuil. Pour ceci, on utilise la couleur et la luminosité pour indiquer les différents degrés d'incertitude. Le deuxième type de représentation cherche à représenter l'intervalle entier pour une α -coupe donnée.

8.2.1 Détermination de la position des résultats par rapport à une valeur seuil

Les décideurs ou utilisateurs des cartes ne sont pas forcément intéressés par la valeur cardinale de la réserve utile ou la teneur en eau au point de flétrissement, mais plutôt par la position de cette valeur relativement à un seuil porteur de signification pour l'application considérée. Pour la réserve utile, ce qui compte, par exemple, est de savoir si la valeur en question est inférieure ou supérieure au seuil correspondant à la limite acceptable pour pra-

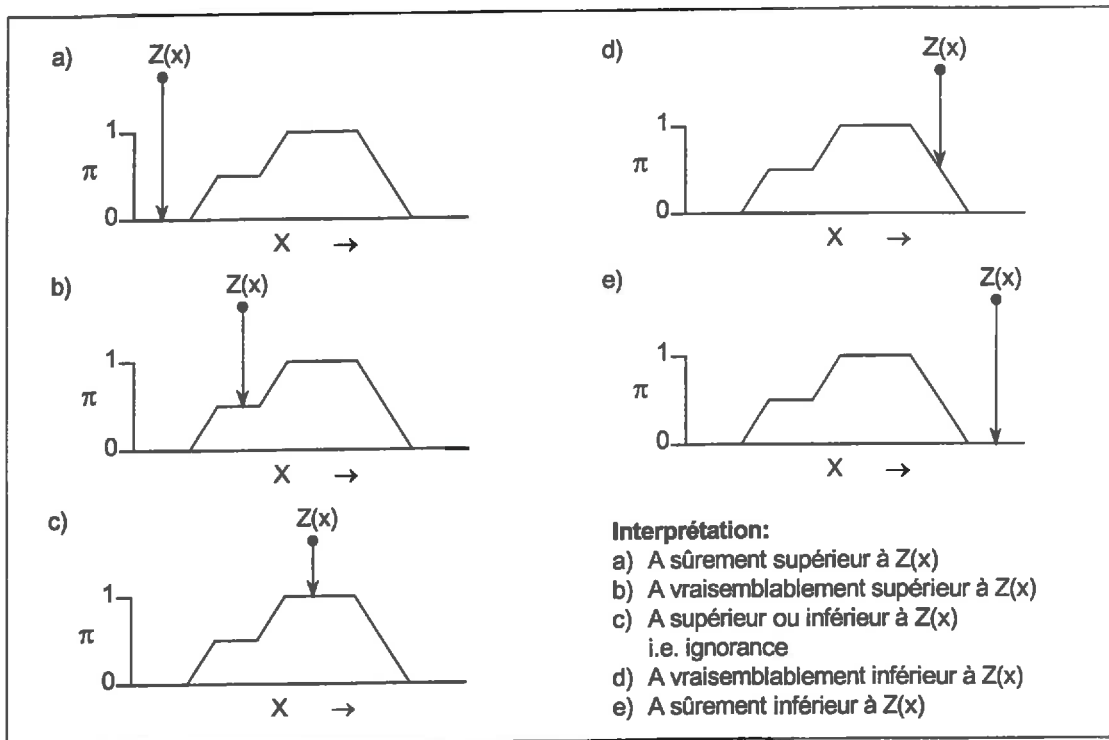


Figure 8.4: Détermination de la position des résultats obtenus par rapport à une valeur seuil.

tiquer certaines cultures. Pour répondre à la question pratique dans cette application, on peut donc se contenter de produire une carte répondant à la requête: est-ce que la variable (i.e. RU ou θ_{pf}) est supérieure ou inférieure au seuil $Z(x)$? L'utilisateur peut rapidement faire, à l'aide d'un SIG, différentes cartes en faisant varier le seuil $Z(x)$ et utiliser des couleurs pour visualiser des niveaux d'incertitude qui sont induits par la position du seuil par rapport à l'estimation obtenue. L'introduction de cette question pratique est donc un moyen permettant d'exprimer l'incertitude sur la connaissance.

En faisant varier le seuil, on peut distinguer 5 situations différentes. Ces cinq situations sont exposées dans la Figure 8.4 où le résultat obtenu pour la variable V en un point dans la zone d'étude est défini par la distribution de possibilité A . La valeur seuil est définie par $Z(x)$. La première situation possible est le cas 8.4a. Dans ce cas, on peut dire que V est sûrement $>$ à $Z(x)$; ce qui signifie que la borne inférieure du support de la distribution de V est supérieure à $Z(x)$. Pour la deuxième situation (la situation 8.4b), on peut dire que V est vraisemblablement $>$ à $Z(x)$; ce qui signifie que la borne inférieure du noyau de la distribution est supérieure à $Z(x)$, mais que la borne inférieure du support ne l'est pas. Dans le cas de la situation 8.4c nous pouvons dire qu'il est autant possible que V soit supérieure à $Z(x)$ ou qu'elle soit inférieure (le noyau contient la valeur seuil). Par rapport à la question que l'on pose, on est mathématiquement incapable de dire que l'on est supérieur ou inférieur à la valeur seuil choisie ($\Pi(A) = \Pi(Ac) = 1$). On est donc typiquement dans un cas d'ignorance (cf. § 3.4.1).

Une sixième alternative est nécessaire à introduire qui correspond aux situations où aucune UTS a été trouvée complètement possible. Cette situation a été appelée "non-renseignée" dans les cartes obtenues.

Les Figures 8.5 et 8.6 montrent 4 exemples de représentation cartographique par rapport à une valeur seuil. La Figure 8.5 montre la détermination de θ_{pf} pour la zone de test ACM par rapport aux valeurs seuil 7.5 et 17.5 g/100g. La Figure 8.6 montre le résultat cartographique obtenu pour la RU en utilisant deux valeurs seuil 150 et 300. Les valeurs seuil ont été choisies pour obtenir des contrastes entre les cartes. L'Annexe XI montre la cartographie de la réserve utile pour la zone d'étude VHOL, en fixant quatre valeurs seuil différentes.

L'ensemble des cartes des Figures 8.5, 8.6 et l'Annexe XI permet de conclure sur deux idées principales. Premièrement, la réponse à la question posée est fortement influencée par la valeur seuil fixée. Le nombre de cas d'ignorance (nombre pixel sur la carte) varie ainsi au cas par cas. Donc, même si les estimations spatiales ont été considérées comme imprécises et généralement peu informatives (cf. le paragraphe précédent), en fixant une valeur seuil, il est possible de répondre avec différents degrés d'incertitude.

Deuxièmement, les cartes permettent d'identifier où un manque d'information a été constaté (i.e. les zones indiquées comme situations 'non-renseignés'). Une distribution de possibilité normalisée n'a pas pu être déterminée, c'est-à-dire une estimation adéquate de la propriété hydrique du sol cherchée n'est donc pas possible. Cette information peut donc être utilisée pour définir des zones où des études pédologiques supplémentaires sont nécessaires.

8.2.2 Représentation de l'intervalle par α -coupe dans un seul pixel

Une alternative à la représentation des résultats par rapport à une valeur seuil a été dérivée d'un type de représentation cartographique proposée par Pebesma & De Kwaadsteniet (1997). Ils proposent la représentation d'un intervalle de confiance à l'aide d'un seul pixel en divisant ce pixel en deux triangles. Ils affectent ensuite les valeurs des bornes supérieures et inférieures aux triangles. Pour les distributions de possibilité obtenues dans la présente étude, une telle approche est également possible pour les bornes inférieures et supérieures par α -coupe. Ceci est illustré par la Figure 8.7 pour les estimations de RU dans la zone de test ACM pour les α -coupes 0.1 et 1.0.

Une telle carte est peut-être plus difficile à interpréter que les cartes présentées dans le paragraphe précédent, mais elle est plus riche en information. La représentation de l'information par rapport à une valeur seuil contient déjà une interprétation par rapport à une question précise. L'alternative présentée ici est une représentation de l'information brute et l'utilisateur ou le décideur doivent effectuer leur propre interprétation. Le premier type de représentation est alors plus adapté que le deuxième, car les décideurs ne sont pas forcément

familiarisés avec les notions théoriques issues de la théorie de possibilité (comme par exemple les α -coupes).

8.3 Conclusion

Ce chapitre a présenté la cartographie de la réserve utile et de θ_{pf} . Cette cartographie est le résultat de l'intégration (par combinaison disjonctive) de l'estimation de la propriété hydrique par UTS (cf. chapitre 6) et la cartographie des UTS (cf. chapitre 7). Nous avons pu constater que les estimations ainsi obtenues sont imprécises, mais elles englobent la plupart (>90%) des mesures de terrain, en ce qui concerne les estimations spatiales de θ_{pf} .

Pour la représentation cartographique des estimations spatiales, nous avons proposé deux alternatives. La première permet d'évaluer l'incertitude sur la connaissance en déterminant la position de l'estimation par rapport à une valeur seuil. Les cartes ainsi obtenues montrent bien qu'il existe des situations différentes en fonction de la valeur seuil. De plus, ce type de représentation permet d'identifier les lieux où une étude complémentaire est nécessaire pour espérer une réponse plus informative. La représentation des résultats en découpant les pixels en deux triangles est considérée comme une alternative satisfaisante pour la représentation de l'information brute, si une question pratique précise n'a pas été formulée.

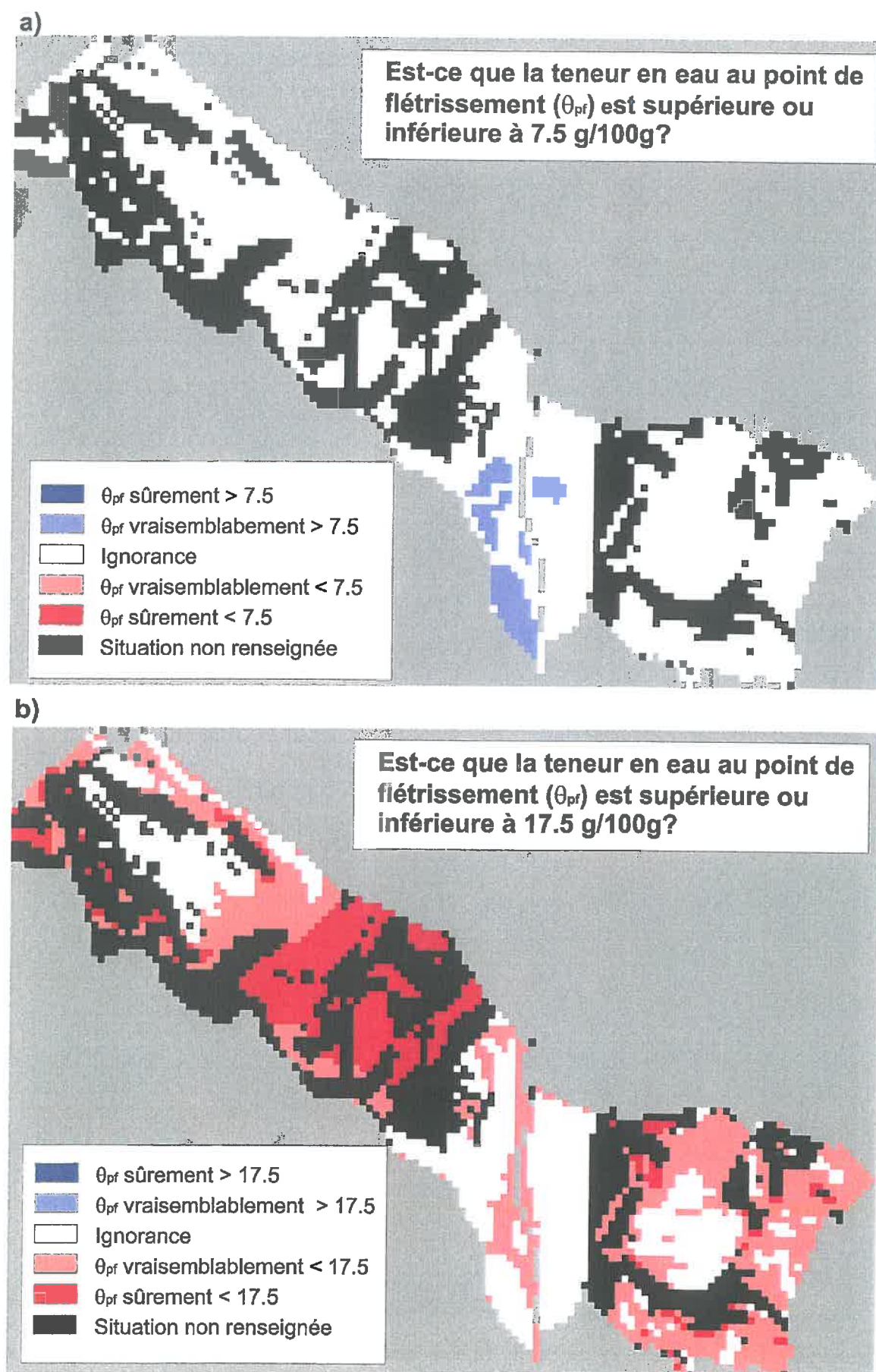


Figure 8.5: Cartographie de θ_{pf} dans le secteur ACM: détermination de la position de la distribution de possibilité de chaque cellule par rapport à une valeur seuil.

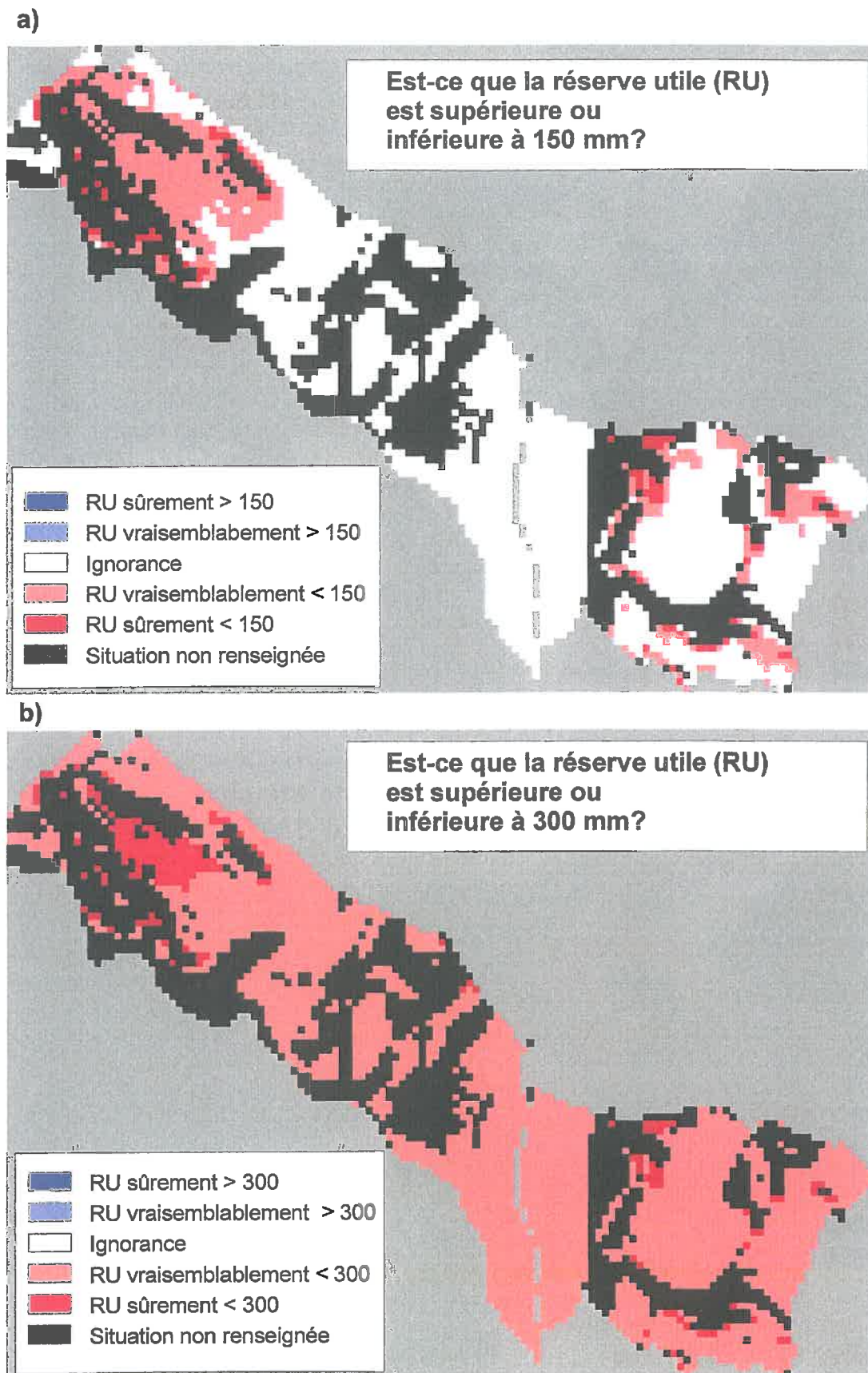
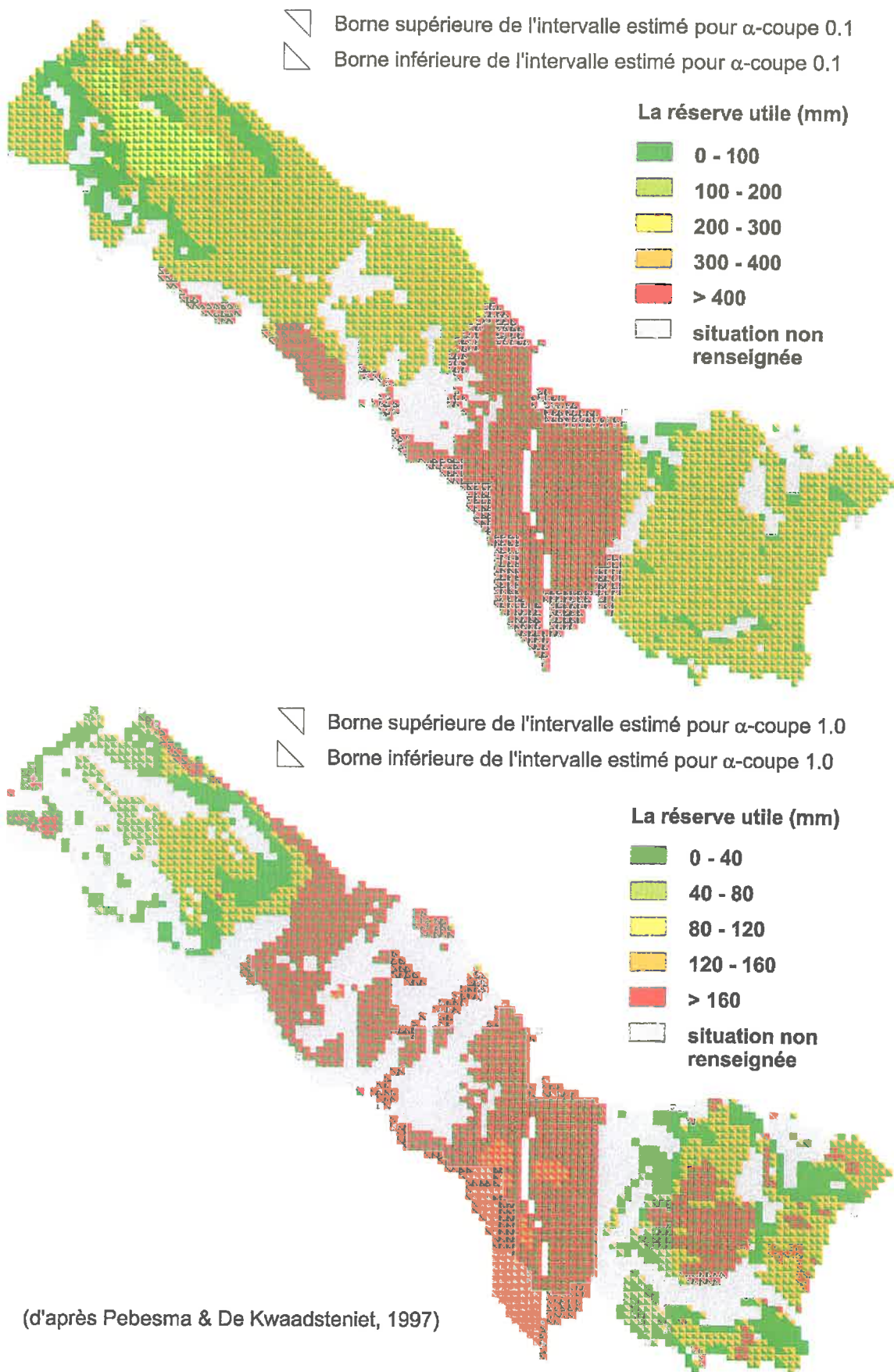


Figure 8.6: Cartographie de la RU dans le secteur ACM: détermination de la position de la distribution de possibilité de chaque cellule par rapport à une valeur seuil.



(d'après Pebesma & De Kwaadsteniet, 1997)

Figure 8.7: Représentation de l'intervalle estimé pour deux α -coupes (0.1 et 1.0) à l'aide d'une division du pixel en deux triangles.

Conclusion générale

L'étude présentée dans ce mémoire a été consacrée à l'utilisation de l'information issue des bases de données sol afin de pouvoir cartographier les propriétés hydriques des sols sur de vastes surfaces. Les cartes pédologiques à petite échelle constituent généralement la seule information disponible sur la variabilité des sols à l'échelle régionale. L'utilisation de ces cartes pour l'estimation des propriétés hydriques est difficile pour deux raisons essentielles. D'une part, l'information attachée à chaque unité typologique de sol (UTS) se présente sous la forme de descriptions qualitatives ou sous la forme d'intervalles de valeurs donnant une estimation d'une propriété de sol. D'autre part, les unités cartographiques de sols des cartes pédologiques à petite échelle sont généralement complexes, c'est-à-dire qu'elles regroupent plusieurs unités typologiques de sols. En conséquence, l'appartenance d'un site donné à une UTS n'est pas connue avec précision à partir de ces seules données.

L'objectif du travail présenté dans ce mémoire est de cartographier la réserve utile et la teneur en eau au point de flétrissement des sols sur un vaste espace en exploitant, malgré leur imprécision, les descriptions des propriétés et des caractéristiques environnementales des unités de sols d'une étude à petite échelle. La démarche générale retenue pour cette cartographie comprend 3 étapes essentielles: (i) l'estimation des propriétés hydriques (en appliquant des fonctions de pédotransfert) pour chaque UTS définie dans la zone à partir de la description des UTS, (ii) la cartographie des UTS dans la zone d'étude et (iii) l'intégration des résultats issus des deux premières étapes.

Une analyse approfondie des trois étapes a permis de distinguer un premier problème méthodologique:

- *Comment représenter une information géographique imprécise dans un contexte où les métadonnées caractérisant l'imprécision sont elles-mêmes imprécises et souvent absentes?*

Pour traiter ce problème, nous avons décidé d'utiliser la théorie des possibilités pour trois raisons principales. Premièrement, ce cadre théorique, basé sur la théorie des sous-ensembles flous, est particulièrement adapté à la représentation et au traitement de l'information imprécise lorsque celle-ci est exprimée sous la forme d'intervalles de valeurs, ce qui est le cas pour l'information pédologique contenue dans une base de données régionale de sol. Deuxièmement, une approche possibiliste permet, de coder de façon absolue

l'ignorance totale ce qui n'est pas le cas pour une approche probabiliste plus classique. Enfin, une approche possibiliste limite la propagation d'éventuelles erreurs contenues dans les données initiales grâce à l'utilisation des opérateurs min max à la place du produit et de la somme. Seule la notion d'ordre a réellement de l'importance.

L'interprétation de l'information et la traduction de la connaissance en distribution de possibilités ont été effectuées selon des voies différentes. Tout d'abord, pour interpréter l'information issue de la description des UTS (la description qualitative et description par profil de référence), nous avons proposé de traduire cette information avec des règles précises. Cette information a été traduite en distribution de possibilité trapézoïdale. Ensuite, pour pouvoir utiliser des données statistiques (par ex. les données exprimant les erreurs des fonctions de pédotransfert), nous proposons une procédure de transformation de cette information probabiliste sous la forme d'un histogramme en distribution de possibilité. Enfin, nous avons également proposé une représentation des limites de cartes choroplèthes compatible avec la théorie des possibilités pour pouvoir tenir compte de l'imprécision sur le positionnement des limites.

Le deuxième problème méthodologique rencontré dans la présente étude a été défini de la manière suivante:

- *Comment estimer, par opérations arithmétiques, les propriétés hydriques des unités typologiques de sol en propageant l'incertitude sur chaque variable intervenant dans l'estimation et en tenant compte des interactions entre ces variables?*

L'estimation de la propriété hydrique par UTS a été considérée comme un calcul d'intervalle impliquant des variables interactives. Le calcul de la variable recherchée peut être effectué en utilisant un principe général de raisonnement, dit de combinaison/projection. La technique utilisée dans la présente étude pour effectuer efficacement la combinaison/projection est la résolution des Problèmes de Satisfaction de Contraintes (CSP) par le solveur de contraintes CON'FLEX. Il est alors nécessaire de représenter cette interactivité entre variables, c'est-à-dire les contraintes. Nous avons reconnu et exploité deux types de contraintes: contraintes 'intra-horizon' qui exprime les interactions entre propriétés du sol appartenant à un même horizon et contraintes 'inter-horizons' qui expriment les interactions entre propriétés appartenant à deux horizons voisins.

Nous avons évalué les résultats obtenus par l'approche CSP avec deux méthodes d'estimation plus conventionnelles: (i) le calcul d'intervalles flous sans introduction des dépendances entre variables et (ii) l'approche stochastique, les calculs étant effectués par la simulation de Monte Carlo. Une analyse graphique et deux mesures quantitatives exprimant

l'imprécision (largeur de l'intervalle estimé) et l'incertitude (taux de mesures de terrain correctement estimées) ont aussi servi de tests d'évaluation. Ceux-ci montrent que les estimations par l'approche CSP et le calcul d'intervalles flous classique sont imprécises, mais couvrent généralement la variabilité observée. L'approche stochastique fournit des résultats généralement plus précis, mais sous-estiment beaucoup la variabilité observée. L'approche CSP, réduisant notablement l'imprécision d'estimation par rapport au calcul d'intervalles flous, apparaît un très bonne alternative aux méthodes conventionnelles.

Le troisième problème méthodologique rencontré dans la présente étude était le suivant:

- *Comment localiser, par combinaisons logiques, les unités typologiques de sol au sein des unités cartographiques de sols en tenant compte de la description des unités typologiques de sol et de l'incertitude sur l'information géographique supplémentaire utilisée?*

Pour la cartographie des UTS nous avons utilisé des informations géographiques supplémentaires décrivant l'environnement des unités des sols. Trois variables sont bien adaptées à la procédure proposée (géologie, pente et occupation du sol). Pour chaque UTS prise séparément, la carte exprimant le degré de possibilité d'apparition est obtenue en effectuant (i) l'appariement entre la classe donnée par la description textuelle de l'UTS et la valeur observée en chaque point et pour chaque variable et (ii) la combinaison conjonctive des cartes intermédiaires pour chaque variable géographique. Ces calculs ont été réalisés à l'aide des systèmes d'information géographique qui permettent d'effectuer des opérations logiques sur un sous-ensemble de couches d'informations.

Ayant mis en place la méthodologie et les techniques nécessaires et ayant interprété et représenté les connaissances disponibles, la cartographie des propriétés hydriques a pu être mise en œuvre en suivant la démarche générale adoptée. Nous avons considéré qu'il était essentiel d'analyser l'influence de l'utilisation de différents types de données dans la cartographie des propriétés hydriques. Une série d'expérimentations a été mise en place pour effectuer cette analyse.

Une première expérimentation a été effectuée pour pouvoir évaluer l'utilisation du type d'information (la description textuelle et la description par profil de référence) pour l'estimation de la réserve utile par unité de sol. Les différences moyennes en termes d'imprécision et d'incertitude entre les estimations issues du profil de référence et de la description textuelle demeurent très réduites. Ce résultat est surprenant car une estimation plus précise était attendue lorsque les profils de référence étaient utilisés. Une analyse plus

approfondie des résultats montre que les profils de référence ne s'avèrent utiles que dans les cas où les horizons sont bien différenciés et leurs épaisseurs connues avec précision. Pour la construction des bases de données sols, ceci implique que la saisie des données de terrain qui décrivent un profil de référence doit s'effectuer avec beaucoup de rigueur permettant une exploitation ultérieure des données dans les meilleures conditions. Cela signifie également que dans un grand nombre de cas, une description textuelle de l'unité, plus facile à fournir par le pédologue, est suffisante à l'échelle de l'étude considérée pour estimer de façon optimale les propriétés hydriques des sols.

L'ensemble des UTS choisies pour les expérimentations a permis d'analyser les différences de performances d'estimation entre les descriptions des UTS d'une carte pédologique à grande échelle et celles issues d'une étude à petite échelle. Nous avons pu vérifier que les estimations issues d'une carte à grande échelle sont plus précises sans qu'elles soient pour autant plus incertaines. C'est également pour les UTS de cartes à grande échelle que l'utilisation de profils de référence amène le gain de précision le plus important.

En ce qui concerne la cartographie des UTS, nous avons constaté que l'introduction des variables géographiques engendre un gain de précision, c'est-à-dire que le nombre d'UTS prédites en chaque point de l'espace diminue généralement avec le nombre de variables introduites. En revanche, l'introduction des variables géographiques provoque une augmentation de l'incertitude, la localisation estimée ne correspondant pas à la localisation observée sur une carte à plus grande échelle (i.e. l'information de référence). Nous avons constaté, dans la région d'étude considérée que la variable "géologie" est discriminante dans la localisation des UTS, tandis que la variable "occupation du sol" l'est assez. Quant à la variable "pente", son introduction provoque généralement un léger gain de précision, tandis qu'elle introduit plus d'incertitude. Les combinaisons de variables qui constituent le meilleur compromis en termes de précision et d'incertitude sont la géologie seule et la géologie en combinaison avec l'occupation du sol.

Ces résultats démontrent bien l'intérêt des descriptions de l'environnement des UTS des bases de données sol régionales, associées aux informations géographiques correspondantes, pour localiser les UTS dans les UCS. Il est ainsi important de renseigner les UTS à l'aide de certaines variables géographiques précises; ce qui permet ensuite une meilleure exploitation des données stockées dans les bases de données sols régionales: mieux vaut renseigner de manière précise un nombre de variables limité que beaucoup de variables de manière incomplète et non systématique.

Afin d'obtenir une cartographie des propriétés hydriques sur la région, nous avons intégré les résultats issus des deux premières étapes de la démarche proposée, c'est-à-dire, d'une part une estimation de la propriété hydrique pour chaque UTS et de l'autre part la cartographie des UTS. Une série de données de validation acquise dans la zone de test a

permis de comparer les estimations aux données mesurées sur le terrain. On s'aperçoit que les estimations sont généralement imprécises, mais elles englobent entre 90 et 98 % des mesures de terrain suivant l' α -coupe considérée.

Le travail présenté dans ce mémoire apporte des perspectives importantes et originales sur les recherches concernant l'utilisation d'une base de données sol régionale en vue de la cartographie des propriétés hydriques des sols.

Tout d'abord, les méthodes et techniques développées dans la présente étude sont transposables à d'autres types de base de données géographiques contenant des informations imprécises.

Ensuite, nous pouvons envisager certaines améliorations pour l'outil développé. Premièrement, nous avons constaté que les supports des distributions de possibilité ("les marges floues") représentant les propriétés du sol ont été choisie selon une règle empirique et uniforme. Une amélioration possible serait de définir expérimentalement les supports en utilisant de grandes bases de données pédologiques existantes pour tester les déterminations des propriétés de sol par le pédologue. Deuxièmement, une amélioration méthodologique ou technique est nécessaire pour pouvoir utiliser plus largement les outils de résolution des problèmes de satisfaction de contraintes pour lesquels nous avons rencontré des problèmes de temps de calculs. Ceci nécessite, d'un côté, des ordinateurs encore plus rapides ou, de l'autre côté, des techniques de résolution encore plus performantes. Ces deux aspects ne concernent cependant pas le domaine de la science du sol, mais plutôt celui de l'informatique et de la mathématique.

Dans la présente étude, nous avons utilisé le résolveur de contraintes comme outil permettant d'estimer les propriétés hydriques par UTS. Il est également envisageable d'utiliser cet outil pour l'analyse de sensibilité de l'estimation des propriétés hydriques aux variations des variables d'entrée. Cette analyse s'effectue en modifiant l'intervalle d'une seule variable d'entrée à la fois, suivi d'une résolution de CSP. Cette analyse de sensibilité permettrait d'identifier les variables critiques dont la précision serait à améliorer pour répondre à une requête donnée et donc de mieux raisonner les stratégies d'enrichissant des bases de données sol.

Dans le système actuel, les estimations des propriétés hydriques sont effectuées en une série de points dans l'espace. Il serait cependant plus intéressant de pouvoir fournir des estimations sur des objets géographiques qui constituent les unités de décision pour les utilisateurs et aménageurs de l'espace géographique (par exemple une parcelle agricole, une commune ou un bassin versant). Si l'on voulait fournir des résultats pour ce type d'objet géographique, il serait nécessaire d'agréger les résultats obtenus par point. Ce problème du changement d'échelle est un problème fréquemment rencontré dans ce type d'application.

Bibliographie

- AFES-INRA 1992. Référentiel Pédologique – principaux Sols d'Europe. Coord: Baize, D. & Girard, M.C. INRA Editions, 222 p.
- Altman, D. 1994. Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *Int. J. Geographical Information Systems* 8 (3), 271-289.
- Armstrong, A.C., Legros, J.P. & Voltz, M. 1996. ACCESS-II: A detailed model for crop growth and water conditions. *International Agrophysics* 10, 171-184.
- Badia, J. & Martin-Clouaire, R. 1989. Choise under imprecision: a simple possibility-theory-based technique illustrated with a problem of weed identification. *Comput. Electron. Agric.* 4, 103-120.
- Baize, D. 1986. Couvertures pédologiques, cartographie et taxonomie. *Science du Sol* 24 (3), 227-243.
- Baize, D. 1988. Guide des analyses courantes en pédologie. INRA, Paris, 172 p.
- Baize, D. & Jabiol, B. 1995. Guide pour la description des sols. INRA Editions Paris, 375p.
- Bardossy, A., Bogardi, I., Kelly, W.E., 1990a. Kriging with imprecise (fuzzy) variograms. I: Theory. *Math. Geol.* 22 (1), 63-79.
- Bardossy, A., Bogardi, I., Kelly, W.E., 1990b. Kriging with imprecise (fuzzy) variograms. II: Application. *Math. Geol.* 22 (1), 81-94.
- Barrière, J. 1971a. Limites d'utilisation des paléosols pour la reconstitution de paléoclimats. *C.R. Acad. Sc. Paris* 272, 2425-2429.
- Barrière, J. 1971b. Utilisation des paléosols comme éléments de datation des formations quaternaires. *C.R. Acad. Sc. Paris* 273, 310-313.
- Bascomb, C.L. & Jarvis, M.G. 1967. Variability in three areas of the Denchworth soil map unit. I. Purity of the map unit and property variability within it. *Journal of Soil Science* 27, 420-37.
- Bastet, G., Bruand, A., Quélin, P. & Cousin I. 1998. Estimation des propriétés de rétention en eau des sols à l'aide de fonctions de pédotransfert (FPT): une analyse bibliographique. *Etude et Gestion des Sols* 5 (1), 7-28.
- Beckett, P.H.T. & Webster, R. 1971. Soil variability: a review. *Soils and Fertilizers* 34 (1), 1-15.
- Bezdek, J.C. 1981. Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York.
- Bie, S.W. & Beckett, P.H.T. 1971. Quality Control in Soil Survey. Introduction : I. The choice of Mapping Unit. *Journal of Soil Science* 22 (1), 32-49.
- Blalock, H.M. 1988. Social statistics. Revised second edition. McGraw-Hill.
- Bolstad, P.V., Gessler, P. & Lillesand, T.M. 1990. Positional uncertainty in manually digitized map data. *Int. J. Geographical Information Science* 4 (4), 399-412.
- Bonfils, P. 1993. Carte pédologique de la France au 1/100.000, Feuille de Lodève (notice + carte). INRA SESCOF Orléans.
- Bornand, M. & Falipou, P. 1996. La réserve utile en eau des sols des plaines en Languedoc-Roussillon, première évaluation. Tentative de spatialisation. Laboratoire Science du Sol, Institut National de la Recherche Agronomique. Montpellier.
- Bornand, M., Barthès, J.P. & Bonfils, P. 1992. Carte des pédopaysages du Languedoc-Roussillon à l'échelle du 1/250.000. Carte + Légende. Laboratoire Science du Sol, Institut National de la Recherche Agronomique. Montpellier.

- Bornand, M., Legros, J.P. & Rouzet, C. 1994. Les banques régionales de données-sols, Exemple du Languedoc Roussillon. *Etude et Gestion des Sols* 1 (1), 67-82.
- Bornand, M., Arrouays, D., Baize, D. & Jamagne, M. 1989. Cadre méthodologique d'une cartographie régionale des sols au 1/250000. *Science du Sol* 27 (1), 17-20.
- Bouchon-Meunier, B. 1993. *La Logique Floue*. Edition Que sais-je ?. Presses universitaires de France.
- Bouma, J. & Van Lanen, J.A.J. 1987. Transfer functions and threshold values: From soil characteristics to land qualities. In: *Proceedings of the International Workshop on Quantified Land Evaluation*. Beek, K.J. Burrough, P.A. & McCormack, D.E. (eds.) *Proceedings of a Workshop*. ISSS and SSSA, Washington, DC. 27 Apr.-2 May 1986. Int. Inst. Aerospace Surv. Earth Sci. Publ. No. 6 ITC. Publ., Enschede, The Netherlands.
- Bregt, A.K & Beemster, J.G.R. 1989. Accuracy in predicting moisture deficits and changes in yield from soil maps. *Geoderma* 43, 301-310.
- BRGM 1981. Carte géologique de la France au 1/50.000, Feuille de Pézenas. Bureau de Recherche Géologique et Minières. Orléans.
- Brown, D.G. 1998. Classification and boundary vagueness in mapping presettlement forest types. *Int. J. Geographical Information Science* 12 (2), 105-129.
- Brown, K.W.M., Gauld, J.H., Smith, B.F.L., Bain, D.C., Burrige, J.C. & Inkson, R.H.E. 1987. Design of a database for Scottish soils. *Journal of Soil Science* 38, 267-277.
- Bruin, S. de & Stein, A. 1998. Soil-landscape modelling using fuzzy c-means clustering of attribute data derived from a Digital Elevation Model (DEM). *Geoderma* 83, 17-33
- Burrough, P.A. 1989. Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science* 40, 477-492.
- Burrough, P.A. & McDonnell, R.A. 1998. *Principles of Geographical Information Systems*. Spatial Information Systems and Geostatistics Series. Oxford University Press, Oxford.
- Burrough, P.A., Van Gaans, P.F.M. & Hootsmans, R. 1997. Continuous classification in soil survey: spatial correlation, confusion and boundaries. *Geoderma* 77, 115-135.
- Cayrol, M., Dubois, D. & Prade, H. 1982. Fuzzy pattern matching. *Kybernetes* 11, 103-116.
- Cazemier, D.R., Lagacherie, P. & Martin-Clouaire, R. 1998. A fuzzy constraint approach for handling uncertainty in regional soil maps. In: *Data quality in geographic information: from error to uncertainty*. Goodchild, M. & Jeansoulin, R. (eds.) Editions HERMES, Paris.
- Cazemier, D.R., Lagacherie, P. & Martin-Clouaire, R. 1998. Application de la théorie des possibilités à la cartographie de propriétés hydriques du sol. In: *Actes de LFA-98 Rencontres francophones sur la logique floue et ses Applications*, Rennes 18-19 Nov. 1998.
- CEC 1985. Soil map of the European Communities 1:1.000.000. Commission of the European Communities, DG VI (agriculture) Luxembourg. 124 p., 7 maps.
- Clochard, P., Leenhardt, D. & Lefay, O. 1991. Cartes des sols du secteur de référence d'Adissan-Montagnac (Moyenne Vallée de l'Hérault). Carte au 1/10.000. Laboratoire Science du Sol, Institut National de la Recherche Agronomique. Montpellier.
- Dalal-Clayton, D.B. & Robinson, D.A. 1992. An assessment of the success of a geomorphologically based reconnaissance soil survey in Eastern Zambia. *Geoderma* 54, 213-230.
- Dou, C., Woldt, W., Bogardi, I. & Dahab, M. 1995. Steady state groundwater flow simulation with imprecise parameters. *Water Resources Research* 31 (11), 2709-2719.
- Droesen, W.J. & Geelen, L.H.W.T. 1993. Application of fuzzy sets in ecohydrological expert modelling. In: *HydroGIS93: Application of geographic Information Systems in hydrology and water resources* (Proceedings of the Vienna Conference, April 1993) IAHS Publ. No. 211

- Dubois, D. & Prade, H. 1988. Possibility Theory: an Approach to Computerized Processing of Uncertainty. Plenum Press, New York.
- Dubois, D., Prade, H. 1994. Possibility theory and data fusion in poorly informed environments. *Control Eng. Practice*, 2 (5), 811-823.
- Dubois, D. & Prade, H. 1998. Possibility theory: qualitative and quantitative aspects. In: Handbook of defeasible reasoning and uncertainty management systems. Vol. 1, Gabbay, D.M. & Smets, Ph. (eds.), pp. 169-226. Kluwer (NL).
- Dubois, D., Fargier, H. & Prade, H. 1993. The calculus of fuzzy restrictions as a basis for flexible constraint satisfaction. Proc of the 2nd IEEE Conference on Fuzzy Sets, San Fransisco, CA, March 1993.
- Dubois, D., Prade, H. & Sandri, S. 1993. On possibility/probability transformations. In: Fuzzy Logic - State of the Art. Lowen, R. & Roubens, M. (eds.), Kluwer pp. 103-112.
- Dunn, R., Harrison, A.R. & White, J.C. 1990. Positional accuracy and measurement error in digital databases of land use: An empirical study. *Int. J. Geographical Information Systems* 4 (4), 385-398.
- FAO-UNESCO 1974. Soil map of the world at 1:5.000.000 + legend. Unesco, Paris, France.
- Fargier, H., Martin-Clouaire, R. & Schiex, T. 1992. Satisfaction de contraintes souples. In: Journées nationales des applications des ensembles flous. Nîmes, France. pp. 183-191.
- Favrot, J.C. 1989. Une stratégie d'inventaire cartographique à grande échelle: la méthode des secteurs de référence. *Science du Sol* 27 (4), 351-368.
- Favrot, J.C., Arrouays, D., Bornand, M., Girard, M.C. & Hardy, R. 1994. Informatisation et spatialisation de la ressource sol: le programme Inventaire, Gestion et Conservation des Sols. *Cahiers Agricultures* 3, 237-246.
- Fischer, P.F. 1992. First experiments in viewshed uncertainty: simulating fuzzy viewsheds. *Photogrammetric Engineering & Remote Sensing* 58, 345-352.
- Frank, A. 1998. Metamodels for data quality description. In: Data quality in geographic information: from error to uncertainty. Goodchild, M. & Jeansoulin, R. (eds.) Editions HERMES, Paris.
- Gaultier, J.P., Legros, J.P., Bornand, M., King, D., Favrot, J.C. & Hardy, R. 1992. L'organisation et la gestion des données pédologiques spatialisées: le projet DONESOL. *Revue de géomatique* 3 (3), 235-253.
- Genuchten, M.Th. van & Leij, F.J. 1992. On estimating the hydraulic properties of unsaturated soils. pp. 1-14 In: Indirect methods for estimation the hydraulic properties of unsaturated soils. Genuchten, M.Th. van, Leij, F.J. & Lund, L.J. (eds.) Proceedings of the International Workshop on Indirect Methods for estimating Hydraulic Properties of Unsaturated Soils. Riverside, California, 11-13 October 1989.
- Girard, M.C. 1983. Recherche d'une modélisation en vue d'une représentation spatiale de la couverture pédologique. Application à une région des plateaux jurassiques de Bourgogne. Thèse d'Etat. Sols n°12 430p. INA Paris-Grignon.
- Girard, M.C. 1984. Analyse spatiale de la couverture pédologique. *Cartographie Cartogénèse. AFES n° spécial cinquantenaire*, 153-166.
- Goodchild, M. & Gopal, S. (eds.) 1989. The accuracy of spatial databases. Taylor & Francis, London.
- Goodchild, M.F., Guoqing, S. & Shiren, Y. 1992. Development and test of an error model for categorical data. *International Journal of Geographical Information Systems* 6 (2), 87-104.
- Gopal, S. & Woodcock, C. 1994. Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering & Remote Sensing* 60 (2), 181-188.
- Gruijter, J.J. de, Walvoort, D.J.J & Gaans, P.F.M. van 1997. Continuous soil maps – a fuzzy set approach to bridge the gap between aggregation levels of process and distribution models. In: Gruijter, J.J. de, McBratney, A.B. & McSweeney, K. (eds.) Fuzzy sets in soil science. *Geoderma* 77, 169-196.
- Gupta, S.C. & Larson, W.E. 1979. Estimating soil water retention characteristics from particle size distribution, organic matter percent and bulk density. *Water Resour. Res.* 15 (6), 1633-1635.

- Harrison, P.A., Butterfield, R.E. & Downing, T.E. (eds.) 1995. Climate Change and Agriculture in Europe. Assessment of impacts and adaptation. Research Report No. 9. Environmental Change Unit, Oxford, UK.
- Heineke, H.J., Eckelmann, W., Thomasson, A.J., Jones, R.J.A., Montanarella, L. & Buckley, B. (eds.) 1998. Land Information Systems: Developments for planning the sustainable use of land resources. European Soil Bureau Research Report No. 4 EUR 17729 EN 546 pp. Office for Official Publications of the European Communities, Luxembourg.
- Heuvelink, G.B.M. & Bierkens, M.F.P. 1992. Combining soil maps with interpolations from point observations to predict quantitative soil properties. *Geoderma* **55**, 1-15.
- Heuvelink, G.B.M. & Burrough, P.A. 1993. Error propagation in cartographic modelling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems* **7** (3), 231-246
- Heuvelink, G.B.M. 1993. Error propagation in quantitative spatial modeling: applications in geographical information systems. *Knag/Faculteit Ruimtelijke Wetenschappen Universiteit Utrecht* 160pp. Utrecht.
- Hodgson, J.M. 1974. Soil survey field handbook. Describing and sampling soil profiles. Technical monograph No 5. Soil Survey of England and Wales.
- Hodgson, J.M., Hollis, J.M., Jones, R.J.A & Palmer, R.C. 1976. A comparison of field estimates and laboratory analyses of the silt and clay contents of some West Midlands soils. *Journal of Soil Science* **27**, 411-419.
- Hootsmans, R. M. 1996. Fuzzy sets and series analysis for visual decision support in spatial data exploration. *Utrecht Knag/ Faculteit Ruimtelijke Wetenschappen Universiteit Utrecht*. PhD Thesis.
- Horn, B.K. 1981. Hill shading and the reflectance map. *Proc. I.E.E.E.* **69**, 14.
- Hubrechts, L., Poorten, K. vander, Vanclooster, M. & Deckers, J. 1998. From soil survey to quantitative land evaluation in Belgium. In: Heineke, H.J., Eckelmann, W., Thomasson, A.J., Jones, R.J.A., Montanarella, L. & Buckley, B. (eds.) 1998. Land Information Systems: Developments for planning the sustainable use of land resources. European Soil Bureau Research Report No. 4 EUR 17729 EN 546 pp. Office for Official Publications of the European Communities, Luxembourg.
- Hudson, B.D. 1990. Concepts of soil mapping and interpretation. *Soil Survey Horizons* **31**, 63-72.
- Hudson, B.D. 1992. The soil survey as paradigm-based science. *Soil Sci. Soc. Am. J.* **56**, 836-841.
- Jamagne, M. 1993. Evolution dans les conceptions de la cartographie des sols. *Pédologie* **43** (1), 59-115.
- Jamagne, M., Hardy, R., King, D. & Bornand, M. 1995. La base de données géographiques des sols de France. *Etude et Gestion des Sols* **2** (3), 153-172.
- Jenny, H. 1941. Factors of soil formation. Mac Graw Hill, New York, 281 pp.
- Johnson, M.E. 1987. Multivariate Statistical Simulation. Wiley, New York.
- Jones, L. 1989. Some results comparing Monte Carlo simulation and first order Taylor series approximation for steady groundwater flow. *Stochastic Hydrology and Hydraulics* **3**, 179-190.
- Journel, A.J. & Huijbregts, C.J. 1978. Mining Geostatistics. Academic Press.
- King, D., Burrill, A., Darousin, J., Le Bas, C., Tavernier, R. & Van Ranst, E. 1995. The EU soil geographic database. In: European Land Information Systems for Agro-Environmental Monitoring. King, D., Jones R.J.A. & Thomasson A.J. (eds.) EUR 16232 EN, 43-60. Office for the Official Publications of the European Communities, Luxembourg.
- King, D., Daroussin, J. & Tavernier, R. 1994. Development of a soil geographic database from the soil map of the European Communities. *Catena* **21**, 37-56.
- Klir, G.J. & Yuan, B. 1995. Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, New Jersey.
- Klir, G.J. & Parviz, B. 1992. Probability-possibility transformations: a comparison. *Int. J. General Systems* **21**, 291-310.

- Kraak, M.J., Müller, J.C. & Ormeling, F.J. (1995) GIS-Cartography: visual decision support for spatio temporal data handling. *International Journal of Geographical Information Systems* **9**, 637-645.
- Lagacherie, P. 1992. Formalisation des lois de distribution des sols pour automatiser la cartographie pédologique à partir d'un secteur pris comme référence. Thèse Doctorat, Académie de Montpellier, Université Montpellier II, Montpellier, France
- Lagacherie, P., Legros, J.P. & Burrough, P.A., 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. *Geoderma* **65**, 283-301
- Lagacherie, P., Andrieux P. & Bouzigues, R. 1996. Fuzziness and uncertainty of soil boundaries: from reality to coding in GIS. In: *Geographic objects with indeterminate boundaries*. Burrough, P.A. and Frank, A.U. (eds.) GISDATA Series II. Taylor & Francis, London. pp. 275-302.
- Lagacherie, P., Cazemier, D.R., Van Gaans, P.F.M. & Burrough, P.A. 1997. Fuzzy k-means clustering of fields in an elementary catchment and extrapolation to a larger area. In: Gruijter, J.J. de, McBratney, A.B. & McSweeney, K. (eds.) *Fuzzy sets in soil science*. *Geoderma* **77**, 197-216.
- Le Bas, C., & Jamagne, M. 1996. Soil databases to support sustainable development. INRA-SESCPF, Joint Research Centre-IRSA, Orléans.
- Le Bas, C., King, D., Daroussin, J., Nicoullaud, B. & Ngongo, M. 1998. Impacts of errors in the assessment of agronomic constraints to crop production at European scale. *Proceedings of the International Congress of Soil Science 20-26 août 1998, Montpellier. Symp. 35 n° 810*.
- Leenhardt, D. 1991. Spatialisation du bilan hydrique. Propagation des erreurs d'estimation de caractéristiques du sol au travers des modèles de bilans hydriques. Cas du blé dur d'hiver. Thèse Doctorat, Ecole Nationale Supérieure Agronomique de Montpellier, Montpellier, France.
- Leenhardt, D., Voltz, M. & Bornand, M. 1994. Propagation of the error of spatial prediction of soil properties in simulating crop evapotranspiration. *European Journal of Soil Science* **45**, 303-310.
- Leenhardt, D., Voltz, M., & Rambal, S. 1995. A survey of several agroclimatic soil water balance models with reference to their spatial application. *Eur. J. Agron.* **4** (1), 1-14.
- Leenhardt, D., Voltz, M., Bornand, M. & Webster, R. 1994. Evaluating soil maps for prediction of soil water properties. *European Journal of Soil Science* **45**, 293-301.
- Leenheer, L. de, Appelmans, F. & Caestecker, K. de. 1954. Sur la valeur objective des évaluations de la texture sur le terrain, et sur le rôle de certaines fractions granulométriques dans ces évaluations. *Trans. 5th int. Congr. Soil Sc.* **4**, 374-382.
- Legros, J.P. 1978a. Etude bibliographique. Recherche et contrôle numérique de la précision en cartographie pédologique I. Précision dans la délimitation des sols. *Ann. agron.* **29** (5), 499-519.
- Legros, J.P. 1978b. Etude bibliographique. Recherche et contrôle numérique de la précision en cartographie pédologique II. Précision dans la caractérisation des unités de sols. *Ann. agron.* **29** (6), 583-601.
- Legros, J.P. 1996. *Cartographies des sols*. Presses Polytechniques et Universitaires Romandes. Lausanne.
- Legros, J.-P. & Nortcliff S. 1990. Conception d'un vocabulaire pour la description du milieu naturel et des sols. *Pedologie*, **40** (2), 195-213.
- Leung, Y. & Leung, K.S. 1993a. An intelligent expert system shell for knowledge-based Geographical Information Systems: 1. The tools. *Int. J. Geographical Information Systems* **7** (3), 189-199.
- Leung, Y. & Leung, K.S. 1993b. An intelligent expert system shell for knowledge-based Geographical Information Systems: 2. Some applications. *Int. J. Geographical Information Systems* **7** (3), 201-213.
- Leung, Y. 1987. On the imprecision of Boundaries. *Geographical Analysis* **19** (2), 125-151.
- Lillesand, T.M. & Kiefer, R.W. 1994. *Remote Sensing and Image Interpretation*. 3rd edition Wiley, New York.
- Magaldi, D. 1995. Soil mapping in EU countries: short review. In: *European Land Information Systems for Agro-Environmental Monitoring*. King, D., Jones R.J.A. & Thomasson A.J. (eds.) EUR 16232 EN, 71-84. Office for the Official Publications of the European Communities, Luxembourg.

- Mallants, D., Mohanty, B.P., Jacques, D. & Feyen, J. 1996. Spatial variability of hydraulic properties in a multi-layered soil profile. *Soil Science* **161** (3), 167-181.
- Manly, B.F.J. 1991. *Randomization and Monte Carlo Methods in Biology*. Chapman and Hall.
- Marsman, B.A. & De Gruijter, J.J. 1986. Quality of soil maps. A comparison of surveys in a sandy area. *Soil Survey Paper*, 15 Netherlands Soil Survey Institute, Wageningen.
- Martin-Clouaire, R., Cazemier, D.R. & Lagacherie, P. 1998. Representing and processing uncertain soil information in a fuzzy constraint network. In: Proc. 3rd IFAC/CIGR Workshop on "Artificial Intelligence in Agriculture" Makuhari, Japan, April 24-26, 1998.
- Matheron, G. 1965. *Les variables régionalisées et leur estimation*. Masson Paris. 305 p.
- McBratney, A.B. & Odeh, I.O.A. 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. In: Gruijter, J.J. de, McBratney, A.B. & McSweeney, K. (eds.) *Fuzzy sets in soil science*. *Geoderma* **77**, 85-113.
- McBratney, A.B. & Gruijter, J.J., de 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. *Journal of Soil Science* **43**, 159-175.
- Météo France 1992. *Statistiques climatiques de l'Hérault*. Centre départemental de la météo, Aéroport Fréjorgues – Montpellier.
- Moulènes, D. 1993. *Caractérisation hydrodynamique des sols du Languedoc-Roussillon: recherche de fonctions de pédotransfert*. DAA ENSA-Montpellier.
- Moussa, R., Lagacherie, P., Voltz, M., Andrieux, P. & Bouzigues, R. 1998. Hydrological modelling of a farmed Mediterranean catchment. *Proceedings of the International Congress of Soil Science 20-26 août 1998, Montpellier*. Symp. 3 n° 668.
- Ngongo, L., King, D., Nicoulaud, B., Brisson, N. & Ruget, F. 1995. The accuracy of methods for the selection of soil data in estimating regional moisture deficit from soil maps. In: *European Land Information Systems for Agro-Environmental Monitoring*. King, D., Jones R.J.A. & Thomasson A.J. (eds.) EUR 16232 EN, 155-165. Office for the Official Publications of the European Communities, Luxembourg.
- Odeh, I.O.A., Bratney, A.B. & Chittleborough, D.J. 1992. Soil pattern recognition with fuzzy-c-means: application to classification and soil-landform interrelationships. *Soil Sci. Soc. Am. J.* **56**, 505-516.
- Pebesma, E.J. & De Kwaadsteniet, J.W. 1997. Mapping groundwater quality in the Netherlands. *Journal of Hydrology* **200**, 364-386.
- Proctor, M.E., Siddons, P.A., Jones, R.J.A., Bellamy, P.H. & Key, C.A. 1998. Landis – a Land Information System for the UK. In: Heineke, H.J., Eckelmann, W., Thomasson, A.J., Jones, R.J.A., Montanarella, L. & Buckley B. (eds.) 1998. *Land Information Systems: Developments for planning the sustainable use of land resources*. European Soil Bureau Research Report No. 4 EUR 17729 EN 546 pp. Office for Official Publications of the European Communities, Luxembourg.
- Prosser, P. 1993. Hybrid algorithms for the constraint satisfaction problem. *Computational Intelligence* **9** (3), 269-299.
- Ranst, E. van, Vanmechelen, L., Thomasson, A.J., Daroussin, J., Hollis, J.M., Jones, R.J.A., Jamagne, M. & King, D. 1995. Elaboration of an extended knowledge database to interpret the 1:1.000.000 EU soil map for environmental purposes. In: *European Land Information Systems for Agro-Environmental Monitoring*. King, D., Jones R.J.A. & Thomasson A.J. (eds.) EUR 16232 EN, 71-84. Office for the Official Publications of the European Communities, Luxembourg.
- Rawls, W.J., Brakensiek, D.L. & Saxton, K.E. 1982 Estimation of soil water properties. *Trans. ASAE* **25**, 1316-1320.
- Rellier, J.P. & Vardon, F. 1998. *CON'FLEX*. Manuel de l'utilisateur, Version 1.2. Institut National de la Recherche Agronomique, Laboratoire de Biométrie et d'Intelligence Artificielle, Toulouse. Manuel + exécutable du logiciel sont disponibles sur <http://www-bia.inra.fr>.

- Robbez-Masson, J.M. 1994. Reconnaissance et délimitation des motifs d'organisation. Application à la cartographie des pédopaysages. Thèse Doctorat. Ecole Nationale Supérieure Agronomique de Montpellier. 161 p. + ann.
- Rogowski, A.S. 1996. Quantifying soil variability in GIS applications: II. Spatial distribution of soil properties. *Int. J. Geographical Information Systems* 10 (4), 455-475.
- Roo, A.P.J. de, Hazelhoff, L. & Heuvelink, G.B.M. 1992. Estimating the effects of spatial variability of infiltration on the output of a distributed runoff and soil erosion model using Monte Carlo methods. *Hydrological Processes* 6, 127-143.
- Rounsevell, M., Armstrong, A. Audley, E., Brown, O., Evans, S., Gylling, M., Lagacherie, P., Margaris, N., Mayr, T., De la Rosa, D. Rosato, P. & Simota, C. 1998. The IMPEL project: integrating biophysical and socio-economic models to study land use change in Europe. Proceedings of the International Congress of Soil Science 20-26 août 1998, Montpellier. Symp. 35 n° 1279.
- Roux, L. & Desachy, J. 1997. Multisources information-fusion application for satellite image classification. In: *Fuzzy Information Engineering*. Dubois, D., Prade, H. & Yager, R. (eds.) Wiley, New York.
- Saint-Joan, D. & Desachy, J. 1995. A fuzzy expert system for geographical problems: an agricultural application. In proceedings of Fourth IEEE International Conference on Fuzzy Systems 20-24 mars, Yokohama Japan 2, 469-476.
- Sandri, S.A. 1997. Elicitation, pooling and assessment of expert opinion in the possibilistic framework. In: *Fuzzy Information Engineering*. Dubois, D., Prade, H. & Yager, R. (eds.) Wiley, New York.
- Scheinost, A.C., Sinowski, W. & Auerswald, K. 1997a. Regionalization of soil water retention curves in a highly variable soilscape, I. Developing a new pedotransfer function. *Geoderma* 78, 129-143.
- Scheinost, A.C., Sinowski, W. & Auerswald, K. 1997b. Regionalization of soil water retention curves in a highly variable soilscape, II. Comparison of regionalization procedures using a pedotransfer function. *Geoderma* 78, 145-159.
- Schiex, T. 1997. Problèmes de satisfaction de contraintes. Chapitre 14 dans: *Intelligence Artificielle et Informatique Théorique*
- Schulz, K. & Huwe, B. 1997. Water flow modeling in the unsaturated zone with imprecise parameters using a fuzzy approach. *Journal of Hydrology* 201, 211-229.
- Shannon, R.E. 1977. *Systems simulation, the art and science*. Prentice Hall. New Jersey.
- Skidmore, A. 1989. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *Int. J. Geographical Information Systems* 3, 323-334.
- Smith, R.E. & Freeze, R.A. 1979. Stochastic analysis of steady state groundwater flow in a bounded domain. 2. Two-dimensional simulations. *Water Resources Research* 15, 1543-1559.
- Smith, R.E. & Hebbert, R.H.B. 1979. A Monte Carlo analysis of the hydrologic effects of spatial variability of infiltration. *Water Resources Research* 15, 419-429.
- Soil Survey Division Staff 1993. *Soil Survey Manual*, USDA, Handbook No. 18.
- Stassopoulou, A., Petrou, M. & Kittler, J. 1998. Application of a Bayesian network in a GIS based decision making system. *Int. J. Geographical Information Science* 12 (1), 23-45.
- Statistica 1995. *Statistica for Windows. General Conventions and Statistics I* (2nd Edition), Statsoft Inc. Tulsa. OK.
- Statistical Sciences 1993. *S-PLUS. Guide to Statistical and Mathematical Analyses*, Version 3.2 MathSoft Inc. Seattle.
- Tang, H. & Ranst, E. van. 1992. Testing of fuzzy set theory in land suitability assessment for rainfed grain maize production. *Pédologie* 42 (2), 129-147.

- Tassinari, C. 1998. Contribution à l'étude de sols méditerranéens (Languedoc, France): caractérisation de leurs nappes et des indicateurs morphologiques associés. Thèse Doctorat, Ecole Nationale Supérieure Agronomique de Montpellier.
- Tomasella, J. & Hodnett M.G. 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Science* **163** (3), 190-202.
- Tsang, E. P. K. 1993. *Foundations of Constraint Satisfaction*. London, Academic Press Ltd.
- Van Hentenryck, P., Michel, L. & Deville Y. 1997. *Numerica : A Modeling Language for Global Optimization*. MIT Press, Cambridge, MA.
- Vereecken, H., Maes, J., Feyen & Darius, P. 1989. Estimating the soil moisture retention characteristics from texture, bulk density and carbon content. *Soil Science* **148** (6), 389-403.
- Voltz, M. & Andrieux, P. 1994. Etudes des flux d'eau et de polluants en milieu méditerranéen viticole : le programme Allegro-Roujan. Institut National de la Recherche Agronomique, Laboratoire de Science du Sol, Montpellier, 1994, 15 p.
- Voltz, M. & Goulard, M. 1994. Spatial interpolation of soil moisture retention curves. *Geoderma* **62**, 109-123.
- Voltz, M., Lagacherie, P. & Louchart, X. 1997. Predicting soil properties over a region using sample information from a mapped reference area. *European Journal of Soil Science* **48**, 19-30.
- Vossen, P. & Meyer-Roux, J. 1995. Crop monitoring and yield forecasting activities of the MARS project. In: *European land information systems for agro-environmental monitoring*. King, D. R.J.A. Jones & Thomasson, A.J. (eds.) EUR 16232 EN, 11-29. Office for the Official Publications of the European Communities, Luxembourg.
- Walter, C. 1990. Estimation de propriétés du sol et quantification de leur variabilité à moyenne échelle. Thèse Doctorat, Université de Paris VI n° SDS 436, 172 p.
- Wang, F. & Hall, G.B. 1996. Fuzzy representation of geographical boundaries in GIS. *Int. J. Geographical Information Systems* **10** (5), 573-590.
- Webster, R. & Burrough, P.A. 1972. Computer-based soil mapping of small areas from sample data: I. Multivariate classification & ordination. *Journal of Soil Science* **23**, 210-221.
- Webster, R. & Oliver, M.A. 1990. *Statistical Methods in Soil and Land Resource Survey*. Oxford University Press, 316 p.
- Wösten, J.H.M., Finke, P.A. & Jansen, M.J.W. 1995. Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics.
- Zadeh, L.A. 1965. Fuzzy sets. *Information and Control* **8** (3), 338-353.
- Zadeh, L.A. 1978. A theory of approximate reasoning. In: *Machine Intelligence*, 9. Hayes, J.E., Michie, D. & Mikulich, L.I. (eds.), Elsevier, 149-194.
- Zadeh, L.A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* **1**, 3-28.
- Zhu, A.X., Band, L., Dutton, B. & Nimlos, T.J. 1996. Automated soil inference under fuzzy logic. *Ecological Modelling* **90**, 123-145.
- Zhu, A.X., Band, L., Vertessy, R. & Dutton, B. 1997. Derivation of soil properties using a soil land inference model (SoLIM). *Soil Sci. Soc. Am. J.* **61**, 523-533.

Glossaire

θ_{cr}	Teneur en eau à la capacité de rétention (en g/100g).
θ_{pf}	Teneur en eau au point de flétrissement (en g/100g).
ACM	Secteur de référence Adissan-Cazouls-Montagnac.
Arg	Taux en argile du sol (%).
BRGM	Bureau de Recherches Géologiques et Minières.
CIF	Calcul d'Intervalles Flous.
CON'FLEX	Outil informatique permettant de résoudre des problèmes à satisfaction de contraintes.
CSP	Constraint Satisfaction Problem (Problèmes à Satisfaction de Contraintes).
Da	Densité apparente du sol.
E	Profondeur du sol ou l'épaisseur d'un horizon (en cm ou en mm.).
ENSAM	Ecole Nationale Supérieure d'Agronomie de Montpellier.
FAO	Organisation des Nations Unies pour l'Agriculture et l'Alimentation.
FPT	Fonction de pédotransfert: relation mathématique entre un ensemble de propriétés du sol permettant d'estimer une propriété sur laquelle on dispose de peu d'information (par ex. la teneur en eau au point de flétrissement)
Horizon	Unité de base de la caractérisation verticale de la couverture pédologique.
IGN	Institut Géographique National.
INRA	Institut National de la Recherche Agronomique.
Lim	Taux en limon du sol (%).
Ensemble flou	Généralisation d'un sous-ensemble classique défini par une fonction d'appartenance contenant des éléments avec un degré d'appartenance entre 0 et 1.
MNA	Modèle numérique d'altitude.
Mo	Taux de la matière organique du sol (%).
Nécessité	Mesure d'incertitude pour la réalisation d'un événement issue de la théorie des possibilités.
Noyau	Partie de la distribution de possibilité où la possibilité est égale à 1.0.
PEC	Proportion des estimations correctes.
PCC	Pourcentage complètement possible.
Possibilité	Mesure d'incertitude pour la réalisation d'un événement issue de la théorie des possibilités.
Probabilité	Mesure d'incertitude pour la réalisation d'un événement issue de la théorie des probabilités.
Profil de référence	Description d'une unité typologique de sol généralement à l'aide d'intervalles de valeurs pour les différentes propriétés de sol, les

	bornes de l'intervalles étant fournis par un petit nombre d'observations de terrain.
Profil représentatif	Profil de sol réellement observé à l'intérieur de l'unité typologique de sol et considéré par le pédologue pour représenter au mieux les caractéristiques moyennes de l'unité typologique de sol.
RU	Réserve utile en eau du sol (en mm).
Sab	Taux en sable du sol (%).
SIG	Système d'Information Géographique.
SMC	Simulation de Monte Carlo.
STIPA	Système de Transfert de l'Information Pédologique et Agronomique.
Support	La partie d'une distribution de possibilité où la possibilité est supérieur à 0.0.
Théorie des possibilités	Théorie permettant de formaliser des incertitudes de nature non-probabiliste sur des événements. Cette théorie est en particulier adaptée à la représentation des informations imprécises.
Théorie des probabilités	Théorie permettant de formaliser des incertitudes sur des événements où la connaissance est basée sur la distribution de fréquences mesurées.
UCS	Unité cartographique du sol: unité de légende sur une carte pédologique contenant une ou plusieurs unités typologiques de sol dépendant de l'échelle cartographique.
UPP	Unité de pédopaysage: nom de l'entité conceptuel de la carte des pédopaysages. Elle constitue l'unité cartographique de sol et elle contient généralement plusieurs unités typologiques de sol.
UTS	Unité typologique du sol: l'objet conceptuel qui rend compte d'un ensemble d'observations de sol considéré comme homogène.
VHOL	Zone d'étude des Vallées de l'Hérault, Orb et Libron.

Liste des figures

- Figure 1.1:** Schéma d'organisation des entités cartographiques pour une carte pédologique à l'échelle régionale.
- Figure 1.2:** La démarche proposée pour l'estimation des propriétés hydriques à partir d'une carte pédologique régionale. Trois étapes principales ont été définies (a, b et c). Voir texte pour plus d'explications.
- Figure 3.1:** Sous-ensemble flou sous la forme d'une fonction trapézoïdale.
- Figure 3.2:** La définition des α -coupes.
- Figure 3.3:** L'évaluation de la compatibilité du sous-ensemble flou F , exprimant une observation d'un pédologue avec la distribution de possibilité π , exprimant le sous-ensemble flou de référence 'Sol Profond'.
- Figure 4.1:** Localisation de la zone d'étude et du secteur de référence.
- Figure 4.2:** Localisation des profils pour lesquels la réserve utile est déterminée. L'arrière plan de cette carte est la carte de sols du secteur au 10.000^{ème} (voir Annexe II).
- Figure 4.3:** Localisation des observations de la teneur en eau au point de flétrissement.
- Figure 4.4:** Partie de la zone d'étude utilisée pour l'évaluation de la cartographie des UTS au sein des UCS.
- Figure 5.1:** Comparaison entre l'appréciation de la texture par le pédologue sur le terrain et la composition granulométrique mesurée. Exemples de textures (a) Limono-Argilo-Sableuses et (b) Argilo-Limoneuses.
- Figure 5.2:** Nuage de points exprimant la relation entre le θ_{cr} et θ_{pf} .
- Figure 5.3:** Nuage de points exprimant la relation entre le taux de limon de l'horizon inférieur et le taux de limon de l'horizon supérieur.
- Figure 5.4:** Détermination de la distribution de possibilité des résidus de la régression multiple.
- Figure 5.5:** Représentation des limites floues dans le cadre de la théorie des possibilités.
- Figure 5.6:** Deux limites géographiques exprimant la même limite entre deux unités cartographiques.
- Figure 5.7:** Détermination de la distribution de possibilité d'apparition pour les unités cartographiques de la carte géologique (a) et la carte des pédopaysages (b). Les distributions de possibilité définissent la possibilité d'apparition d'une unité cartographique à l'extérieur de ses limites.
- Figure 5.8:** Détermination des distribution de possibilité exprimant l'incertitude sur la pente.
- Figure 6.1:** Création d'une fonction de répartition à partir d'une fonction de densité de probabilité trapézoïdale.
- Figure 6.2:** Représentation de l'ensemble des estimations issues des trois types de calcul différents pour les quatre UTS d'évaluation et les deux variables considérées.
- Figure 6.3:** Comparaison graphique entre les estimations obtenues par la simulation de Monte Carlo et les mesures de terrain.
- Figure 6.4:** Estimations de θ_{pf} issues des deux types de description des UTS différents (profil de référence et description textuelle) avec la distribution des mesures de terrain.

- Figure 6.5:** Estimations de la réserve utile issues des deux types de description des UTS différents (profil de référence et description textuelle) avec la distribution des mesures de terrain.
- Figure 7.1:** Combinaison logique des variables géographiques en vue de localiser les UTS.
- Figure 7.2:** Cartes floues exprimant la localisation des UTS 102 et 974 pour le secteur de référence ACM.
- Figure 7.3:** Comparaison entre la localisation estimée (par différentes α -coupes) et la localisation observée (information non-floue).
- Figure 7.4:** Proportion des points où l'UTS est correctement prédite pour les différentes combinaisons de variables et quatre α -coupes différentes.
- Figure 7.5:** Nombre d'UTS moyen estimé sur l'ensemble des points de la zone de validation pour chaque combinaison de variables géographiques et pour trois α -coupes différentes.
- Figure 7.6:** Proportion de la zone d'évaluation où au moins une UTS atteint la possibilité 1 (PCC = pourcentage complètement compatible).
- Figure 8.1:** Combinaison disjonctive de deux sources d'information.
- Figure 8.2:** Pourcentage des points de validation dont la valeur de θ_{pf} se situe à l'extérieur de l'intervalle estimé.
- Figure 8.3:** Largeur moyenne (plus ou moins l'écart type) de l'intervalle estimée pour θ_{pf} sur l'ensemble des points de validation.
- Figure 8.4:** Détermination de la position des résultats obtenus par rapport à une valeur seuil.
- Figure 8.5:** Cartographie de θ_{pf} dans le secteur ACM: détermination de la position de la distribution de possibilité de chaque cellule par rapport à une valeur seuil.
- Figure 8.6:** Cartographie de la RU dans le secteur ACM: détermination de la position de la distribution de possibilité de chaque cellule par rapport à une valeur seuil.
- Figure 8.7:** Représentation de l'intervalle estimé pour deux α -coupes (0.1 et 1.0) à l'aide d'une division du pixel en deux triangles.

Liste des tableaux

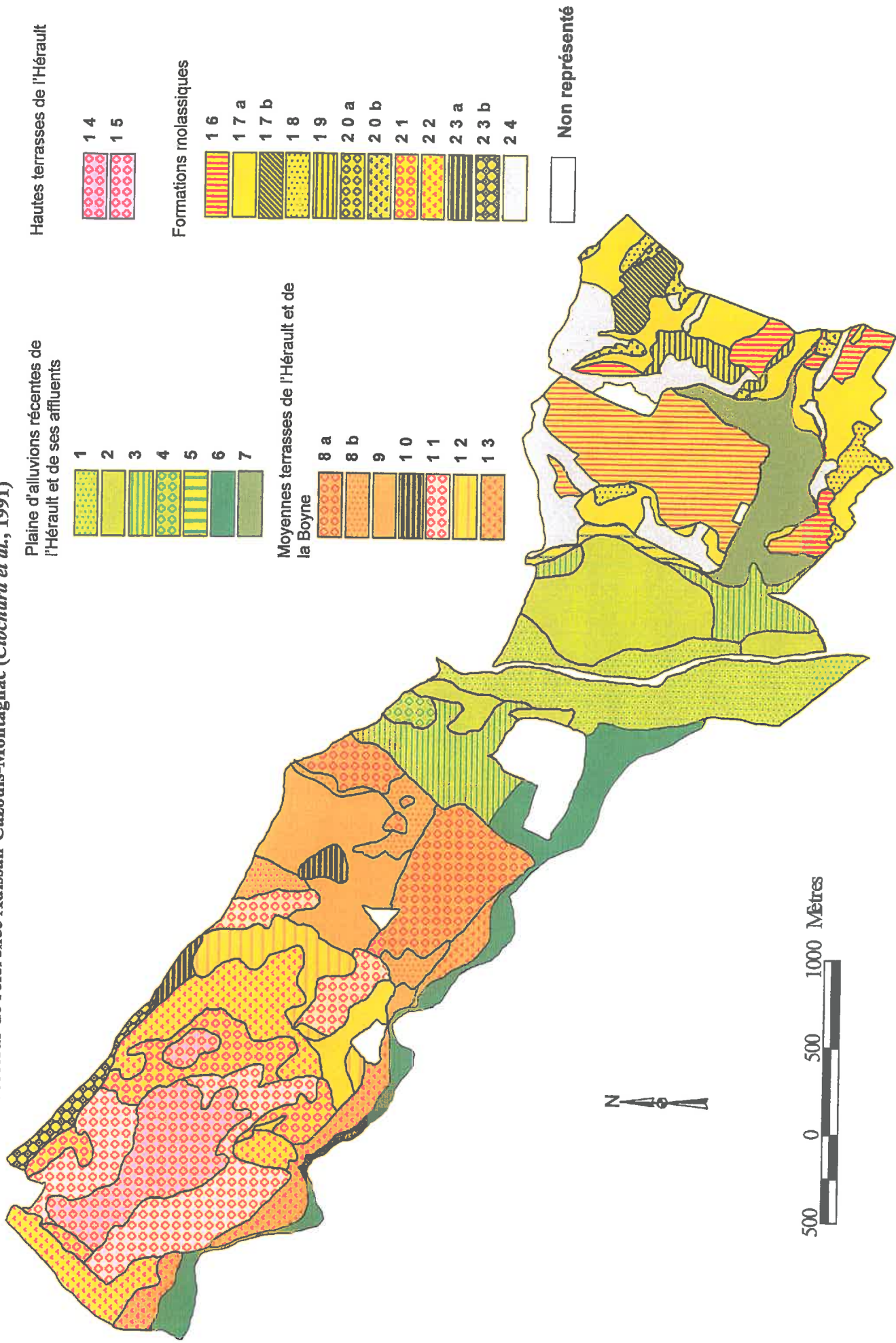
- Tableau 2.1:** La description textuelle concernant les données de l'environnement pour 4 UTS dans la base de données sol au 1/250.000 du Languedoc-Roussillon (Bornand *et al.* 1992).
- Tableau 2.2:** La description textuelle concernant les données purement pédologiques pour 4 UTS dans la base de données sol au 1/250.000 du Languedoc-Roussillon (Bornand *et al.* 1992).
- Tableau 2.3:** Un extrait de la base de données sol au 1/250.000 du Languedoc-Roussillon. L'ensemble des horizons pour une UTS constitue le profil de référence.
- Tableau 3.1:** Les opérations ensemblistes principales.
- Tableau 3.2:** Propriétés principales des opérations ensemblistes avec des sous-ensembles flous.
- Tableau 3.3:** Quelques propriétés simples des mesures de possibilité et de nécessité.
- Tableau 3.4:** Les quatre opérations arithmétiques sur des intervalles fermés.
- Tableau 3.5:** Les propriétés principales des opérations arithmétiques sur des intervalles flous.
- Tableau 4.1:** Variabilité interannuelle de la pluviométrie mesurée au site expérimental du Bassin Versant du Roujan (d'après Tassinari, 1998).
- Tableau 4.2:** Le nombre d'UPP et d'UTS par zone d'étude.
- Tableau 4.3:** Les fonctions de pédotransfert utilisées pour l'estimation de θ_{cr} et θ_{pf} .
- Tableau 5.1:** Pourcentage des observations, issues de la base de profil STIPA, dont le taux d'Argile, Limon ou Sable (mesuré au laboratoire) n'est pas inclus dans la marge d'erreur de 10 % à chaque côté du noyau (le noyau étant défini par la projection des bornes des classes sur les axes du triangle de texture GEPPA).
- Tableau 5.2:** Classes de profondeur d'enracinement (d'après Soil Survey Division Staff, 1993).
- Tableau 5.3:** Classes de pierrosité (d'après Soil Survey Division Staff, 1993).
- Tableau 5.4:** Distributions de possibilité exprimant les différentes classes de teneur en matière organique (exprimées en %) (d'après Legros, 1996).
- Tableau 5.5:** Indicateurs statistiques exprimant les distributions de densité apparente par groupe de matériaux.
- Tableau 5.6:** Distribution de possibilité pour la densité apparente appliquée sur toute la zone d'étude.
- Tableau 5.7:** Les régressions linéaires entre θ_{cr} et θ_{pf} .
- Tableau 5.8:** Les contraintes inter-horizons pour les composants granulométriques.
- Tableau 5.9:** Distribution de possibilité pour les différentes classes de pente.
- Tableau 5.10:** Illustration de la traduction de la description textuelle d'une UTS pour les quatre variables géographiques
- Tableau 6.1:** Le jeu des variables d'entrées dans la définition de CSP afin d'estimer la RU et θ_{pf} .
- Tableau 6.2:** Les pas de discrétisation utilisé pour les variables utilisées dans la présente étude.
- Tableau 6.3:** Origines des descriptions des UTS et nombre d'horizons décrits dans les deux types de descriptions.
- Tableau 6.4:** Ensemble des expérimentations effectuées pour les différentes UTS choisies.

- Tableau 6.5:** Largeur des intervalles estimés pour les trois approches évaluées et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).
- Tableau 6.6:** Taux des bonnes estimations pour les trois approches évaluées et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).
- Tableau 6.7:** Largeur des intervalles estimés pour les deux types d'information évalués et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).
- Tableau 6.8:** Taux des bonnes estimations pour les deux types d'informations évaluées et les deux variables respectives (θ_{pf} à 40 cm et la réserve utile sur 120 cm).
- Tableau 6.9:** Moyennes des largeurs d'intervalles pour les deux groupes d'unités typologiques de sol issus de cartes pédologiques à différente échelle pour les deux variables estimées, par α -coupe et par type d'information utilisé.
- Tableau 6.10:** Moyennes des taux de bonnes estimations pour les deux groupes d'unité typologiques de sol issus de cartes pédologiques à différente échelle pour les deux variables estimées, par α -coupe et par type d'information utilisé.
- Tableau 7.1:** 8 combinaisons de variables géographiques différentes utilisées pour l'évaluation de la localisation des UTS.
- Tableau 7.2:** Somme des rangs pour deux critères ($PEC_{1.0}$ et $NB_UTS_moy_{1.0}$) pour les huit combinaisons de variables évaluées.
- Tableau 7.3:** Les résultats obtenus pour les deux combinaisons 'optimales' (cf. Tableau 7.2).

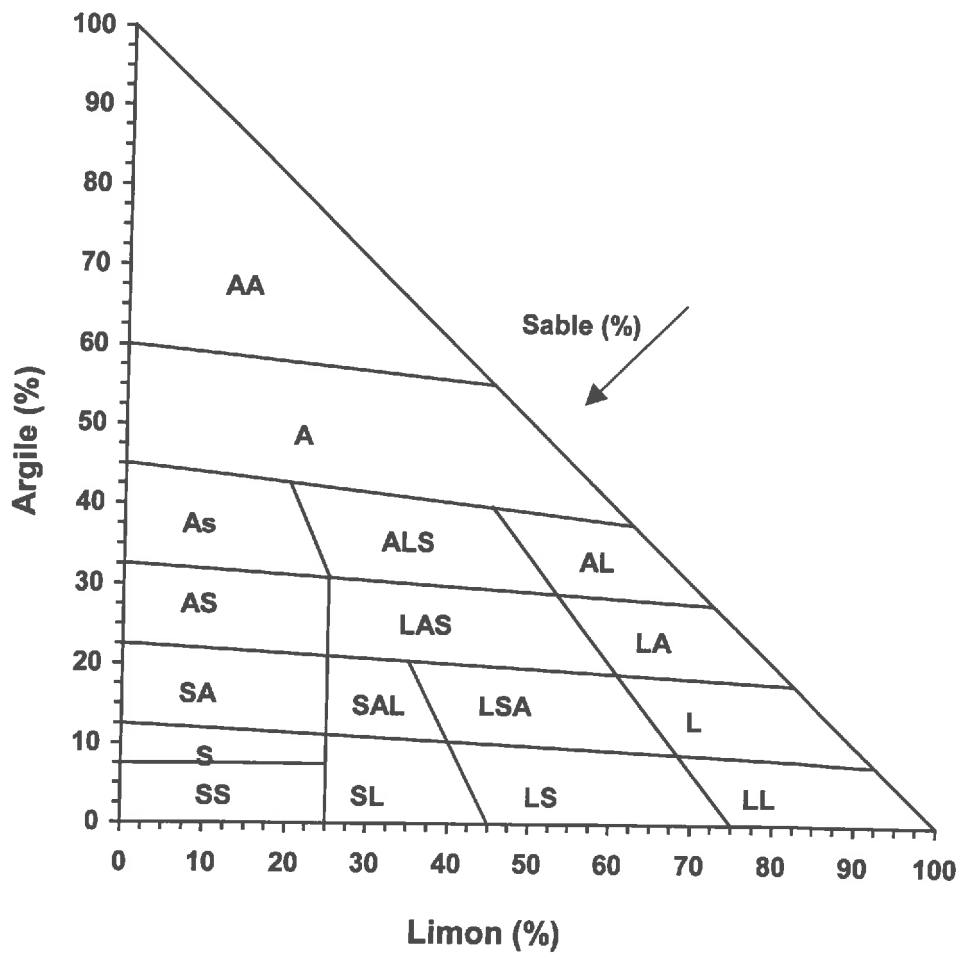
Annexes

- Annexe I:** Extrait de la carte régionale des pédopaysages du Languedoc-Roussillon (M.Bornand *et al.*, 1992). Zone d'étude : vallées de l'Hérault, de l'Orb et du Libron.
- Annexe II:** Carte des sols du secteur de référence Adissan-Cazouls-Montagnac (Clochard *et al.*, 1991).
- Annexe III:** Triangle de texture GEPPA avec les intervalles correspondantes.
- Annexe IV:** Transformation probabilité \rightarrow possibilité pour un histogramme sur un domaine continu.
- Annexe Va:** Exemple de CSP: cas d'un sol composé d'un seul horizon.
- Annexe Vb:** Explications des variables.
- Annexe VI:** Exemple de CSP: cas d'un sol composé de deux horizons.
- Annexe VII:** Estimations de la RU et de θ_{pf} par le CSP et la Simulation de Monte Carlo pour quatre UTS différentes: comparaison avec les mesures de terrain.
- Annexe VIII:** Estimations de la RU et de θ_{pf} par le CSP et le calcul d'intervalles flous pour quatre UTS différentes: comparaison avec les mesures de terrain.
- Annexe IX:** Données utilisées pour l'établissement des fonctions de pédotransfert.
- Annexe X:** Données de validation de la teneur en eau au point de flétrissement pour les points d'observation à l'intérieur de la zone ACM (Série A).
- Annexe XI:** Cartographie de la réserve utile dans la zone d'étude VHOL.
- Annexe XII:** La déclaration des variables pour lmes 8 UTS de validation pour α -coupe 1.

Annexe II: Carte de Sols du secteur de référence Adissan-Cazouls-Montagnac (Clochard *et al.*, 1991)



Annexe III: Triangle de texture GEPPA avec les intervalles correspondantes



Nom	Code	Argile (%)		Limon (%)		Sable (%)	
		min	max	min	max	min	max
argilo-argileux	AA	55	100	0	45	0	40
argileux	A	37.5	60	0	62.5	0	55
argilo-sableux	As/AS	21	45	0	25	37.5	67.5
argilo-limono-sableux	ALS	29	42.5	20	53	15	44
argilo-limoneux	AL	27.5	39.5	45	72.5	0	18
limono-argilo-sableux	LAS	19	31	25	61	18	54
limono-argileux	LA	17.5	29	53	82.5	0	20
sablo-argileux	SA	11	22.5	0	25	54	87.5
sablo-argilo-limoneux	SAL	10	21	25	40	45	64
limono-sablo-argileux	LSA	9	20	35	68	20	50
limoneux	L	7.5	19	61	92.5	0	23
sableux	S	7.5	12.5	0	25	64	92.5
sablo_sableux	SS	0	7.5	0	25	67.5	100
sablo-limoneux	SL	0	11	25	45	50	75
limono-sableux	LS	0	10	40	75	23	55
limono-limoneux	LL	0	9	68	100	0	25

Annexe IV: Transformation probabilité → possibilité pour un histogramme sur un domaine continu

Considérons, $X = \{x_1, x_2, \dots, x_n\}$, $p_i = P(\{x_i\})$, $\pi_i = \Pi(\{x_i\})$ et l'ordre des valeurs de probabilité tel que: $p_1 \geq p_2 \geq \dots \geq p_n$. La solution pour obtenir la distribution de possibilité optimale (i.e. qui perd le moins d'information possible) proposée par Dubois & Prade est la suivante.

$$\forall i = 1, n, \quad \pi_i = \sum_{j=i, n} p_j$$

Par exemple, soit $X = \{x_1, x_2, x_3, x_4\}$ avec $p_1 = 0.4$, $p_2 = 0.3$, $p_3 = 0.2$ et $p_4 = 0.1$. En appliquant la formule 3.21, on obtient $\pi_1 = 1$, $\pi_2 = 0.6$, $\pi_3 = 0.3$ et $\pi_4 = 0.1$.

L'histogramme est une représentation brute de la distribution d'une variable en donnant la fréquence des observations pour un nombre limité de classes. La fréquence relative est ainsi considérée comme valeur de probabilité et un histogramme peut alors être transformé en distribution de possibilité selon la transformation mentionnée ci-dessus. La distribution de possibilité obtenue par cette transformation est influencée par deux aspects affectant la forme de l'histogramme. Premièrement, la forme du histogramme est influencée par la largeur de la classe de fréquence utilisée. Par exemple, la largeur de classe élevée ne permet pas de représenter des variations de probabilité locales au sein de la distribution totale. Deuxièmement, la forme d'un histogramme est influencée par la position des classes sur le domaine de référence.

Pour prendre en compte ces deux facteurs affectant la forme d'un histogramme, il est proposé d'effectuer une transformation d'un histogramme en distribution de possibilité par une série d'opérations:

1. Déterminer la largeur de classe idéale pour représenter les observations. Ceci implique la recherche de l'histogramme uni-modal contenant suffisamment de détail en ce qui concerne la distribution de probabilité.
2. En prenant la largeur déterminée en phase 1, créer une série d'histogrammes avec des positionnements de classes différentes. Dans la présente étude, le nombre d'histogrammes pour une série ne dépassait pas 5.
3. Effectuer pour chaque histogramme de la série la transformation probabilité – possibilité.
4. Déterminer la distribution de possibilité finale en effectuant l'union des distributions de possibilité déterminée dans l'étape 3.

Cette transformation de probabilité → possibilité en ce qui concerne les histogrammes nous a permis d'introduire l'information basée sur une série d'observations dans différents types analyses proposées. Il s'agit ici, par exemple, les données exprimant l'erreur d'estimation des fonctions de pédotransfert ou des données observées et nécessaires pour la validation.

Annexe Va: Exemple de CSP: cas d'un sol composé d'un seul horizon

```
#-----  
# CSP pour l'estimation de la reserve utile et la teneur  
# en eau au point de fletrissement  
#-----  
# UTS 974  
#-----  
# Parametres de requete et options  
#-----  
  
\alpha:                1;  
\filtering:            uf 4;  
\search:               fc first_solution;  
\dynamic_labeling_order: smallest_domain;  
\verbose:              display_solutions;  
  
#-----  
#Declaration des variables  
#-----  
  
\vr : arg      5    [0,0,10.5,21.5];  
\vr : lim      5    [30,40,75,85];  
\vr : sab      5    [13,23,55,65];  
\vr : pier     5    [0,3,25,35];  
\vr : prof     0.5  [2,3,6,7];  
\vr : da       0.1  [1.35, 1.6];  
\vr : w1500    0.5  [0.6,23.6];  
\vr : w10      0.5  [3.6,30.7];  
\vr : ru       5    [0,1000];  
  
#-----  
#CONSTRAINTS  
#-----  
  
\ci: c_somme, arg + lim + sab = 100;  
\ci: c_ordre, w10 > w1500;  
\ci: c_ru , ru = (w10 - w1500) * da * prof * ((100-pier)/100);  
\ci: c_fpt1 , w1500 = 16.20 + 0.20*arg - 6.11*da - 0.05*sab + [-0.5, 2.0];  
\ci: c_fpt2 , w10 = 35.24 + 0.04*lim - 10.53*da - 0.08*sab + [-1.0, 1.5];  
\ci: c_wxw, w10 = 10.46 + 0.772 * w1500 + [-2.5, 0.0];
```

Annexe Vb: Explications des variables

arg	taux d'argile (%)
lim	taux de limon (%)
sab	taux de sable (%)
pier	pierrosité, taux d'éléments grossiers (%)
da	densité apparente
prof	profondeur du sol (en dm)
epais	épaisseur de l'horizon (en dm) (voir annexe VI)
w10	teneur en eau au capacité de rétention (g/100g)
w1500	teneur en eau au point de flétrissement (g/100g)
ru	la réserve utile (mm)

Annexe VI: Exemple de CSP: cas d'un sol composé de deux horizons

```

#-----
# CSP pour l'estimation de la reserve utile et la teneur
# en eau au point de fletrissement
#-----
# UTS 303
# Description par profil de référence
# Alpha coupe : 1.0
#-----
# Parametres de requete et options
#-----

\fuzzy_cut_step =      0.1;
\alpha =              1;
\filtering =          uf 4;
\search =              fc, first_solution 1;
\static_labeling_order = smallest_domain_by_degree;
\dynamic_labeling_order = smallest_domain_by_degree;
\value_order =        edges_first;
\verbose =             display_solutions;

#-----
# Variables
#-----
\vr : sab1      5      [35.0000,60.0000];
\vr : lim1      5      [25.0000,38.0000];
\vr : arg1      5      [11.0000,36.0000];
\vr : pier1     5      [28.0000,55.0000];
\vr : epais1    0.5    [2.5000,6.0000];
\vr : da1       0.1    [1.3500,1.6500];
\vr : w15001    0.5    [0.5,15];
\vr : w101      0.5    [3,23];
\vr : sab2      5      [19.0000,43.0000];
\vr : lim2      5      [22.0000,54.0000];
\vr : arg2      5      [29.0000,59.0000];
\vr : pier2     5      [50.0000,80.0000];
\vr : epais2    0.5    [6.0000,9.5000];
\vr : da2       0.1    [1.3500,1.6000];
\vr : w15002    0.5    [0.5,15];
\vr : w102      0.5    [3,23];
\vr : prof      0.5    [0.0000,10.0000];
\vr : ru        5      [0,1000];

#-----
# Constraints
#-----
\ci : c_sommel , sab1 + lim1 + arg1 = 100;
\ci : c_somme2 , sab2 + lim2 + arg2 = 100;
\ci : c_depth , epais1 + epais2 = prof;
\ci : c_fpt1_1 , w101 = 43.98 - 0.22 * sab1 - 11.12*da1 + [-4.0,4.0];
\ci : c_fpt1_2 , w15001 = 39.9 - 12.3* da1 - 0.20 * sab1 - 0.09 * lim1
                    + [-0.5,2.5];
\ci : c_fpt1_3 , w101 = 7.45 + 1.04 * w15001 + [-2.0,2.5];
\ci : c_fpt2_1 , w102 = 43.98 - 0.22*sab2 - 11.12*da2 + [-4.0,4.0];
\ci : c_fpt2_2 , w15002 = 39.9 - 12.3 * da2 - 0.20 * sab2 - 0.09 * lim2 +
                    [-0.5,2.5];
\ci : c_fpt2_3 , w102 = 7.45 + 1.043 * w15002 + [-2.0,2.5];
\ci : c_int2_2 , da2 >= da1;

```

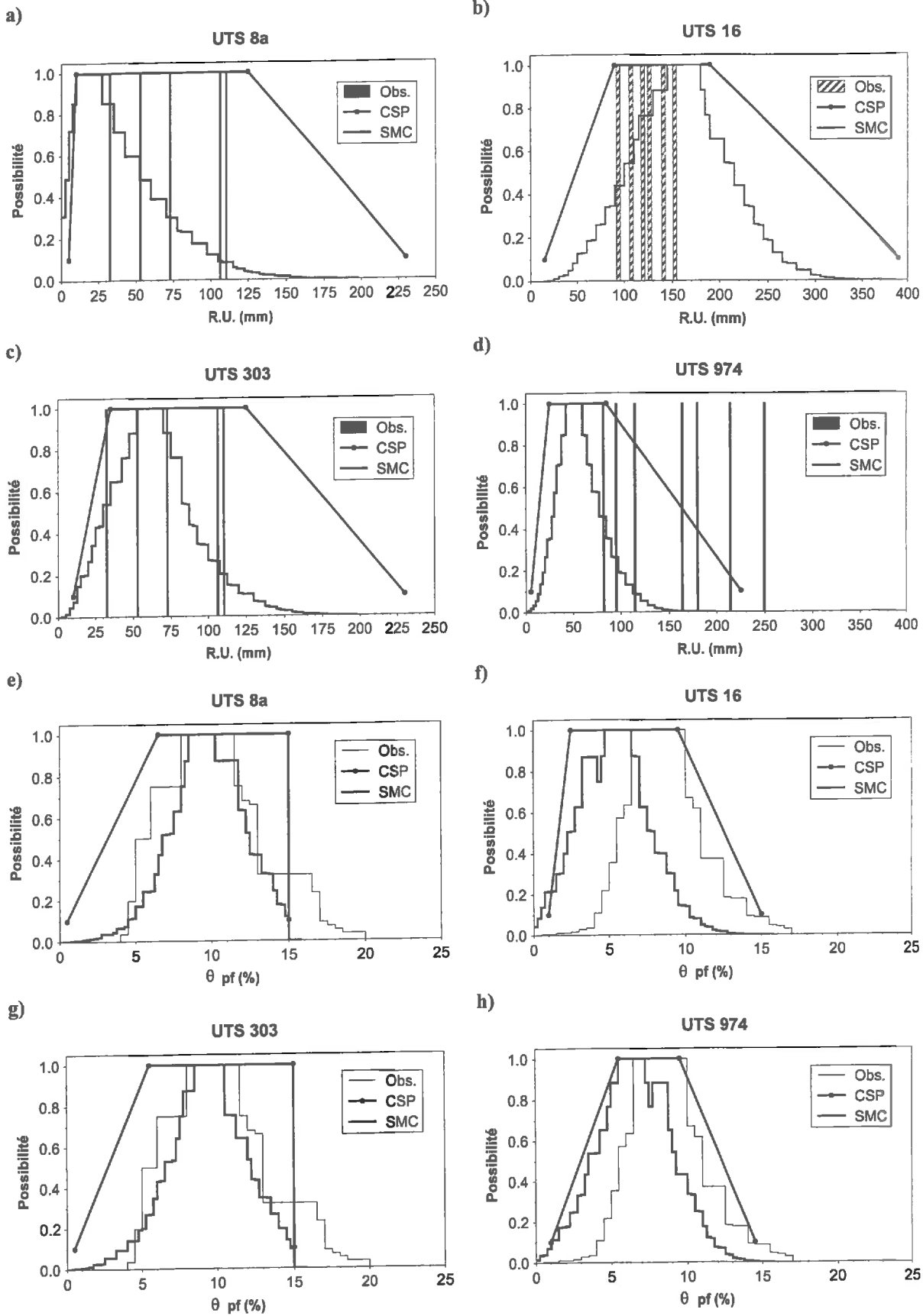
```

\ci : c_int2_6      , sab2 = 4.88 + 0.88 * Sand1 + [-10,7.5];
\ci : c_int2_3      , lim2 = 3.65 + 0.88 * Silt1 + [-10,7.5];
\ci : c_int2_4      , arg2 = 4.53 + 0.85 * Clay1 + [-10,7.5];
\ci : c_int2_5      , w102 = 5.64 + 0.67 * w101 + [-2,2.5];
\ci : c_int2_1      , w15002 = 2.154 + 0.793 * w15001 + [-2,2.5];
\ci : c_w10w1500_1 , w101 > w15001;
\ci : c_w10w1500_2 , w102 > w15002;

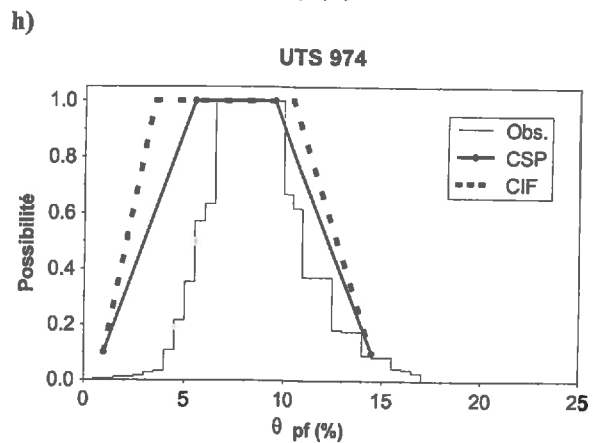
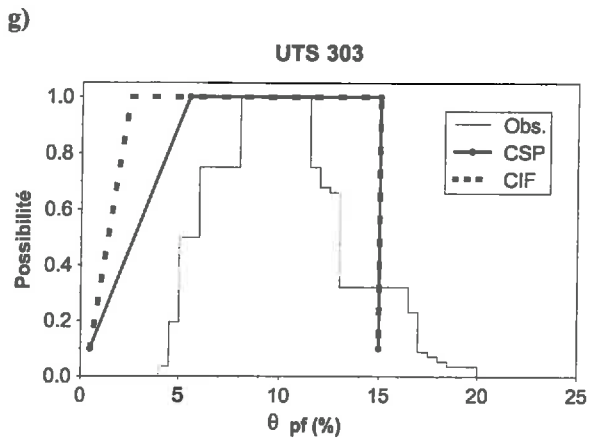
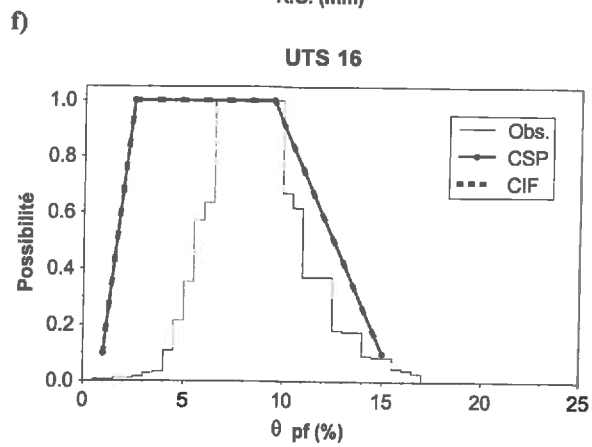
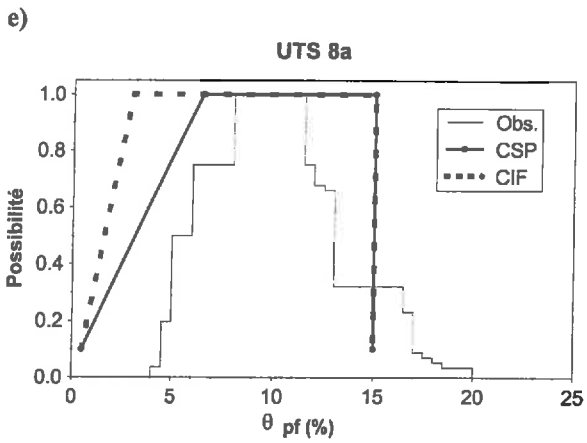
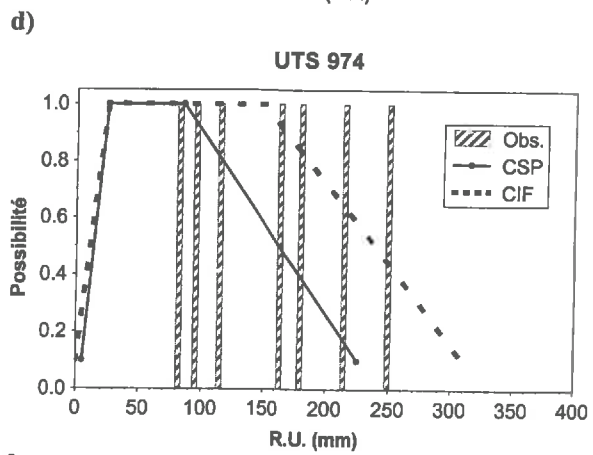
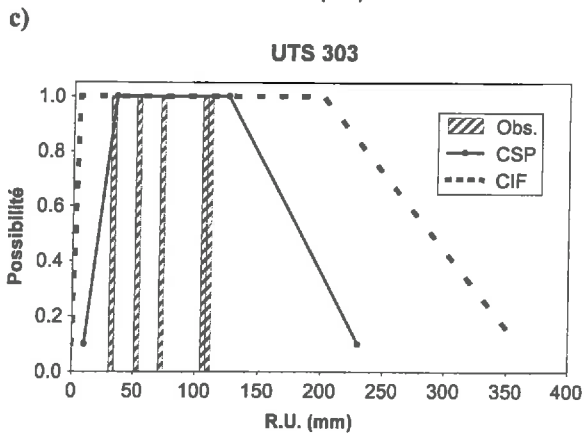
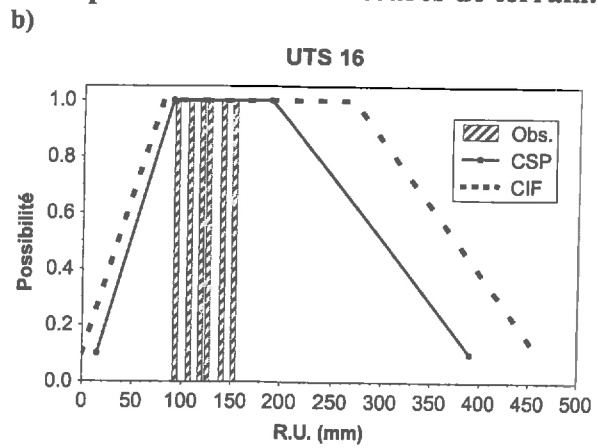
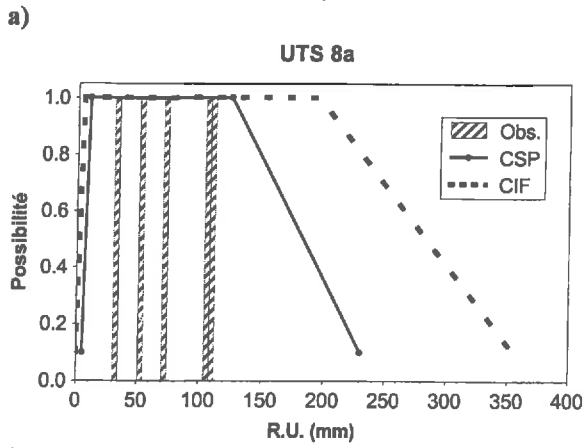
#-----
# Contraintes définies par morceaux pour
# trouver la solution pour la réserve utile
#-----
\cp : CP1
  \for
    \ci : c_ru_2_1 , 12 <= epais1;
  \do
    \ci : c_ru_res_2_1 , ru = (w101-w15001)*12*da1*((100-pier1)/100);
  \for
    \ci : c_ru_2_2 , 12 > epais1;
    \ci : c_ru_2_3 , 12 <= epais1 + epais2;
  \do
    \ci : c_ru_res_2_2 , ru = (w101-w15001) * epais1*da1*((100-pier1)/100)
      + (w102-w15002)*(12-epais1)*da2*((100-
      pier2)/100);
  \for
    \ci : c_ru_2_4 , 12 > epais1 + epais2;
  \do
    \ci : c_ru_res_2_3 , ru = (w101-w15001)*epais1*da1*((100-pier1)/100) +
      (w102-w15002)*epais2*da2*((100-pier2)/100);
;

```

Annexe VII: Estimations de la RU et de θ_{pf} par le CSP et la Simulation Monte Carlo pour quatre UTS différentes: comparaison avec les mesures de terrain.



Annexe VIII: Estimations de la RU et de θ_{pf} par le CSP et le calcul d'intervalles flous pour quatre UTS différentes: comparaison avec les mesures de terrain.

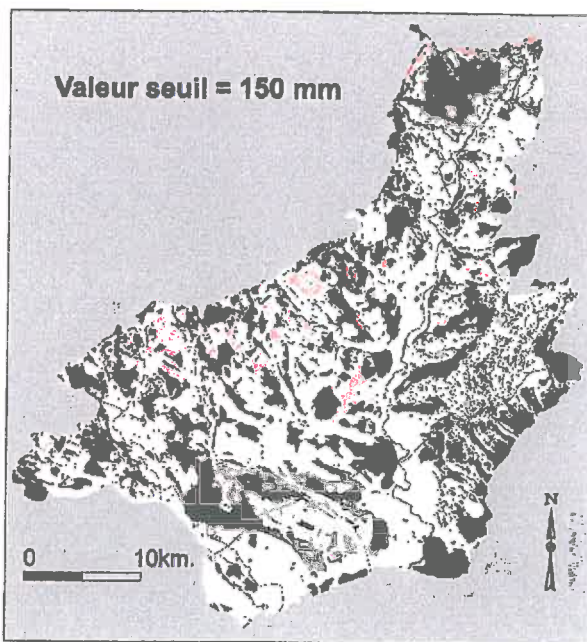
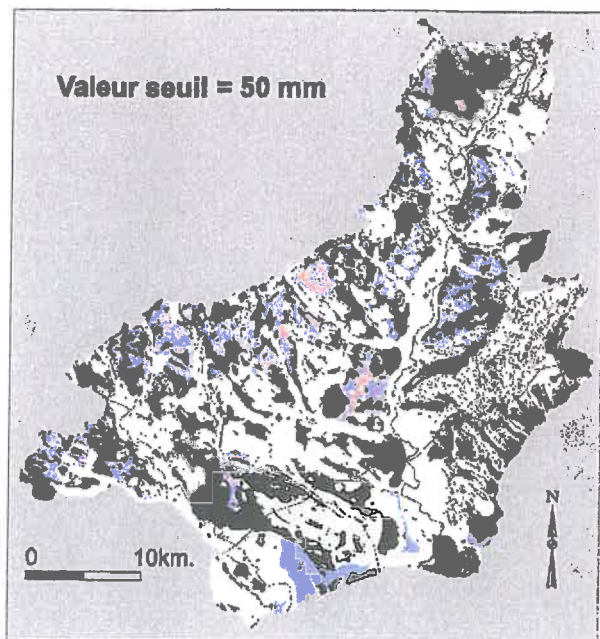


Annexe X: Données de validation de la teneur en eau au point de flétrissement pour les points d'observation à l'intérieur de la zone ACM (Série A)

NUM	XCOORD(m)	YCOORD(m)	θpf(%)	UTS250	NUM	XCOORD(m)	YCOORD(m)	θpf(%)	UTS250
1	687537	136625	4.35	0	125	690657	134305	7.52	0
2	687760	136590	6.93	0	126	691044	133265	5.90	102
3	687646	136730	11.14	0	127	691062	133510	6.36	102
4	687760	136821	7.67	974	128	690660	134522	10.36	0
5	687763	137029	10.22	345	129	690446	134488	6.73	0
6	687540	137031	8.44	974	130	690248	134500	6.42	0
7	687522	136811	11.10	0	131	690242	134720	6.12	0
8	687740	137204	5.70	974	132	690431	135120	8.97	303
9	687943	137435	5.18	974	133	690363	135211	7.67	303
10	687996	137227	15.79	345	134	690481	135362	11.28	303
11	688164	137227	5.26	345	135	690666	135123	12.24	103
12	688156	136991	5.79	345	136	690673	135335	11.67	103
13	687978	137001	5.32	345	137	690767	135236	13.34	103
14	687976	136881	4.13	345	138	690871	135343	9.81	103
15	688151	136800	9.89	345	139	690861	135120	10.35	101
16	688337	137427	15.04	11	140	691078	135129	5.96	101
17	688360	137192	15.65	345	141	690962	135047	8.58	101
18	688362	136800	6.85	345	142	691078	134919	8.41	102
19	688360	136978	11.30	345	143	690956	134799	10.25	101
20	688593	137189	6.07	345	144	691075	134724	9.89	102
21	688457	137089	5.02	345	145	691062	134529	7.81	102
22	688563	137001	5.44	345	146	690959	134620	9.31	101
23	688791	136999	11.27	11	147	690877	134510	10.51	102
24	688801	136788	13.36	11	148	690855	134937	9.10	103
25	688585	136808	16.31	345	149	693521	134304	6.48	974
26	688523	136873	2.17	345	150	693706	134329	7.72	974
27	688783	136832	8.00	345	151	693533	134525	5.71	0
28	688583	136592	17.10	345	152	693336	134295	16.37	0
29	688585	136392	6.72	345	153	693120	134298	6.01	514
30	688791	136382	15.12	345	154	693134	134082	9.41	0
31	687988	136191	7.94	0	155	693235	134197	8.30	0
32	688156	136008	9.12	0	156	693339	134085	6.79	974
33	688365	136001	15.74	0	157	693533	134082	6.51	0
34	688370	136199	13.97	974	158	693331	133900	8.03	974
35	688380	136379	8.49	345	159	693129	133905	16.52	0
36	688304	136475	10.13	345	160	693131	133740	11.87	974
37	688375	136605	4.34	345	161	693134	133485	6.40	974
38	688171	136625	6.29	345	162	693129	133299	9.34	974
39	687961	136610	8.23	345	163	692904	133490	8.82	0
40	687946	136397	4.48	0	164	693011	133619	10.70	974
41	688091	136512	11.15	345	165	692913	133729	9.63	0
42	688171	136389	4.90	345	166	692812	133616	7.83	0
43	688149	136186	8.82	345	167	693011	133201	9.93	974
44	689097	136485	10.79	345	168	693140	133109	9.27	0
45	689001	136595	7.50	345	169	693126	132924	7.76	974
46	688979	136785	12.04	11	170	692910	132912	8.16	974
47	689202	136770	8.25	974	171	692714	132915	6.46	974
48	689423	136572	8.36	974	172	692915	133103	6.30	974
49	689413	136359	8.16	974	173	692901	133308	8.35	974
50	689295	136452	11.55	974	174	692708	133313	8.34	974
51	689197	136585	14.32	974	175	692820	133210	5.79	974
52	689194	136387	5.25	11	176	692722	133109	8.03	974
53	688999	136369	9.65	11	177	692498	133083	6.72	974
54	688590	136206	12.49	11	178	692296	133100	6.82	0
55	688786	136169	13.61	345	179	692091	133117	8.28	974
56	688901	136246	13.30	345	180	691897	133280	7.24	506
57	689215	135544	5.94	974	181	691903	133501	6.90	506
58	689001	135751	5.21	974	182	691795	133607	6.63	506
59	688797	135580	8.58	0	183	692094	133283	7.91	506
60	688782	135775	9.97	974	184	692298	133490	6.78	506
61	688580	135778	8.45	0	185	692113	133482	7.12	506
62	688593	135988	8.52	11	186	691903	133715	6.39	974
63	688914	135884	10.50	974	187	692318	133299	7.97	506
64	688793	135993	16.25	345	188	692509	133313	7.03	506
65	689444	135538	7.50	0	189	692705	133513	5.43	506
66	689633	135544	12.30	303	190	692503	133493	8.31	974
67	689416	135766	7.62	0	191	692708	134716	7.04	514
68	689215	135778	8.01	0	192	692495	134710	9.45	514
69	689335	135902	6.19	0	193	692517	134503	7.72	974
70	689438	135976	8.94	0	194	692716	134500	9.12	974
71	689212	136179	5.93	974	195	692921	134301	6.83	0
72	688996	136179	8.53	11	196	692927	134074	16.00	0
73	689021	135996	7.11	974	197	692719	134304	7.84	974
74	689225	135976	8.94	0	198	692514	134309	7.13	974
75	689400	136181	7.47	974	199	692509	134104	8.92	974
76	689616	136159	8.96	0	200	692315	134318	8.45	974
77	689608	136372	9.61	0	201	692310	134503	6.73	514
78	689816	136367	6.99	0	202	692310	134096	9.27	974
79	689811	136156	6.02	0	203	692203	134427	5.95	974
80	689626	135966	8.83	0	204	692119	134511	7.54	514

81	689824	135966	9.48	0	205	692113	134323	7.33	974
82	689744	135860	9.14	305	206	692108	134099	5.56	514
83	689633	135766	9.50	305	207	692113	133902	6.66	974
84	689856	135775	7.85	305	208	692102	133712	8.09	974
85	689856	135556	9.84	305	209	692216	133614	7.03	974
86	690042	136164	6.22	303	210	692296	133717	7.89	974
87	689944	136041	6.79	303	211	692206	133813	6.86	974
88	690047	135983	11.27	119	212	692318	133914	7.94	974
89	690253	136151	9.99	305	213	692405	134017	8.43	974
90	690666	135943	13.88	303	214	692512	133905	7.07	974
91	690656	135745	7.23	303	215	692714	133919	7.93	974
92	690060	135748	11.59	119	216	692716	133734	7.86	506
93	689961	135652	9.25	305	217	692514	133706	7.99	974
94	690063	135532	13.68	305	218	691499	133497	7.64	103
95	690265	135541	11.76	305	219	691493	133281	7.32	103
96	690301	135778	10.14	305	220	691698	133507	7.33	103
97	690145	135866	10.51	119	221	691496	133721	7.57	103
98	690255	135971	10.03	305	222	691672	133721	6.54	103
99	690473	135751	7.07	305	223	691396	133818	5.66	102
100	690566	135842	12.91	303	224	691506	133922	6.37	101
101	690466	135953	12.77	305	225	691301	134098	6.70	102
102	690482	135577	12.02	305	226	691509	134095	8.02	101
103	690662	135538	9.49	303	227	691691	134095	7.24	101
104	690864	135556	13.09	103	228	691597	134013	6.81	101
105	689405	135342	7.62	303	229	691682	133906	6.29	101
106	689624	135317	5.86	303	230	691798	134000	6.95	101
107	689614	135115	5.70	0	231	691921	134104	5.83	0
108	689634	134933	7.45	0	232	691896	134290	8.18	101
109	689727	135034	9.97	0	233	691993	134419	6.41	0
110	689816	134923	8.88	0	234	691883	134510	5.82	103
111	689823	135140	6.34	303	235	691493	134730	6.07	101
112	689816	135332	6.63	303	236	691679	134708	5.87	101
113	690057	135145	9.75	303	237	691675	134507	7.12	101
114	690141	135221	6.54	303	238	691694	134290	6.34	101
115	690035	135330	9.53	303	239	691597	134214	6.53	101
116	690257	135335	11.18	303	240	691493	134299	6.48	101
117	690274	135135	8.67	303	241	691377	134230	7.87	101
118	690153	135027	10.02	303	242	691276	134315	6.44	102
119	691078	133890	6.15	102	243	691374	134419	8.64	101
120	690845	133909	8.43	0	244	691477	134522	6.60	101
121	690839	134104	8.24	0	245	691270	134510	5.26	101
122	691072	134120	5.92	102	246	688274	136096	12.55	345
123	691069	134321	6.07	102	247	688477	136101	15.43	974
124	690858	134302	8.43	0					

Annexe XI: Cartographie de la réserve utile dans la zone d'étude VHOL.



 RU sûrement > valeur seuil

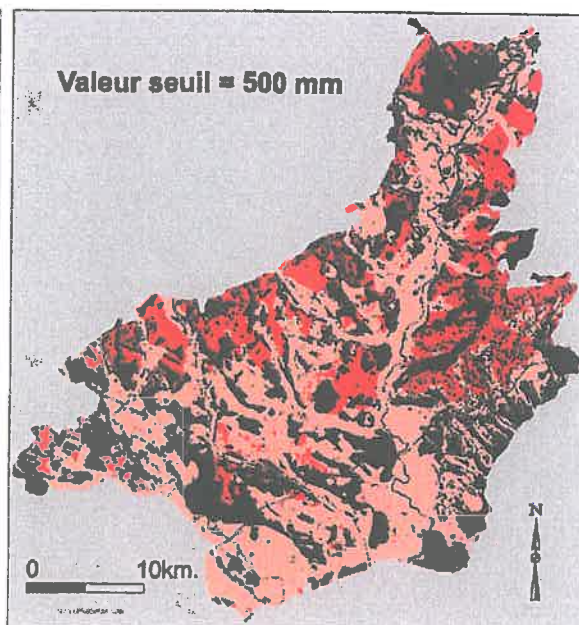
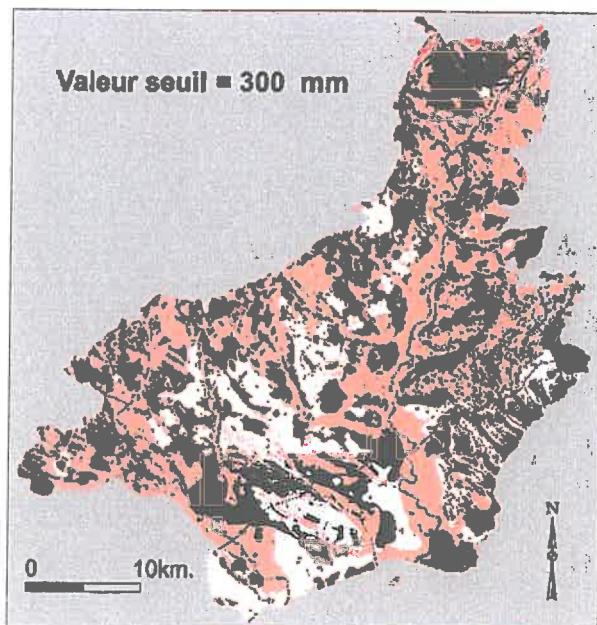
 RU vraisemblablement < valeur seuil

 RU vraisemblablement > valeur seuil

 RU sûrement < valeur seuil

 Ignorance

 Situation non renseignée



Annexe XII: La déclaration des variables pour les 8 UTS de validation pour α -coupe 1 et seulement

UTS 2

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [23.0000,75.0000];  
\vr : Silt1      5      [25.0000,75.0000];  
\vr : Clay1      5      [0.0000,11.0000];  
\vr : Stoniness1 5      [0.0000,0.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [6.0000,8.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.8,19.2];  
\vr : W101       0.5    [3.6,31.9];  
  
\vr : Sand2      5      [0.0000,23.0000];  
\vr : Silt2      5      [61.0000,92.5000];  
\vr : Clay2      5      [7.5000,19.0000];  
\vr : Stoniness2 5      [0.0000,0.0000];  
\vr : Organic2   0.5    [0.0000,1.0000];  
\vr : Thickness2 0.5    [5.0000,7.0000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [0.8,19.2];  
\vr : W102       0.5    [3.6,31.9];  
  
\vr : Depth      0.5    [12.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [39.0000,54.0000];  
\vr : Silt1      5      [37.0000,47.0000];  
\vr : Clay1      5      [9.0000,14.0000];  
\vr : Stoniness1 5      [0.0000,0.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [12.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.8,19.2];  
\vr : W101       0.5    [3.6,31.9];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

UTS 8A

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [18.0000,54.0000];  
\vr : Silt1      5      [25.0000,61.0000];  
\vr : Clay1      5      [19.0000,31.0000];  
\vr : Stoniness1 5      [36.0000,90.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [10.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.5,15];  
\vr : W101       0.5    [3,23];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC         5      [0,1000];  
\vr : VarOut      0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [46.0000,52.0000];  
\vr : Silt1      5      [28.0000,30.0000];  
\vr : Clay1      5      [19.0000,25.0000];  
\vr : Stoniness1 5      [40.0000,50.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [4.0000,5.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.5,15];  
\vr : W101       0.5    [3,23];  
  
\vr : Sand2      5      [44.0000,57.0000];  
\vr : Silt2      5      [27.0000,35.0000];  
\vr : Clay2      5      [15.0000,21.0000];  
\vr : Stoniness2 5      [65.0000,80.0000];  
\vr : Organic2   0.5    [0.0000,1.0000];  
\vr : Thickness2 0.5    [4.0000,8.0000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [0.5,15];  
\vr : W102       0.5    [3,23];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC         5      [0,1000];  
\vr : VarOut      0.5    [0,50];
```

UTS 14

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [0.0000,67.5000];  
\vr : Silt1      5      [0.0000,62.5000];  
\vr : Clay1      5      [31.0000,60.0000];  
\vr : Stoniness1 5      [61.0000,90.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [7.0000,9.0000];  
\vr : Density1   0.1    [0.0000,100.0000];  
  
\vr : W15001     0.5    [1.4,19.5];  
\vr : W101       0.5    [6.2,31];  
  
\vr : Sand2      5      [0.0000,55.0000];  
\vr : Silt2      5      [0.0000,62.5000];  
\vr : Clay2      5      [37.5000,60.0000];  
\vr : Stoniness2 5      [61.0000,90.0000];  
\vr : Organic2   0.5    [0.0000,1.0000];  
\vr : Thickness2 0.5    [3.0000,5.0000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [1.4,19.5];  
\vr : W102       0.5    [6.2,31];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [42.0000,59.0000];  
\vr : Silt1      5      [25.0000,32.0000];  
\vr : Clay1      5      [16.0000,26.0000];  
\vr : Stoniness1 5      [50.0000,60.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [5.0000,6.5000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [1.4,19.5];  
\vr : W101       0.5    [6.2,31];  
  
\vr : Sand2      5      [33.0000,45.0000];  
\vr : Silt2      5      [14.0000,25.0000];  
\vr : Clay2      5      [30.0000,53.0000];  
\vr : Stoniness2 5      [60.0000,60.0000];  
\vr : Organic2   0.5    [0.0000,1.0000];  
\vr : Thickness2 0.5    [3.5000,5.0000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [1.4,19.5];  
\vr : W102       0.5    [6.2,31];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

UTS 16

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [50.0000,75.0000];  
\vr : Silt1      5      [25.0000,45.0000];  
\vr : Clay1      5      [0.0000,11.0000];  
\vr : Stoniness1 5      [0.0000,2.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [10.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.6,23.6];  
\vr : W101       0.5    [3.6,30.7];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [40.0000,54.0000];  
\vr : Silt1      5      [31.0000,50.0000];  
\vr : Clay1      5      [9.0000,16.0000];  
\vr : Stoniness1 5      [0.0000,1.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [5.0000,7.5000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.6,23.6];  
\vr : W101       0.5    [3.6,30.7];  
  
\vr : Sand2      5      [36.0000,43.0000];  
\vr : Silt2      5      [34.0000,59.0000];  
\vr : Clay2      5      [5.0000,23.0000];  
\vr : Stoniness2 5      [0.0000,0.0000];  
\vr : Organic2   0.5    [0.0000,1.0000];  
\vr : Thickness2 0.5    [4.0000,4.5000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [0.6,23.6];  
\vr : W102       0.5    [3.6,30.7];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```


UTS 102

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [23.0000,100.0000];  
\vr : Silt1      5      [0.0000,75.0000];  
\vr : Clay1      5      [0.0000,12.5000];  
\vr : Stoniness1 5      [0.0000,0.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [12.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.8,19.2];  
\vr : W101       0.5    [3.6,31.9];  
  
\vr : Depth      0.5    [12.0000,12.0000];  
\vr : WSC         5      [0,1000];  
\vr : VarOut      0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [38.0000,64.0000];  
\vr : Silt1      5      [28.0000,52.0000];  
\vr : Clay1      5      [6.0000,14.0000];  
\vr : Stoniness1 5      [0.0000,1.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [12.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.8,19.2];  
\vr : W101       0.5    [3.6,31.9];  
  
\vr : Depth      0.5    [12.0000,12.0000];  
\vr : WSC         5      [0,1000];  
\vr : VarOut      0.5    [0,50];
```

UTS 303

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [18.0000,54.0000];  
\vr : Silt1      5      [25.0000,68.0000];  
\vr : Clay1      5      [9.0000,31.0000];  
\vr : Stoniness1 5      [36.0000,60.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [10.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [0.5,15];  
\vr : W101       0.5    [3,23];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [35.0000,60.0000];  
\vr : Silt1      5      [25.0000,38.0000];  
\vr : Clay1      5      [11.0000,36.0000];  
\vr : Stoniness1 5      [28.0000,55.0000];  
#\vr : Organic1  0.5    [0.0000,0.0000];  
\vr : Thickness1 0.5    [2.5000,6.0000];  
\vr : Density1   0.1    [1.3500,1.6500];  
  
\vr : W15001     0.5    [0.5,15];  
\vr : W101       0.5    [3,23];  
  
\vr : Sand2      5      [19.0000,43.0000];  
\vr : Silt2      5      [22.0000,54.0000];  
\vr : Clay2      5      [29.0000,59.0000];  
\vr : Stoniness2 5      [50.0000,80.0000];  
#\vr : Organic2  0.5    [0.0000,0.0000];  
\vr : Thickness2 0.5    [6.0000,9.5000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [0.5,15];  
\vr : W102       0.5    [3,23];  
  
\vr : Depth      0.5    [0.0000,10.0000];  
#\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

UTS 345

Description textuelle

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [0.0000,67.5000];  
\vr : Silt1      5      [0.0000,82.5000];  
\vr : Clay1      5      [17.5000,45.0000];  
\vr : Stoniness1 5      [36.0000,60.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [10.0000,12.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [6.2,31];  
\vr : W101       0.5    [1.4,19.5];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

Description par profil de référence

```
#-----  
# Variables  
#-----  
\vr : Sand1      5      [42.0000,62.0000];  
\vr : Silt1      5      [16.0000,32.0000];  
\vr : Clay1      5      [22.0000,30.0000];  
\vr : Stoniness1 5      [30.0000,60.0000];  
\vr : Organic1   0.5    [0.0000,1.0000];  
\vr : Thickness1 0.5    [4.0000,5.0000];  
\vr : Density1   0.1    [1.3500,1.6000];  
  
\vr : W15001     0.5    [6.2,31];  
\vr : W101       0.5    [1.4,19.5];  
  
\vr : Sand2      5      [35.0000,45.0000];  
\vr : Silt2      5      [15.0000,25.0000];  
\vr : Clay2      5      [30.0000,50.0000];  
\vr : Stoniness2 5      [8.0000,60.0000];  
\vr : Organic2   0.5    [0.0000,1.0000];  
\vr : Thickness2 0.5    [7.0000,8.0000];  
\vr : Density2   0.1    [1.3500,1.6000];  
  
\vr : W15002     0.5    [6.2,31];  
\vr : W102       0.5    [1.4,19.5];  
  
\vr : Depth      0.5    [10.0000,12.0000];  
#\vr : WSC        5      [0,1000];  
\vr : VarOut     0.5    [0,50];
```

UTS 974

Description textuelle

```
#-----  
#      VARIABLES DOMAINS TRANSLATING THE INPUT DATA  
#-----  
\vr : arg 5 [0,0,10.5,21.5];  
  
\vr : lim 5 [30,40,75,85];  
\vr : sab 5 [13,23,55,65];  
\vr : pier 5 [0,3,25,35];  
\vr : prof 0.5 [2,3,6,7];  
\vr : da      0.1   [1.35, 1.6];  
\vr : w1500   0.5   [0.6,23.6];  
\vr : w10     0.5   [3.6,30.7];  
\vr : ru      5     [0,1000];
```

Description par profil de référence

```
#-----  
#      VARIABLE DOMAINS TRANSLATING THE INPUT DATA  
#-----  
\vr : arg1     5     [0,5,30,40] ;  
\vr : arg2     5     [0,5,30,40] ;  
\vr : arg3     5     [0,1,20,30] ;  
\vr : lim1     5     [5,15,60,70] ;  
\vr : lim2     5     [10,20,60,70] ;  
\vr : lim3     5     [10,20,40,50] ;  
\vr : sab1     5     [10,20,80,90] ;  
\vr : sab2     5     [15,25,75,95] ;  
\vr : sab3     5     [40,50,80,90] ;  
\vr : pier1    5     [0,4,30,40];  
\vr : pier2    5     [0,1,45,55];  
\vr : pier3    5     [60,70,95,100];  
  
\vr : epais1   0.5   [0,0.5,5.5,6.5];  
\vr : epais2   0.5   [0,1,4,5];  
\vr : epais3   0.5   [0,0,10,10];  
\vr : prof     0.5   [2,3,6,7];  
  
\vr : da1     0.1   [1.35, 1.6];  
\vr : da2     0.1   [1.35, 1.6];  
\vr : da3     0.1   [1.35, 1.6];  
  
\vr : w101    0.5   [3.6, 30.7];  
\vr : w15001  0.5   [0.6, 23.6];  
\vr : w102    0.5   [3.6, 30.7];  
\vr : w15002  0.5   [0.6, 23.6];  
\vr : w103    0.5   [3.6, 30.7];  
\vr : w15003  0.5   [0.6, 23.6];  
\vr : w150040 0.5   [0.6, 23.6];  
\vr : ru      5     [0,1000];
```

**Annexe I: EXTRAIT DE LA CARTE REGIONALE DES PEDOPAYSAGES
DU LANGUEDOC-ROUSSILLON (M. BORNAND et al., 1992)
Zone d'étude: vallées de l'Hérault, de l'Orb et du Libron**

LEGENDE

Plaines alluviales récentes

Plaines littorales

--	--

**Terrasses d'alluvions anciennes
Glacis plio-quaternaires**

Plateaux et collines tabulaires

Bassins, Collines, Versants

Serres de type cévenol

--	--

