

INSTITUT NATIONAL AGRONOMIQUE EL-HARRACH - ALGER
THÈSE Présentée en vue de l'obtention du diplôme de Magister en Sciences Agronomiques
Spécialité : Science du sol

***Contribution à la détermination des unités
homogènes du sol par différentes
approches de cartographie numérique***

Par :

M. Mohamed HAMOUNI

M. Douaoui A. Maître de Conférences Directeur de thèse Uni. Hassiba Benbouali,
Chlef

03-07- 2006

Soutenue devant le jury composé de : **M. Djili K.** Professeur Président I.N.A, Alger **M. Aidaoui
A.** Professeur Examineur I.N.A, Alger **M. Hartani T.** Maître de Conférences Examineur I.N.A,
Alger **M. Habila M.** Directeur Invité INSID

Table des matières

..	1
Avant-propos .	3
صـخـلـم .	5
Résumé .	7
Summary . .	9
INTRODUCTION .	11
Chapitre I . Présentation du milieu physique . .	13
1-1 Situation géographique de la zone d'étude .	13
1.2 . Milieu naturel . .	14
1.2.1 . Climat de la région de Hmadna . .	15
1.2.2 . Géologie de la région de Hmadna . .	18
1.2.3 . Hydrogéologie . .	19
1.2.4 . Topographie de la zone d'étude .	20
1.2.5 . La végétation . .	20
1.2.6 . Actions anthropiques . .	20
Chapitre II . La cartographie pédologique de la zone d'étude .	23
2.1 . Les contraintes de la cartographie pédologique . .	23
2.2 . Caractérisation pédologique de la zone d'étude .	24
2.2.1 . Le document cartographique de base .	24
2.2.2 . Les caractéristiques pédologiques des unités cartographiques . .	25
Conclusion . .	29
Chapitre III . Matériel et méthodes .	31
3.1 . Matériel d'étude . .	31
3.1.1 . Dispositif expérimental .	31
3.1.2 . Justification du plan d'échantillonnage .	32
3.1.3 . Justification du pas d'échantillonnage . .	32

3.1.4 . La profondeur du sol .	33
3.1.5 . Les variables étudiées .	34
3.2 . Méthodes d'études .	35
3.2.1 . Analyse en composantes principales . .	35
3.2.2 . Classification ascendante hiérarchique . .	36
3.2.3 . Analyse factorielle discriminante .	38
3.2.4 . La géostatistique .	39
Chapitre IV . Description statistique des variables du sol .	43
4.1 . Description de l'hétérogénéité spatiale de la parcelle d'étude . .	43
4.2 . Analyse statistique classique .	44
4.2.1 . Recherche de la nature de la loi de distribution .	44
4.2.2 . Analyse des niveaux de la variabilité spatiale des variables du sol dans la parcelle d'étude .	47
Conclusion . .	48
Chapitre V . Détermination des unités homogènes du sol par l'analyse des composantes principales . .	49
5.1 . Notation et résultats préliminaires .	49
5.1.1 . La matrice des données de base .	50
5.1.2 . Relation entre les variables du sol .	50
5.2 . Caractéristiques des axes des composantes principales .	51
5.3 . Analyse de la relation entre les individus . .	52
5.3.1 . La relation entre les variables du sol .	52
5.3.2 . La relation entre les sondages du sol .	54
Conclusion . .	55
Chapitre VI . Contribution de la classification ascendante hiérarchique à la détermination des unités homogènes du sol . .	57
6.1 . La structure des données de base .	58
6.2 . Résultats de la partition des classes de sondage du sol .	59
6.2.1 . Exploration du dendrogramme de la partition . .	59
6.2.2 . Algorithme de la partition des classes de sondage .	62

6.2.3 . Le choix du nombre de classes de sondage du sol .	64
6.2.4 . Caractéristiques des classes de sondage du sol issues de la partition .	65
6.3 . Représentation cartographique des classes de sondage . .	70
Conclusion . .	71
Chapitre VII . Contribution de l'analyse factorielle discriminante à la détermination des unités homogènes du sol .	73
7.1 . Présentation des résultats préliminaires .	73
7.1.1 . La structure des données de base .	74
7.1.2 . Analyse de la variance . .	74
7.2 . Analyse factorielle discriminante (AFD) proprement dite . .	74
7.2.1 . Résultats de l'AFD .	75
7.2.2. Corrélations inter-classes entre les variables du sol et les axes discriminants .	76
7.2.3 . Détermination du degré de dissimilation entre les classes de sondage .	77
7.2.4 . Affectation des sondages du sol dans leurs classes représentatives .	77
7.3 . Caractéristiques des classes de sondage du sol issues de l'AFD .	78
7.4 . Répartition géographique des classes de sondage du sol . .	81
Conclusion . .	83
Chapitre VIII .La géostatistique multiparamétrique . .	85
8.1. Structure des données de base .	86
8.2. Analyse de la structure spatiale du premier axe discriminant .	87
8.2.1. Calcul du variogramme expérimental .	87
8.2.2. Modélisation du variogramme expérimental . .	89
8.2.3. Tendances anisotropiques du premier axe discriminant .	90
8.3. Résultats de l'analyse variographique des deuxième et troisième axes discriminants .	92
8.4. Interpolation par krigeage . .	94
8.4.1. Mise en oeuvre du krigeage .	94
8.4.2. Résultats du krigeage .	94

8.4.3. La validation croisée . .	100
Conclusion . .	101
Chapitre IX. Analyse de la pureté des unités homogènes du sol .	103
9.1. Pureté partielle et pureté moyenne d'une unité cartographique . .	104
9.2. Analyse de la pureté des unités du sol .	105
Conclusion . .	108
Synthèse et Discussion .	109
Conclusion générale .	119
Références bibliographiques .	123

A ma mère mon père ma grand-mère adorée ma chère aimée mes sœurs et mes frères A la mémoire de notre cher ami et collègue Sider Farid Je dédie ce travail.

Avant-propos

Ce mémoire est le résultat d'un travail entrepris en l'an 2000 au laboratoire régional d'analyse de l'eau et du sol, laboratoire de l'Institut National du Sol de l'Irrigation et du Drainage (INSID). Il a été mené de façon parallèle à une activité professionnelle d'Ingénieur Agronome au sein de cet institut.

Je tiens à remercier en premier lieu les membres du jury qui ont accepté d'évaluer ce travail.

Monsieur **Djili Kadour**, Professeur à l'institut National Agronomique, qui m'a fait l'honneur d'accepter d'évaluer ce travail et de présider ce jury.

Monsieur **Aidaoui Abdellah**, Professeur à l'institut National Agronomique, pour avoir accepté d'examiner ce travail .

Monsieur **Douaoui Abdelkader**, Maître de conférence à l'université de Cheliff ,pour m'avoir encadré et guidé depuis le commencement de ce travail. Son expérience dans le domaine des statistiques et géostatistiques, en qualité de spécialiste, ses conseils, sa collaboration et sa disponibilité ont été d'un apport considérable pour l'aboutissement de ce travail.

Monsieur **Hartani Tarik**, Maître de conférence à l'institut National Agronomique, m'a également fait l'honneur de juger ce travail.

Monsieur **Habila Mohamed**, Directeur Général de l'Institut National des Sols de l'Irrigation et du Drainage (INSID), pour l'aide précieuse qu'il m'a prodigué.

Tous ont lu mon travail avec une très grande attention, malgré des emplois du temps chargés. Leurs remarques et critiques m'ont beaucoup aidé et je saurai en tenir compte dans mes activités de recherche à venir. Qu'ils en soient vivement remerciés.

Je n'oublie pas que ce sont les enseignants de l'Institut National agronomique d'El Harrach qui m'ont donné le goût pour la science du sol. J'aimerais les remercier ici.

Je tiens à remercier mon père, pour la patience qu'il a eu à lire et à corriger mes défaillances linguistiques.

Mes remerciements vont également à M^{elle} Touaf Leila pour son dévouement, son attention scientifique, son aide et ses conseils durant la rédaction de cette thèse.

Mes vifs remerciements vont également à l'égard de l'ensemble du personnel de l'INSID, particulièrement à M^F Kradia Laid Chef du laboratoire régional de l'INSID et Kellou Abdelkhani technicien supérieur au département de cartographie pour toute l'aide qu'ils m'ont procuré.

ص-خ-لم

الهدف من هذا العمل هو البحث عن أفضل تمييز لوحدات التربة في منطقة ذات مساحة أدراها 506,25 هكتار في سهل الخلف السفلي. وقد تم إنجاح في الطريقة الخرائطية المنتهجة في هذا العمل، كلاً من الفضاء الشعاعي و الفضاء الجغرافي وبالعلم المتوسط.

كل مرحلة من مسار هذا العمل مسمية لما قبلها الفضاء الشعاعي الفالحم عن ACP يمثل نقطة ابتداء لأحسن تقرب لتحريات التربة بفضل العزم الأقصى الذي تشكله المركبات الشعاعية الأساسية الثلاثة، يساهم نظام التصنيف التصاعدي CAH بفضل خصوصيته الاندماجية السهولة (دليل وورد)، في اتخاذ قرار اختيار عدد أصناف التحريات الأكثر اندماجاً. تطويل شعاعي مسميز يمثل هو أيضاً مرحلة من مراحل هذا العمل حيث يمكننا من إظهار متغيرات التربة الأكثر تمييزاً لأصناف تحريات التربة الناجمة عن التقسيم.

يدرس التحليل الجزيئي إحصائي قسماً مهماً من نظام التحليل الشعاعي المسميز وذلك لإظهار نوعية الشريط الإنشائي بين وحدات التربة المشجولة وقياس إسهامها ومساحتها الجغرافية.

تم تقدير مستوى صفاء وحدات التربة المستخلصة من كل مرحلة من مسار العمل، بأخذ تحريات مسمية مأخوذة من كل وحدة قرابية.

Résumé

L'objectif de ce travail porte sur la recherche d'une meilleure individualisation et différenciation des unités du sol dans un secteur de 506.25 hectares, situé dans la plaine du Bas-Chélif. L'espace factoriel et l'espace géographique sont tous les deux mêlés à la démarche entreprise dans ce travail, à moyenne échelle.

Chaque étape dans cette démarche est complémentaire à celle qui la précède. L'espace factoriel issu de l'analyse des composantes principales (ACP) constitue le point de départ pour une meilleure individualisation des sondages du sol. Avec un maximum de variance (inertie) formé par les trois premières composantes principales, la classification ascendante hiérarchique (CAH) contribue par son critère d'agrégation, appelé indice de Ward ($\#_Q$) dans la décision du choix du nombre de classes de sondages les plus cohérents. Une analyse factorielle discriminante (AFD) succède par la suite pour ressortir les variables du sol les plus discriminantes entre les classes de sondages issues de la partition. Une partie importante de l'information factorielle discriminante est traitée par la suite par la géostatistique multiparamétrique pour ressortir la nature de la transition entre les unités homogènes du sol et mesurer son étendue et son aire géographique. Une appréciation du niveau de pureté des unités du sol correspondant à chacune des étapes de la démarche a eu lieu par des sondages complémentaires effectués dans chacune des unités du sol.

Mots clés : Analyse des composantes principales (ACP), Classification ascendante hiérarchique (CAH), Analyse factorielle discriminante (AFD), **individualisation**, **Géostatistique multiparamétrique**, **pureté** des unités du sol.

Summary

The objective of this work concerns the search for a better individualization and differentiation of the soil units in a sector of 506.25 hectares, located in the plain of Bas-Chéliff. Factorial space and geographical space are both mixed with the step taken in this work, with an average scale.

Each stage in this step is complementary to that which precedes it. Factorial space resulting from the analysis of the principal components (ACP) constitutes the starting point for a better individualization of the soil surveys. With a maximum of variance (inertia) formed by the first three principal components, the ascending hierarchical clustering (CAH) contributes by its criterion of aggregation, called index of Ward ($\#_Q$) in the decision of the choice of the number of classes of the most coherent surveys. A discriminating factorial analysis (AFD) succeeds there after to arise the most discriminating soil variables between the classes of surveys resulting from the partition. A significant part of discriminating factorial information is treated there after by the geostatistic multiparametric one to arise the nature of the transition between the homogeneous soil units and to measure its extent and its geographical surface. An appreciation of the level of purity of the soil units corresponding to each stage of the step took place by complementary surveys carried out in each soil unit.

Key words : Analyze principal components (ACP), Ascending hierarchical clustering (CAH), Discriminating factorial analysis (AFD), individualization **multiparametric Géostatistic, purity** of the soil units.

INTRODUCTION

La couverture pédologique présente une variabilité dans l'espace géographique. Ceci est connu depuis très longtemps par les pédologues. La délimitation des unités du sol relativement homogènes fait encore le plus souvent appel aux approches traditionnelles et manuelles de cartographie. De ce fait, la mise en relief de ces unités est difficile du fait que les limites de transitions sont souvent progressives.

De nombreux auteurs distinguent l'erreur sur la position des limites et l'erreur sur le contenu des plages (Veregin, 1989). Très longtemps, cet aspect a été passé quasiment sous silence dans les ouvrages traitant de cartographie. Cette carence n'est que rarement compensée dans les rapports par une explication de ce que signifie exactement chaque limite et par la fourniture d'une méthode pour la retrouver sur le terrain. Les lois de délimitation des unités sont donc très difficiles à dégager. Au fait, il est évident qu'une erreur de position génère une erreur de contenu (Chrisman & Lester, 1991). Pour ce faire, il est nécessaire de se baser sur des techniques et méthodes numériques fines qui permettent de cartographier et de quantifier cette diversité avec beaucoup d'apparence et de précision.

Ce travail a pour objectif de rechercher la meilleure individualisation et différenciation qui pourraient exister entre les unités homogènes du sol. Par conséquent, ceci induit, une possibilité de réduire la variabilité intra-unité. A partir d'une variable linéaire issue de l'analyse factorielle discriminante, il devient aisé par l'emploi de la géostatistique de délimiter et de quantifier la transition entre ces unités homogènes du sol.

Le chapitre I de cette étude aborde la description du milieu physique de la région. Le

second chapitre décrit la nature pédologique des sols de la zone d'étude et traite des contraintes de la cartographie pédologique classique. Le chapitre III se réfère à l'approche méthodologique. Enfin, le chapitre IV est consacré à la description statistique des variables du sol jugées prépondérantes dans le comportement du milieu édaphique de la zone d'étude.

La démarche cartographique adoptée dans ce travail est constituée d'un enchaînement par étapes des différentes approches de cartographie numérique. Chacune de ces approches abordées dans un chapitre apporte sa part de contribution à la détermination des unités homogènes du sol. Le chapitre V et le chapitre VI constituent les deux premières étapes de cette démarche. Ils sont consacrés à la recherche d'un meilleur niveau d'individualisation des unités homogènes du sol. C'est aussi par référence à ce niveau d'individualisation que se reporte le nombre de ces unités. Ces deux chapitres se rapportent, pour l'un, à l'analyse des composantes principales (ACP) et pour l'autre, à la classification ascendante hiérarchique (CAH). Dans le chapitre VII, se rapportant à l'analyse factorielle discriminante (AFD), on ressort l'ordre discriminant des variables du sol des unités individualisées. Le chapitre VIII, est consacré à la géostatistique multiparamétrique. Celle-ci permet de ressortir la structure spatiale des unités homogènes du sol et les aires de transitions existantes entre elles. En guise de synthèse, le dernier chapitre dans ce travail traite de la pureté des unités homogènes du sol.

Chapitre I . Présentation du milieu physique

1-1 Situation géographique de la zone d'étude

Le site choisi se situe dans la région du Bas Chellif entre l'Oued Chellif et la route nationale n°4, à 3.2 km au nord ouest de la station expérimentale de l'Institut National de la Recherche Agronomique de Hmadna. De forme carrée, sur une surface plane de plusieurs kilomètres aux alentours, la parcelle prospectée couvre une superficie de 506.25 ha (*fig.1*).

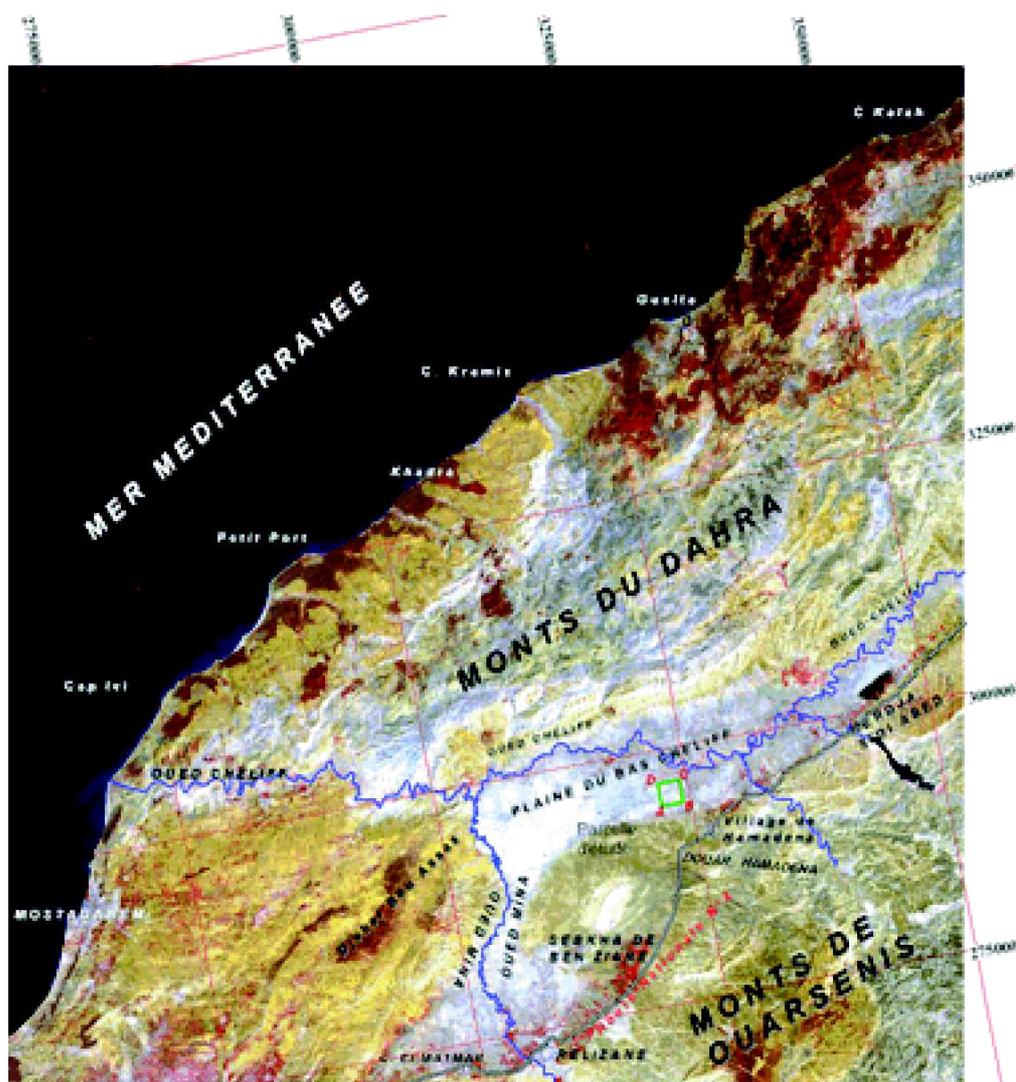


Figure n°1 . Localisation de la zone d'étude par imagerie satellitale.

(Image Landsat 5. Date 03 / 09 / 1994. Composition colorée des canaux : 4-3-1)

Les coordonnées Lambert des quatre points délimitant la parcelle prospectée sont les suivantes :

Points	A	B	C	D
Longitude (mètre), X =	322000	324250	324250	322000
Latitude (mètre), Y =	294500	294500	296750	296750
Altitude (mètre), Z =	50	50	55	50

1.2 . Milieu naturel

1.2.1 . Climat de la région de Hmadna

L'analyse du climat de la région de Hmadna est effectuée sur la base des données de la station climatique de l'Agence Nationale des Ressources Hydriques (ANRH). Ces données climatiques couvrent une période de dix sept ans, allant de 1985 à 2002.

Les coordonnées géographiques de la station climatique de Hmadna sont les suivantes :

- Longitude , X = 00° 47' E
- Latitude , Y = 35° 54' E
- Altitude (mètre), Z = 48

1.2.1.1 . Les précipitations

L'analyse pluviométrique de la région de Hmadna déduite à partir des précipitations (Tab. I) est résumée par les points suivants :

- a)le mois de janvier est le mois le plus pluvieux (33.32 mm) avec des précipitations moins irrégulières (pourcentage des précipitations max / moy des précipitations = 31.09) par rapport aux autres mois de l'année,
- b)le mois de juillet enregistre la plus faible pluviosité dans l'année avec une pluviométrie égale à 1.11 mm,
- c)Bien que la période estivale soit la moins pluvieuse dans l'année, celle ci enregistre des pluies sous formes d'averses de forte intensité. Ces pluies interviennent en période hivernale comme en période estivale, dans le processus de dégradation et d'érosion du sol.

Tableau n°I . Données pluviométriques de la station météorologique de Hmadna (1985-2002)

Mois Paramètres	J	F	M	A	M	J	J	A	S	O	N	D
Moy des précipitations Mensuelle (mm)	33.32	31.49	28.60	24.25	13.34	3.48	1.11	1.81	16.73	17.59	32.80	26.14
Max. des précipitations / j	10.36	10.46	10.71	10.94	7.2	2.07	0.82	0.80	12.22	13.38	13.95	11.12
Nombre de jours de Pluie / mois	7	6	5	4	3	1	1	1	3	4	6	6
Intensité moy (mm/j) des précipitations mensuelles	4.76	5.24	5.72	6.06	4.44	3.48	1.11	1.81	5.57	4.39	5.46	4.34
% des précipitations max. / moy des précipitations	31.09	33.49	37.44	45.11	53.97	59.48	73.87	44.19	73.04	76.06	42.53	42.70

1.2.1.2 . Les températures

Le régime thermique de la région de Hmadna est caractérisé par des températures élevées en été et modérées en hiver. Les températures élevées sont enregistrées aux mois de juillet et août où elles atteignent un maximum de 38.5°C en moyenne. Les basses températures se manifestent au mois de janvier avec une moyenne minimale de 5.15 °C (Tab. II).

Tableau n° II . Température moyenne mensuelle des minimums et maximums en °C de la station météorologique de Hmadna. (1985-2002)

Mois Paramètres	J	F	M	A	M	J	J	A	S	O	N	D
Température moy mensuelle des min. en °C	5	7	9	13	17	20	21	17	13	10	6	5
Température moy mensuelle des max. en °C	15.90	17.70	21.50	24.00	28.90	33.00	37.60	38.50	32.80	27.20	21.40	17.10
Température moyenne mensuelle en °C	18	21	24	29	33	38	38	32	27	21	17	18

Le diagramme ombrothermique de la station de Hmadna, entre 1985 et 2002, montre une longue période de sécheresse qui s'étale sur 6 mois, allant de la mi-avril à la mi-octobre (fig. 2).

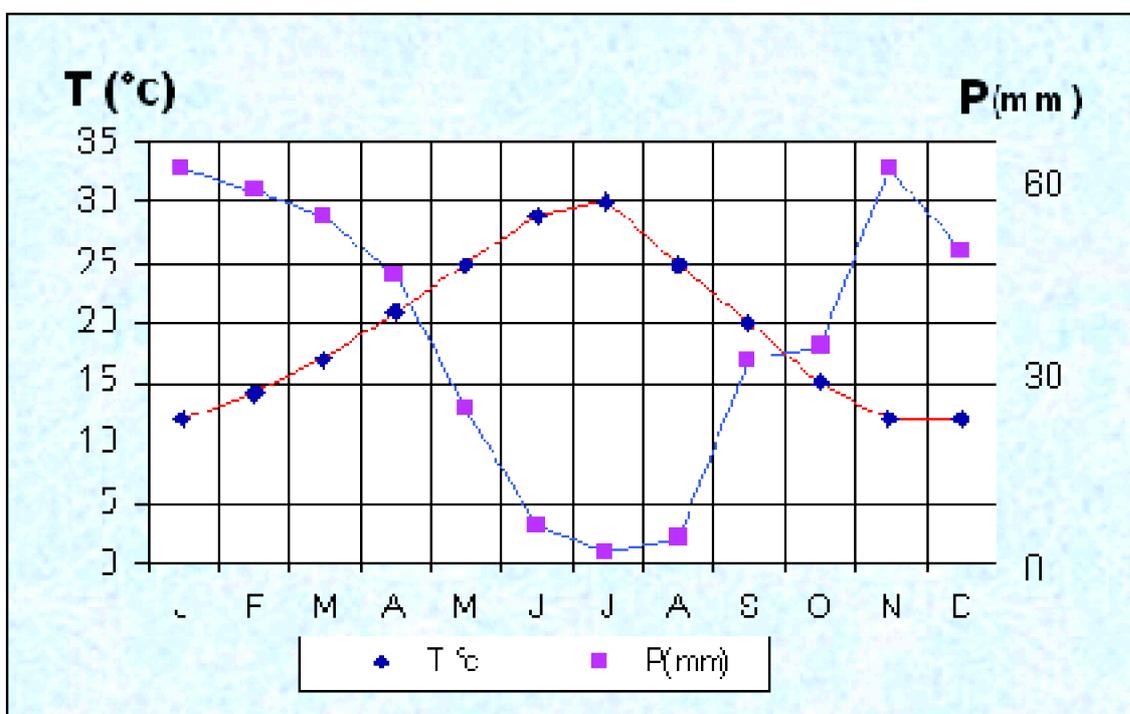


Figure n° 2 . Diagramme ombrothermique de la station climatique de Hmadna.

(1985-2002)

1.2.1.3 . Autres paramètres

- l'humidité relative de l'air

L'humidité relative de l'air est généralement plus élevée le matin pour tous les mois. Elle se situe entre 65 et 75 % en hiver et près de 35 % en été. Le mois le plus humide est le mois de décembre ; par contre, le mois le plus sec est le mois de juillet.

- Les vents

D'après les observations effectuées à la station climatique de Hmadna, il ressort que le vent dominant est le vent d'ouest. Sa vitesse varie entre 13 et 33 km /h.

- L'évapotranspiration potentielle

Tenant compte des données de précipitation et de température moyennes mensuelles, l'évapotranspiration potentielle (ETP) est calculée par la formule de **Thornthwaite** (1948), (Tab. III).

$$ETP = 1.6 \left(10 \left(\frac{T}{I} \right)^a \right)$$

ETP : évapotranspiration potentielle mensuelle.

I : somme des indices thermiques mensuels (pour la région de Hmadna I = 96.22)

T : température moyenne du mois considéré, exprimée en °C.

a : constante qui dépend de la somme des indices thermiques mensuels de l'année.

Pour le cas de la station expérimentale de Hmadna, a = 2.03

Tableau n° III . Evapotranspiration potentielle en mm, établie selon la formule de Thornthwaite (1944). (station météorologique de Hmadna, 1985-2002)

Mois	J	F	M	A	M	J	J	A	S	O	N	D
ETP (mm)	16.70	19.72	36.46	52.86	95.71	135.52	184.70	183.51	115.87	70.90	36.03	19.80

1.2.1.4 . Classification du climat de la région de Hmadna

La valeur de l'indice annuel d'aridité (**la**) de **De Martonne**, déterminée à partir des données de la station climatique de Hmadna, classe la région dans l'étage bioclimatique semi-aride.

- **la** = Pluviométrie annuelle en (mm) / (Température annuelle moyenne en °C +10)
- **la** = 230.56 / (19.21+10) = 7.89

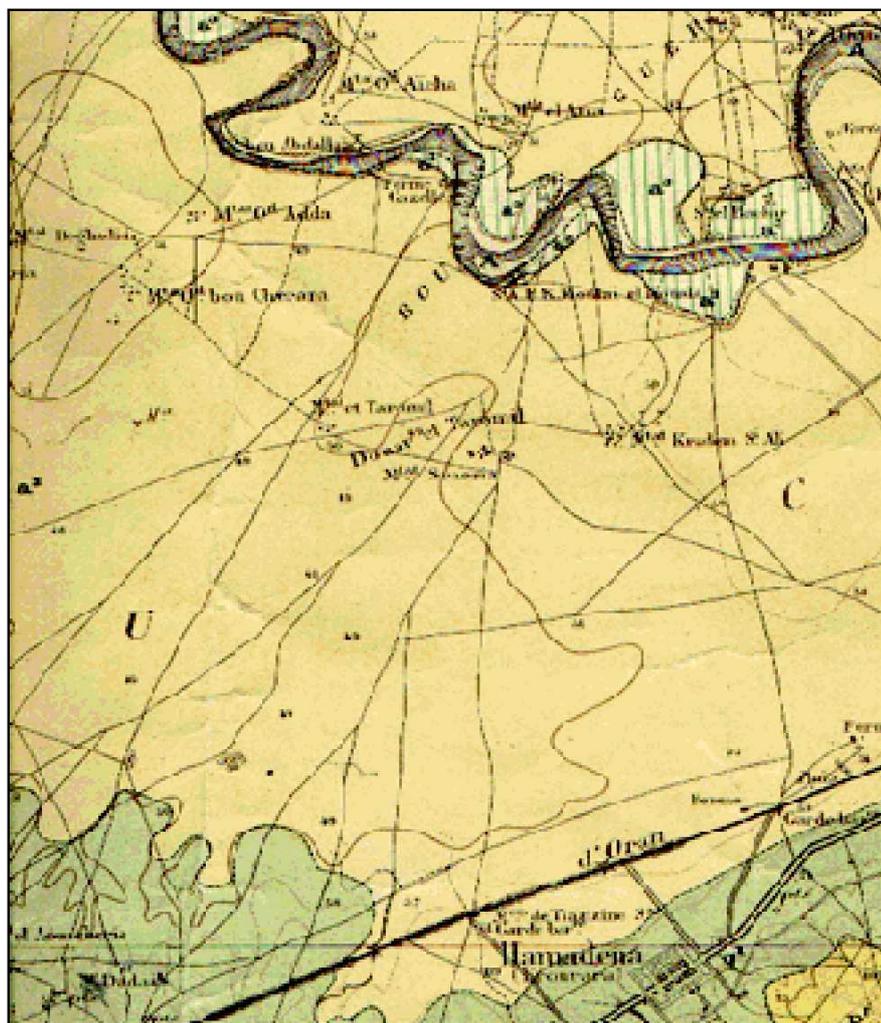
La correspondance entre les fortes températures et la faible pluviométrie indique le caractère méditerranéen du climat de la région de Hmadna.

La grande différence entre les températures moyennes de l'été et celles de l'hiver

traduit bien la continentalité du climat de la région.

1.2.2 . Géologie de la région de Hmadna

Le site d'étude fait partie de la plaine du Bas Chellif. Cette plaine est un synclinal comblé d'alluvions du quaternaire (Rharbien) (S. N. REPAL, 1952). Elle est limitée au Nord et au Sud par des collines de l'âge tertiaire (miocène et pliocène) d'où proviennent directement, par érosion, les alluvions qui la composent. Ces alluvions sont d'une épaisseur de plusieurs dizaines de mètres. Elles sont constituées par des matériaux originaires des bassins versants du Chellif, de l'Oued Rhiou et l'Oued Djediouia où dominent des schistes et des marnes souvent riches en sels (fig. 3).



Légende

- alluvions récentes
- poudingues et grès (pliocène supérieur)
- alluvions anciennes (niveau inférieur)

Figure 3 . carte géologique d'une partie de la plaine du Bas Chellif.

(Service géologique de l'Algérie, 1923) ; échelle : 1 / 65.000

Le forage effectué par les services de l'hydraulique de la wilaya de Relizane près de

la station expérimentale de Hmadna dévoile un faciès à strates relativement peu épaisses. Ces strates marneuses sont différenciées par leur couleur et l'abondance des éléments détritiques (fig. 4).

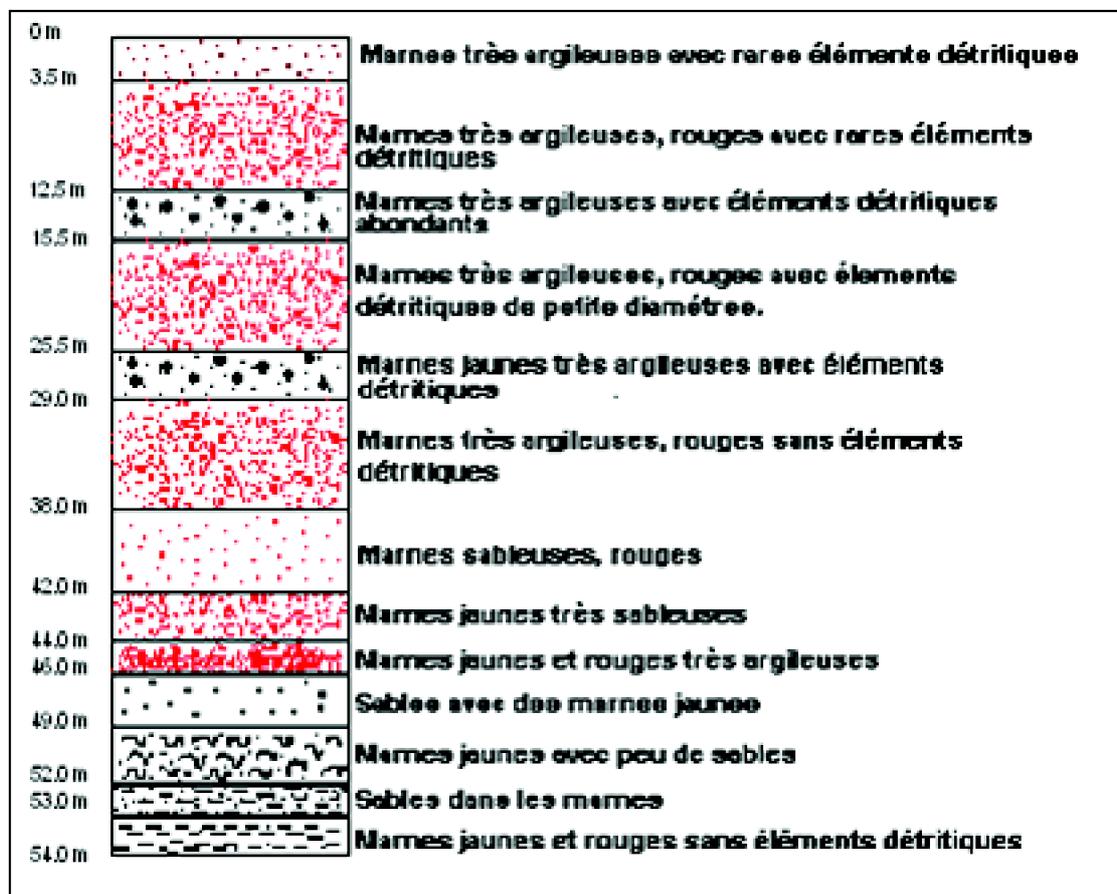


Figure n° 4 . Coupe d'un sondage effectué par les services de l'hydraulique près de la zone d'étude (Année 1965. Forage n° 3 de la station agricole de Hmadna).

Le réseau hydrographique est constitué principalement par l'oued Chellif et ses affluents comme l'oued Rhiou, l'oued Djediouia et l'oued Ouarizane. Les lits d'oueds sont encastrés dans la terrasse alluviale du Rharbien à un niveau inférieur à 10m avec des bords escarpés et très abrupts. Ces oueds sont pratiquement à sec pendant la période estivale.

1.2.3 . Hydrogéologie

Deux sortes de nappes superposées existent dans le quaternaire ; les nappes superficielles libres contenues dans les formations du quaternaire récent et les nappes de profondeurs captives contenues dans celles du plioquaternaire.

Les aquifères responsables des nappes superficielles sont pratiquement étendus sur toute la plaine du Chellif. Les nappes profondes captives sont renfermées le plus souvent dans un système de lentilles d'éléments détritiques attribuées soit au pliocène supérieur continental soit au quaternaire ancien.

La salinité des nappes superficielles est généralement plus élevée que celle des nappes profondes. Elle atteint les 22 g/l dans l'axe Chellif. Le faciès chimique des eaux des nappes superficielles et des nappes profondes est bicarbonaté à chloruro-sodique (Benziane, 1983).

1.2.4 . Topographie de la zone d'étude

Le paysage de la zone d'étude présente une morphologie très légèrement ondulée. L'altitude de celle-ci varie entre 50 et 60 mètres avec des pentes dépassant très rarement les 2 %.

1.2.5 . La végétation

1.2.5.1 . La végétation naturelle

La strate arborescente du genre myrtacée (eucalyptus) a presque disparu du site suite au déboisement qui a sévi durant ces dernières décennies.

La strate arbustive est inexistante dans la zone d'étude.

Pour la strate herbacée, les espèces les plus représentatives et les plus abondantes sont les suivantes :

Cichorium intybus, Ziziphus lotus, Chrozophora tinctoria, Suaeda fructicosa, Atriplex halimus, Malva parviflora, Anacyclus clavatus, Polypogon monspeliensis, Xanthium hispanicus.

1.2.5.2 . Les cultures

Dans cette région et plus particulièrement au niveau de la zone d'étude, les terres labourées sont essentiellement occupées par des céréales et des cultures maraîchères comme l'artichaut, le melon ou la pastèque.

1.2.6 . Actions anthropiques

La zone d'étude fait partie d'un ancien périmètre irrigué. Les parcelles affectées autrefois à la culture du riz sont aujourd'hui occupées par une végétation spontanée halophyte où prédomine la Suaeda Fructosa. La surface du sol est caractérisée partiellement par une structure squameuse ou particulaire avec une accumulation des sels d'apparence blanchâtre durant la période des grandes chaleurs. Les parcelles cultivées sont caractérisées en surface par une structure pouvant être parfois fragile et instable entraînant la formation d'une croûte de battance.

Les canaux d'irrigation longeant les extrémités des parcelles sont endommagés et inutilisables. Ils témoignent de l'intense activité agricole qui régnait dans la zone. Contrairement à leur destination initiale qui consistait à évacuer les excès d'eau, les drains à ciel ouvert servent aujourd'hui comme lieu de stockage de l'eau de pompage et

de pluie par obturation partielle du réseau.

Chapitre II . La cartographie pédologique de la zone d'étude

L'objet de ce chapitre est de montrer la nature des sols rencontrés dans la zone d'étude en décrivant les principaux horizons jugés représentatifs des unités cartographiques.

Le peu de données réellement quantifiées, lié directement à la faible densité de sondage, reste en deçà de nos connaissances sur le degré d'homogénéité de ces unités. D'autant plus que la couverture pédologique telle qu'elle est décrite, donne à travers les profils types, un aperçu assez sommaire pour les désigner synthétiquement. Elle sert, au regard d'un observateur, comme un exemple illustrant les contraintes qui réduisent sa représentation dans l'approche cartographique classique.

2.1 . Les contraintes de la cartographie pédologique

L'objet de la cartographie pédologique réside dans la définition et la délimitation des unités cartographiques au sein desquelles la variabilité des propriétés du sol est moindre que la variabilité générale observée au sein d'un paysage (Walter & Arousseau, 1992).

La mise en œuvre pratique de ce cadre conceptuel a longtemps interpellé le pédologue tant dans son illustration graphique de la réalité du couvert pédologique que dans la caractérisation de ses unités qu'il considère homogènes.

Confronté d'abord à la réalité du terrain, le cartographe se trouve souvent contraint d'être imprécis dans la représentation spatiale de la couverture pédologique. Cette contrainte est due à la complexité du paysage, caractérisée par une certaine mosaïque très dense et très variée des unités simples que les spécialistes du domaine regroupent sous le nom d'unités complexes.

En second, la caractérisation verticale de la couverture pédologique par des profils types (Jamagne, 1967) laisse suggérer implicitement que l'unité cartographique est un espace homogène à contour délimité.

A l'observation immédiate, la couverture pédologique ne se présente pas sous la forme d'objets individualisés mais comme une réalité continue à variations progressives (King & Duval, 1989). Toute stratification du milieu s'opère donc par combinaison d'un certain nombre de critères sélectionnés. Ce découpage se fait nécessairement sous l'influence déterminante des caractéristiques les plus visibles. Mais les modifications de ces critères peuvent être tellement progressives que les limites introduites auront parfois un caractère artificiel (Bourennane, 1997).

Ce bref aperçu sur les quelques insuffisances dans la cartographie pédologique classique a été évoqué par de nombreux auteurs. Dans ce cas, nous pouvons citer les travaux de Becket & Bie (1975), Girard (1983), Feeny (1988), Ruellan & al. (1989) qui se réfèrent au problème du tracé des limites et ceux de Wilding & al. (1965), Ragg & Henderson (1980), Wang (1982), Edmonds & Lentner (1986), Marsman & Gruijter (1986), Iris (1986), Thomas & al. (1989) concernant la variabilité des propriétés du sol au sein des unités cartographiques.

Ainsi, l'efficacité de la démarche cartographique classique en pédologie reste dépourvue de tous les atouts de sa précision. Elle se trouve complètement dépendante des capacités habituelles d'observation d'un pédologue, c'est à dire de son expérience acquise sur le terrain. Pour plus de rigueur scientifique quant à l'expression spatiale des unités cartographiques, une partie de la plaine du Bas Chellif faisant partie d'une zone cartographiée au préalable sert comme champ d'application aux différentes méthodes d'étude de la variabilité spatiale des sols.

2.2 . Caractérisation pédologique de la zone d'étude

2.2.1 . Le document cartographique de base

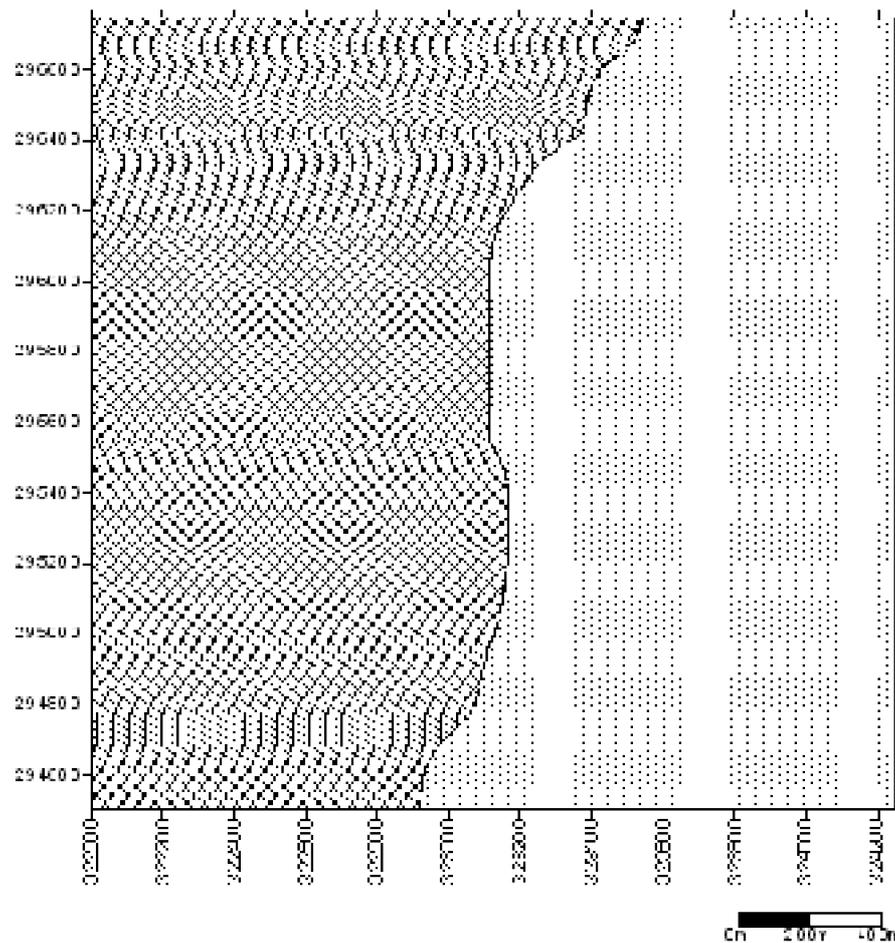
La caractérisation pédologique de la zone d'étude est extraite d'une étude des ressources en sol dont le thème s'intitule "Etude de l'avant projet détaillé des extensions de Guerouaou et de Sebket Benziane et du réaménagement du bas Chellif ". Cette étude agropédologique qui est orientée vers l'aménagement hydraulique de la région a été réalisée par le bureau d'étude Sir M Mac Donald et Partners Limited en association avec WS Atkins International Limited et le Bureau National d'Etudes pour le Développement

Rural en 1989, à une échelle de 1/25.000.

2.2.2 . Les caractéristiques pédologiques des unités cartographiques

Selon le système français de la classification des sols (CPCS, 1967) adopté pour cette étude, la parcelle d'investigation (*fig. 5*) est constituée de deux unités cartographiques relevant chacune de la classe des sols suivants :

- classe des sols sodiques, appelés actuellement salsodiques,
- classe des sols calcimagnésiques.



Légende
 [Motif à triangles pointant vers le haut] Classe des sols sodiques
 [Motif à triangles pointant vers le bas] Classe des sols calcimagnésiques

Figure n° 5. Carte agropédologique de la zone d'étude

Description morphologique du profil type n°1

- Date : été, 1989

- Temps : ensoleillé
- Coordonnées Lambert (Km) : X = 319.250Y = 296.125
- Altitude (m) : Z = 50
- Substratum : alluvions
- Géomorphologie : plaine
- Pente : 0 %
- Occupation du sol : ancienne rizière
- Drainage externe : mauvais
- Etat de surface : encroûtement moyen avec présence de fentes de retrait en Surface.
- Végétation : Suaéda fructosa

Profondeur

(cm)

0 - 23 Ap	Sec, couleur 2.5Y 5/3, texture argileuse, structure poudreuse à polyédrique fine, peu dur, porosité moyenne, présence de quelques fentes larges de 2 à 3 cm, racines peu nombreuses fines éparpillées, faible activité biologique, effervescence moyenne, transition distincte et ondulée.
23 - 67 (B) ₁	Sec, couleur 2.5Y 5/3, texture argileuse, structure prismatique à tendance angulaire, extrêmement dur, faible porosité, présence de fentes larges de 1 cm, pas de racines, effervescence moyenne, transition graduelle et régulière.
67 - 135 (B) ₂	Sec, à très peu humide, couleur 2.5Y 4/3, texture argileuse, structure angulaire fortement développée, compact et ferme, très faible porosité, peu de fentes, pas de racines, activité biologique nulle, effervescence moyenne, transition graduelle et régulière.
> 135 B _g	Humide, couleur 2.5Y 5/1 avec présence de taches de rouilles et des accumulations noires ferromagnésiques, texture argilo-limoneuse structure non apparente, compact, absence de pores, pas de fentes, pas de racines, absence d'activité biologique, effervescence vive.

- Classification CPCS (1967)
 - Classe : sols salsodiques
 - S/classe : structure dégradée
 - Groupe : salés alcalins

- S/groupe : halomorphes, vertiques, hydromorphes.

Synthèse sur les caractéristiques de l'unité cartographique

Cette unité couvre 55 % de la surface totale de la zone d'étude. Elle est occupée en grande partie par une végétation halophyte (*Suaeda fructosa*) moyennant un taux de recouvrement végétatif de 65 % dans toute la partie Sud Ouest et de 30 à 35 % dans la partie Nord Ouest.

Un encroûtement de battance est observé en surface mais non continu sur toute l'unité. Dans bon nombre d'endroits, les sels s'accumulent en surface dans un aspect blanchâtre et confèrent à la surface du sol une structure poudreuse.

Le sol est modérément calcaire le long du profil (d'après GEPPA). Il enregistre une moyenne de 16.97 %. La teneur en matière organique est faible ; son gradient diminue de la surface vers la profondeur pour atteindre une valeur de 0.55 %. Contrairement à la matière organique, la salinité du sol est élevée sur l'ensemble du profil ; son gradient augmente de la surface vers la profondeur pour atteindre une valeur de 17.55 dS/m à 140 cm. Le pH_{eau} du sol est alcalin sur tout le profil (Tab. IV).

Horizon		H ₁	H ₂	H ₃	H ₄
Profondeur (cm)		0 - 23	23 - 67	67-135	> 135
Symbole de l'horizon		Δ ₁	(B) ₁	(E) ₂	B _g
Calcaire total CaCO ₂ (%)		15.16	16.20	15.60	21.00
pH _{eau} (1/25)		7.50	7.86	7.60	7.64
Conductivité électrique, CE (dS/m)		7.34	13.73	14.04	17.55
Caractéristiques Biochimiques Du sol	C (%)	1.01	0.61	0.45	1.32
	N (%)	0.13	0.09	0.08	1.06
	MO (%)	1.73	1.04	0.77	1.55
	C/N	7.76	6.77	5.62	5.33
Granulométrie	Argile (%)	60	59	56	45
	Limon fin (%)	36	33	32	43
	Limon grossier (%)	1	3	0	2
	Sable fin (%)	2	3	0	2
	Sable grossier (%)	1	0	0	1
Classe granulométrique (diagramme de texture Selon Jamagne)		Argileux	Argileux	Argileux	Argilo - Limoneux

Tableau IV . Fiche analytique du profil représentant l'unité n°1.

Unité cartographique n°2

Description morphologique du profil type n°2

- Date : été, 1989
- Temps : ensoleillé
- Coordonnées Lambert (Km) : X = 326.950Y =300.120
- Altitude (m) : Z = 50

- Substratum : alluvions
- Géomorphologie : plaine
- Pente : 0 %
- Occupation du sol : céréales
- Drainage externe : mauvais
- Etat de surface : faible encroûtement
- Végétation : chaumes de blé dur

Profondeur
(cm)

0 - 28 Ap	Sec, couleur 2.5Y 5/3, texture argilo-limoneuse, structure polyédrique moyenne à fortement développée, bonne porosité, racines peu nombreuses et fines, effervescence moyenne, transition distincte et ondulée.
28 - 59 (B)	Peu humide, couleur 2.5Y 6/4, texture limono-argileuse, structure subangulaire, moyennement développée, bonne porosité, racines peu nombreuses et fines, effervescence moyenne, transition distincte et régulière.
59 - 95 IC	Peu humide, couleur 2.5Y 6/4, texture sablo-limoneuse, friable, pas de structure, bonne porosité, peu de racines, effervescence moyenne, transition graduelle et régulière.
95 - 130 IIC _x	Humide, couleur 2.5Y 5/1, présence de taches de rouilles, texture argilo-limoneuse, structure massive, très faible porosité, absence de racines, effervescence moyenne.

- Classification CPCS (1967)
 - Classe : sols calcimagnésiques
 - S/classe : sols carbonatés
 - Groupe : sols bruns calcaires
 - S/groupe : hydromorphe

Synthèse sur les caractéristiques de l'unité cartographique

Cette unité cartographique représente 45 % de la surface totale de la zone d'étude. La majeure partie des parcelles qui s'y trouvent sont cultivées en sec (blé dur et blé tendre) ou laissées en jachère. Les parcelles destinées aux cultures maraîchères (melon,

pastèque et artichaut) sont irriguées à partir d'une eau puisée des fosses de drainage. Le restant de cette unité est occupé par une végétation spontanée à dominance hallophyte (*Suéda fructosa*).

Un encroûtement de battance est observé fréquemment en surface. Le sol est développé dans un matériau argilo-limoneux surmontant une couche friable, sablo-limoneuse. Cette dernière repose sur une couche argileuse de couleur grise, imprégnée de taches de rouille, signe d'hydromorphie.

L'analyse physico-chimique du sol (Tab.V) montre une teneur modérée en calcaire total sur le long du profil, moyennant une valeur de 17.10 %. Le taux de matière organique est faible sur tout le profil ; son gradient diminue de la surface vers la profondeur pour atteindre une valeur de 0.12 %. La salinité du sol présente une forte variation le long du profil en restant faible à moyennement salée dans l'ensemble des horizons. Le pH_{eau} du sol est alcalin sur tout le profil.

Horizon		H ₁	H ₂	H ₃	H ₄
Profondeur (cm)		0 - 28	28 - 59	59 - 95	>95
Symbole de l'horizon		Ap	(B)	C ₁	C _{2g}
Calcaire total CaCO ₃ (%)		21.20	22.80	18.00	18.60
pH _{eau} (12.5)		7.66	7.84	7.90	7.86
Conductivité électrique, CE (ds/m)		3.78	1.71	1.58	3.47
Caractéristiques Rinchimiques du sol	C (%)	0.71	0.40	0.12	0.06
	N (%)	0.09	0.09	0.03	0.08
	MO (%)	1.22	0.84	0.20	0.12
	C/N	7.88	6.28	3.75	4.82
Granulométrie	Argile (%)	32	37	9	47
	Limon fin (%)	33	44	16	47
	Limon grossier (%)	16	9	16	3
	Sable fin (%)	15	5	59	1
	Sable grossier (%)	3	3	0	1
Classe granulométrique (diagramme de texture Selon Jarnagne)		Argilo- Limoneux	Argilo- Limoneux	Sablo- Limoneux	Argileux

Tableau V. Fiche analytique du profil représentant l'unité n°2

Conclusion

L'étude agropédologique de la zone d'étude montre l'existence de deux unités cartographiques : les sols salsodiques à structure dégradée, situés à l'Ouest de la parcelle et les sols calcimagnésiques, situés à l'Est de la parcelle.

La confrontation des données morphologiques des deux profils types, révèle la présence d'une différenciation texturale et structurale dans l'ensemble des horizons (profil 1 et 2). Cette différenciation apparaît également dans les données analytiques et concerne la teneur en argile, limon grossier, sable fin et modérément la matière

organique.

Chapitre III . Matériel et méthodes

3.1 . Matériel d'étude

3.1.1 . Dispositif expérimental

La parcelle d'investigation couvre une superficie de 506.25 ha et forme un carré de 2250 m de côté. Le dispositif expérimental est constitué de deux plans d'échantillonnage suivants :

- plan d'échantillonnage n°1

Ce plan d'échantillonnage est réalisé selon un quadrillage régulier aux nœuds d'une grille à maille carrée de 150m de coté (*fig. 6a*). L'ensemble des nœuds correspond à un nombre de sondages égal à 256.

- plan d'échantillonnage n°2

Ce plan d'échantillonnage effectué sur la même parcelle consiste en la réalisation de deux transects T1 et T2. Ces transects ont été effectués selon deux directions différentes Sud-Nord, Est-Ouest (*fig. 6b*). Le nombre de sondages alignés dans chacun d'eux est de

46, d'intervalle équidistant de 50 m.

3.1.2 . Justification du plan d'échantillonnage

Le plan d'échantillonnage prédéterminé à maille carrée a été adopté parce qu'il assure un semis régulier des sondages sur l'ensemble de la parcelle. Il est également le plus simple à réaliser avec un minimum d'erreurs par rapport à la grille triangulaire qui est l'optimum selon Voltz (1986).

Le second plan défini par les deux transects a été choisi dans le souci de répondre aux critères de la géostatistique telle que la micro-régionalisation(Douaoui, 1993).

3.1.3 . Justification du pas d'échantillonnage

Le choix du pas d'échantillonnage est défini par les lois de la géostatistique lesquelles exigent d'abord une quantification de la corrélation spatiale entre les différents points de sondage pour chaque variable étudiée.

Cette corrélation est déterminée par une fonction aléatoire (variogramme) qui délimite la distance d'autocorrélation, appelée la portée. Au-delà de cette portée le pas d'échantillonnage n'est plus en mesure de représenter les variations continues d'une variable dans l'espace. Il en est de même pour la discontinuité

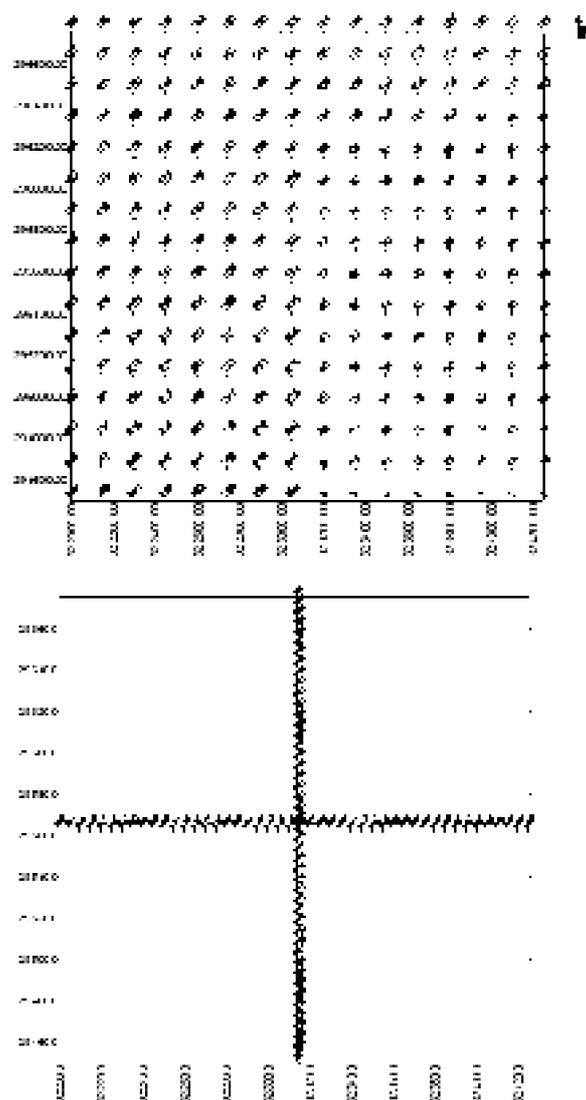


Figure 6 : Schéma du plan d'échantillonnage de la parcelle (échelle 1 / 22.500)

6 a: Semis régulier à maille carrée. 6 b: Semis régulier le long des transects (T1, T2)

de la fonction aléatoire à son origine qui doit être minimale. Ainsi, pour éviter cette situation, le plan d'échantillonnage de 150m à maille carrée a été renforcé par un second plan (transects) avec un pas d'échantillonnage réduit à 50 m. Burgess, al. (1981) recommandent au préalable ce type d'échantillonnage préventif. De même, Atteia & al. (1994) ont proposé une stratégie d'échantillonnage analogue qui consiste à compléter une grille régulière par des échantillons aux courtes distances placés aléatoirement autour des nœuds de cette grille. Un travail réalisé au préalable par Benbouali (1999) près de la zone d'étude vient conforter ce choix du pas d'échantillonnage de 150 m. Il dévoile par l'étude de la variabilité spatiale de certaines de nos variables une portée largement supérieure au pas d'échantillonnage retenu.

3.1.4 . La profondeur du sol

La profondeur d'investigation correspondant à l'horizon de surface est de 0,30 m. Le fait de se limiter à une telle profondeur n'affecte en rien le but assigné à notre objectif ; bien qu'une caractérisation complète d'une unité cartographique nécessite une extension des prélèvements aux horizons inférieurs. Il en ressort par ce choix, une suggestion de représenter la couverture pédologique en une épure où chaque horizon est étudié en tant qu'individu de base et comme un élément fondamental de l'organisation de la couverture pédologique (Baise, 1986 ; Girard, 1983).

3.1.5 . Les variables étudiées

3.1.5.1 . Choix des variables du sol

Les variables choisies pour la caractérisation des sols de la parcelle sont :

- le potentiel hydrogène du sol (pH_{eau}). C'est un paramètre facile d'accès. Il est la synthèse de toutes les interactions chimiques dans le sol.
- la conductivité électrique (CE). Elle exprime le degré de la salinité du sol. La salinité est la principale caractéristique des sols de la région (Boulaine, 1957 ; Durant, 1958 ; Aubert, 1976 ; Douaoui, 1995).
- le calcaire total (CaCO_3) par son rôle favorable dans la conservation de la structure du sol (Harris al., 1966 ; Mbagwu Bazzoffi, 1998). Il est présent dans les sols de la région sous la forme diffuse au niveau des horizons de surface.
- la matière organique (MO) a un rôle essentiel dans l'amélioration de la structure du sol. Cette influence est expliquée par le rôle que joue la MO dans la liaison entre les particules minérales du sol d'une part (Chenu, 1989) et l'hydrophobie des surfaces des particules minérales d'autre part, ce qui a pour effet de ralentir la vitesse d'humectation des agrégats et donc de réduire l'éclatement (Boulaine, 1957 ; Monnier, 1965 ; Guckert al., 1975 ; Tisdal Oades, 1982 ; Churchman Tate, 1987 ; Jouany al., 1992 ; Haynes, 1993 ; Le Bissonnais Le Souder, 1995 ; Le Bissonnais Arrouays, 1997).
- la texture du sol. Elle est une des caractéristiques fondamentales qui influe sur ses propriétés. Elle est liée à l'ensemble des comportements qui résultent de la taille et de la nature de ses constituants. Cette analyse inclut les fractions granulométriques du sol suivantes: argile (A), limon fin (Lf), limon grossier (Lg), sable fin (Sf) et sable grossier (Sg).
- l'indice d'instabilité structurale (I_o), (Héni Monnier, 1956). Cet indice permet d'apprécier la stabilité structurale du sol. Les sols de la région présentent souvent une stabilité structurale moyenne à médiocre (Heddadj, 1970 ; Derdour, 1981 ; Daoud, 1983 ; Saidi, 1992 ; Douaoui & al., 2004). La dégradation de leur état structural est particulièrement marquée au niveau de l'interface sol-atmosphère par l'apparition d'une croûte de battance. Cette croûte résulte essentiellement de la désagrégation des mottes de terre (Boiffin, 1984 ; Le Bissonnais, 1988 ; Le Souder, 1990).

- l'indice de percolation (K), (Héniñ & al., 1969). L'introduction de ce second indice pour apprécier également la stabilité structurale du sol tient du fait que les différents facteurs intervenant dans la stabilité structurale n'agissent pas de la même façon vis-à-vis de ces deux derniers tests. C'est ainsi que la présence des sels influe plus fortement sur les résultats du test de percolation que ceux de l'indice d'instabilité structurale. Par contre, la matière organique modifie davantage les résultats de l'indice d'instabilité structurale que ceux du test de percolation (Henin al., 1958).

3.1.5.2 . Les méthodes d'analyses du sol

Les déterminations correspondant aux différentes variables étudiées caractérisent chacun des échantillons prélevés dans la parcelle (tab. VI).

Tableau n° VI . Tableau récapitulatif des méthodes des analyses du sol.

Déterminations	Abréviation	Méthodes	Unité
- Préparation des échantillons de sol - Conductivité électrique - pH _{eau} - Calcaire total	- CE total pH _{eau}	- Séchage à l'air , tamisage à 2 mm - Rapport : sol / eau = 1/2.5. - Rapport : sol / eau = 1/2.5. - Calcimètre de BERNARD.	- dS/m - %

Suite du tableau VI.

- Matière organique - Indice de percolation (tamisage du sol : 0.2 – 2 mm) - Indice d'instabilité structural (tamisage du sol : 0.2 – 2 mm) <u>Granulométrie</u> Argile (< 2 um) Limon fin (2 - 20 um) Limon grossier (20–50 um) Sable fin (50 – 200 um) Sable grossier(200 – 2000 um)	MO K lo A L Méthode	Méthode ANNE : oxydation à chaux par mélange sulfochromique. - Méthode Henin. - Méthode Henin Monnier - Méthode internationale sans décarbonatation et après destruction de la matière organique.	% cm/h - % % % % %
---	---------------------	---	--------------------

3.2 . Méthodes d'études

3.2.1 . Analyse en composantes principales

L'analyse en composantes principales (A.C.P) est une méthode statistique ancienne. C'est Pearson (1901) qui en a entrevu les idées essentielles. On doit la première

publication sur ce sujet à Hotteling (1933) ; Tomassone (1976) l'abordât dans une présentation générale des méthodes d'analyse multidimensionnelle. Cependant, ce n'est que depuis 1980 avec l'apparition de la micro-informatique que l'ACP est de plus en plus utilisée (Philippeau, 1986).

Plusieurs auteurs ont déjà utilisé cette méthode en pédologie. Nous pouvons citer dans ce cas les travaux de Webster (1977) ; Girard (1983) ; Oliver (1984) ; Voltz (1986) ; Arrouays (1987) ; Girard King (1988) et Douaoui (1993).

L'ACP est une méthode statistique essentiellement descriptive. Son but est de présenter sous la forme graphique, le maximum de l'information contenue dans un tableau de données (Philippeau, 1986). Ce tableau est constitué en ligne par des individus (n) et en colonne par des variables (p) quantitatives. En effet, la réflexion sur l'organisation des données sous la forme d'un tableau statistique de type « individus-variables » (par ex : horizon-caractère, Girard King, 1988) est une nécessité avant tout traitement. Le tableau définit alors deux espaces :

- l'espace des individus (i) où chacun d'eux est représenté par un point de coordonnées $(A_{i1}, A_{i2}, \dots, A_{ip})$;
- l'espace des variables (j) où chacune d'elles est représentée par le point de coordonnées $(A_{1j}, A_{2j}, \dots, A_{nj})$.

Pour faciliter la visualisation des individus, l'ACP réduit les dimensions de cet espace à une dimension (axe principal) ou à deux dimensions (plan à deux axes). Les axes issus de cette réduction ne sont pas réalisés avec les variables initiales mais avec des indices synthétiques obtenus par combinaison linéaire des variables initiales. Ces axes sont appelés composantes principales ou axes principaux. Les individus sont donc représentés dans un espace à P dimensions.

Pour mieux observer les individus, il ne reste plus qu'à construire le plan à partir de deux composantes principales qui sont de variance maximale sous contrainte de non corrélation entre elles.

3.2.2 . Classification ascendante hiérarchique

La classification ascendante hiérarchique (C.A.H) est une des techniques de classification automatique destinée à produire une partition d'individus (n) décrits par un certain nombre de variables (p). Cette partition définit des classes nettement individualisées, c'est à dire que chacune d'elle forme un tout cohérent et en même temps elles sont toutes distinctes les unes des autres.

Les techniques de classification automatique des sols, s'inspirent de la méthode utilisée implicitement par le pédologue cartographe Rayner (1966) proche de la méthode dite des nuées dynamiques (Diday, 1971 ; Diday, 1972). Dans ce cas, on peut citer les travaux de Norris, Loveday (1971), Webster, Burrough (1972 a et b), Martin (1976), Banfield, Bascomb (1976) qui utilisent essentiellement comme variables les caractéristiques chimiques et texturales du sol.

La technique de classification ascendante hiérarchique est basée sur le principe des algorithmes ascendants appelés encore agglomératifs. Ces algorithmes procèdent à la construction des classes par agglomération successive des individus ou regroupement des individus deux à deux, décrivant une hiérarchie de leur partition par étapes. L'algorithme fournit une hiérarchie de partition se présentant sous la forme d'un arbre appelé dendrogramme contenant $n - 1$ classes de partitions.

La CAH est particulièrement aisée à mettre en œuvre lorsque l'agrégation est effectuée après une analyse en composantes principales, où les individus à classer étant repérés dans l'espace factoriel par leurs coordonnées sur les premiers axes principaux.

Les individus à classer constituent un nuage de points dans un espace des variables à P dimensions. Chaque point x_i (vecteur à P composantes) est muni d'une masse m_i .

$$\text{La masse totale du nuage de points } M = \sum_{i=1}^n m_i \quad (1)$$

L'inertie totale I du nuage de points est la quantité

$$I = \sum_{i=1}^n m_i \|x_i - G\|^2 \quad (2)$$

Le carré de la distance entre le point i et le centre de gravité G est noté :

$$\|x_i - G\|^2 = d^2(x_i, G) \quad (3)$$

où

$$G = 1/M \sum_{i=1}^n m_i x_i \quad (4)$$

La partition de l'ensemble des individus en Q classes à la $q^{\text{ème}}$ partition a pour masse :

$$m_q = \sum_{i \in q} m_i \quad (5)$$

et pour centre de gravité :

$$G_q = 1/m_q \sum_{i \in q} m_i x_i \quad (6)$$

La relation de Huygens fournit une décomposition de la quantité d'inertie en inertie intra-classe et inter-classe suivant la formule :

$$I_{\text{totale}} = I_{\text{intra-classe}} + I_{\text{inter-classe}}$$

$$I_{\text{totale}} = \sum_{i=1}^n m_i \|G_i - G\|^2 + \sum_q \sum_{i \in q} m_i \|x_i - G_q\|^2 \quad (7)$$

La stratégie d'agrégation appliquée à cette méthode est fondée sur le critère de perte d'inertie minimale inter-classe dit critère de Ward (1963). Ainsi à chaque étape de partition, l'inertie inter-classe diminue et l'inertie intra-classe augmente de cette même

quantité. Ceci revient à considérer la perte d'inertie inter-classe à chaque étape de partition comme des indices de dissimilarité, appelés indices de niveau (Δ_Q).

$$\Delta_Q = I_{\text{inter-classe}}(Q-1) - I_{\text{inter-classe}}(Q) \quad (8)$$

Q : partition à Q classes

$Q-1$: partition à $Q-1$ classes

La somme des indices de niveau dans la hiérarchie est égale, à la fin de la partition, à l'inertie totale (I_{total}) du nuage de points.

$$I_{\text{total}} = \sum_{Q=1}^{Q=n-1} \Delta_Q = \sum_{Q=1}^{Q=n-1} [I_{\text{inter-classe}}(Q-1) - I_{\text{inter-classe}}(Q)] \quad (9)$$

Il existe aussi de nombreux autres critères d'agrégation ou de regroupement des individus se ressemblant dont on peut trouver une énumération dans Cailliez, Pages (1976), Webster (1977) et Girard (1983).

3.2.3 . Analyse factorielle discriminante

L'analyse factorielle discriminante (A.F.D) est une méthode statistique essentiellement descriptive issue de l'analyse des données (Tomassone, 1988). Son but est de discriminer les groupes d'individus définis à partir d'une classification a priori, à l'aide d'un ensemble de P variables quantitatives, différentes de celles ayant déjà servi à la classification (Voltz, 1986).

L'A.F.D a été utilisée par les cartographes pédologues dans le but de tester une classification préalablement effectuée ou pour une distribution des profils à des classes pédologiques prédéterminées (Webster, Burrough, 1974 ; Burrough, Webster, 1976 ; Henderson, Ragg, 1980 ; Douaoui & al., 2001).

L'A.F.D est une extension de l'analyse de la variance. Dans cette dernière, on dispose d'une seule variable avec deux niveaux différents de variabilités :

- l'un est dû à la variabilité à l'intérieur de chaque groupe ; cette variabilité est caractérisée par une variance résiduelle (on dit aussi variance d'erreur) ;
- l'autre est dû à la variabilité que l'on attribue aux différences entre les groupes.

Tout consiste à savoir si, compte tenu de la première variabilité, la seconde est suffisamment grande pour être jugée significative. Le test qui en découle est un test de Fischer-Snedecor.

Dans le cas d'une A.F.D, nous disposons de plusieurs variables (x_i) avec les mêmes niveaux de variabilités que celles de l'analyse de la variance. On cherche une combinaison linéaire de ces variables de départ tel que le rapport des deux variabilités soit au maximum. On a la combinaison linéaire suivante :

$$Z = a_1 x_1 + a_2 x_2 + \dots + a_i x_i$$

et il s'agit de déterminer les coefficients a_i .

On démontre que Z est le vecteur propre associé à la plus grande valeur de $W^{-1} B$ où :

- W est la matrice des variances intra-groupes dont la traduction géométrique est le nuage moyen des points d'un groupe d'individus.
- B est la matrice des variances inter-groupes. Celle-ci fait intervenir géométriquement les seuls points moyens des groupes avec un poids proportionnel au nombre d'individus de chacun des groupes.

On va trouver un nombre de vecteurs propres égal au plus petit des deux nombres p et $k-1$ où p est égal au nombre de variables et k au nombre de groupes. Ces vecteurs propres correspondent aux axes factoriels discriminants qui ont, comme en analyse des composantes principales, la propriété d'être non corrélés entre eux.

Les différents axes factoriels discriminants permettent de différencier les groupes d'individus par rapport à leurs centres de gravité correspondants, en affectant chaque individu mal classé, de son groupe d'origine vers le groupe d'affectation. Le critère de différenciation étant la distance qui sépare chaque individu au centre de gravité du groupe le plus proche (distance de Mahalanobis).

3.2.4 . La géostatistique

La géostatistique est une méthode d'interpolation relativement récente pour étudier la variabilité spatiale. Elle est basée sur la théorie des variables régionalisées, développée par Krige (1951), Matheron (1965), Journé, Huijbregts (1978), Cressie (1991).

Les techniques géostatistiques ont été reprises par de nombreux auteurs en pédologie pour l'estimation des propriétés du sol et la quantification de leur variabilité spatiale. Nous pouvons citer les travaux de Nortcliff (1978), Mc Bratney, al. (1982), Wopereis, al. (1982), Webster, Burgess (1983), Oliver (1984), Gascuel-Oudou (1984), Walter (1990), Chéry (1995) et Douaoui (2005) qui se réfèrent à la géostatistique monoparamétrique et ceux de Webster (1977), Goulard, al. (1978), Rambal (1978), Mc Bratney, Webster (1981) et Douaoui (1993) sur la géostatistique multiparamétrique.

L'intérêt de la géostatistique est de permettre, à la fois, une estimation des propriétés du sol dans l'espace et une mesure de la qualité de cette estimation. Elle a surtout été employée à grande échelle, en sciences du sol, où elle a contribué à la compréhension de la variabilité spatiale de ses propriétés intrinsèques.

Les techniques géostatistiques procèdent en deux étapes qui sont les suivantes :

- l'analyse de la structure spatiale de la variable étudiée grâce au variogramme ;
- l'emploi de la technique de prédiction spatiale par le krigeage.

3.2.4.1 . Le variogramme

L'analyse du variogramme expérimental $M^*(h)$ a pour objectif de mettre en évidence les caractéristiques structurales majeures de la variable régionalisée (VOLTZ AL., 1989).

Pour mesurer la variabilité à différentes échelles d'une variable régionalisée $Z(x)$, on réalise une mesure de dissemblance entre deux données Z_1 et Z_2 situées en deux points x_1 et x_2 d'un domaine spatial (D). Cette dissemblance entre deux valeurs, désignée par M^* , vaudra :

$$M^* = (Z_2 - Z_1)^2 / 2 \quad (10)$$

C'est à dire la moitié du carré de la différence entre les deux valeurs.

On fait dépendre la dissemblance M^* de la distance et de l'orientation d'une paire de points décrits par le vecteur $x_2 - x_1$, indifféremment de la position de la paire de points dans le domaine étudié.

$$M^*(h) = 1/2 [Z(x_1+h) - Z(x_1)]^2 \quad (11)$$

En formant la moyenne des distances $M^*(h)$ entre les valeurs pour les paires de points reliées par le vecteur h donné pour une maille donnée (avec, le cas échéant, une certaine tolérance sur la longueur et l'angle du vecteur), on obtient la notion du variogramme expérimental :

$$M^*(h) = 1 / 2n_h \sum_{\alpha=1}^{n_h} [Z(x_1+h) - Z(x_1)]^2 \quad (12)$$

Habituellement, on observe que la dissemblance des valeurs augmente en moyenne en fonction de l'éloignement spatial des points de mesure et atteint fréquemment un palier de variance pour une distance appelée portée. Lorsque la pente du variogramme change abruptement, on peut penser à des paliers intermédiaires.

Le comportement aux faibles distances près de l'origine du variogramme est d'une importance capitale. Celui-ci est un indicateur du degré de continuité de la variable régionalisée à savoir : différentiable, continue mais non différentiable ou carrément discontinue. Dans ce dernier cas, on aura affaire à une variable régionalisée donnant lieu à un effet pépite, symptôme des valeurs changeant abruptement à très petites distances par rapport à la maille.

Lorsque la dissemblance moyenne des valeurs est constante pour toutes les distances $\square h \square$, il y a une absence complète de la structuration spatiale des valeurs. A l'inverse, une pente non nulle du variogramme près de l'origine indique une structuration des données. Un changement soudain de la pente du variogramme indique le passage à une structuration des valeurs de nature différente.

La généralisation de l'équation (12) suppose deux conditions connues sous le terme

d'hypothèse intrinsèque :

- l'espérance de Z est constante quelle que soit la position de x dans le domaine D ;

$$E [Z(\mathbf{x})] = \text{constante} \quad (13)$$

- pour toute distance h , la différence $[Z(x) - Z(x + h)]$ a une variance finie, indépendante des coordonnées x .

$$\text{VAR} [Z(\mathbf{x}) - Z(\mathbf{x}+h)] = E\{ [Z(\mathbf{x}) - Z(\mathbf{x}+h)]^2 \} = 2 M (h) \quad (14)$$

Une fois le variogramme expérimental calculé, on essaie de lui ajuster un modèle qui rend compte des caractéristiques observées. C'est l'inférence du variogramme. Quatre principaux modèles peuvent lui être ajustés : exponentiel, gaussien, sphérique ou linéaire.

Le variogramme expérimental modélisé est dit variogramme moyen ou omnidirectionnel. Il analyse la structure spatiale d'une variable régionalisée sans tenir compte de la présence d'une éventuelle anisotropie de la variabilité spatiale. Le phénomène d'anisotropie se présente quand la variabilité spatiale n'est plus la même dans toutes les directions c'est à dire qu'il y a des directions privilégiées vis à vis du phénomène étudié (Freycon & Sebastien, 1991). L'orientation préférentielle de ces directions incite à rechercher une éventuelle anisotropie dans la variabilité spatiale des propriétés intrinsèques du sol.

Le calcul de l'anisotropie nécessite la construction de deux variogrammes directionnels modélisés. Le premier est construit dans la direction principale de l'anisotropie et le second dans la direction perpendiculaire à la première. Le rapport entre les valeurs des deux portées relatives aux deux variogrammes directionnels donne la valeur de l'anisotropie, étape qui permet de passer par la suite, au krigeage et donc à l'interpolation.

3.2.4.2 . Le krigeage

Le krigeage est une méthode d'estimation d'un phénomène connu en un certain nombre de points et a l'avantage parmi d'autres techniques d'interpolation d'utiliser la structure spatiale de la variable pour l'estimation (Delhomme, 1978). Ses avantages sont les suivants :

- le krigeage intègre la connaissance émanant de l'analyse préalable de la structure par le variogramme ;
- le krigeage est un interpolateur exact. Si un point d'échantillonnage coïncide avec un nœud de grille, la solution du krigeage est égale à la valeur de l'échantillon ;
- le krigeage livre une interpolation sur l'erreur d'estimation traduite par la variance de krigeage.

Plusieurs auteurs reconnaissent la robustesse de cette méthode (Laslett al., 1987 ; Voltz Al., 1989).

Le variogramme modélisé est utilisé pour déterminer les points \square_i nécessaires à l'interpolation. La variable estimée en un point x_o correspond à la somme pondérée des valeurs des variables mesurées (Beucher-Darricaú, 1981 ; Meylan, 1986).

$$Z_o^* = \sum_{i=1}^n \lambda_i Z_i \quad (15)$$

$$\sum_{i=1}^n \lambda_i = 1 \quad (16)$$

λ_i est le poids associé à la variable mesurée (Z_i) pour l'estimation de Z_o^* .

Les poids λ_i sont choisis de manière à ce que la valeur estimée Z_o^* ne soit pas biaisée ; ceci veut dire que l'erreur moyenne d'estimation doit être nulle (Beucher-Darricaú, 1981; Gascuel-Odoux, 1987) :

$$E (Z_o^* - Z_o) = 0 \quad (17)$$

et la variance de l'estimation est minimale,

$$VAR (Z_o^* - Z_o) = E (Z_o^* - Z_o)^2 \quad (18)$$

Lors du krigeage, si toutes les valeurs de Z_i sont utilisées pour kriger en un point donné, on parle de krigeage simple. Si, seul un sous ensemble de données de Z_i les plus proches est utilisé, on parle alors de voisinage glissant (krigeage ordinaire).

Chapitre IV . Description statistique des variables du sol

Ce chapitre vise à décrire et quantifier l'hétérogénéité spatiale de chacune des variables du sol mesurée en un grand nombre de points de la parcelle d'étude. La taille de l'échantillon utilisée pour l'analyse statistique correspond au nombre de points de prélèvements sur le maillage de la parcelle d'étude plus ceux des deux transects ; soit un total égal à 348 points de sondage.

4.1 . Description de l'hétérogénéité spatiale de la parcelle d'étude

L'allure de l'histogramme de fréquence d'une variable permet de décrire l'hétérogénéité spatiale d'une parcelle selon la classification proposée par Freeze (1975). Il ressort de celle-ci une distribution des fréquences de type unimodale pour toutes les variables, caractérisant ainsi la zone d'étude comme un milieu homogène, non uniforme (*fig. VII*). La considération du milieu d'étude au sens statistique, comme homogène non uniforme, à partir d'une seule réalisation impose l'introduction des deux hypothèses suivantes (Becket al., 1981; Meirvenę Hofman, 1989) :

- hypothèse de stationnarité qui suppose que la loi de probabilité est la même en tout point, donc invariante par translation et cela pour l'ensemble des variables ;
- hypothèse d'ergodicité, qui suppose que la réalisation unique suffit à rendre compte de l'ensemble des réalisations éventuelles.

4.2 . Analyse statistique classique

L'analyse statistique classique unidimensionnelle permet de caractériser statistiquement une propriété mesurée en un grand nombre de points dans l'espace (Gascuel-Oudou, 1984). Elle consiste essentiellement en la détermination de la loi de distribution des variables non uniformes et homogènes et de leurs principaux paramètres statistiques, à savoir la moyenne, l'écart type et le coefficient de variation (Voltz, 1986).

4.2.1 . Recherche de la nature de la loi de distribution

La recherche de la loi de distribution des variables est indispensable lorsqu'on désire traiter les données en terme de valeurs moyennes ou la probabilité d'avoir une valeur inférieure à une certaine limite (Vauclin, 1982). Notons que si les hypothèses d'un milieu homogène (stationnarité et ergodicité) ne sont pas vérifiées, la recherche de l'ajustement de la distribution de nos variables à une distribution théorique est inutile.

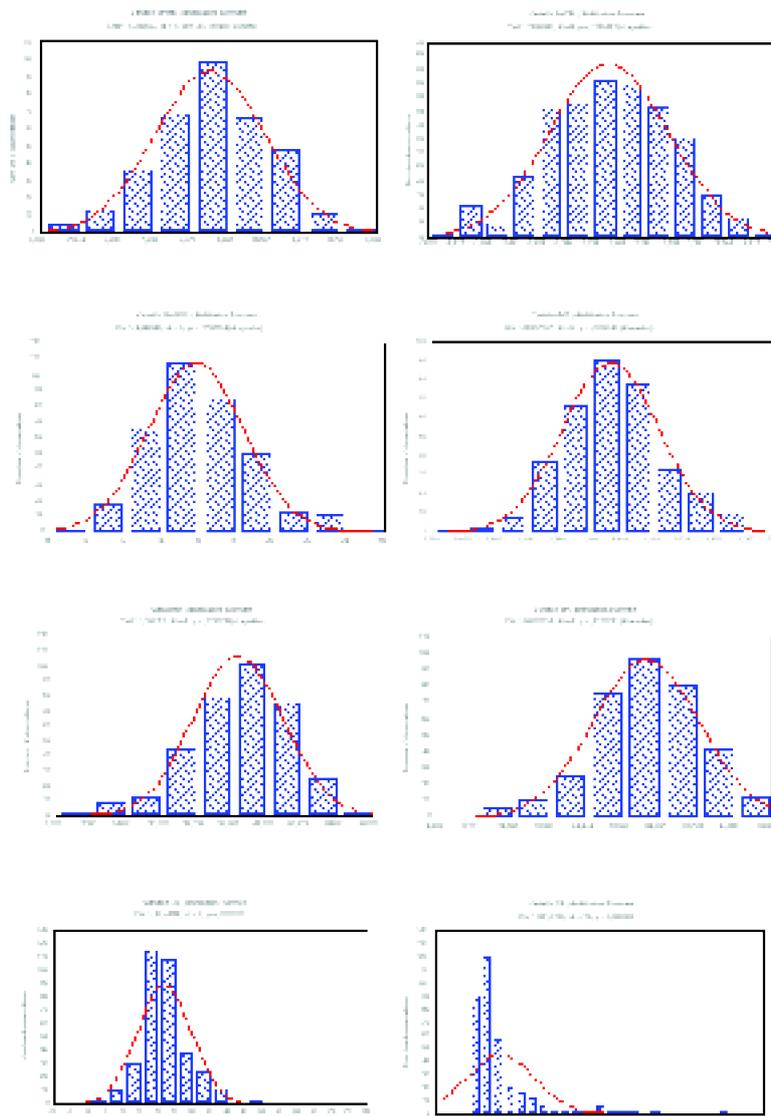


Figure n° 7 . Allure des histogrammes expérimentaux des 11 variables du sol étudiées et leur ajustement à la loi normale.

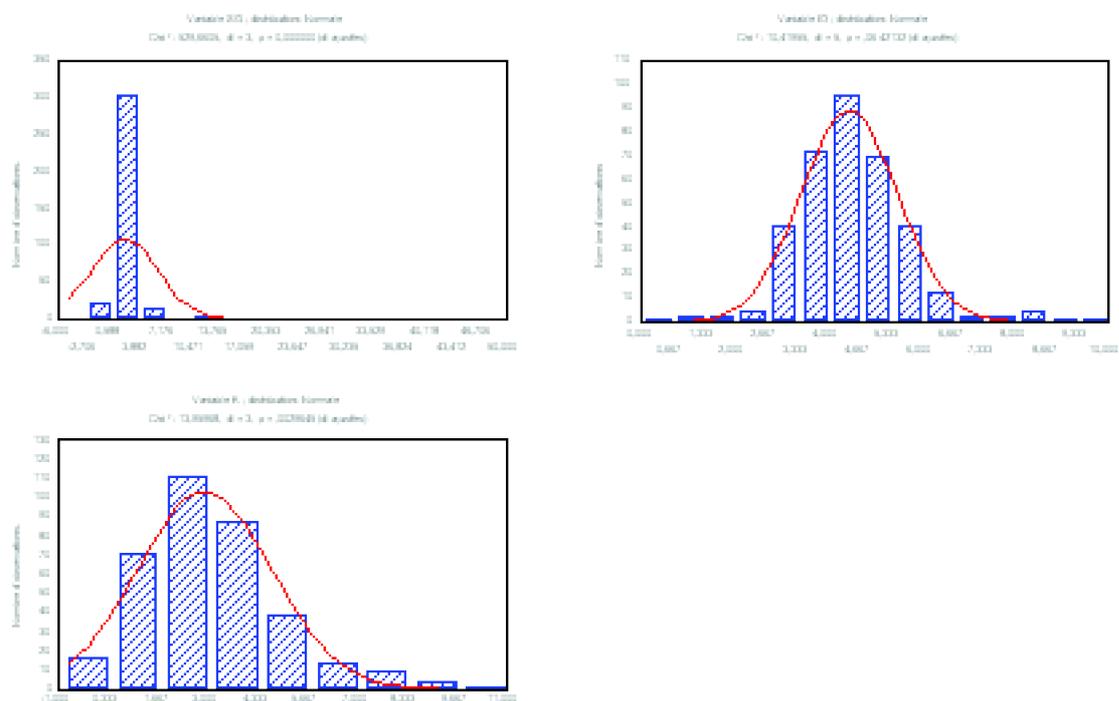


Figure n° 7 (suite) . Allure des histogrammes expérimentaux des 11 variables du sol étudiées et leur ajustement à la loi normale.

Pour chacune des variables étudiées on a émis l'hypothèse a priori que notre échantillon d'observations provenait d'une variable aléatoire de distribution normale ; cette hypothèse n'a pas été rejetée par le test de χ^2 au seuil $\alpha = 0.05$ et cela pour la variable pH_{eau} , $CaCO_3$ total, MO, A, Lf et le I_o . La variable CE a vu son test d'ajustement répondre à la loi log-normale au même seuil de probabilité. Par contre, les variables du sol Lg, Sf, Sg et le K ont vu leur test d'ajustement à la loi normale ou log-normale rejeté par le test de χ^2 (Tab.VII). On utilise ici la distribution de probabilité normale puisque l'on peut faire appel au théorème central de la limite qui suppose une distribution normale d'une moyenne lorsque la taille de l'échantillon est ≥ 30 . Ceci permet d'estimer dans ce cas la moyenne et l'écart-type de ces variables comme dans le cas de la distribution normale.

La qualité d'ajustement de chacune des variables à la loi normale a été estimée par le calcul du coefficient d'aplatissement et d'asymétrie. Rappelant que la tendance d'une distribution des fréquences d'une variable à une loi normale correspond aux valeurs des coefficients les plus convergents à la valeur 0. Il s'avère donc que le meilleur ajustement à la loi normale se confirme pour le pH_{eau} , LnCE, $CaCO_3$ total, MO, A et Lf, alors que l'ajustement qui reste à vérifier est celui du Lg, Sf, Sg et le K (Tab.VII).

Tableau n° VII . Nature de la loi de distribution des fréquences des variables par le test de χ^2 .

Variable	N	Nature de la loi	ddl	χ^2 expérimental	χ^2 théorique	Prob. $\chi^2 =$	Coef.. d'aplatissement	Coef.. d'asymétrie
pH _{eau}	348	LN	4	4.75	9.49	0.31	-0.27	-0.11
CE	348	LnN	8	12.96	15.50	0.11	-0.48	-0.16
CaCO ₃ total	348	LN	3	4.94	7.81	0.17	0.03	0.41
MO	348	LN	5	8.31	11.10	0.13	-0.22	-0.08
A	348	LN	4	7.25	9.49	0.12	-0.16	-0.33
Lf	348	LN	4	5.60	9.49	0.23	0.47	-0.86
Lg	348	\square	-	-	-	-	4.73	1.31
Sf	348	\square	-	-	-	-	9.37	2.78
Sg	348	\square	-	-	-	-	66.76	7.77
lo	348	LN	5	10.51	11.10	0.06	3.77	0.94
K	348	\square	-	-	-	-	1.17	0.99

4.2.2 . Analyse des niveaux de la variabilité spatiale des variables du sol dans la parcelle d'étude

Le pH_{eau} est la variable du sol la plus centrée autour de sa valeur moyenne comme le témoigne la faible valeur de son CV égal à 4 %. Elle constitue donc la variable du sol la plus homogène dans la zone d'étude. Par contre, la CE est une variable du sol des plus dispersée, compte tenu de la valeur élevée qu'affiche son CV égal à 104.86 %.

D'autre part, pour les variables liés à la constitution du sol, le CaCO₃ total, la MO, l'A, le Lf et le Lg ont des coefficients de variation qui montrent une faible dispersion de leurs teneurs par rapport à leurs moyennes. Les valeurs de leurs CV respectifs sont : 16 %, 27 %, 24 %, 22 % et 36 %. Par ailleurs, pour les paramètres physiques liés à la stabilité structurale, le lo a une faible variabilité, indiquée par un CV égal à 23 %. Pour le K, le CV de 63 % est plutôt élevé pour montrer une grande variabilité de ce paramètre par rapport à sa moyenne.

Les moyennes des variables du sol situées dans les intervalles de confiances, au risque de 5 % (tab. VIII) font apparaître que le sol de la parcelle d'étude est de nature basique d'après son pH_{eau} égal à 7.90 (AFES-INRA, 1988). Avec une moyenne de la CE égale à 3.29 dS/m, le sol de la zone d'étude est salé (Richards, 1954). Il est organo-minérale d'après sa teneur en MO égale à 2.69 % (AFES, 1995). Il est modérément calcaire avec une teneur en CaCO₃ total de 15.91 % (d'après GEPPA). Sa texture est généralement argilo-limoneuse (Jamagne, 1967). Il présente aussi une très faible stabilité structurale d'après la valeur moyenne de l'indice d'instabilité (I_o) égale à 4.42 (Monnier Stengel, 1982), (Tab. VIII).

Tableau n° VIII . Paramètres statistiques descriptives des variables d'étude.

Variable Du sol	Moyenne Estimée	Écart-type Estimé (m)	Coef.. de Variation (%)	Inter.. de Conf.. (m - \square x) (1- \square = 0.95)	Inter..de Conf.. (m + \square x) (1- \square = 0.95)
-----------------	-----------------	-----------------------	-------------------------	---	--

				%)	%)
pH _{eau}	7.90	0.36	4	7.86	7.94
CE	3.29	3.45	104.86	2.95	3.64
CaCO ₃ total	15.91	2.56	16	15.63	16.18
MO	2.69	0.73	27	2.61	2.77
A	38.01	9.49	24	37.01	39.01
Lf	32.78	7.40	22	32.00	33.56
Lg	21.84	8.03	36	20.99	22.68
Sf	5.37	6.05	112	4.74	6.01
Sg	2.15	4.20	195	1.70	2.59
lo	4.42	1.03	23	4.32	4.53
K	2.88	1.79	62	2.70	3.07

Conclusion

Les variables du sol étudiées au niveau de la parcelle sont toutes homogènes et non uniformes (unimodales), sous hypothèse de stationnarité et d'ergodicité.

Relativement à la moyenne et sous hypothèse d'une distribution normale de chaque variable du sol, il en ressort d'après les valeurs des CV des niveaux de variabilité différents. Le pH_{eau} basique, est de très faible variabilité. La salinité du sol est élevée et très variable. Le sol, qualifié d'organo-minérale, a une faible variabilité de sa teneur en MO de même pour sa teneur modérée en CaCO₃ total. Il présente dans l'ensemble une texture argilo-limoneuse. Ses paramètres granulométriques A, Lf et Lg ont une très faible variabilité, contrairement au Sf et au Sg qui sont très variables. Il a aussi une très faible stabilité structurale d'après le I₀ avec une faible variabilité de ce dernier. Pour le K, les résultats ont montré une variabilité élevée.

Chapitre V . Détermination des unités homogènes du sol par l'analyse des composantes principales

L'objet de ce chapitre est essentiellement descriptif dans le sens où la relation entre les individus est déduite relativement à leur position dans l'espace décrit par l'ACP. Notons au préalable que la notion de l'individu introduite dans ce chapitre est attribuée à la fois aux variables et aux sondages du sol.

Nous examinerons la relation entre les individus en fonction de l'objectif à atteindre. Ainsi la description de la relation entre les variables est faite selon leur direction. Cette relation permet de répondre à la question : comment les variables sont-elles structurées ? Par contre, celle relative aux 348 sondages, elle est en rapport avec la distance de rapprochement entre celles-ci. Par cette relation de distance on essaye de repérer d'éventuels groupements de sondages qui forment des ensembles pédologiques cohérents. La visualisation de l'ensemble des unités homogènes au niveau de la parcelle d'étude est faite par attribution d'un caractère identifiant l'appartenance de chaque sondage à son ensemble pédologique.

5.1 . Notation et résultats préliminaires

5.1.1 . La matrice des données de base

L'organisation préliminaire des données de base correspond à un tableau à double entrée dans lequel les lignes de ce tableau (i) représentent les sondages et les colonnes (j) les propriétés physiques et chimiques du sol (Tab. IX). Ce tableau présenté sous cette forme est appelé, matrice des données de base. Chaque élément $x_{i,j}$ de cette matrice représente la valeur de la variable j pour un sondage correspondant à la ligne i du tableau ; ainsi par exemple l'élément $x_{3,4} = 2.69$ représente la teneur de la MO du troisième sondage du sol.

Tableau n° IX . Matrice représentant la structure des données de base

Variables (j) Sondages (i)	pH _{eau}	CE	CaCO ₃	MO	A	Lf	Lg	Sf	Sg	I ₀	K
1	37,50	48,01	187,67	9,25	3,25	1,28	1,50	1,32	1,40	1,60	1,50
2	37,50	48,01	187,67	9,25	3,25	1,28	1,50	1,32	1,40	1,60	1,50
3	37,50	48,01	187,67	9,25	3,25	1,28	1,50	1,32	1,40	1,60	1,50
4	37,50	48,01	187,67	9,25	3,25	1,28	1,50	1,32	1,40	1,60	1,50
5	37,50	48,01	187,67	9,25	3,25	1,28	1,50	1,32	1,40	1,60	1,50
- - - 346	37,50	48,01	187,67	9,25	3,25	1,28	1,50	1,32	1,40	1,60	1,50

5.1.2 . Relation entre les variables du sol

Les résultats de l'ACP étant influencés par les différences d'unités des variables (Philippeau, 1986), nous allons donc donner, dans ce travail, le même poids aux différentes variables d'où l'utilité de travailler sur des données centrées et réduites, c'est à dire sur des valeurs des écarts à la moyenne, divisées par leurs écart-types (valeurs standardisées).

Les coefficients de corrélations linéaires (r) entre les variables du sol sont faibles dans l'ensemble (Tab.X). L'application du test de conformité du coefficient de corrélation sur l'ensemble des variables révèle des corrélations significatives entre certaines variables du sol au seuil critique de signification α égal à 5 %.

Ces corrélations sont négatives, respectivement en décroissance entre les couples de variables du sol : A-Sf, pH_{eau}-CE, Lf-Lg, A-Lg, CE-K, CE-CaCO₃, A-Sg, MO-pH_{eau}, CaCO₃-MO, Lf-Sg, CaCO₃-A, MO-K, Lf-Sf, K-I₀, A-Lf, Lg-Sg, I₀-Sg, pH_{eau}-I₀, MO-Sg, Lf-K. Elles sont positives, respectivement en décroissance entre les couples de variables du sol : pH_{eau}-K, CE-MO, CE-I₀, CaCO₃-K, CaCO₃-Sf, Sf-Sg, pH_{eau}-CaCO₃, Lg-K, Sg-K.

Tableau n° X . Coefficients de corrélations linéaires (r) entre les variables du sol.

Variable	pH _{eau}	CE	CaCO ₃	MO	A	Lf	Lg	Sf	Sg	K	I ₀
pH _{eau}	1										
CE	-0.50	1									
CaCO ₃	0.15	-0.24	1								
MO	-0.23	0.33	-0.23	1							
A	0.00	0.00	-0.21	0.01	1						
Lf	0.05	-0.09	0.01	-0.00	-0.18	1					
Lg	-0.03	0.07	0.06	0.05	-0.43	-0.43	1				
Sf	-0.08	0.04	0.22	-0.01	-0.57	-0.20	0.00	1			
Sg	0.06	-0.05	0.02	-0.11	-0.24	-0.22	-0.17	0.19	1		
K	0.36	-0.30	0.27	-0.21	-0.05	-0.10	0.12	-0.03	0.12	1	
I ₀	-0.12	0.32	0.06	-0.02	0.03	0.02	-0.01	0.03	-0.13	-0.20	1

En gras, corrélations significatives au seuil $\alpha = 5\%$. (la valeur critique est de 0.10 pour 346 degré de liberté).

5.2 . Caractéristiques des axes des composantes principales

La matrice de corrélation représentée par les valeurs des coefficients de corrélations entre les variables du sol prises deux à deux est une matrice carrée et symétrique. Ladiagonalisation de celle-ci fournit une matrice diagonale des valeurs propres et une matrice des vecteurs propres.

La valeur propre ou variance exprime le degré de dispersion des variables ou des sondages du sol par rapport à leur point d'origine sur l'axe d'une composante principale. La variance expliquée par chaque axe est exprimée en pourcentage par rapport à la variance totale.

Dans notre cas (*Tab. XI*), la première composante principale a une variance de 2,30. Elle explique 20,90 % de la variance totale. La variance de la deuxième composante principale est égale à 1,90 . Elle exprime 17,30 % de la variance totale. Celle de la troisième composante principale est de 1,37, représentant 12,50 % de la variance totale.

Les différentes composantes principales étant non corrélées entre elles. La part de l'information expliquée par deux axes du même plan est égale à la simple addition des parts des variances expliquées par chacun d'eux. Ainsi, le plan formé par les deux premiers axes explique 38,20 % de la variance totale. Le plan formé par le premier et le troisième axe explique 33,40 % de la variance totale. Les deux plans formés par les trois premiers axes correspondent à un total de variance égal à 5,57. Ils représentent un total de 50,70 % de la variance totale. Ce faible taux de variation expliqué par les trois premières composantes principales est dû à la faible corrélation entre les paramètres.

Tableau n° XI . Variances des axes principaux.

Numéro De l'axe	Variance	% total de la variance	Cumul des variances	Cumul des Variances en %
1	2.30	1.90	1.37	20.19
2	1.90	1.57	2.94	38.20
3	1.90	1.57	4.51	50.70
4	1.90	1.57	6.08	61.80
5	1.90	1.57	7.65	71.70

5.3 . Analyse de la relation entre les individus

5.3.1 . La relation entre les variables du sol

L'ACP, par son traitement multiparamétrique, permet de mettre en évidence les interactions entre les variables du sol. En effet, la projection de toutes les variables dans l'espace factoriel décrit par l'ACP est exprimée par le coefficient de corrélation (r) de chacune de ces variables (Tab. XII). La valeur de ce coefficient élevée au carré représente la part de la variation d'une variable expliquée par une composante principale.

L'ensemble des variables du sol qui contribuent fortement à l'axe 1 sont les variables qui ont les valeurs du coefficient au carré le plus élevé ; il s'agit en l'occurrence et par ordre décroissant de la CE, pH_{eau} , K, MO, CaCO_3 total et I_0 d'après leurs valeurs respectives de r^2 égales à 0.53, 0.43, 0.43, 0.28, 0.27 et 0.12.

La deuxième et la troisième composantes principales sont qualifiées d'axes des variables granulométriques vu les valeurs élevées de leurs r^2 . Ainsi, la formation de la deuxième composante principale résulte de la contribution de la variable A et Sf comme le témoigne leurs valeurs élevées de r^2 respectivement égales à 0.57 et 0.52 . Par contre, les variables qui contribuent le plus à la formation de la troisième composante principale sont le Lf et le Lg d'après leurs valeurs respectives de r^2 égales à 0.54 et 0.36.

La part de la variation d'une variable dans un plan est également expliquée en faisant la somme de ses r^2 correspondant aux deux axes qui le composent. Ainsi, les variables qui contribuent le plus à la formation du plan 1-2 sont : la CE, A, Sf et le pH_{eau} . Ceux qui contribuent à la formation du plan 1-3 sont: Lf, CE, K, pH_{eau} , Lg et la MO . La figure 8a présente le plan 1-2 porteur de plusieurs significations. La CE, la MO et le I_0 sont situés sur la même direction que le pH_{eau} et le K. Ces deux groupements de variables expriment une relation d'opposition entre eux par rapport au point d'origine. Cette opposition se manifeste également entre les fractions fines A et Lf et les fractions grossières Lg, Sf, Sg. La figure 8b représentant le plan 1-3, montre une autre forte opposition ; celle relative à I_0 et K, pH_{eau} et I_0 , CE et CaCO_3 total, MO et CaCO_3 total, Lf et Lg et enfin entre A et Sf.

VARIABLES	Axe 1		Axe 2		Axe 3		Plan 1-2	Plan 1-3
	r_1	r_1^2	r_2	r_2^2	r_3	r_3^2	$r_1^2+r_2^2$	$r_1^2+r_3^2$
pH _{eau}	-0.66	0.43	-0.31	0.09	-0.05	0.00	0.53	0.43
CE	0.73	0.53	0.32	0.10	-0.02	0.00	0.64	0.53
CaCO ₃	-0.52	0.27	0.21	0.04	0.23	0.05	0.31	0.26
MO	0.53	0.28	0.12	0.01	-0.14	0.02	0.30	0.30
A	0.28	0.07	-0.75	0.57	-0.32	0.10	0.64	0.17
Lf	0.04	0.00	-0.37	0.14	0.73	0.54	0.14	0.54
Lg	-0.08	0.00	0.55	0.30	-0.60	0.36	0.31	0.36
Sf	-0.18	0.03	0.72	0.52	0.34	0.11	0.55	0.14
Sg	-0.29	0.08	0.26	0.06	0.09	0.00	0.15	0.08
K	-0.65	0.43	-0.05	0.00	-0.29	0.08	0.43	0.51
I ₀	0.35	0.12	0.11	0.01	0.24	0.06	0.13	0.18

Tableau n° XII . Coefficients de corrélations entre variables du sol et les trois premiers axes principaux.

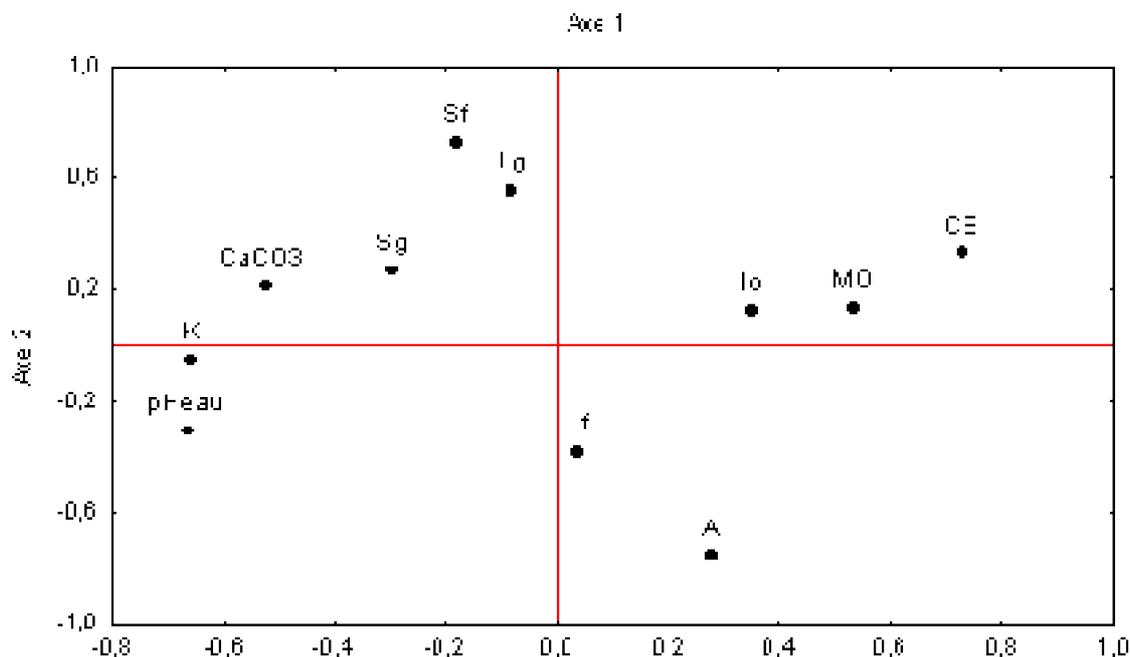


figure n°8a. Projection des 11 variables du sol dans le plan de l'axe 1 et l'axe 2

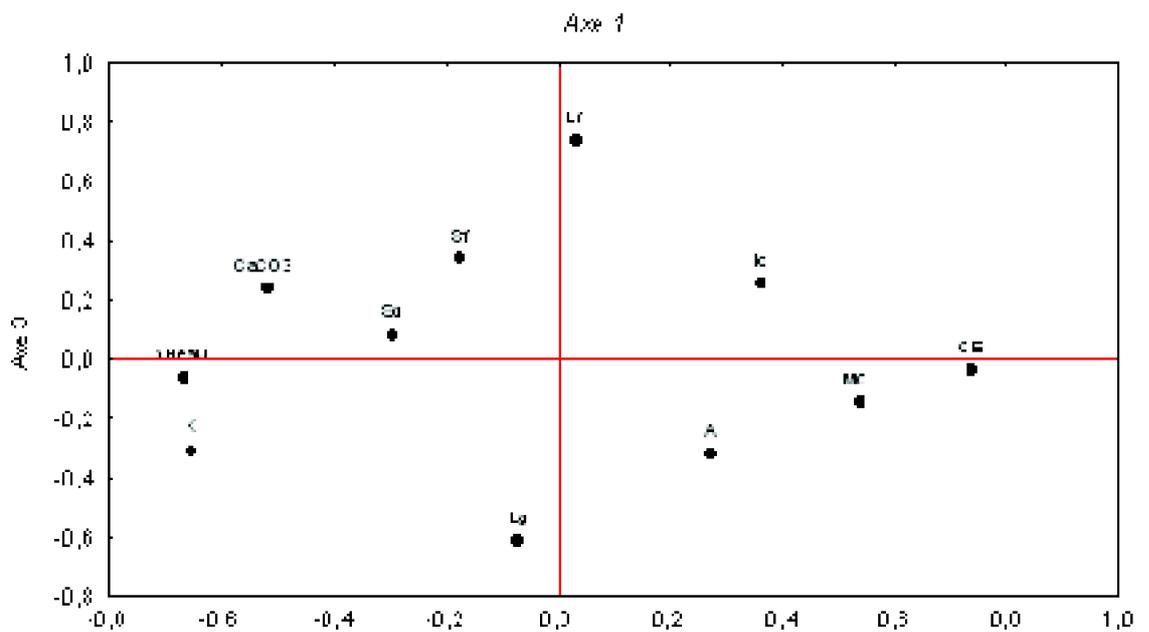


figure n°8b. Projection des 11 variables du sol dans le plan de l'axe 1 et l'axe3

5.3.2 . La relation entre les sondages du sol

Les deux plans illustrés par les figures 9 et 10 sont issus des trois premières composantes principales. Ces plans permettent de mieux repérer l'emplacement des sondages ainsi que l'existence d'éventuels groupements entre eux. Ces sondages sont repérés dans ces plans par leurs coordonnées sur les axes des composantes principales. Il est également à signaler que le degré de dispersion de ces sondages par rapport à leurs points d'origine est analogue à celui des variables. Cette information permet de déduire comme pour ces dernières, une faible qualité de représentation des sondages dans le plan 1-2 et le plan 1-3. On conséquence on peut dire que la prise de vue séparée de ces deux plans nous éloigne de la réalité sur la distance réelle entre les sondages du sol.

La forte densité de regroupement des sondages au milieu des plans et la proximité élevée d'un grand nombre d'entre eux rend encore plus difficile la distinction ou la séparation des groupes de sondages (*fig.9 et 10*). L'exemple est donné par les sondages 223 et le 227 qui sont plus proches l'un de l'autre sur le plan 1-2 mais très éloignés entre eux sur le plan 1-3.



Figure n°9. Vue de la projection des 348 sondages du sol dans le plan 1-2

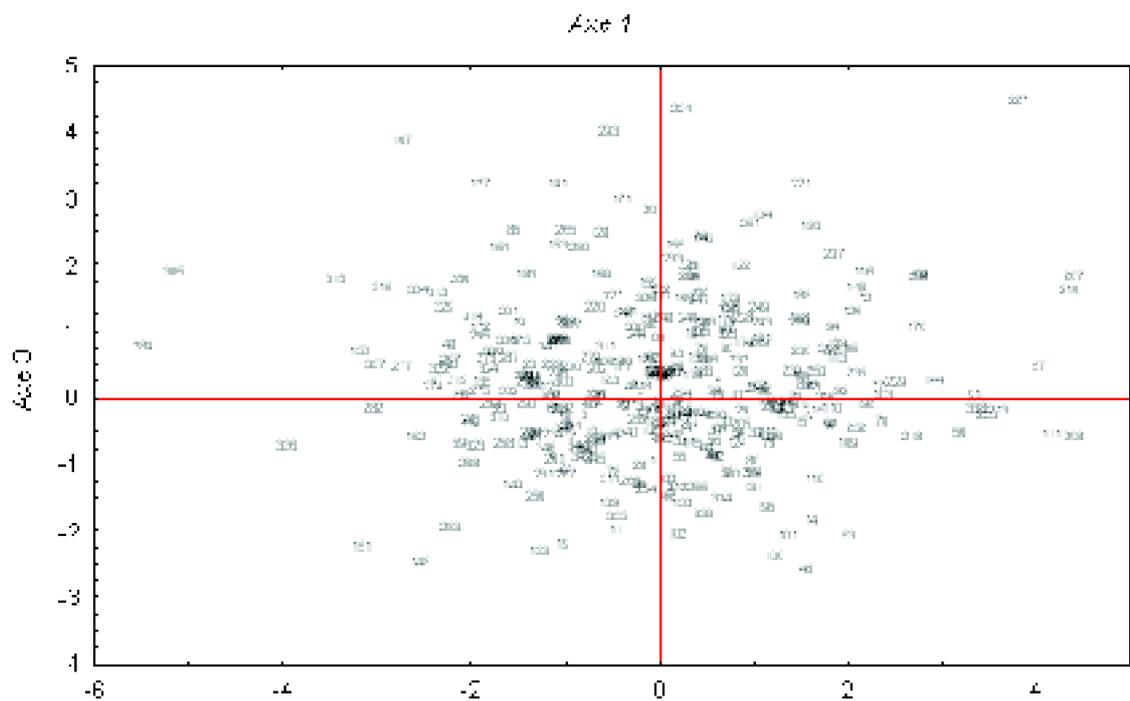


Figure n°10. Vue de la projection des 348 sondages du sol dans le plan 1-3

Conclusion

Globalement les relations entre les variables du sol sont faibles. La vue de la projection des sondages du sol dans les deux plans factoriels issus des composantes principales ne montre pas de groupements distincts pouvant servir à la constitution d'ensembles pédologiques cohérents.

Enfin, notons que les résultats statistiques obtenus par l'ACP n'ont pas été très conséquents, vu le faible pourcentage de l'information expliquée par les plans des variables et des sondages du sol.

On montre dans le chapitre qui suit comment ressortir dans cet espace décrit par l'ACP, les ensembles de sondages pédologiques les plus cohérents en utilisant la classification ascendante hiérarchique. Celle-ci, permet de comprendre d'avantage l'organisation des sondages du sol par combinaison de l'information expliquée par les deux plans.

Chapitre VI . Contribution de la classification ascendante hiérarchique à la détermination des unités homogènes du sol

Dans ce chapitre, on constitue les groupements des ensembles pédologiques les plus cohérents en utilisant la classification ascendante hiérarchique (CAH) dont le déroulement de l'algorithme impose au préalable le choix des critères suivants :

- un critère de ressemblance entre les sondages du sol. Il est représenté ici par la distance euclidienne. A chaque étape de partition, l'algorithme agrège les deux classes de sondage les plus proches en une seule entité fictive qu'on surnomme également classe. Dans ce chapitre, on attribue l'expression "classe de sondage " à un ou plusieurs sondages exprimant une entité homogène.
- un critère d'agrégation appelé indice de Ward. Celui-ci mesure le gain d'inertie intra-classe résultant de la fusion de deux classes de sondage à chaque étape d'une partition. C'est un critère d'aide à la décision utilisé dans le choix du nombre de classes les plus cohérentes.

6.1 . La structure des données de base

Les données de base se référant à ce chapitre sont issues de la transformation orthogonale de la matrice $X_{i,j}$, présentée dans le chapitre précédent (*tab. IX*). Elles représentent la matrice $Y_{i,p}$ des coordonnées des 348 sondages du sol dans l'espace décrit par les trois premiers axes des composantes principales. Rappelons que ces axes sont perpendiculaires l'un par rapport à l'autre et sont porteurs d'un maximum d'informations en terme de variance (50.70 %). Chacun des éléments $y_{i,p}$ de la matrice représente une valeur d'un sondage projeté sur l'axe p (*tab.XIII*).

Tableau n° XIII . Matrice des coordonnées des sondages du sol ($Y_{i,p}$) projetés dans l'espace factoriel de l'ACP.

Axe p Sondage i	1	2	3
1	-0.08	0.59	-1.20
2	0.83	0.53	-0.78
3	-0.18	-0.54	-0.14
4	0.63	-1.01	0.01
5	-0.02	-1.32	-0.45
6	-0.81	0.66	-0.52
7	-0.91	1.03	-0.85
8	-1.13	0.40	-1.09
-	-	-	-
346	-1.20	-1.02	-1.01
347	-1.02	-1.28	-0.71
348	0.15	-0.71	-0.82

L'initialisation dans la CAH consiste à réaliser un tableau des distances entre les 348 sondages pris deux à deux (*tab. XV*). Ce tableau représente une matrice diagonale $D_{i,i}$ dont les valeurs de ses éléments diagonaux sont toutes nulles. Celle ci est déduite d'une autre matrice diagonale $\square_{i,i}$ (*tab. XIV*) telle que :

$\square_{i,i} = Y_{i,p} \cdot Y'_{p,i}$; où $Y'_{p,i}$ représente la transposée de la matrice des coordonnées des sondages $Y_{i,p}$. Ainsi, les distances de départ entre tous les sondages du sol pris deux à deux sont déduites à partir des éléments de la matrice $\square_{i,i}$ comme suit :

$$D_{i,i'} = (\square_{i,i} + \square_{i',i'} - 2\square_{i,i'})^{1/2} ; \text{ où } i, i' \text{ représentent deux des 348 sondages projetés dans l'espace de l'ACP.}$$

Tableau n° XIV . Matrice des paramètres de distance ($\square_{i,i}$) d'une partie des 348 sondages.

n°	1	2	3	4	5	6	7	8	9	10	-	-	346	347	348
1	1.80	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	1.19	1.59	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-0.13	-0.33	0.35	-	-	-	-	-	-	-	-	-	-	-	-
4	-0.67	-0.02	0.43	1.42	-	-	-	-	-	-	-	-	-	-	-
5	-0.24	-0.37	0.79	1.31	1.95	-	-	-	-	-	-	-	-	-	-
6	1.09	0.09	-0.13	1.19	0.62	1.37	-	-	-	-	-	-	-	-	-
7	1.72	0.46	-0.26	-1.63	-0.95	1.88	2.64	-	-	-	-	-	-	-	-
8	-0.75	-1.82	-1.09	-2.22	-1.59	0.10	0.37	0.36	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
346	0.70	-0.75	0.93	0.27	1.84	0.81	0.90	2.05	0.65	-0.09	-	-	3.52	-	-
347	0.17	-0.97	0.99	0.64	2.05	0.34	0.22	1.43	0.74	-0.12	-	-	3.27	3.21	-
348	0.54	0.38	0.48	0.81	1.30	0.17	0.17	0.43	0.67	0.67	-	-	1.38	1.35	1.20

A titre d'exemple, le calcul de la distance entre le sondage 1 et 348 dans l'espace décrit par les trois premières composantes principales se fait comme suit : $D_{1, 348} = (1.80 + 1.20 - 2(0.54))^{1/2} = 1.38$

Tableau n° XV . Représentation partielle de la matrice des distances ($D_{i,i}$) entre quelques sondages du sol .

n°	1	2	3	4	5	6	7	8	9	10	-	-	346	347	348
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	1.008	0	-	-	-	-	-	-	-	-	-	-	-	-	-
3	1.562	1.619	0	-	-	-	-	-	-	-	-	-	-	-	-
4	2.141	1.752	0.957	0	-	-	-	-	-	-	-	-	-	-	-
5	2.064	2.074	0.849	0.862	0	-	-	-	-	-	-	-	-	-	-
6	0.999	1.670	1.417	2.280	2.141	0	-	-	-	-	-	-	-	-	-
7	1.001	1.820	1.879	2.709	2.551	0.506	0	-	-	-	-	-	-	-	-
8	1.080	2.000	1.646	2.524	2.153	0.709	0.707	0	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-
346	1.980	2.512	1.417	2.098	1.334	1.806	2.084	1.434	1.630	2.133	-	-	0	-	-
347	2.161	2.601	1.254	1.830	1.094	1.973	2.324	1.734	1.475	2.079	-	-	0.430	0	-
348	1.386	1.422	0.775	1.007	0.736	1.710	2.047	1.729	0.545	0.837	-	-	1.399	1.311	0

6.2 . Résultats de la partition des classes de sondage du sol

6.2.1 . Exploration du dendrogramme de la partition

Le dendrogramme de la partition des 348 sondages est défini comme un mode de représentation des résultats de la CAH (*fig.11*). Pour interpréter le dendrogramme, il faut considérer ce dernier comme un objet mobile. En conséquence, l'ordre de regroupement des 348 sondages sur le dendrogramme ne constitue qu'une projection dans un hyper-plan d'un ensemble d'axes qui de fait, est en trois dimensions.

Les paramètres dendrographiques constituant la hiérarchie de cette partition sont représentés par le " numéro du nœud " et de " l'indice de niveau ". Chaque regroupement de deux classes dans l'espace en trois dimensions est représenté dans le dendrogramme par un nœud dont le numéro suit exactement celui du regroupement précédent. Le numéro du nœud indique également la fin d'une étape de partition pour laquelle est attribué un indice de niveau exprimant l'importance de la différenciation ou d'individualisation des classes de sondage formées après chaque étape d'une partition.

On attribue à l'inertie inter-classe la notion de différenciation qui exprime une dispersion entre les classes formées et à l'inertie intra-classe la notion d'individualisation exprimant leur compacité. Prenons le 345^{ème} nœud relatif à la 345^{ème} partition (*tab. XVI*). A la fin de cette étape, les 348 sondages du sol se retrouvent regroupés en trois classes différenciées entre elles par une quantité d'inertie inter-classe égale à 33.02 %. Autrement dit, une individualisation de ces trois classes donnant une quantité d'inertie intra-classe égale à 66.97 % de l'inertie totale.

Le dendrogramme de partition coupé à un même niveau d'inertie intra-classe parallèlement à l'axe des sondages permet de voir que tout ce qui est à droite de la coupure, appartient à une même classe. En le parcourant de la droite vers la gauche on a les segmentations suivantes :

- a) la segmentation à deux classes . Elle est issue à un niveau de coupure réalisé entre 346^{ème} et le 347^{ème} nœud. Elle aboutit à la formation des classes de sondage A et B. La classe A est plus hétérogène que la classe B, vu son niveau d'inertie intra-classe supérieur à celui de la classe B.

(Inertie intra-classe) (niveau) (noeud) 1000

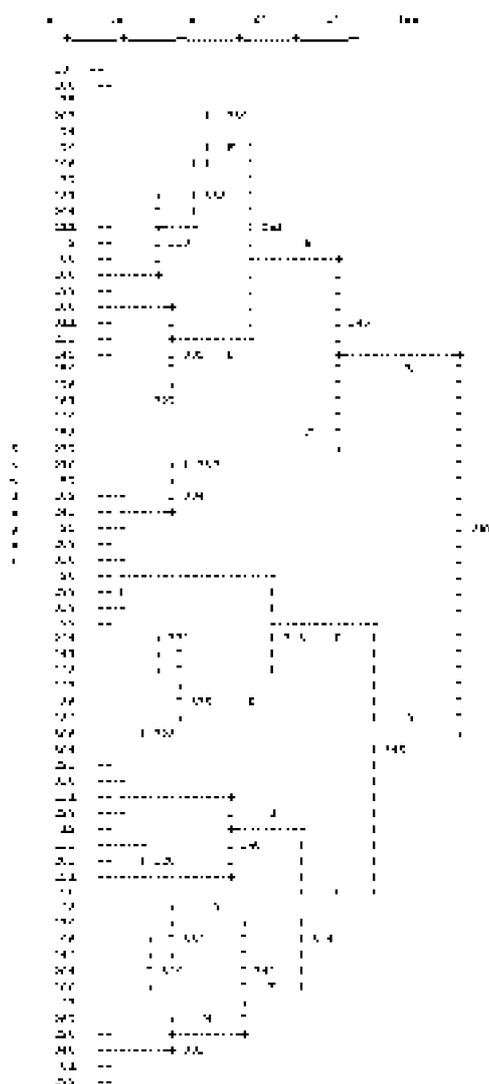


Figure n° 11 . Dendrogramme de partition, des 348 sondages, constitué par la CAH.

- b) la segmentation à trois classes est engendrée à un niveau de coupure situé entre le 345^{ème} et le 346^{ème} noeud. Celle-ci aboutit à la formation des classes A, B et C. La classe C devient davantage homogène que la classe A, précédente, puisque s'en est détaché la classe D qui participait grandement à la valeur élevée de la distance intra-classe de A. La distance intra-classe de B reste la même que dans la précédente partition mais elle est plus élevée que celle de C et de D. A cette étape de partition l'hétérogénéité est beaucoup plus apparente dans la classe B, suivie de C puis D.
- c) la segmentation à quatre classes C, D, E et J se forme par la coupure à un niveau d'inertie intra-classe situé entre le 344^{ème} et 345^{ème} noeud. Les classes C et D gardent les mêmes caractéristiques que dans l'étape précédente. C'est donc l'ancienne classe B qui se segmente en deux classes E et J. La classe E est hétérogène par rapport à la classe J d'après son indice de niveau élevé par rapport à

cette dernière.

- d) la segmentation à cinq classes est issue de la coupure située entre le 343^{ème} et le 344^{ème} nœud. A ce niveau, les classes D , E et J restent non affectées par la segmentation. Elles gardent les mêmes caractéristiques que dans la partition précédente. C'est donc la classe C qui se segmente en deux classes G et F. La classe J reste la plus homogène puisque ayant le plus faible niveau d'inertie intra-classe dans l'ensemble.

6.2.2 . Algorithme de la partition des classes de sondage

Le nombre des étapes de partition de l'algorithme étant élevé (347), on résume dans les paragraphes qui suivent, l'algorithme de la CAH à quatre étapes. Elles sont composées des deux premières et des deux dernières étapes. L'algorithme (*fig. 11*) dispose avant l'opération de partition de 348 sondages du sol; chacun d'eux représente sa propre classe. A cette étape, l'inertie intra-classe est nulle et l'inertie inter-classe est égale à 5.57, valeur relative à 100 % de l'inertie totale. Par la suite, l'algorithme opère par étapes de la façon suivante :

- dans la 1^{ère} étape de partition, il recherche dans la matrice des distances de départ (*tab. XV*), les deux sondages ayant la plus petite distance entre eux et décide alors de les regrouper en une seule entité fictive. Il s'agit du sondage 30 et 108 distancés de 0.04. A la fin de cette étape, l'algorithme de partition dispose de 347 classes dont une classe fictive et le reste réel. Cette étape génère une faible quantité d'inertie intra-classe par rapport à l'inertie totale, égale à 10^{-5} % ;

Tableau n° XVI . Inertie inter-classe, intra-classe et totale par étapes de partition.

Etape de partition	Nombre de Classes	Inertie inter-classe	Inertie Intra-classe	Inertie Inter-classe (%)	Inertie Intra-classe (%)
347	1	0	5,57	0	100
346	2	1,18	4,39	21,21	78,78
345	3	1,84	3,73	33,02	66,97
344	4	2,27	3,30	40,71	59,28
343	5	2,76	2,81	49,49	50,50
342	6	3,07	2,49	55,18	44,81
341	7	3,29	2,28	59,12	40,87
340	8	3,47	2,10	62,29	37,70
339	9	3,79	1,78	67,97	32,02
338	10	3,91	1,65	70,24	29,75
-	-	-	-	-	-
328	20	4,61	0,96	82,67	17,32
-	-	-	-	-	-
308	40	5,10	0,47	91,51	8,48
-	-	-	-	-	-
288	60	5,30	0,27	95,11	4,88
-	-	-	-	-	-
268	80	5,39	0,18	96,68	3,31
-	-	-	-	-	-
248	100	5,44	0,13	97,66	2,33
-	-	-	-	-	-
208	140	5,50	0,06	98,76	1,23
-	-	-	-	-	-
168	180	5,54	0,03	99,36	0,63
-	-	-	-	-	-
128	220	5,56	0,01	99,69	0,30
-	-	-	-	-	-
98	250	5,56	0,00	99,83	0,16
-	-	-	-	-	-
68	280	5,57	0,00	99,92	0,07
-	-	-	-	-	-
38	310	5,57	0,00	99,97	0,02
-	-	-	-	-	-
8	340	5,57	0,00	99,99	0,00
7	341	5,57	0,00	99,99	0,00
6	342	5,57	0,00	99,99	0,00
5	343	5,57	0,00	99,99	0,00
4	344	5,57	0,00	99,99	0,00
3	345	5,57	0,00	99,99	0,00
2	346	5,57	0,00	99,99	0,00
1	347	5,57	0,00	100,00	0,00
0	348	5,57	0,00	100,00	0,00

- dans la 2^{ème} étape de partition, l'algorithme recherche de nouveau les deux sondages les plus proches. Cette recherche se fait dans une nouvelle matrice des distances résultant de l'agrégation précédente. Pour cela, il calcule d'abord les nouvelles distances entre la classe fictive et les sondages réels restants. Les autres distances restent inchangées. Il s'agit du sondage 119 et 194 distancés de 0.05, que l'algorithme agrège, en une seule classe fictive. A la fin de cette étape, il reste 346 classes, individualisées par une quantité d'inertie intra-classe légèrement supérieure à celle de l'étape précédente. Cette inertie est égale à 5.10^{-5} % de l'inertie totale ;
- on arrivant à la 346^{ème} étape de partition, l'algorithme dispose de 3 classes fictives B, C et D. Il procède par la suite à la formation de la classe A en agrégeant les deux classes de sondage fictives C et D ayant la plus faible distance dans l'ensemble. A la fin de cette étape, il ne reste que 2 classes fictives A et B avec une quantité d'inertie intra-classe égale à 78.78 % de l'inertie totale ;
- à la 347^{ème} étape de partition, l'algorithme réitère le même processus que celui de l'étape précédente on agrégeant les deux dernières classes fictives A et B en une seule classe regroupant les 348 sondages de départ. Cette étape marque la fin de l'opération d'agrégation par égalisation d'une part de la quantité d'inertie intra-classe à l'inertie totale et d'autre part, par une annulation de la quantité d'inertie inter-classe.

6.2.3 . Le choix du nombre de classes de sondage du sol

On utilise l'indice de Ward pour choisir la partition qui tend à disperser au mieux les classes. Rappelons que l'indice de Ward est un gain d'inertie intra-classe résultant de l'agrégation de deux classes à chaque étape d'une partition. Plus cet indice est grand et plus est apparente la dispersion entre les classes. Le choix du niveau de la coupure doit être fait après les agrégations correspondant à des valeurs peu élevées de l'indice regroupant les classes les plus proches les unes des autres. Autrement dit, le niveau de coupure se fait avant les agrégations correspondant à des valeurs élevées de cet indice, relatives à une dissociation des classes bien distinctes.

On remarque dans la figure 12 que plus on se rapproche du sommet de l'arbre et plus l'écart du niveau d'inertie intra-classe entre deux nœuds successifs (indice de Ward) devient grand. Il prend des sauts importants à partir de la 339^{ème} partition et atteint son maximum à la fin, entre la 346^{ème} et la 347^{ème} partition. La coupure est réalisée au niveau de la 343^{ème} et la 344^{ème} partition. Ce choix se justifie par le saut relativement important du niveau d'inertie intra-classe par rapport aux sauts précédents. Ce niveau de coupure génère sur le dendrogramme cinq classes de sondage que l'on identifie par les lettres alphabétiques E, J, D, G, et F.

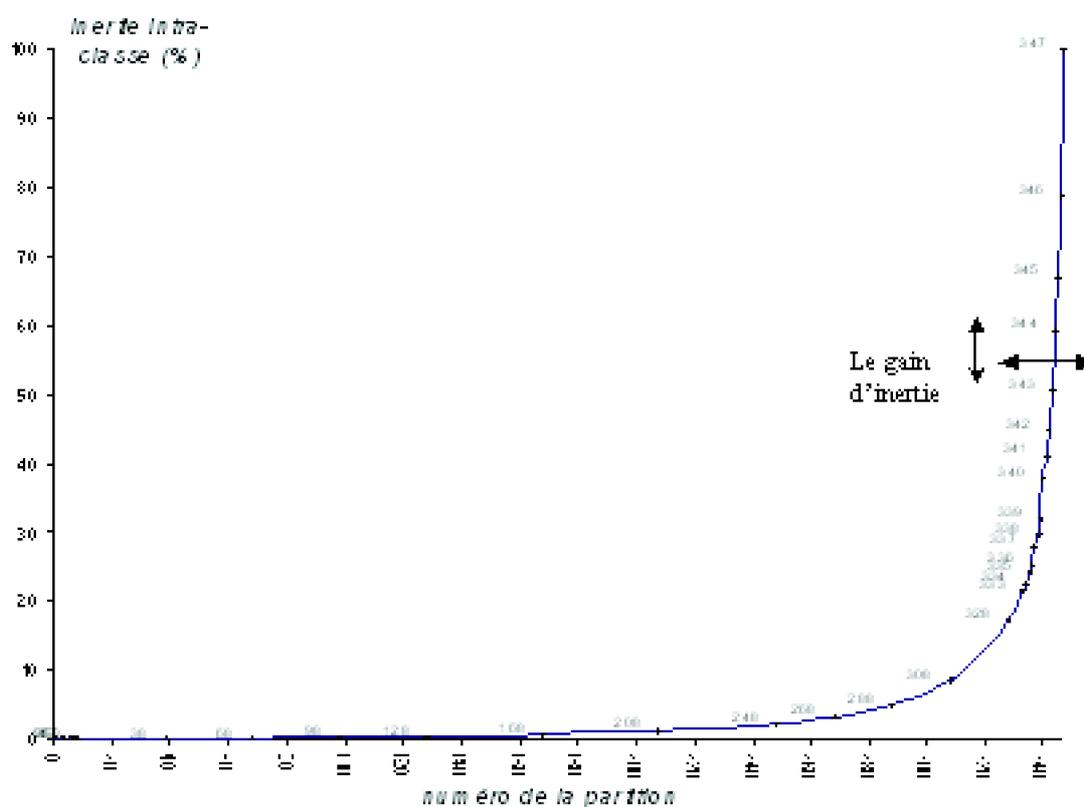


Figure n°12 : Courbe des indices de niveau in tra-classe des étapes des partition

6.2.4 . Caractéristiques des classes de sondage du sol issues de la partition

La comparaison des cinq classes de sondage pour chacune des variables du sol (test de Kruskal et Wallis) confirme l'existence de différences significatives entre elles au seuil de probabilité de 5 %. Le χ^2 observé supérieur au χ^2 théorique à un degré de liberté égal à 4, en est le témoin (Tab. XVII).

Tableau n° XVII . Résultats du test de comparaison entre les cinq classes de sondage du sol (test de Kruskal et Wallis).

Contribution à la détermination des unités homogènes du sol par différentes approches de cartographie numérique

Variables du sol	χ^2 2 obs	χ^2 2 thé	ddl	Probabilité
pH _{eau}	110.82	9.49	4	0.000
CE	102.42	9.49	4	0.000
CaCO ₃	111.66	9.49	4	0.000
Mo	61.02	9.49	4	0.000
A	112.33	9.49	4	0.000
Lf	69.92	9.49	4	0.000
Lg	71.00	9.49	4	0.000
Sf	93.22	9.49	4	0.000
Sg	22.07	9.49	4	0.000
lo	33.39	9.49	4	0.000
K	93.92	9.49	4	0.000

Une représentation graphique (*fig.13*) indique les valeurs extrêmes, les quartiles supérieurs et inférieurs et la médiane de chacune des variables du sol. De cette illustration graphique on ressort également les variables impliquées dans la différenciation et l'indifférenciation entre les classes de sondage.

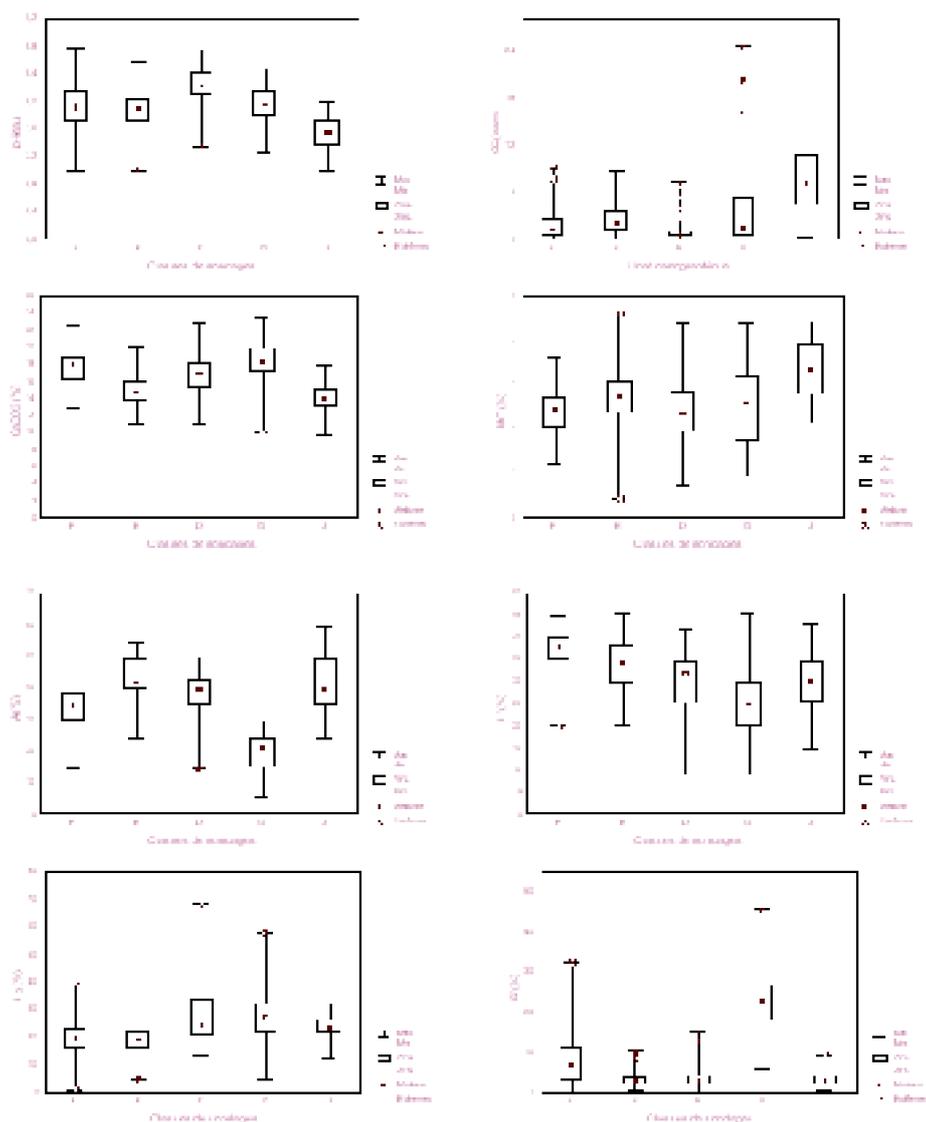


Figure n° 13. Comparaison graphique de la variabilité des cinq classes de sondage du sol issues de la CAH.

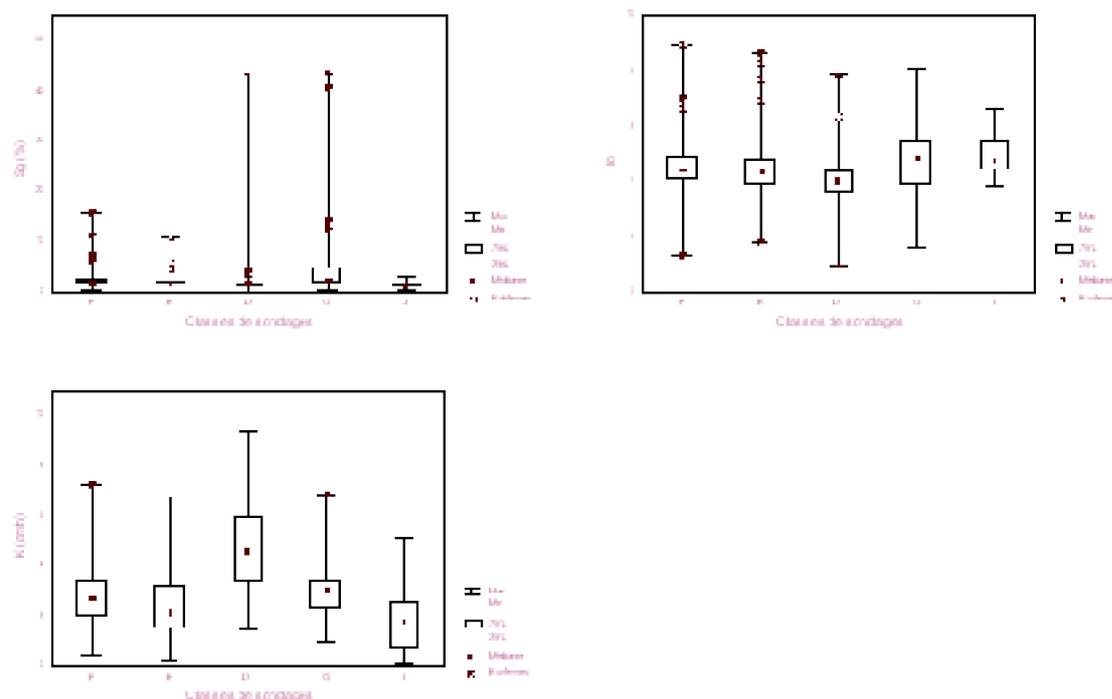


Figure n° 13 (suite). Comparaison graphique de la variabilité des cinq classes de sondage du sol issues de la CAH.

Les différents paramètres statistiques correspondant à la médiane, la moyenne, l'écart-type, le minimum et le maximum permettent d'évaluer les niveaux de variabilité dans chacune des classes de sondage (Tab. XVIII).

Pour toutes les classes de sondage du sol, les plus faibles variabilités intra-classes concernent le pH_{eau} et la MO d'après les faibles valeurs de leurs écart-types. Il en est de même, mais d'une façon moindre, pour le CaCO₃ total et les variables physiques liées à la stabilité structurale représentées par la variable I₀ et K. Les variabilités les plus élevées correspondent aux variables granulométriques A, Lf et Lg, d'après les valeurs élevées de leurs écart-types (Tab. XVIII).

La faible variabilité de la CE est particulièrement enregistrée dans les classes de sondage F, E et D respectivement d'après les valeurs de leurs écart-types égales à 1.98, 1.87 et 1.42. Les variabilités les plus élevées de la CE sont enregistrées dans les classes de sondage G et J respectivement d'après leurs écart-types égaux à 7.00 et 4,80.

Le Sf, est de forte variabilité dans les classes de sondage F et G respectivement d'après les valeurs de leurs écart-types égales à 6.10 et 10.18. Par contre, les faibles variabilités du Sf touchent les classes de sondage E, D et J relativement aux faibles valeurs de leurs écart-types de 1.76, 2.76 et 1.88. Pour le Sg, les variabilités les plus élevées sont notées dans les classes de sondage D et G d'après les valeurs de leurs écart-types respectivement égales à 5.09 et 12.13. Par contre, les faibles variabilités de ce paramètre se retrouvent dans les classes de sondage F, E et J. Leurs écart-types, des plus faibles dans l'ensemble des classes, sont respectivement égaux à 2.37, 1.34 et 0.60.

Variable	pH _{eau}	CE	CaCO ₃	MO	A	Lf	Lg	Sf	Sg	lo	K
Ecart-type											
F	0.39	1.98	2.15	0.57	6.52	4.74	6.59	6.10	2.37	1.04	1.28
E	0.28	1.87	1.80	0.66	6.67	5.72	5.07	1.76	1.34	0.97	1.32
D	0.29	1.42	2.57	0.71	6.63	8.42	10.18	2.76	5.09	0.94	2.01
G	0.32	7.00	3.08	0.92	6.82	9.14	12.02	9.05	12.13	1.55	1.47
J	0.24	4.80	1.92	0.61	9.42	6.83	4.43	1.88	0.63	0.70	1.29
Médiane											
F	7.30	1.17	17.89	2.46	34.45	38.0	19.65	6.85	1.53	4.36	2.7
E	7.89	2.11	14.71	2.75	41.53	34.35	19.23	2.70	1.43	4.31	2.13
D	8.23	0.60	16.85	2.38	39.45	31.75	24.05	3.75	1.29	3.95	4.58
G	7.95	1.45	10.26	2.50	20.02	25.30	27.00	22.78	1.95	4.82	3.03
J	7.54	7.25	13.90	3.33	39.45	29.76	23.38	2.85	1.17	4.72	1.76
Moyenne											
F	7.91	1.89	17.68	2.37	32.95	37.45	19.51	7.97	2.19	4.46	2.75
E	7.36	2.69	14.86	2.76	42.32	34.23	18.72	3.16	1.78	4.38	2.32
D	8.21	1.14	16.89	2.45	37.64	29.27	27.87	3.56	1.85	4.07	4.80
G	7.91	4.59	18.02	2.58	20.10	25.26	25.97	21.60	7.08	4.86	3.17
J	7.54	7.95	14.09	3.35	41.38	30.52	23.57	3.38	1.30	4.85	1.79
Minimum											
F	7.31	0.11	13.00	1.25	14.45	20.30	1.26	0.40	0.30	1.30	3.45
E	7.31	0.12	11.02	0.43	24.45	20.30	4.55	0.70	0.05	1.74	3.15
D	7.34	0.10	11.02	0.70	14.45	9.15	13.71	0.20	0.13	0.92	1.49
G	7.25	0.24	10.14	0.95	5.45	9.13	4.51	1.60	0.53	1.56	3.93
J	7.31	0.41	9.75	2.15	24.45	14.44	12.48	0.81	0.45	3.80	3.12
Maximum											
F	8.79	9.27	22.58	3.65	49.45	45.30	29.25	22.45	15.50	8.96	7.25
E	8.58	8.66	20.2	4.62	54.45	45.50	22.25	10.40	10.90	8.62	6.73
D	8.76	7.32	22.74	4.4	49.45	42.30	68.35	15.00	43.53	7.84	7.31
G	8.48	24.50	23.62	4.39	29.3	45.20	37.95	45.33	43.85	8.13	6.83
J	8.58	9.79	22.74	4.40	49.45	40.30	68.35	29.80	43.53	8.13	9.31

Tableau n°XVIII. Paramètres statistiques des 11 variables du sol des cinq classes de sondage

Une différenciation notable marque certaines classes de sondage relativement à certaines variables du sol (fig.13). Elle concerne la classe de sondage J représentée par des valeurs faibles en pH_{eau}, CaCO₃ total et K et élevées en CE et MO. De même pour la classe D qui se distingue par ses valeurs élevées en pH_{eau} et K et faibles en lo. On distingue également la classe G beaucoup plus par ses faibles teneurs en A et Lf et élevées en Sf et Sg .

Une situation d'indifférenciation, partielle et totale, apparaît entre certaines classes de sondage en relation avec les variables suivantes :

- le pH_{eau} et le K par rapport aux classes de sondage F, E et G ;
- la CE et la MO entre les classes de sondage F, E, D et G ;
- le CaCO₃ total pour les classes de sondage F, D et G ;

- l'A relativement entre les classes de sondage F, E, D et J ;
- le Lf d'une part entre le couple de classes de sondage F et E et d'autre part avec les classes de sondage D et J ;
- le Lg par rapport aux classes de sondage F et E de même pour l'ensemble des classes de sondage D, G et J ;
- le Sf pour l'ensemble des classes E, D et J ;
- le Sg entre les classes de sondage F, E, D et J ;
- le I₀ communément entre les classes de sondage F, E, G et J. De même pour les classes de sondage F, E, D et G.

6.3 . Représentation cartographique des classes de sondage

La représentation cartographique des cinq classes de sondage du sol dans la zone d'étude est abordée par la constitution d'un ensemble de surfaces élémentaires de forme géométrique carrée correspondant chacune d'elle à un point de sondage repéré par ses coordonnées Lambert. Par la suite, on reporte pour chacune de ces surfaces son identifiant dégagé par le processus de la CAH en le représentant par son identifiant (couleur) marquant son appartenance à l'une des cinq classes issue de la partition. Les points voisins appartenant à deux classes sont séparés par une limite qui passe à mi-distance entre eux. Ceux, de la même classe sont englobés dans la même plage.

Il s'ensuit, d'après la représentation de ces classes de sondage du sol sur la carte, une dispersion de celles-ci en de petites et grandes plages, contraignant ainsi à une interprétation géographique d'une manière disparate de chacune d'entre elles dans la zone d'étude (*fig. 14*).

La classe E est la plus dominante dans la zone d'étude puisqu'elle regroupe le plus grand taux d'effectif, soit 140 sondages sur les 348, correspondant à 40.22 % de l'effectif total. Elle occupe la majeure partie du sud-est, du centre et du nord-est de la zone d'étude.

La classe D est composée de 70 sondages, soit 20.11% de l'effectif total. Elle est présente dans sa majorité dans la partie nord-ouest et le reste se disperse en de petites plages dans la partie nord-est et sud-est ainsi qu'à la périphérie du centre.

La classe F occupe dans sa totalité la partie sud-est et nord-ouest. Elle reste absente dans les deux extrémités nord-ouest et sud-est ainsi que près du centre. Le total des sondages du sol composant cette unité est de 68, soit 19.54 % de l'effectif total.

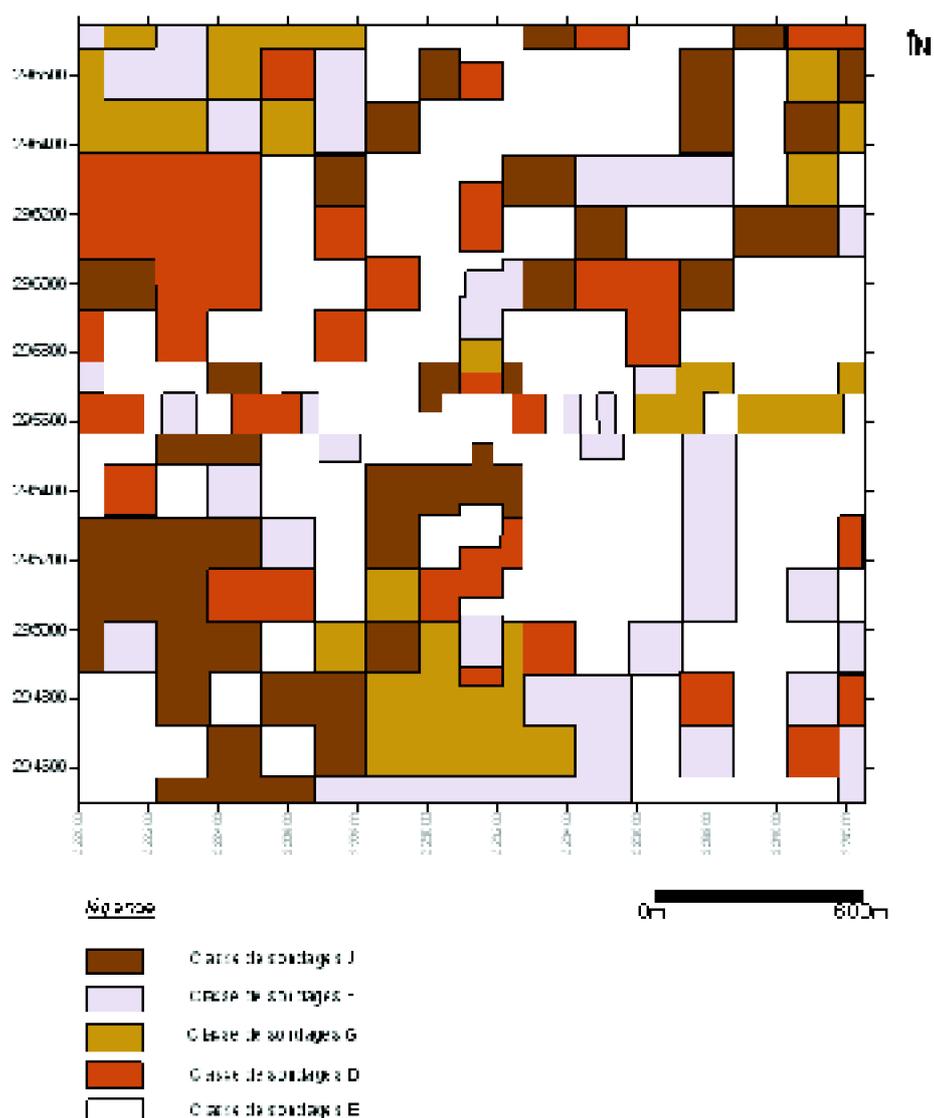


figure 14 . représentation cartographique des cinq classes de sondages du sol constituées par la CAH.

La classe J est formée de 48 sondages sur les 348, soit 13.79 % de l'effectif total. Elle est beaucoup plus représentée dans sa majorité dans la partie sud-ouest ; le reste a tendance à se répartir en de petites plages au nord-est et près du centre. Elle est absente dans presque la totalité de la partie sud-est et l'extrémité nord-ouest.

La classe G est la moins représentée dans la zone vu le faible taux de son effectif égal à 22 sondages, soit 6.32 % de l'effectif total. Elle se concentre dans sa majorité dans l'extrémité nord-ouest et nord-est. Le peu restant se retrouve dispersé dans la partie centrale et au sud de la parcelle.

Conclusion

Globalement la faible décision dans le choix du nombre de classes de sondage est justifiée par l'absence d'un écart important entre les indices de niveaux comme le prévoit une bonne partition. Tout de même, la décision de retenir cinq classes de sondage du sol est admise par le saut relativement élevé de l'inertie intra-classe, entre la 343^{ème} et la 344^{ème} partition. Cette information montre, dans ce cas, combien l'individualisation des cinq classes de sondage du sol est faible malgré la possibilité pour la CAH de travailler dans l'hyperplan défini par les trois premières composantes principales.

Le résultat de la représentation spatiale des cinq classes de sondage du sol issues de la CAH dévoile une répartition fragmentaire de celles-ci en de grandes et petites plages, allant jusqu'aux surfaces élémentaires représentées par un seul sondage du sol.

Il y a une nette domination en superficie de la classe E. Cependant, cette classe de sondage ne présente pas une différenciation notable en terme de variables dans l'ensemble. Elle est suivie en superficie par la classe D caractérisée par des pH_{eau} et K les plus élevés et des I_0 faibles. La classe de sondage F, comme la classe E, est indifférenciable dans l'ensemble des classes. Elle s'avère néanmoins, moins argileuse et plus élevée en CaCO_3 total, par rapport à la classe E. Elle occupe la troisième place en superficie. La quatrième place revient à la classe J, marquée dans l'ensemble par ses faibles valeurs en pH_{eau} , CaCO_3 total et K et élevées en CE et MO. La classe de sondage G qui se distingue compte à elle par ses teneurs faibles en A et Lf et élevées en Sf et Sg est la plus petite en superficie relativement à son plus faible effectif.

Chapitre VII . Contribution de l'analyse factorielle discriminante à la détermination des unités homogènes du sol

L'objet de ce chapitre consiste d'abord à déterminer l'ordre discriminant des 11 variables du sol entre les cinq groupes de sondage constitués préalablement par l'emploi de la classification ascendante hiérarchique (CAH). On cherche par la suite, l'axe factoriel ou la formule linéaire qui discrimine au mieux, non pas l'ensemble des 348 sondages du sol, mais l'ensemble des cinq classes de sondages du sol : il est clair que la discrimination est d'autant plus aisée que les classes de sondage sont plus distinctes les unes des autres et que les sondages d'une même classe sont proches les uns des autres dans l'espace factoriel.

7.1 . Présentation des résultats préliminaires

Cette partie du chapitre est un élément clef du programme, puisqu'elle permet de dire si réellement la variable Sf suffit à elle seule à discriminer l'ensemble des cinq classes de sondage ou bien c'est l'ensemble des variables du sol qui contribue le mieux à cette discrimination. Le programme consiste en premier lieu à réaliser l'opération de diagonalisation de la matrice carrée issue du rapport de la matrice de variance inter-classe sur la matrice totale. Les résultats de cette diagonalisation présentés ci-dessous, sont des éléments d'informations sur les axes discriminants des cinq points constituant les centres de gravité des cinq classes de sondage.

7.2.1 . Résultats de l'AFD

- La plus grande valeur des pseudo F correspondant au premier axe discriminant est égale à 237.54 (Tab. XXI). Cette valeur est très largement supérieure à la plus grande valeur de F de Fischer Snedecor du Sf égale à 117.94. L'écart important entre ces deux valeurs montre tout l'intérêt qu'il y a pour faire une analyse factorielle discriminante pour la simple raison qu'aucune des variables du sol ne peut expliquer à elle seule la différence entre les cinq classes de sondage. A titre d'exemple, la valeur du pseudo F relative au premier axe discriminant se calcule comme suit :

$$F_{\text{pseudo } F_1} = \{(x - g) / (g - 1)\} V_1 = \{(348 - 5) / (5 - 1)\} \cdot 2.77 = 237.54$$

- n : nombre de sondages du sol,
- g : nombre de groupes de sondage du sol,
- v_1 : valeur propre de l'axe discriminant 1.

On applique la même formule pour calculer le Pseudo F du deuxième axe factoriel discriminant dont la valeur propre (V_2) est égale à 1.03. Il en est de même pour l'axe 3 et 4.

- Les valeurs propres sont des indices du pouvoir discriminant de chaque axe factoriel par rapport à l'ensemble des cinq classes de sondage. Ainsi, la plus grande discrimination relative au premier axe discriminant est égale à 2.77 et représente une part de discrimination égale à 61.7 % (inertie). Le second axe discriminant contribue par une valeur égale à 1.03 correspondant à une part de discrimination de 23 %, suivi de loin par le troisième axe factoriel avec une part discriminante égale à 13 %. Les trois premiers axes factoriels perpendiculaires l'un par rapport à l'autre, cumulent un pouvoir discriminant égal à 97.7 %. Le reste de la discrimination concernant le quatrième axe factoriel est égal à 2.30 %.
- On présente les cinq classes de sondage dans l'hyperplan défini par les trois premiers axes factoriels discriminants. C'est la meilleure représentation possible de ces classes pour que celles-ci soient le plus possible séparées. Le choix de cette dimension est confirmé par les résultats de la statistique de Wilks. En effet, le test

Contribution à la détermination des unités homogènes du sol par différentes approches de cartographie numérique

d'hypothèse d'absence de différence entre les classes a été rejeté pour les quatre axes discriminants à un seuil de signification de 5 %. La statistique de Wilks est égale à 879.76 pour l'axe 1. Elle est suivie de loin par celle de l'axe 2 avec une valeur de 429.87 puis de celle de l'axe trois avec une valeur de 189.75 et en dernier lieu par celle de l'axe quatre avec une valeur de 33.99. Ces valeurs correspondent toutes à des probabilités inférieures ou égales à 0.01 % pour les quatre axes.

Tableau n° XXI . Résultats de la diagonalisation.

n° de l'axe (i)	Valeur Propre (V)				Inertie (I) (%)	Pseudo f	Wilks	ddl	Proba (%)	Corrélation (r ²)			
1 2 3 4	2.77	1.03	0.58	0.17	23.00	123.05	2.88	37.59	0.75	0.73	0.50	0.36	0.09

- Les coefficients de corrélation canonique (r^2) expriment la relation existante entre les cinq classes de sondage et les axes discriminants. C'est sur le 1^{er} axe discriminant que l'individualisation des cinq classes de sondage est la meilleure d'après la valeur élevée de r^2 égale à 0.73. Cette individualisation diminue avec l'axe discriminant 2 puis de l'axe 3, d'après leurs valeurs respectives de r^2 égales à 0.50 et 0.36. Elle s'atténue par la suite avec le 4^{eme} axe discriminant pour atteindre une valeur de r^2 égale à 0.09.

7.2.2. Corrélations inter-classes entre les variables du sol et les axes discriminants

Sur le 1^{er} axe discriminant, les cinq classes de sondage du sol se distinguent entre elles par les variables A, CaCO₃ total, Sf, Sg, MO et K. Ces variables présentent des corrélations élevées dans l'ensemble. Elles sont respectivement égales en valeur absolue à 0.96, 0.88, 0.87, 0.85, 0.62 et 0.52 (*tab. XXII*). Il est à rappeler d'abord que c'est sur cet axe là que l'individualisation des cinq classes de sondage du sol est la meilleure. Il en est de même lorsqu'il s'agit de leur discrimination à partir des variables citées ci dessus.

Sur le deuxième axe discriminant, ce sont les lo, pH_{eau}, K et CE qui discriminent le mieux entre les cinq classes de sondage d'après leurs corrélations élevées exprimées en valeurs absolues et égales respectivement à 0.95, 0.85, 0.82 et 0.72 .

Le troisième axe discriminant permet de différencier les cinq classes de sondage du sol par les variables Lf, Lg, CE et MO. Ces variables du sol présentent avec cet axe des corrélations élevées dans l'ensemble. Elles sont respectivement égales en valeurs absolues à 0.88, 0.75, 0.58 et 0.56.

Tableau n° XXII . Corrélations inter-classes entre les variables du sol et es axes discriminants.

n° de l'axe	pH _{eau}	CE	CaCO ₃ total	MO	A	Lf	Lg	Sf	Sg	lo	K
1 2 3 4	-0.47	0.85	0.72	0.88	0.52	0.62	0.29	0.29	0.82	0.85	0.95
	-0.17	-0.12	-0.17								

7.2.3 . Détermination du degré de dissimilation entre les classes de sondage

La dissimilation entre les cinq classes de sondage prises deux à deux est exprimée par la distance de Mahalanobis. Elle représente l'éloignement entre les centres de gravités des cinq classes dans l'hyperplan formé par les trois premiers axes discriminants. Ainsi, dans l'ensemble, la dissimilation est plus importante entre les couples de classes de sondage E-G, G-J, D-G d'après leurs distances de Mahalanobis élevées, égales respectivement à 3.44, 3.41 et 3.25. Par contre, les plus faibles distances égales à 1.46 et 1.27 sont enregistrées respectivement dans les couples de classes E-J et F-E. Les faibles distances de Mahalanobis traduisent une faible différenciation entre ces classes dans chaque couple de sondage (*Tab. XXIII*).

Tableau n° XXIII . Distances de Mahalanobis entre les cinq classes de sondage du sol

Classe de sondage	F	E	D	G	J
F	0	-	-	-	-
E	1.27	0	-	-	-
D	1.94	1.81	0	-	-
G	2.69	3.44	3.25	0	-
J	2.36	1.46	2.46	3.41	0

7.2.4 . Affectation des sondages du sol dans leurs classes représentatives

A présent, on aborde le reclassement de tous les sondages du sol dans les cinq classes en affectant chacun d'eux à l'une des classes la plus proche. C'est une étape décisionnelle. Le critère de choix d'affectation pour un point de sondage du sol à l'une des classes représente la distance la plus courte qui le sépare du barycentre d'une des cinq classes de sondage. Le programme donne un pourcentage des sondages bien classés de 83.04. Ceci signifie que sur 348 sondages de sol constituant les cinq classes issues de la CAH, 59 d'entre eux sont concernés par une affectation dans l'AFD. Ainsi, le résultat de ce classement donne un taux d'effectif par classe de sondages comme suit :

- pour la classe E, le nombre de sondage est de 135,
- pour la classe D, le nombre de sondages est de 76,
- pour la classe G, le nombre de sondages est de 23,
- pour la classe F, il est égal à 66,
- et enfin, pour la classe J, il est égal à 48 sondages.

7.3 . Caractéristiques des classes de sondage du sol issues de l'AFD

Le test de comparaison des cinq classes de sondage, pour chacune des variables du sol (test de Kruskal et Wallis), confirme l'existence de différences significatives entre elles au seuil de probabilité de 5 %. En effet le χ^2 observé ressort largement supérieur au χ^2 théorique pour toutes les variables du sol à un degré de liberté égal à 4 (Tab. XXIV).

Pour toutes les classes de sondage, les faibles variabilités sont relatives aux variables du sol pH_{eau}, MO, lo, K et CaCO₃ total, d'après les faibles valeurs de leurs écart-types. Elles traduisent des valeurs très groupées autour des moyennes. A l'inverse, les fortes variabilités concernent les variables granulométriques A, Lf et Lg d'après les valeurs élevées de leurs écart-types (Tab.XXV).

Tableau n° XXIV . Comparaison multiple des valeurs moyennes des variables du sol, des cinq classes de sondage (test de Kruskal et Wallis).

Variables du sol	χ^2 obs	χ^2 thé	ddl	Probabilité
pH _{eau}	100.91	9.49	4	0.000
CE	142.98	9.49	4	0.000
CaCO ₃ total	114.17	9.49	4	0.000
MO	76.73	9.49	4	0.000
A	110.34	9.49	4	0.000
Lf	74.25	9.49	4	0.000
Lg	52.71	9.49	4	0.000
Sf	118.57	9.49	4	0.000
Sg	31.33	9.49	4	0.000
lo	28.95	9.49	4	0.000
K	106.91	9.49	4	0.000

La faible variabilité de la CE est particulièrement enregistrée dans les classes de sondage F, E et D respectivement d'après les valeurs de leurs écart-types égales à 1.52, 1.77, 0.98. Les variabilités les plus élevées de la CE concernent les classes de sondage G et J d'après les valeurs de leurs écart-types respectivement égales à 6.01 et 4.51.

Le Sf, est de forte variabilité dans les classes de sondage F et G respectivement d'après les valeurs de leurs écart-types égales à 4.55 et 7.59. Par contre, les faibles variabilités du Sf touchent les classes de sondage E, D et J relativement aux faibles valeurs de leurs écart-types égales respectivement à 1.56, 2.74, 2.11.

Pour le Sg, les variabilités les plus élevées sont notées dans les classes de sondage D et G d'après les valeurs de leurs écart-types respectivement égales à 4.90 et 11.96. Par contre, les faibles variabilités de ce paramètre se retrouvent dans les classes de sondage

F, E et J. Leurs écart-types sont respectivement égaux à 1.70, 1.08 et 1.55.

Une différenciation notable marque les classes de sondage relativement à certaines variables du sol (*fig.15*). Elle est énumérée par couple de classes en ordre décroissant de dissimilation, établit par l'AFD.

La classe de sondage E se diffère de la classe G par ses teneurs faibles en CaCO₃ total, Sf et Sg et élevées en A. Ces deux classes demeurent par contre indifférenciables par leurs pH_{eau}, CE, MO, Lf, Lg, lo et K.

La classe de sondage J se diffère de la classe G par ses faibles valeurs en pH_{eau}, CaCO₃ total, Sf et Sg. Ces deux classes de sondage sont indifférenciables par le Lf, Lg, lo et K.

variable	pH _{eau}	CE	CaCO ₃	MO	A	Lf	Lg	Sf	Sg	lo	K
Ecart-type											
F	0.35	1.52	2.08	0.55	6.33	4.12	5.74	4.55	1.73	1.00	1.14
E	0.30	1.77	1.74	0.63	6.35	5.66	5.11	1.56	1.08	1.02	1.20
D	0.30	0.98	2.61	0.66	9.46	8.24	9.83	2.74	4.97	0.90	1.93
G	0.30	6.01	2.98	0.83	7.62	7.81	11.03	7.59	11.96	1.57	1.45
J	0.26	4.51	1.95	0.64	9.50	7.74	8.31	2.11	1.55	0.72	1.33
Médiane											
F	7.89	1.03	18.01	2.38	34.45	38	20.1	6.75	1.63	4.41	2.30
E	7.89	2.07	14.54	2.75	41.35	34.35	19.05	2.60	1.41	4.28	2.15
D	8.22	0.63	16.64	2.47	39.45	31.75	23.94	2.42	1.17	4.10	4.69
G	7.87	1.81	17.53	2.76	34.15	25.00	25.45	23.80	1.97	4.86	2.92
J	7.64	7.63	13.99	2.53	43.27	29.46	22.65	3.1	1.30	4.73	2.08
Moyenne											
F	7.89	1.66	17.89	2.30	32.63	37.55	20.16	7.56	2.07	4.48	2.42
E	7.85	2.58	14.73	2.75	41.96	34.57	18.98	3.03	1.68	4.35	2.19
D	8.21	1.00	16.84	2.46	37.92	29.75	27.46	3.34	1.75	4.09	4.94
G	7.87	4.88	17.54	2.03	21.03	25.86	21.82	23.38	7.47	4.82	3.11
J	7.58	8.88	14.25	2.48	42.28	29.34	23.32	3.57	1.61	4.90	2.11
Minimum											
F	7.01	0.11	13.49	0.43	14.45	24.35	5.57	0.81	0.05	2.41	0.34
E	7.01	0.22	10.12	1.16	24.45	20	4.55	0.8	0.05	1.30	0.15
D	7.01	1.1	11.02	0.70	14.45	9.15	13.71	0.20	0.11	0.93	1.47
G	7.29	0.24	10.14	0.95	5.45	9.13	1.26	9.00	0.57	1.56	0.93
J	7.01	1.05	0.75	2.15	24.15	10.00	8.46	0.81	0.18	3.80	0.12
Maximum											
F	8.65	5.91	23.00	3.57	47.45	45.20	34.15	22.30	11.19	8.96	4.78
E	8.79	8.65	20.20	4.28	54.56	45.50	32.25	9.30	6.21	8.63	5.11
D	8.76	5.57	22.74	3.98	54.45	45.00	58.35	15.00	43.53	7.84	9.31
G	8.48	24.50	23.62	4.30	33.90	40.00	40.15	45.53	43.85	8.15	6.83
J	8.68	9.79	22.74	4.39	34.45	45.00	58.30	32.45	43.33	8.15	9.31

Tableau n°XXV. Paramètres statistiques des 11 variables du sol des cinq classes de sondage.

La classe de sondage G se diffère de la classe D par ses valeurs faibles en pH_{eau},

A, et K et élevées en CE, Sf et Sg. Elles sont indifférenciables par le CaCO_3 total, MO, Lf, Lg et I_0 .

La classe G se diffère de la classe F par ses faibles teneurs en A et Lf et élevées en Sf. Ces deux classes sont indifférenciables par le pH_{eau} , CE, CaCO_3 total, MO, Lg, Sg, lo et K.

La classe de sondage J marque sa différenciation avec la classe D par ses faibles valeurs en pH_{eau} , CaCO_3 total et K et élevées en CE et MO. Elles sont, par contre, indifférenciables par leurs constitutions granulométriques.

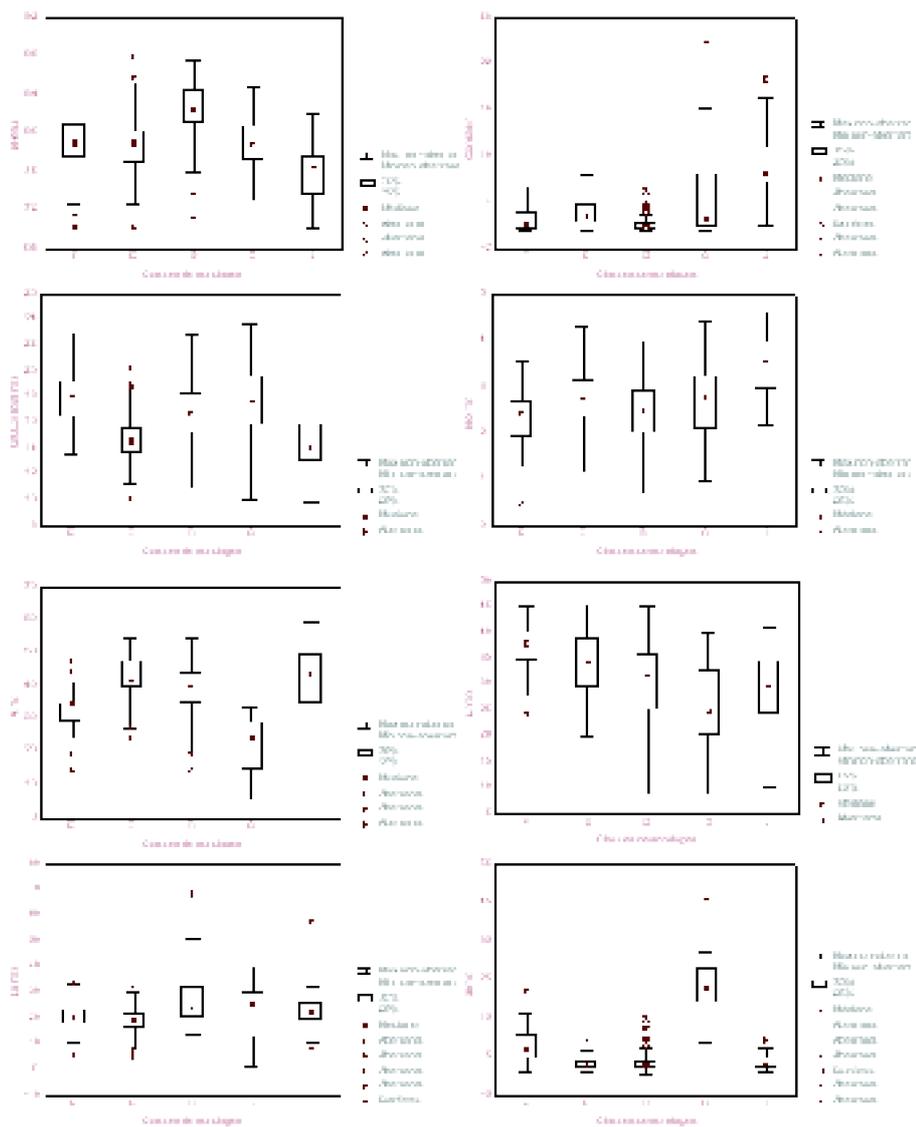


Figure n° 15. Représentation graphique de la distribution des cinq classes de sondage du sol issues de AFD.

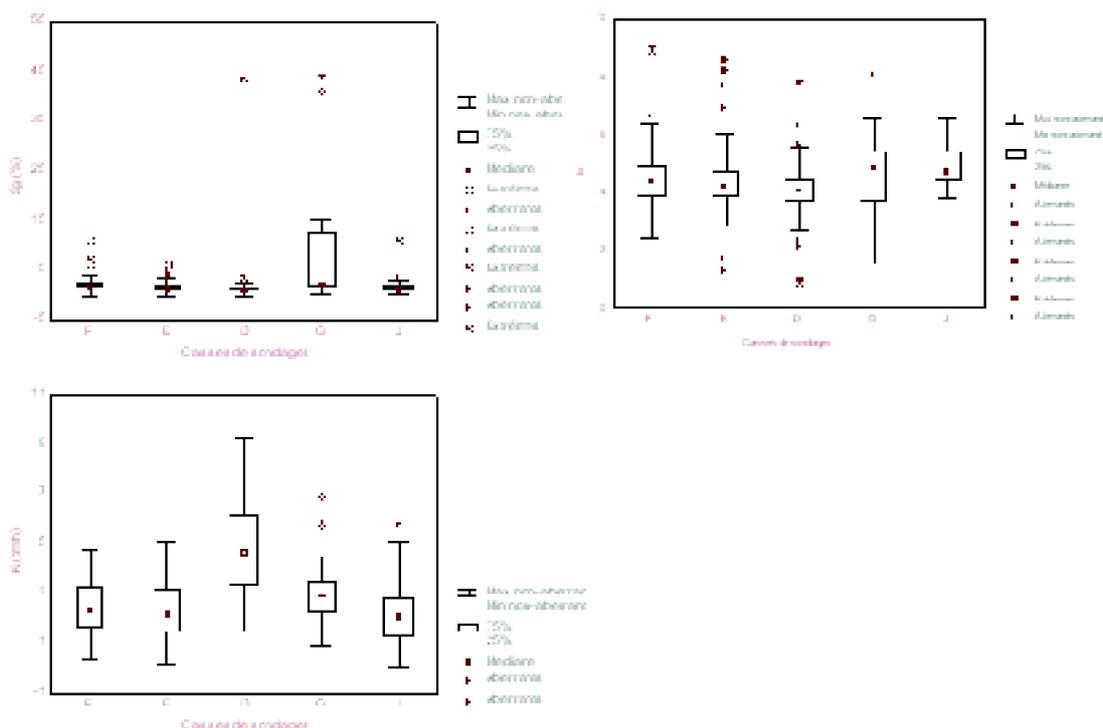


Figure n° 15 (suite). Représentation graphique de la distribution des cinq classes de sondage du sol issues de AFD.

La classe de sondage J se diffère de la classe F par ses faibles valeurs en pH_{eau} , CaCO_3 total et Sf et élevées en CE, MO, A et Lf. Elles restent, par contre, unies par leurs teneurs en Lg, et Sg ainsi que par les paramètres de stabilité structurale (I_0 et K).

La classe de sondage F diffère de la classe D par ses faibles valeurs en pH_{eau} , A, Lg et K et élevées en Lf et Sf. Elles restent, par contre, indifférenciables en CE, CaCO_3 total, MO, Sg et I_0 .

La classe de sondage E se diffère de la classe D par ses faibles valeurs en pH_{eau} , CaCO_3 total, Lg et K et élevées en CE. Elles sont indifférenciables en MO, A, Lf, Sg et I_0 .

La classe E se diffère de la classe J par ses faibles valeurs en CE et MO et élevées en pH_{eau} . Elles sont indifférenciables par leurs compositions granulométriques ainsi que par le CaCO_3 total et K.

On reconnaît la classe F de la classe E par ses faibles teneurs en MO et A et élevées en CaCO_3 total et Sf. Elles sont, par contre, indifférenciables en pH_{eau} , CE, Lf, Lg, Sg, I_0 et K.

7.4 . Répartition géographique des classes de sondage du sol

Le reclassement partiel des sondages du sol par l'AFD entraîne quelques modifications

dans la configuration géographique de l'ensemble des plages en gardant l'aspect fragmentaire de celles-ci (fig. 16).

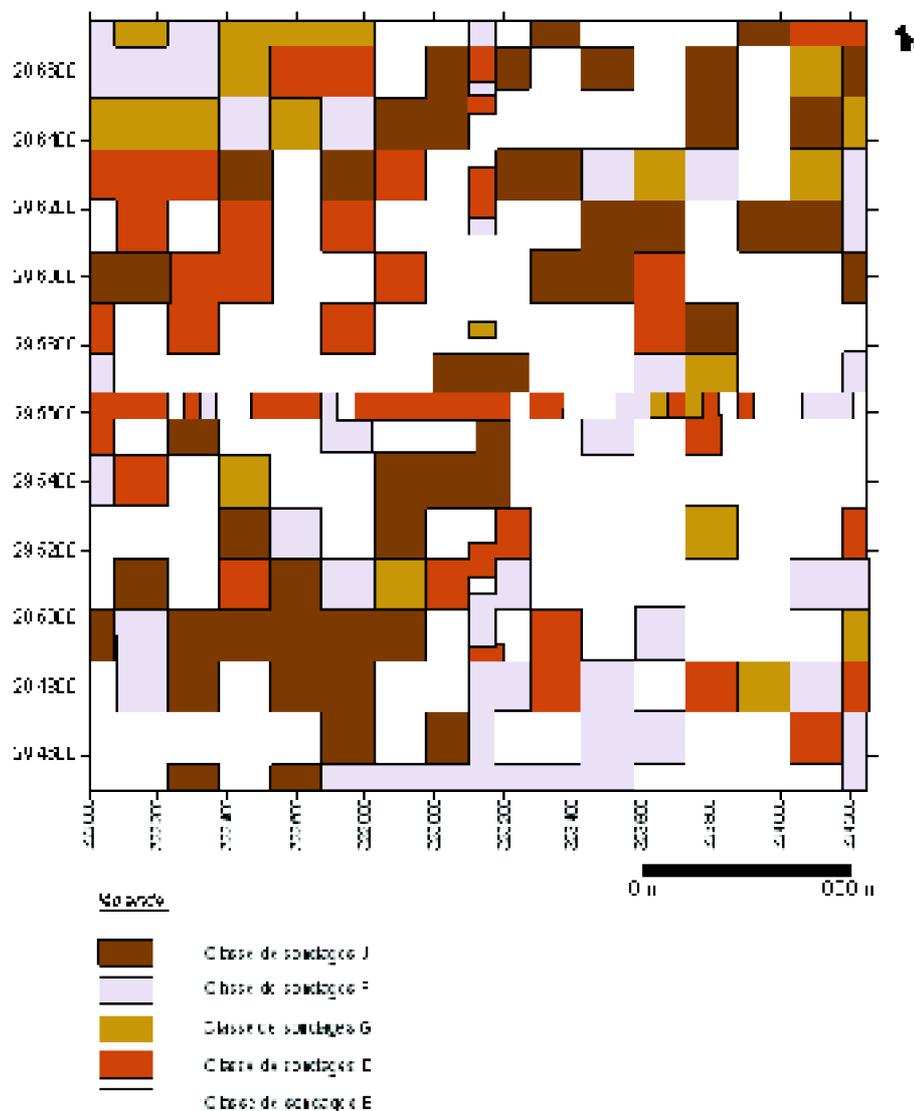


Figure 16. Représentation cartographique des cinq classes de sondage du sol constituées par l'AFD.

Ainsi, la classe E reste la plus dominante dans l'ensemble. Elle prend par endroit, de l'espace à la classe F et J dans la partie sud-est et sud-ouest. La classe G subit quelques légères modifications par ses apparitions soudaines dans l'ensemble de la zone à l'exception de la partie nord-ouest où elle garde ses mêmes extensions comme dans la CAH. La classe D gagne en surface dans la plupart des plages y afférentes en conquérant des surfaces élémentaires limitrophes, appartenant déjà à la classe F ou J. La classe J reste absente dans toute la partie sud-est. Elle se concentre dans sa majorité dans la partie sud-ouest.

Conclusion

L'AFD aboutit à une meilleure différenciation entre les cinq classes de sondage et cela par son pouvoir d'allier l'ensemble des variables du sol à cette tâche. Son meilleur espace de discrimination est attribué à son 1^{er} axe factoriel représentant 61.7 % de l'information sur le pouvoir discriminant. Il le doit, respectivement et par ordre décroissant, aux variables du sol A, CaCO₃ total, Sf, Sg, MO et K. Le 2^{ème} axe factoriel, avec 23 % sur la variance totale, est de faible pouvoir discriminant. Il doit cette part de discrimination respectivement et par ordre décroissant aux variables du sol I₀, pH_{eau}, K et CE. Le 3^{ème} axe factoriel détient 13 % sur l'information totale. La discrimination entre les cinq classes de sondage dans cet axe est attribuée aux variables du sol Lf, Lg, CE et MO.

Dans l'hyper-plan formé par ces trois axes discriminants, l'AFD ressort des niveaux de dissemblance entre les classes de sondage. Les couples de classes de sondage du sol les plus différenciables sont respectivement en ordre décroissant E-G, G-J, D-G, G-F. Par contre, le couple de classes de sondage E-J et E-F est le moins différenciable. L'AFD a effectué également un changement d'affectation de quelques 58 sondages du sol par rapport à leurs classes d'origine. Elle l'applique à tous les sondages par référence à leurs distances de rapprochement aux cinq points de sondage type représentant les barycentres des cinq classes de sondage.

Chapitre VIII .La géostatistique multiparamétrique

Ce chapitre permet de voir la manière et les proportions qu'implique la géostatistique multiparamétrique pour examiner la part de la corrélation de l'information factorielle contenue dans chaque axe discriminant. Plus précisément, on essaye de vérifier si les sondages du sol sont d'autant plus semblables qu'ils sont proches spatialement. Si cela est exact, il sera alors possible de prédire par interpolation la valeur d'une variable régionalisée en tout point dans la zone, en tenant compte des valeurs des sondages du sol attribuées par l'AFD dans le voisinage.

On aborde ici l'espace discriminant d'une façon partielle où chaque part de l'information reflète une vue de l'organisation des cinq unités homogènes du sol sur chacun des axes discriminants. L'expression d'unité homogène du sol est attribuée ici à chacune des classes de sondages marquant ainsi le passage de l'espace factoriel discriminant vers l'espace géographique.

Le traitement géostatistique est réalisé moyennant deux logiciels ; le Variowin (*Pannatier, 1996*) et le Surfer 8.0 (*Hohl, 1997*). Le Variowin offre certains programmes qui sont utilisés pour l'analyse spatiale des données ainsi que la création et la modélisation du variogramme relatif à chacun des espaces factoriels (axes discriminants). Le surfer 8.0 sert, quant à lui, à l'interpolation spatiale et à la réalisation des cartes des courbes d'isovaleurs.

8.1. Structure des données de base

La structure des données de base est formée d'un espace géographique et d'un espace factoriel, constituant ensemble les 348 variables régionalisées (Tab. XXVI).

Variable régionalisée (n° du sondage)	<i>Espace géographique</i>		<i>Espace factoriel</i>		
	Coordonnées Lambert (m)		Coordonnées des points de sondages sur les 3 premiers axes discriminants		
	Longitude (X)	Latitude (Y)	1 ^{er} axe	2 ^{ème} axe	3 ^{ème} axe
1	324250	294500	-0.366	0.811	-1.232
2	324100	294500	0.076	0.902	-3.448
3	323950	294500	0.439	0.232	-3.607
4	323800	294500	0.894	0.059	3.208
.
346	324150	295625	-0.255	-0.683	-3.783
347	324200	295625	-0.111	-0.719	-3.789
348	324250	295625	0.197	-0.058	-3.408

Tableau n° XXVI. Structure des données de base relative aux variables régionalisées

Ainsi, chacun des sondages du sol se retrouve localisé dans ces espaces comme suit :

- par ses coordonnées Lambert dans l'espace géographique. Ces coordonnées permettront par la suite de déterminer les distances (h) entre les points de sondage;
- par sa valeur relative à sa projection dans l'un des axes factoriels issu de l'AFD. C'est la variable qui caractérise chacun des sondages du sol.

Comme exemple d'une variable régionalisée, celle du premier point de sondage projeté sur le 1^{er} axe discriminant. Elle se calcule pour les variables réduites, de la façon suivante :

$$i = 11$$

$$X_1 = \sum_{i=1}^{i=11} ((m_i - x_i)/\sigma_i) (a_i) \quad \text{où,}$$

$$i = 1$$

- i : variable initiale,

- x_i : valeur de la variable initiale i ,
- σ_i : écart type de la variable initiale i ,
- m_i : moyenne correspondant à la variable initiale i ,
- a_i : vecteur propre de la variable initiale i .
- X_1 : valeur du premier sondage égale à sa valeur de projection sur le premier axe discriminant.

$$X_1 = (-0.249) (-1.12) + (-0.062) (-0.623) + \dots + (-0.153) (-0.367) = -0.366$$

Tableau n° XXVII. Exemple de calcul d'une coordonnée de projection du premier sondage du sol sur le premier axe discriminant.

VARIABLES INITIALES (i)	VARIABLES CENTRÉES ET RÉDUITES ($(m_i - x_i)/\sigma_i$)	VECTEURS PROPRES (a_i)
pH _{eau}	$(7.50 - 7.905) / 0.361 = -1.12$	-0.249
CE	$(0.81 - 3.071) / 3.62 = -0.623$	-0.062
CaCO ₃ total	$(17.25 - 15.911) / 2.56 = 0.522$	-0.274
MO	$(2.56 - 2.695) / 0.73 = -0.184$	0.155
A	$(33.35 - 38.015) / 9.48 = -0.491$	-0.296
Lf	$(36.00 - 32.783) / 7.39 = 0.434$	-0.359
Lg	$(17.65 - 21.802) / 8.01 = -0.517$	-0.566
Sf	$(11.30 - 5.378) / 6.04 = 0.980$	-0.917
Sg	$(1.70 - 2.149) / 4.19 = -0.107$	-0.331
lo	$(4.16 - 4.430) / 1.03 = -0.261$	-0.029
K	$(2.23 - 2.889) / 1.79 = -0.367$	-0.153

8.2. Analyse de la structure spatiale du premier axe discriminant

8.2.1. Calcul du variogramme expérimental

Le variogramme expérimental relatif à cette variable (1^{er} axe discriminant) est représenté par une courbe discontinue de points (fig.17) liant les valeurs de la dissemblance $\gamma(h)$ à leurs distances respectives (h). La dissemblance $\gamma(h)$ liée à la distance (h) est calculée dans toutes les directions d'où l'appellation du variogramme expérimental moyen de variogramme omnidirectionnel.

Le tableau XXVIII montre les résultats obtenus pour la réalisation de ce variogramme omnidirectionnel déterminé sur une distance de 1000.17m.

Tableau n° XXVIII. Résultats de calcul du variogramme expérimental robuste relatif au premier axe

discriminant.

n° de la classe de sondages (<i>i</i>)	Nombre de couple de sondages	Distance moyenne (m)	Dissemblance $\gamma(h)$
1	200	54.01	0.504
2	406	95.52	0.711
3	1384	149.38	0.713
4	1324	210.45	0.785
5	858	251.30	0.895
6	1446	301.54	0.844
7	2236	339.55	0.860
8	1918	407.15	0.866
9	3242	461.22	0.898
10	852	501.79	1.021
11	2842	543.90	0.824
12	3156	607.19	0.921
13	3156	655.16	0.881
14	1676	696.35	0.893
15	4456	753.34	0.919
16	2496	803.20	0.983
17	3592	855.95	0.992
18	3280	903.48	1.074
19	3580	952.77	0.930
20	3200	1000.17	0.949

Cressie & Hawkins (1980) et Cressie (1984) sont parmi les premiers à avoir proposé l'application d'un variogramme robuste pour l'estimation d'un variogramme théorique. Le choix du variogramme robuste de chaque variable discriminante se fait sur la base de critères exigés pour l'estimation correcte de ce type de variogramme. Ces critères sont :

- un nombre minimum de classes de distance, de l'ordre de 15 à 20 ;
- un nombre de couples de sondages du sol par classe supérieur à 50 ;
- un variogramme calculé sur une distance inférieure à la moitié de la distance existante entre les deux points les plus éloignés (Journel & Huijbregts, 1978).

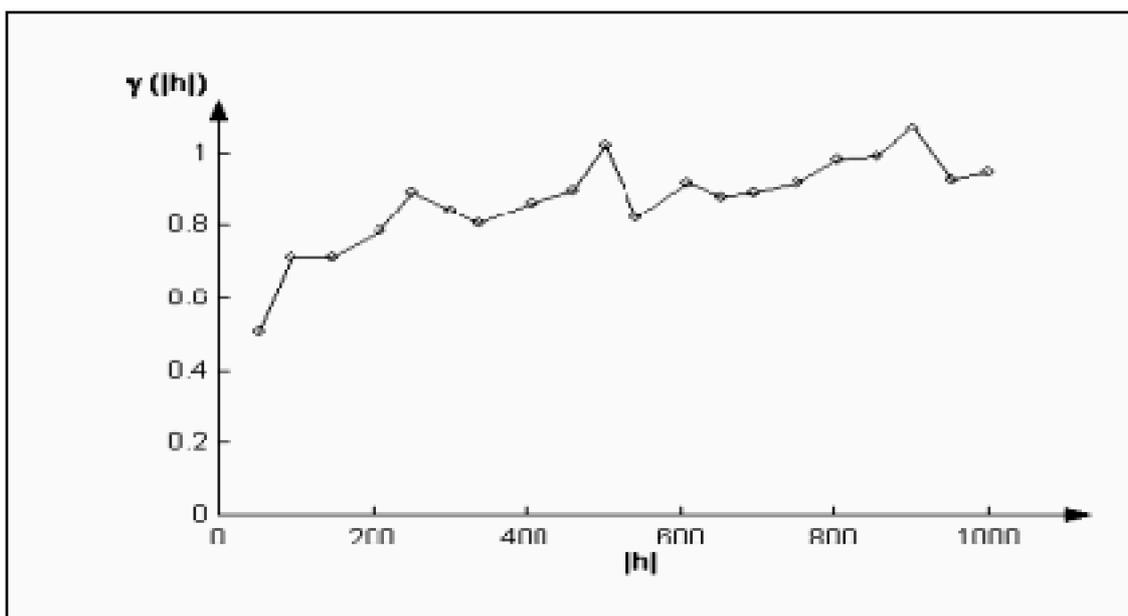


Figure n° 17. Variogramme expérimental robuste de la première variable discriminante.

8.2.2. Modélisation du variogramme expérimental

Cette étape consiste en la recherche d'un modèle structural théorique du variogramme expérimental. Il est à rappeler que le modèle théorique est une expression analytique que l'on tente d'ajuster le mieux possible aux points du variogramme expérimental. Beaucoup d'efforts ont été investis dans l'étude des estimateurs des variogrammes et dans le perfectionnement de leur mode d'estimation (Davis & Borgman, 1982 ; Myers, 1985, 1991). L'inférence de la courbe expérimentale à un modèle théorique est réalisée grâce à l'IGF (Indice Goodness Fit) ; indice proposé par Pannatier (1994). Cet indice oriente l'opérateur vers le bon ajustement en convergeant vers la valeur 0 à chaque variation de la valeur d'un des paramètres de la courbe théorique, à savoir ; le palier de la fonction (sill variance), la portée de la fonction (range) ou le pépite de la fonction (nugget variance). L'IGF est calculé selon la formule suivante :

$$IGF = \frac{1}{N} \sum_{k=1}^N \sum_{i=0}^{n(k)-1} \left(\frac{P(i)}{\sum_{i=0}^{n(k)-1} P(i)} \right) \left(\frac{r_k(i)}{r_k(0)} \right) \left(\frac{r_k(i) - \mu(i)}{\sigma} \right)^2, \text{ ou}$$

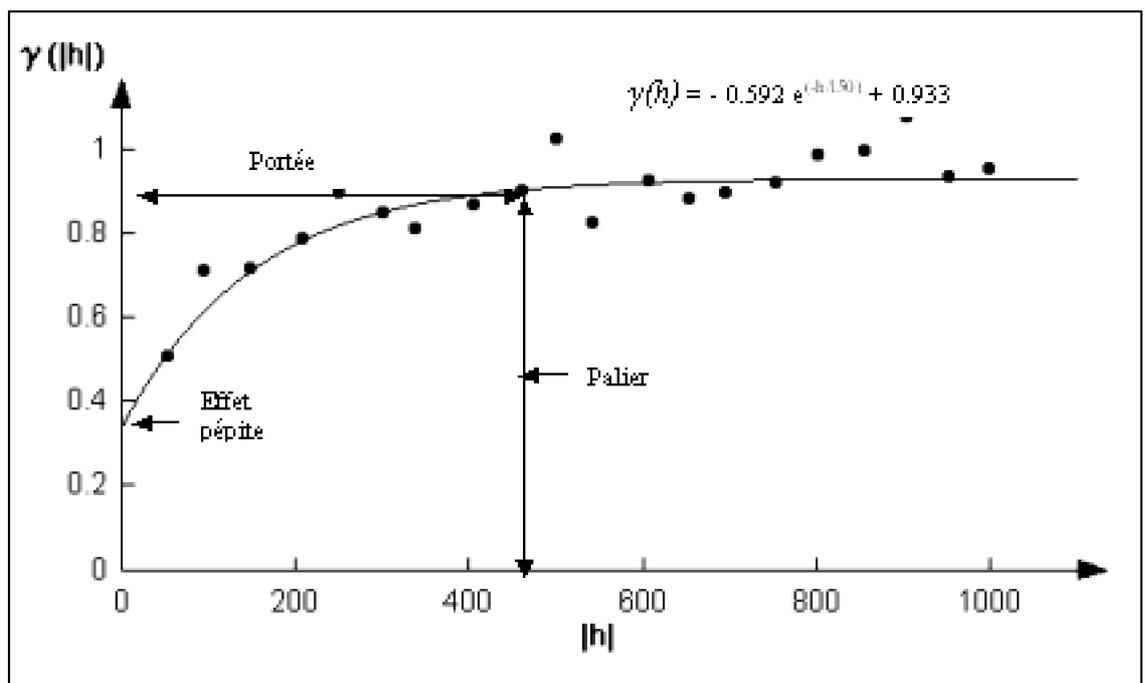
- N : nombre de variogrammes directionnels,
- $n(k)$: nombre de classes de distances relatif au variogramme k ,
- $P(i)$: nombre de paires d'individus composant la classe i du variogramme k ,
- $D(k)$: distance maximale relative au variogramme k ,
- $d(i)$: moyenne de la distance par classe du variogramme k ,

- $\gamma(i)$: dissemblance expérimentale de la continuité spatiale de la classe de sondage i ,
- $\gamma(i)'$: dissemblance du modèle de la continuité spatiale de $d(i)$,
- σ^2 : variance du variogramme standard.

Le variogramme théorique (fig.18), s'ajustant le mieux au variogramme expérimental omnidirectionnel retenu, correspond à un modèle de type exponentiel donnant un faible IGF de 0.0054119. Il est caractérisé par :

- un palier atteint à une distance (h) égale à 450m. C'est la distance d'autocorrélation qu'on surnomme la portée en deçà de laquelle, il y a corrélation entre les points de sondage, au delà il y a indépendance entre ceux-ci. La valeur de la dissemblance $\gamma(h)$ relative à la distance d'autocorrélation (palier de la courbe) , est égale à 0.933. Elle exprime la variance totale ;

- un effet pépîte correspondant à une valeur de dissemblance à l'origine ($h=0$) égale à 0.341. Cette valeur de dissemblance $\gamma(h_0)$ représente la variance locale. Elle exprime soit des erreurs de mesure, soit des hétérogénéités spatiales existant à une échelle plus détaillée que celle optée pour l'échantillonnage. Elle représente 36.54 % de la variance totale, ce qui indique une variabilité non négligeable aux faibles distances.



8.2.3. Tendence anisotropique du premier axe discriminant

Le variogramme surfacique proposé dans le Variowin permet l'identification des directions principales de l'anisotropie du comportement spatial de la variable régionalisée. En d'autres termes, il permet de voir dans quelles directions la continuité spatiale de la

variable est la meilleure.

Pour détecter la direction de plus forte continuité spatiale, il faut considérer le variogramme surfacique comme un cercle trigonométrique avec comme centre la cellule centrale du variogramme. Chacune des cellules carrées reflète par sa couleur une valeur de la dissemblance $\gamma(h)$ calculée pour une direction et une distance relatives à sa position par rapport à la cellule centrale du variogramme surfacique.

Le variogramme surfacique du premier axe discriminant (*fig. 19*) dévoile un comportement anisotrope de la variable régionalisée. Cette anisotropie se manifeste par des valeurs de la dissemblance en croissance dans la direction Sud- Est , Nord-Ouest (135°).

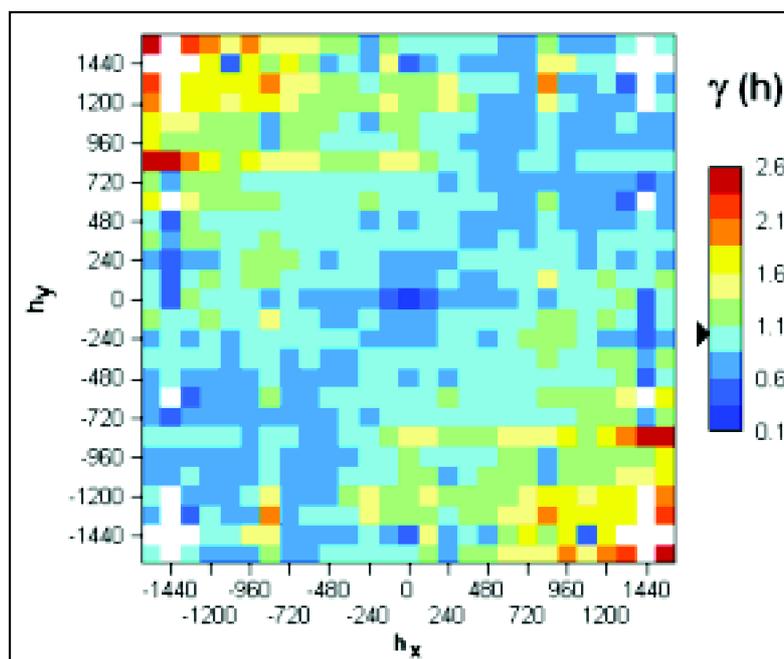


Figure n° 19. Variogramme surfacique relatif au premier axe discriminant

Grâce au variogramme surfacique, il est maintenant possible de réaliser un ajustement commun des deux variogrammes directionnels, nécessaires au calcul de l'anisotropie. La valeur de l'anisotropie est incluse dans la modélisation du semi-variogramme. Elle est égale au rapport de deux portées relatives aux deux variogrammes directionnels. Le premier variogramme est réalisé dans la direction principale de l'anisotropie relative à 45° avec une tolérance de fenêtre angulaire de 30° . Le second variogramme est calculé dans la direction perpendiculaire à la première, équivalant à un angle de 135° et de même fenêtre angulaire.

L'opération de l'ajustement simultané des deux variogrammes directionnels aux variogrammes théoriques aboutit à un modèle sphérique avec un IGF de 0.020792, une anisotropie de 0.53, un effet pépite de 0.417, un palier de 0.890 et une portée d'une distance de 275 m (*fig.20*). Cette anisotropie qui témoigne de la présence d'une dérive reste limitée aux faibles distances comme le montre les directions de l'accroissement des deux variogrammes directionnels qui se fait dans la même direction.

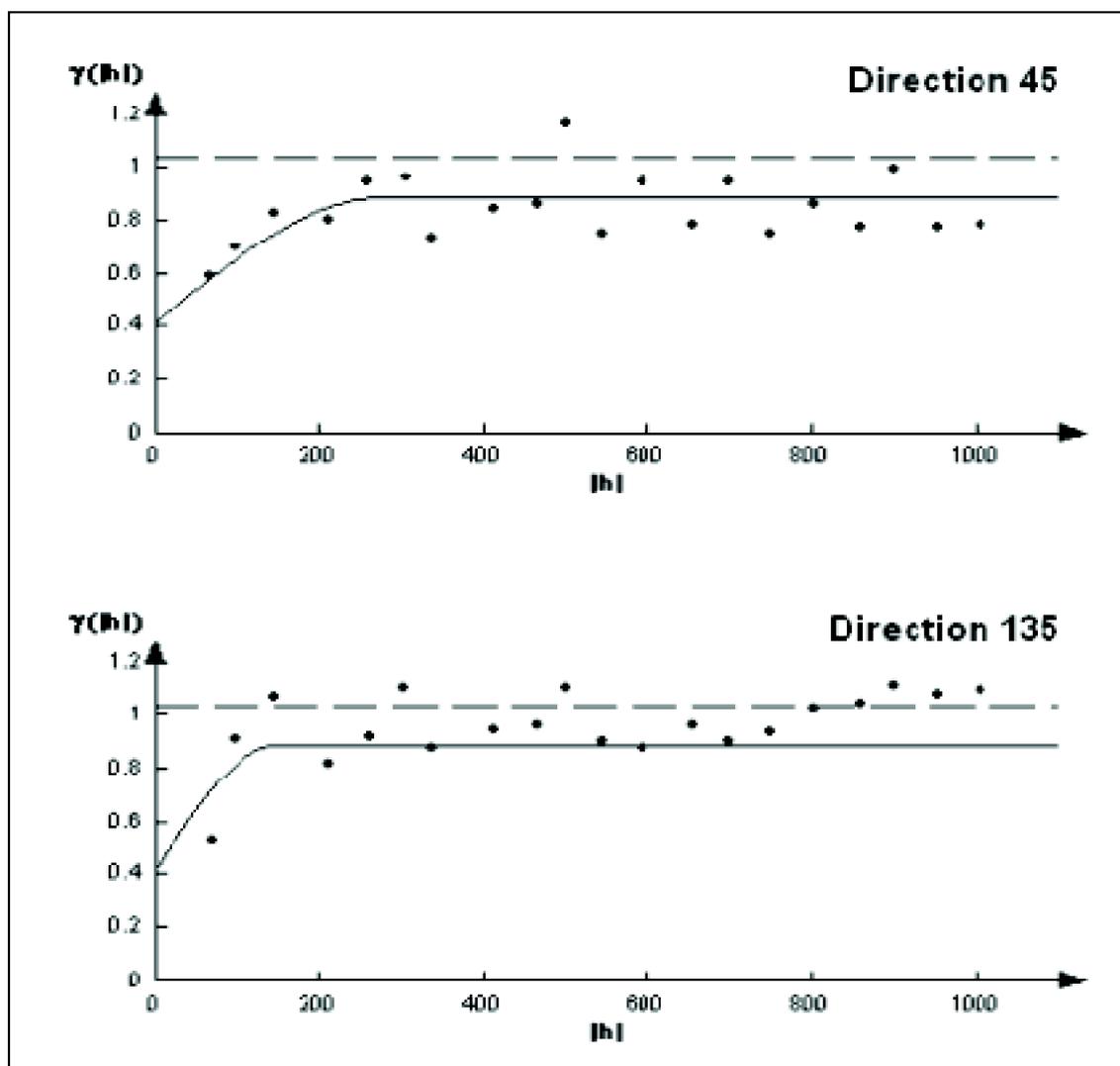


Figure n° 20. Ajustement commun des deux variogrammes directionnels du 1^{er} axe discriminant, au modèle théorique (sphérique).

8.3. Résultats de l'analyse variographique des deuxième et troisième axe discriminants

Une analyse variographique est réalisée également pour le deuxième et le troisième axe discriminants. Le calcul de leurs variogrammes expérimentaux est réalisé sur une distance de 1600 m. L'opération d'ajustement de ces deux variogrammes aux modèles théoriques aboutit aux résultats suivants (Tab. XXIX) :

- un modèle sphérique correspondant au deuxième variogramme expérimental (fig. 21). Il atteint un palier de 1.009 pour une portée de 560 m. Ce variogramme montre une forte variance locale de 0.650 soit, 64.42 % de la variance totale. Dans ce cas la variabilité aux faibles distances est très élevée. Aucune tendance anisotropique n'a été décelée lors du calcul du variogramme surfacique pour cette deuxième variable discriminante.

-un modèle exponentiel correspondant au troisième variogramme expérimental (fig. 22). Il est caractérisé par un palier égal à 1.033 relatif à une portée de 448 m. Ce variogramme montre une variance locale de 0.384 soit, 37.17 % de la variance totale. Comme pour le deuxième variogramme surfacique, la troisième variable discriminante ne montre aucune tendance anisotropique.

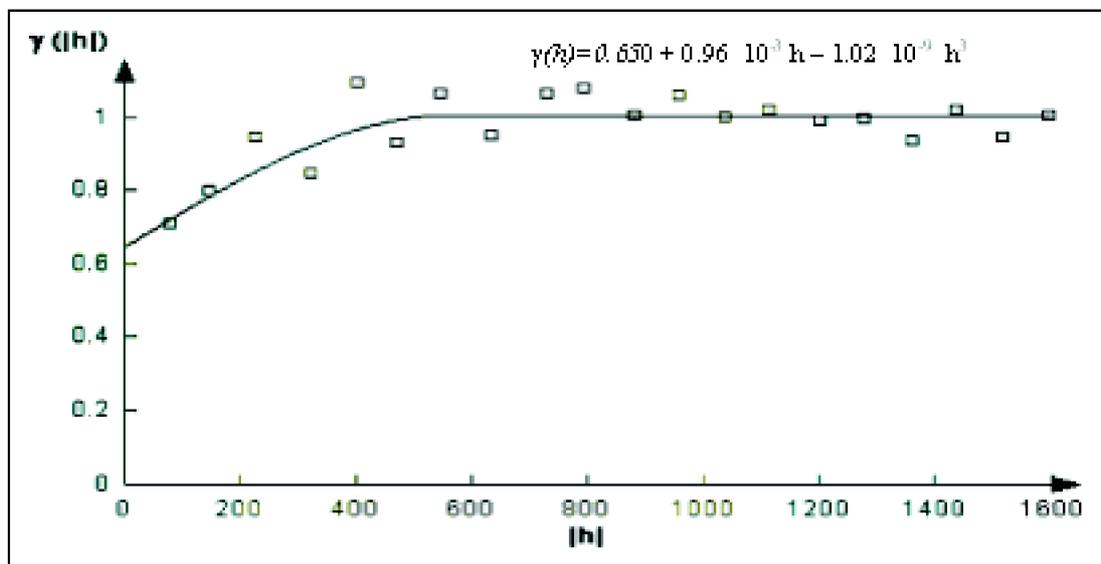


Figure n° 21. Variogramme expérimental ajusté au variogramme théorique (Modèle sphérique) relatif au deuxième axe discriminant.

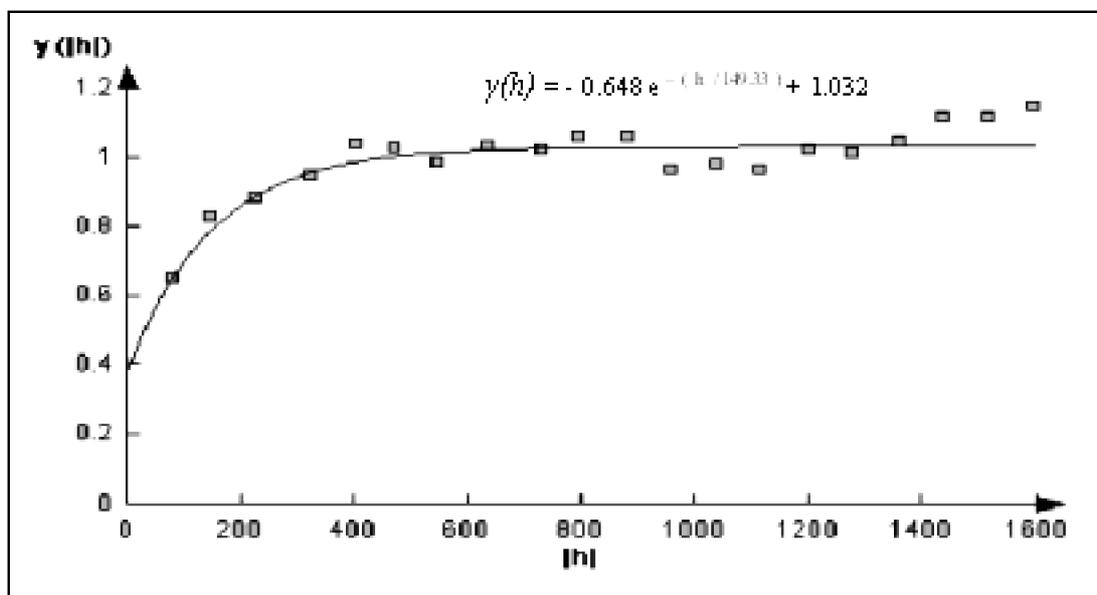


Figure n° 22. Variogramme expérimental ajusté au variogramme théorique (Modèle exponentiel) relatif au troisième axe discriminant.

Tableau n° XXIX . Caractéristiques structurales de la deuxième et troisième variables discriminantes des cinq unités homogènes du sol .

Paramètres de la fonction	2 ^{ème} axe discriminant	3 ^{ème} axe discriminant
---------------------------	-----------------------------------	-----------------------------------

$\gamma(h)$		
Effet pépité	0.650	0.384
Palier	1.009	1.033
Portée	560 m	448 m
Modèle	Sphérique	Exponentiel

8.4. Interpolation par krigeage

8.4.1. Mise en oeuvre du krigeage

L'analyse variographique a permis de modéliser le variogramme expérimental pour chacun des axes discriminants. Ce modèle devient par la suite un outil indispensable dans l'estimation d'une variable régionalisée X en tout point non échantillonné dans la zone d'étude. Dans un domaine d'intervalle relatif à sa portée, la structure du modèle intervient dans le calcul des coefficients λ_j , appelés poids d'interpolation.

La méthode d'interpolation retenue est le krigeage ordinaire à voisinage glissant. Il est le plus utilisé en science du sol (*Trangmar & al., 1985 ; Voltz, 1986 ; Douaoui, 1992*). Son choix se justifie notamment par la présence d'une dérive aux faibles distances reflétée par une tendance anisotrope de la 1^{er} variable discriminante.

La mise en œuvre du krigeage doit également tenir compte de certaines déterminations pour atteindre une précision acceptable. Ces déterminations sont :

- la distance du voisinage ; elle exprime la zone de recherche des points de sondage expérimentaux sur lesquels se base le krigeage pour estimer la valeur d'un sondage mesuré. Le rayon qui couvre cette zone doit être inférieur ou égal à la portée du modèle. Cette condition permet de prendre lors du krigeage uniquement les sondages expérimentaux qui sont corrélés entre eux. En présence d'une anisotropie, on adopte une zone de recherche elliptique parallèle à la direction de la meilleure continuité spatiale. Par contre, en présence d'une isotropie, on délimite une zone de recherche à rayon constant autour de chaque point de sondage à estimer ;
- le voisinage du krigeage ; il représente un sous ensemble de points de sondage expérimentaux les plus proches du sondage du sol à estimer. Ces points de sondage sont utilisés pour kriger. Leur nombre est fixé à un maximum de 10 points pour notre cas ;
- la grille du krigeage ; c'est une grille régulière, aux nœuds correspondant aux points de sondage à estimer. La distance entre ces points de sondages est fixée à 30 m. Ainsi, la grille d'estimation des valeurs de tous les points estimés totalise un nombre de points égal à 5625.

8.4.2. Résultats du krigeage

Chacune des cartes est le résultat du krigeage des valeurs des sondages correspondant à leur projection sur l'un des trois axes discriminants. Les lignes des isovaleurs correspondent aux limites des intervalles de localisation des classes de sondages du sol sur l'axe factoriel discriminant. Le choix de ces limites permet une reproduction fidèle de la réalité des relations qui existent entre les unités homogènes du sol tant dans leur espace factoriel discriminant que géographique.

L'absence d'une individualisation totale des cinq unités homogènes du sol dans leur espace discriminant signifie l'apparition de parties communes où communiquent partiellement les unités homogènes du sol entre elles. On qualifie ces parties communes d'unités d'association. Elles sont au nombre de sept dans chacune des cartes. Les deux unités restantes, figurant individuellement mais d'une façon partielle dans chacune des cartes, font référence à leur propre domaine d'apparition sur l'axe discriminant correspondant.

Dans les trois cartes, l'ensemble des sept unités d'association occupent la surface la plus importante de la parcelle d'étude. Elles se comportent par endroit comme zones de transition progressive vers chacune des unités homogènes.

Rappelant d'après les résultats de l'AFD (*Chap.VII*) que les nouvelles unités issues de l'association se placent dans une position intermédiaire entre deux unités homogènes. Cette situation se traduit obligatoirement sur les trois cartes par la rencontre de toutes les unités d'association le long du parcours reliant les deux unités homogènes. L'AFD sert également à mieux différencier les unités homogènes en fonction des variables du sol les plus discriminantes. Celles-ci sont reconnues par leurs coefficients de corrélation inter-unités les plus élevés dans l'ensemble.

8.4.2.1. Krigeage de la première variable discriminante

Pour cette première variable discriminante, la zone de krigeage retenue autour de chaque point à estimer est de forme elliptique en raison de la présence d'une tendance anisotropique. Le voisinage du krigeage est fixé à un maximum de 10 points de sondages expérimentaux. La fonction analytique $\gamma(h)$ du semi-variogramme utilisée pour le calcul des poids du krigeage (λ_j) est la suivante :

$$\gamma(h) = -0.592 e^{(-h/150)} + 0.933 \text{ où } h < 450m$$

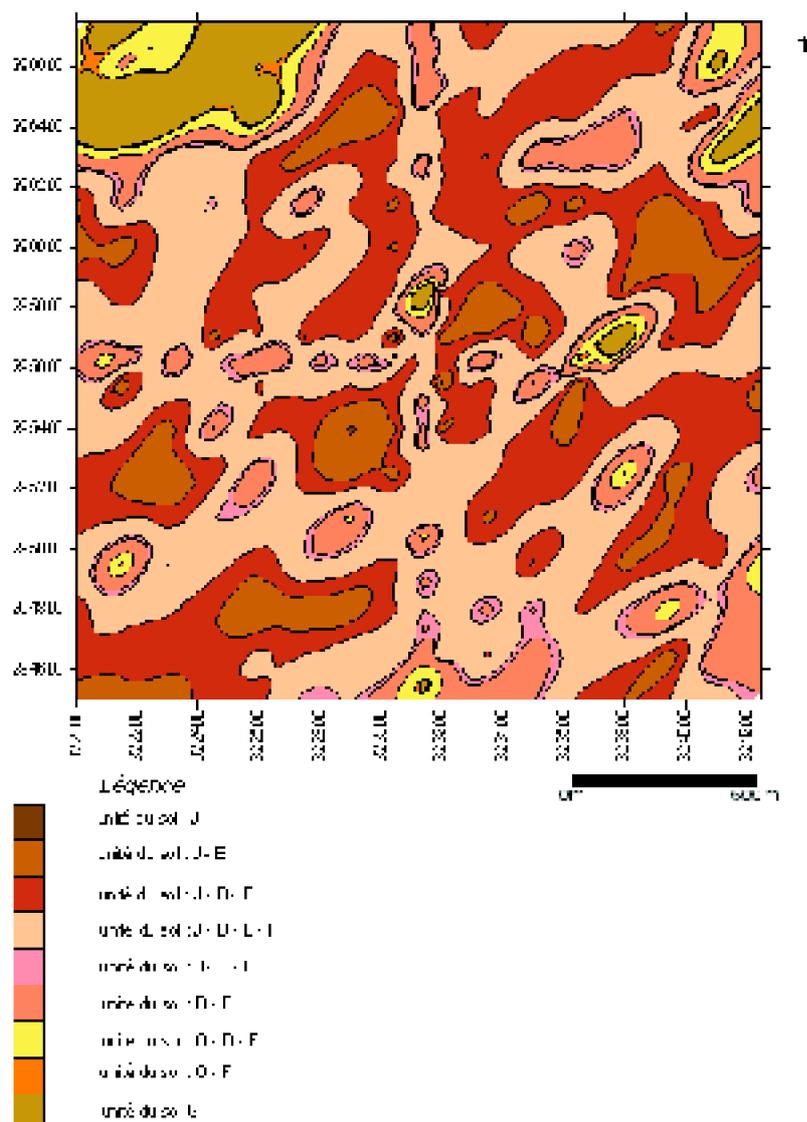


Figure n° 23. Carte de la première composante discriminante des unités homogènes du sol.

C'est dans cette carte qu'on discrimine mieux entre les unités du sol d'après les résultats de l'AFD. Les unités homogènes du sol G et J sont les plus discriminantes comme l'atteste l'écart de distance important qui les oppose sur le premier axe discriminant. L'unité J se distingue plus de l'unité G par ses teneurs élevées en A et MO et faibles en CaCO_3 total, Sf et Sg. Alors que les valeurs des variables du sol A et MO baissent progressivement le long des unités d'associations respectives J-E, J-D-E, J-D-E-F, J-D-F, D-F, G-D-F et G-F, celles du CaCO_3 total, Sf et Sg augmentent en revanche. L'ensemble des unités d'associations couvrent une surface plus importante dans la parcelle d'étude.

Sur la première carte relative au 1^{er} axe discriminant (fig. 23), l'unité du sol homogène G occupe individuellement l'extrémité Nord-Ouest de la parcelle. Elle se présente également en de petites plages dans l'extrême partie Nord-Est, au centre et au Sud de la parcelle. Elle délimite graduellement autour d'elle des zones de transition en

communiquant respectivement avec les unités homogènes du sol F et D. Ces zones de transition sont respectivement G-F et G-F-D. Une deuxième unité homogène du sol apparaît individuellement dans la parcelle d'étude ; il s'agit de l'unité J. Elle est la moins apparente dans la parcelle où on la localise au centre en de petites plages. Cette unité délimite autour d'elle plus de zones de transition que l'unité G. Ces zones de transition sont formées respectivement des unités d'association J-E, J-D-E, J-D-E-F et de J-D-F. Elles couvrent une surface plus importante dans la parcelle d'étude.

8.4.2.2. Krigeage de la deuxième variable discriminante

Pour la deuxième variable discriminante, la zone de krigeage retenue autour de chaque point à estimer est de forme circulaire en raison de la tendance isotropique de cette variable. Le voisinage du krigeage est composé d'un maximum de 10 points de sondages expérimentaux. La fonction analytique $\gamma(h)$ du semi-variogramme utilisée pour le calcul des poids du krigeage (λ_j) est la suivante :

$$\gamma(h) = 0.650 + 0.96 \cdot 10^{-3} h - 1.02 \cdot 10^{-9} h^3 \text{ où } h < 560 \text{ m}$$

Il est à rappeler, d'après les résultats de l'AFD que la différenciation entre les unités du sol est due beaucoup plus aux variables I_o , K, pH_{eau} et CE. Elles différencient beaucoup plus les unités homogènes du sol D et G mais aussi leurs premières zones de transition respectives D-F et G-F. Alors que les variables du sol CE et le I_o s'accroissent le long du passage des unités respectives D, D-F, D-E-F, J-D-E-F, J-G-D-E-F, J-G-E-F, J-G-F, G-F, G, les variables du sol pH_{eau} et K baissent en revanche. Les ensembles d'unités d'association couvre une surface plus importante dans la parcelle d'étude.

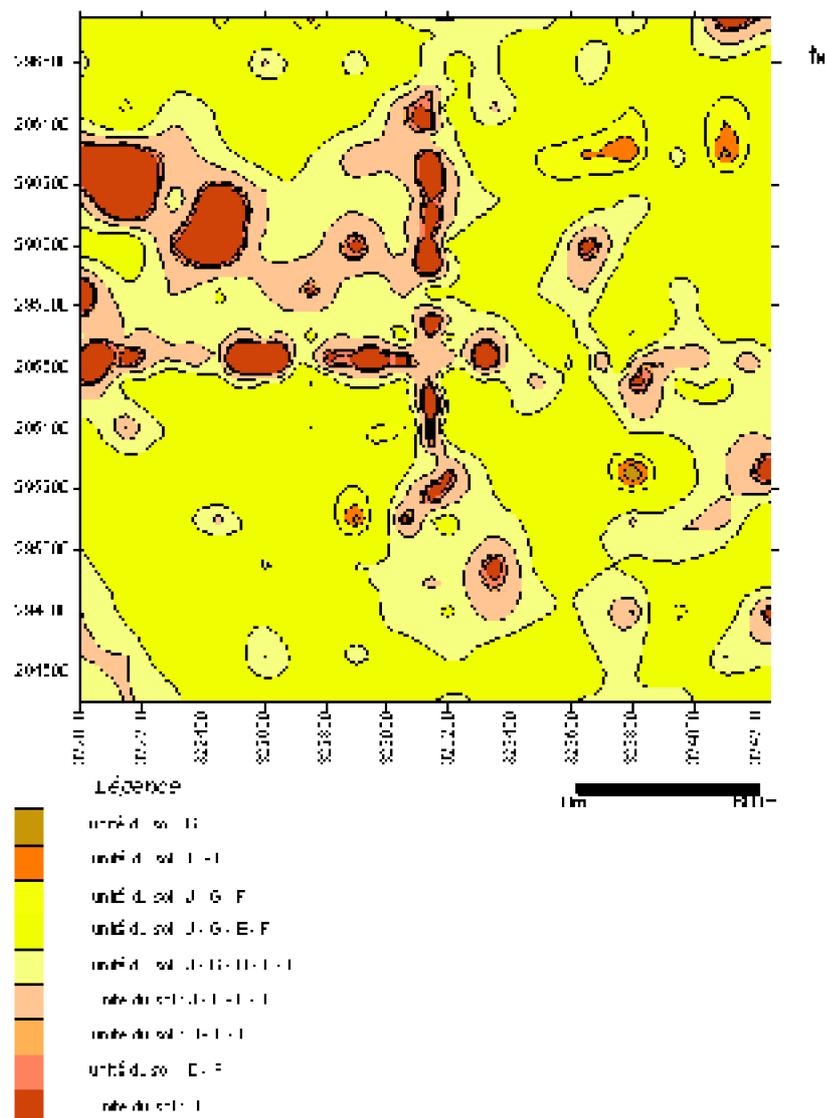


figure n°24. Carte de la deuxième composante des unités homogènes du sol.

Sur la deuxième carte (fig. 24) issue de l'opération de krigeage, les deux unités homogènes du sol D et G apparaissent individuellement. Elles partagent communément leurs premières limites de transition avec l'unité du sol F. L'unité homogène D apparaît beaucoup plus dans la partie Nord-Ouest de la parcelle. On la retrouve également à l'extrémité Nord-Est. L'unité G est la moins apparente. Elle est représentée par de petites plages au Nord-Est et au Sud de la parcelle d'étude.

8.4.2.3. Krigeage de la troisième variable discriminante

La zone de krigeage retenue pour la troisième variable discriminante est de forme circulaire et cela pour sa tendance isotropique. Le voisinage du krigeage est fixé à un maximum de 10 points de sondages expérimentaux. Ce sont les points de sondage les plus proches du point à estimer. La fonction analytique $\gamma(h)$ du semi-variographique utilisée pour le calcul des poids du krigeage (λ_j) est la suivante :

$$\gamma(h) = - 0.648 e^{- (h / 149.33) } + 1.032 \text{ où } h < 448 \text{ m}$$

Les variables contribuant le plus à la différenciation des unités du sol sont par ordre décroissant le Lf, Lg, CE et la MO d'après l'AFD. Cette différenciation est beaucoup plus importante entre les unités homogènes du sol J et F mais aussi entre leurs zones de transition respectives J-G et F-E. Alors que la teneur en Lf augmente le long du passage des unités respectives J-G, J-G-D, J-G-D-E, G-D-E, F-E-G-D, F-E-G, F-E, celles du Lg et de la MO baissent. Il en est de même pour la CE qui tend à diminuer dans ce sens là.

Sur la troisième carte (fig. 25) représentant le troisième axe discriminant, les deux unités homogènes du sol J et F apparaissent individuellement mais d'une façon partielle. Elles occupent moins d'espace dans la parcelle d'étude. Elles partagent respectivement leurs premières limites de transition avec les unités du sol G et E. L'unité homogène du sol J apparaît dans une petite plage dans la partie Sud-Ouest de la parcelle d'étude et l'unité F au méridien Est de la parcelle.

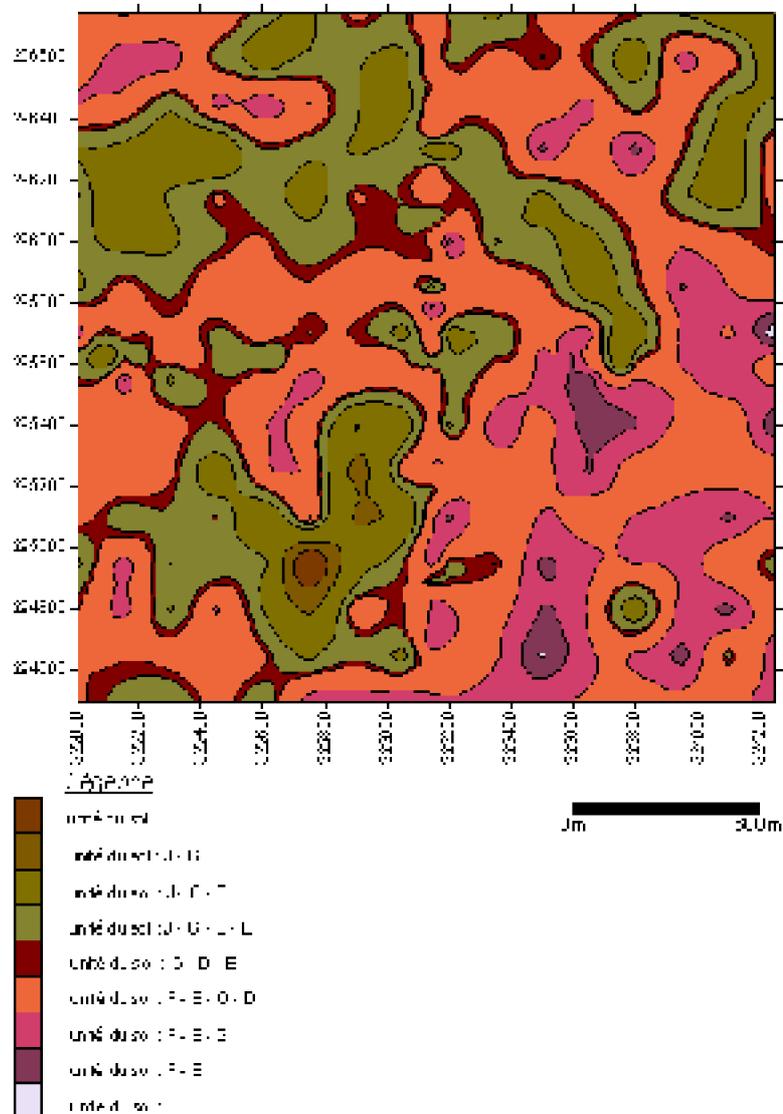


figure n°25. Carte de la troisième composante discriminante des unités homogènes du sol.

8.4.3. La validation croisée

La validation croisée permet de valider les paramètres de l'analyse spatiale (Issaks & Srivartava, 1989). Elle résume en tout l'importance du bon ajustement du variogramme théorique au variogramme expérimental ainsi que le voisinage choisi pour le krigeage. On applique le test de la validation croisée pour chacune des variables discriminantes des cinq unités homogènes du sol. Il est réalisé sur un nombre de sondages du sol n égal à 35, correspondant à 10 % de l'effectif total. Les sondages de la validation croisée sont pris au hasard sur un programme conçu dans le logiciel Surfer 8.0.

La procédure de la validation croisée consiste à éliminer à tour de rôle un sondage déterminé au préalable puis de l'estimer par la suite à l'aide de ses voisinages. Ainsi, en chaque point de sondage expérimental on a une valeur vraie (X_i) et une valeur estimée (\hat{X}_i) de la variable discriminante.

On juge également de la qualité de l'estimation par l'application du test des rangs relatif aux cas des échantillons associés par paires (test de Wilcoxon, 1945). Ce test est basé sur le calcul de la valeur U_{obs} ;

$$U_{obs} = \frac{|Y - n(n+1)/4|}{\{(n(n+1)(2n+1))/24\}^{1/2}}$$

Une valeur de $U_{obs} \geq U_{1-\alpha/2}$ conclut au rejet de l'hypothèse d'égalité des résultats entre les valeurs réelles et celles estimées par le krigeage, au risque α .

La valeur de $U_{obs} \leq U_{1-\alpha/2}$ constatée pour chacune des variables discriminantes (tab. XXX) au risque de signification $\alpha = 0.05$, conclut à l'égalité entre les valeurs réelles et celles estimées par le krigeage donc à la bonne approximation de la variable régionalisée.

Tableau n° XXX : Résultats du test des rangs appliqué sur les valeurs réelles et estimées des 35 sondages du sol.

Variable discriminante	Y	U_{obs}	$U_{1-\alpha/2}(\alpha=0.05)$	Niveau p
1 ^{ère} variable discriminante	236	0.79	1.96	0.42
2 ^{ème} variable discriminante	230	0.35	1.96	0.72
3 ^{ème} variable discriminante	207	1.06	1.96	0.28

Le calcul de l'erreur moyenne (EM) et la racine carrée de l'erreur quadratique moyenne concluent à une meilleure estimation de la 2^{ème} variable discriminante relativement à la 1^{ère} et la 3^{ème} variables. En effet, son EM et son RMSE égales respectivement à -0.089 et 0.459 sont les plus faibles dans l'ensemble (tab.XXXI).

Variable discriminante	Paramètres de la validation croisée	
	Erreur moyenne (EM)	Racine carrée de l'erreur quadratique moyenne (RMSE)
1 ^{ère} variable discriminante	0.222	0.750
2 ^{ème} variable discriminante	-0.089	0.459
3 ^{ème} variable discriminante	0.120	0.536

Tableau n° XXXI. Résultats de la validation relatifs à l'erreur moyenne (EM) et à la racine carrée de l'erreur quadratique moyenne (RMSE).

L'erreur moyenne est égale à $\frac{1}{n} \sum (X_i^* - X_i)$. La racine carrée de l'erreur quadratique moyenne est égale à $\sqrt{\frac{1}{n} \sum (X_i^* - X_i)^2}$.

Conclusion

L'analyse variographique caractérisant la structure spatiale de chacune des variables discriminantes révèle l'existence d'une dépendance spatiale entre les points de sondage du sol. Grâce à cette dépendance spatiale, il a été possible de procéder par la suite à l'exécution du modèle d'interpolation relatif au krigeage à voisinage glissant.

Pour les trois variables discriminantes (axes factoriels discriminants), la distribution spatiale des cinq unités homogènes du sol montre l'existence de larges zones de transition. Elles représentent des parties communes où s'alternent les unités du sol entre elles. Ces parties communes occupent des surfaces plus importantes dans la zone d'étude.

L'apparence individuelle des unités homogènes reste faible et partielle dans la zone d'étude et cela pour les trois variables discriminantes. Elle concerne les unités homogènes du sol G et J pour la première variable discriminante. Pour la deuxième variable discriminante, cette individualisation partielle est représentée par les unités homogènes du sol D et G. Pour la troisième variable discriminante, l'individualisation partielle concerne les unités homogènes du sol J et F.

Le test des rangs appliqué aux résultats de la validation croisée dans les 3 variables discriminantes conclut à la bonne approximation de celles-ci au risque de signification $\alpha = 5\%$. Dans l'ensemble, les faibles valeurs de EM et de RMSE de la 2^{ème} variable discriminante indiquent de faibles biais de son estimation.

Chapitre IX. Analyse de la pureté des unités homogènes du sol

Ce chapitre est considéré comme une finalité en soi du fait que l'objectif final pour toute cartographie pédologique, comme le souligne Walter & Aurousseau (1992), réside dans la définition et la délimitation des unités cartographiques au sein desquelles la variabilité des propriétés du sol est moindre que la variabilité générale observée au sein d'un paysage.

En adoptant les concepts proposés par Marsman & Gruijter (1986) de mesure de la pureté partielle et pureté moyenne, on décrit la variabilité des 11 paramètres du sol au sein de chacune des unités.

Ce chapitre dresse un constat sur le niveau de pureté au sein de la parcelle d'étude pour les unités suivantes :

- unités cartographiques issues de la carte des sols, se référant à l'étude agropédologique réalisée par le Bureau d'Etude Sir M Mac Donald et Partners Limited en association avec le BNEDER, à une échelle de 25 000^{ème},
- unités homogènes du sol issues de la Classification Ascendante Hiérarchique (CAH),
- unités homogènes du sol issues de l'analyse factorielle discriminante (AFD),
- unités du sol, après une analyse géostatistique, de la première variable discriminante.

9.1. Pureté partielle et pureté moyenne d'une unité cartographique

La pureté d'une unité cartographique correspond à la mesure de la concordance entre la définition d'une unité cartographique et les observations faites de façon indépendante en son sein. De ce fait, un nombre d'échantillons du sol est prélevé au sein de chacune d'elles (Tab. XXXII).

Pour chaque variable du sol j et pour un échantillon du sol indépendant, l'idée est de déterminer, comme le préconise Wilding & al. (1965), le nombre de fois où la valeur de la variable du sol j se retrouve dans l'intervalle de la moyenne m ($m_k \pm 1.96 (\sigma/(N_k)^{1/2})$) de tous les sondages de l'unité du sol k , au risque de 5 %.

Origine des unités	Unités du sol								
	Unité cartographique					Unité sectorielle			
Unités cartographiques (cartes pédologiques)									
N	172					176			
Sondages du sol indépendants (n)	17					18			
Unités de sol issues de la CAH	F	E			D	C	B		
N	56	140			76	22	46		
Sondages du sol indépendants (n)	7	14			7	5	7		
Unités de sol issues de AFD	F	E			D	C	B		
N	56	135			76	23	46		
Sondages du sol indépendants (n)	7	14			8	5	5		
Unités du sol de la 1 ^{ère} variable du canononic	J	J-E	J-D-E	J-D-E-F	J-D-F	D-F	C-D-F	C-F	G
N	1	41	72	93	23	62	25	0	16
Sondages du sol indépendants (n)	-	7	3	5	7	8	3	-	5

Tableau N° XXXII. Nombre de sondages constituant les unités du sol.

Pour cela, considérons une fonction binaire P telle que pour tout sondage du sol de vérification i ,

$$\begin{cases} P_{jk} = 1 & \text{si le sondage du sol } i \text{ satisfait à la définition de la variable } j \text{ pour l'unité du sol } k, \\ P_{jk} = 0 & \text{si le sondage du sol } i \text{ ne satisfait pas à la définition de la variable } j \text{ pour l'unité du sol } k. \end{cases}$$

On définit, de la sorte, la pureté partielle de l'unité du sol k pour la variable j notée P_{jk} :

$$P_{ik} = 1 / n^* \sum_{i=1}^n P_{ijk}$$

De la sorte, P_{ik} est égale à 1 si les valeurs d'une variable du sol j de tous les sondages de vérification n se situent dans l'intervalle de la moyenne m_k . Elle est égale à 0, si celles-ci se retrouvent toutes en dehors de cet intervalle.

On définit la pureté moyenne d'une unité du sol k par la moyenne arithmétique des puretés partielles des 11 variables du sol. Avec ces 11 variables cartographiques, la pureté moyenne \bar{P}_{mk} de l'unité du sol k est calculée par :

$$\bar{P}_{mk} = 1/11 \left(\sum_{j=1}^{j=11} P_{jk} \right).$$

n^* : le nombre de sondages de vérification effectués au sein de l'unité du sol k ,

N^* : le nombre de sondages qui constituent l'unité du sol k .

9.2. Analyse de la pureté des unités du sol

Le calcul des puretés, partielle et moyenne, est réalisé pour chacune des deux unités cartographiques issues de l'étude des ressources en sol. Il s'agit de l'unité cartographique appartenant à la classe des sols calcimagnésiques et l'unité du sol appartenant à la classe des sols salsodiques (CPCS, 1967). Le tableau XXXII/1 regroupe l'ensemble des résultats.

Pour les variables chimiques pH_{eau} , CE, CaCO_3 total et MO, les puretés partielles sont comprises entre 10 % et 34 %. Le taux de pureté partielle le plus élevé est enregistré dans l'unité du sol calcimagnésique, relativement à la variable MO.

Pour les variables granulométriques A, Lf, Lg, Sf et Sg les puretés partielles sont comprises entre 10 % et 45 %. Le taux de pureté partielle le plus élevé est enregistré dans l'unité du sol calcimagnésique, relativement à la variable argile (A).

Pour les variables du sol I_0 et K, lesquelles sont liées à la stabilité structurale, les taux de pureté partielle sont compris entre 16 % et 49 %. Le taux de pureté partielle le plus élevé est celui de l'unité du sol sodique relativement à la variable I_0 .

La pureté moyenne correspond à la moyenne des 11 puretés partielles. Elle est égale à 24 % pour l'unité calcimagnésique et de 25 % pour l'unité sodique.

Contribution à la détermination des unités homogènes du sol par différentes approches de cartographie numérique

Unité cartographique (k)	Pureté partielle (%)											pureté moy (%)
	pH _{eau}	CE	CaCO ₃	MO	A	Lf	Lg	Sf	Sg	I ₀	K	
-Sol calcimagnésique	10	10	10	04	45	04	10	22	03	22	10	24
-Sol salinodun	27	21	21	21	10	10	38	21	43	40	1E	25

Tableau XXXIII. Puretés partielles et puretés moyennes des 2 unités cartographiques du sol.

Les résultats du calcul de la pureté partielle pour les 11 variables du sol de chacune des 5 unités homogènes issues de la CAH sont regroupés dans le tableau XXXIV.

Pour les variables chimiques pH_{eau}, CE, CaCO₃ total et MO, les puretés partielles sont comprises entre 22 % et 69 %. Le taux de pureté partielle le plus élevé est enregistré dans l'unité du sol J relativement à la MO. Il est de 69 %.

Pour les variables granulométriques A, Lf, Lg, Sf et Sg les puretés partielles sont comprises entre 23 % et 91 %. Les puretés partielles les plus élevées sont enregistrées dans les unités du sol F, G et D, relativement au Sg. Elles sont respectivement égales à 68 %, 79 % et 91 %.

Pour les variables du sol I₀ et K, lesquelles sont liées à la stabilité structurale, les puretés partielles sont comprises entre 29 % et 56 %. Le taux de pureté partielle le plus élevé est celui de l'unité G relativement à la variable K. Il est égal à 56 %.

La pureté moyenne est comprise entre 34 % et 48 % pour l'ensemble des unités homogènes. C'est l'unité du sol G qui détient la valeur de la pureté moyenne la plus élevée.

Unité du sol (k)	Pureté partielle (%)											pureté moy (%)
	pH _{eau}	CE	CaCO ₃	MO	A	Lf	Lg	Sf	Sg	I ₀	K	
J	36	42	38	69	40	40	46	32	34	8	44	42
F	41	33	30	31	30	44	46	33	68	9	42	40
G	51	54	51	38	23	46	51	28	79	6	56	48
D	42	31	31	38	53	33	26	35	91	1	29	42
E	33	22	23	34	26	52	35	33	36	3	29	34

Tableau XXXIV. Puretés partielles et puretés moyennes des 5 unités du sol issues de la CAH.

Comme pour la CAH, un calcul de la pureté partielle et pureté moyenne est effectué

pour chacune des 5 unités du sol après une AFD. Le tableau XXXV regroupe l'ensemble de ces résultats.

Pour les variables chimiques pH_{eau} , CE, CaCO_3 total et MO, les puretés partielles sont comprises entre 27 % et 61 %. C'est l'unité du sol G qui en détient les taux de pureté partielle les plus élevés dans l'ensemble, relativement à son pH_{eau} et à son taux de CaCO_3 total.

Pour les variables granulométriques A, Lf, Lg, Sf et Sg les puretés partielles sont comprises entre 26 % et 93 %. Une pureté partielle de 93 % en Sg, exceptionnellement élevée, est enregistrée dans l'unité du sol D.

Pour les variables du sol I_0 et K exprimant la stabilité structurale, les valeurs de la pureté partielle, sont comprises entre 32 % et 70 %. C'est l'unité du sol G qui en détient le taux le plus élevé, relativement à la variable du sol K. Il est égal à 70 %.

Les puretés moyennes des cinq unités homogènes du sol sont situées entre 35 % et 56 %. L'unité du sol G a le taux de pureté moyenne le plus élevé.

Unité du sol (k)	Pureté partielle (%)										pureté moy (%)	
	pH_{eau}	CE	CaCO_3	MO	A	Lf	Lg	Sf	Sg	I_0		K
J	52	48	34	31	47	48	50	41	66	5	45	46
F	43	35	38	32	32	50	42	43	47	2	37	40
G	61	47	61	47	66	56	52	56	47	2	70	56
D	42	40	45	45	42	32	26	40	93	8	40	45
F	41	27	38	37	31	42	37	31	32	5	32	35

Tableau XXXV. Puretés partielles et puretés moyennes des 5 unités du sol issues de l'AFD.

La pureté partielle et la pureté moyenne sont calculées également pour les unités du sol issues de l'analyse géostatistique, en référence à la première variable discriminante (Tab.XXXVI). En rappelant toutefois que c'est cette variable qui discrimine le mieux entre les cinq unités homogènes du sol. Elle génère, en tout, 9 unités du sol ; parmi celles-ci sept d'entre elles représentent des unités d'associations correspondant sur le champ d'étude aux zones de transition.

En raison du très faible nombre de sondages constituant les unités du sol J et G-F,

les puretés partielles et moyennes de celles-ci ne sont pas calculées.

Pour les variables chimiques pH_{eau}, CE, CaCO₃ total et MO, les puretés partielles sont comprises entre 22 % et 67 %. Le taux de pureté partielle le plus élevé correspond à l'unité du sol G relativement à la variable MO.

Pour les variables granulométriques A, Lf, Lg, Sf et Sg, les valeurs de la pureté partielle sont comprises entre 17 % et 92 %. Les puretés partielles les plus élevées sont notées pour les unités du sol G-D-F et G par rapport à la variable Sg. Elles sont égales respectivement à 92 % et 89 %. Il en est de même pour l'unité du sol G, relativement au Sf. La valeur de sa pureté partielle est égale à 74 %.

Pour les variables du sol liées à la stabilité structurale ; I₀ et K, les puretés partielles sont comprises entre 26 % et 74 %. L'unité du sol G a le taux de pureté partielle le plus élevé. Il correspond à la variable K. Il est égal à 74 %.

Les puretés moyennes des 7 unités du sol sont comprises entre 28 % et 57 %. C'est l'unité du sol G qui détient le taux moyen le plus élevé.

Variable régionalisée (n° du sondage)	<i>Espace géographique</i>		<i>Espace factoriel</i>		
	Coordonnées Lambert (m)		Coordonnées des points de sondages sur les 3 premiers axes discriminants		
	Longitude (X)	Latitude (Y)	1 ^{er} axe	2 ^{ème} axe	3 ^{ème} axe
1	324250	294500	-0.366	0.811	-1.232
2	324100	294500	0.076	0.902	-3.448
3	323950	294500	0.439	0.232	-3.607
4	323800	294500	0.894	0.059	3.208
.
346	324150	295625	-0.255	-0.683	-3.783
347	324200	295625	-0.111	-0.719	-3.789
348	324250	295625	0.197	-0.038	-3.408

Tableau XXXVI. Puretés partielles et puretés moyennes des unités du sol issues après analyse géostatistique de la 1^{ère} variable discriminante.

Conclusion

Il est difficile de juger dans l'absolu des indices de pureté, car on ne dispose pas de références sur ce que peut être un niveau de pureté fiable (Walter, 1990). Le jugement sur le niveau de pureté obtenu n'est que relatif à l'ensemble des unités du sol issues de chaque approche cartographique.

En passant de la cartographie pédologique à l'analyse factorielle discriminante (*tab. XXXII, XXXIII, XXXIV*) on observe, à une exception près, une nette amélioration de la pureté partielle tant pour les variables chimiques que pour les variables granulométriques

et de stabilité structurale. Il en est de même pour la pureté moyenne relativement aux unités du sol.

Il est important de signaler que l'échelle de travail adoptée pour la réalisation de la carte pédologique classique n'est pas la même que celle adoptée pour ce travail. De ce fait, une réserve s'impose lors de la comparaison entre la pureté obtenue dans la cartographie classique et celle issue de la cartographie factorielle.

Synthèse et Discussion

La zone d'étude fait partie d'une plaine alluviale d'âge géologique récent issu du quaternaire. Elle constitue le milieu d'investigation des unités homogènes du sol. Cette investigation s'opère par la recherche d'une meilleure individualisation spatiale des unités homogènes intégrant à la fois un ensemble de paramètres du sol susceptibles de faire cette différenciation.

Le critère de choix des variables du sol est lié d'une part, à leur signification agronomique et d'autre part, à l'influence de certaines d'entre elles sur les propriétés physiques des sols de la région et à l'accessibilité facile à certaines d'entre elles. Il s'agit du pH_{eau} , de la conductivité électrique (CE), du calcaire total (CaCO_3), de la matière organique (MO), de l'argile (A), du limon fin (Lf), du limon grossier (Lg), du sable fin (Sf), du sable grossier (Sg), de l'indice d'instabilité structurale (I_0) et enfin, de l'indice de percolation (K).

En se référant à la fois à la moyenne (m) et au coefficient de variation (CV), la variabilité de chacune des variables du sol se dimensionne. De ce fait, le niveau de variabilité de chacune d'entre elles est évalué dans son propre contexte, c'est-à-dire, uniquement au niveau de la zone d'étude.

Le pH_{eau} de moyenne égale à 7.90 est alcalin. Cette variable est très homogène comme l'atteste son très faible CV (4 %). Le pH_{eau} est connu pour sa faible variabilité (Cipra & al., 1972 ; Compbell, 1978 ; Luxmore & Shama, 1980 ; Gajem & al., 1981 ; Vauclin, 1982 ; Laslett & al., 1987 ; Saidi & al., 1999 ; Benbouali & al., 2003).

La salinité du sol exprimée par la CE est élevée. Elle affiche une moyenne de 3.29 dS/m pour l'extrait dilué au 1/2.5. Le CV, égal à 104.86 %, traduit une variation importante de cette variable au niveau de la zone d'étude. La salinité est connue pour sa variabilité élevée dans la région (Saidi & al., 1999 ; Benbouali & al., 2003). De même, la variabilité élevée de ce paramètre est déjà citée par Wagenet & Jurinak (1978), Hajrasulika & al. (1980).

La teneur en CaCO_3 total du sol est modérée. La moyenne de cette variable est égale à 15.91 %. Le CaCO_3 total présente une très faible variabilité avec un CV égal à 16 %. Cette variabilité est deux fois plus grande que celle trouvée dans la région par Saidi & al. (1999) et Benbouali & al. (2003). Elle reste, tout de même, très faible en référence aux appréciations de Bascomb & Jarvis (1980).

Le sol est classé comme organo-minéral. Les résultats affichent une teneur en moyenne de la MO égale à 2.69 %. Sa variabilité a le même ordre de grandeur que celle citée par Cipra & al. (1972), Vauclin (1982), ainsi que Williot (1995) ou celle déjà trouvée dans la région par Saidi & al. (1999) et Benbouali & al. (2003). Le CV de la MO égal à 27 %, indique une faible variabilité de cette dernière.

La texture du sol est à dominance argilo-limoneuse. Les fractions granulométriques A, Lf et Lg sont de faible variabilité, respectivement, d'après leurs CV de 24 %, 22 % et 36 %. Les valeurs des CV de ces variables sont en accord avec celles compilées par Warrick & Nielsen (1980), Vauclin (1982). En revanche, les éléments grossiers Sf et Sg ont une dispersion très élevée par rapport aux normes citées par ces mêmes auteurs, ou celles trouvées dans la région par Saidi & al. (1999) et Benbouali & al. (2003).

La valeur moyenne de l'indice d'instabilité structurale est de 4,42. Ce niveau correspond à une faible stabilité structurale, confirmant ainsi les résultats auxquels ont abouti les travaux effectués par Saidi & al. (1999) dans la région. Le coefficient de variation calculé est de 23 %. Il révèle une faible variabilité de cet indice.

Le coefficient de variation permet de juger de la dispersion globale des variables du sol pH_{eau} , CaCO_3 total, MO, A, Lf et I_0 sous hypothèse d'une distribution normale de chacune d'entre elles, pour un seuil α égal à 5 %. Il induit, par contre, une erreur sur l'évaluation de la dispersion globale des variables du sol Lg, Sf, Sg et K vu le rejet de l'hypothèse de distribution normale de celles-ci. Le rejet de la distribution normale de ces variables est exprimé par allure dissymétrique de leurs histogrammes de fréquences et enfin des valeurs élevées de leurs indices d'asymétrie. Ainsi, pour peu de valeurs extrêmes asymétriques parmi l'ensemble des valeurs dans chacune de ces variables, on obtient des coefficients de variation élevés, ce qui n'est pas évident. Les quelques valeurs extrêmes rencontrées dans ces variables peuvent constituer des cas particuliers sur terrain et ne peuvent être représentatives de leurs environnements immédiats. Cette situation est de nature fréquente dans l'échantillonnage systématique, adopté pour la démarche, où l'emplacement du prélèvement du sol est déjà fixé au préalable. Il est également important de signaler qu'il est difficile de comparer des CV obtenus à partir des études réalisées à différentes échelles d'observation. Ce problème, conduit certains auteurs à décrire la variation des propriétés des sols comme un objet fractal (Burrough, 1983, a, b, c). Il ne faut pas omettre que le CV intègre également une part de variabilité due aux erreurs de mesure. Il induit par conséquent une part d'incertitude sur l'estimation de la moyenne et de la variabilité spatiale pour chacune des variables du sol.

Dans l'analyse structurale de la variabilité spatiale, seule les lois normale et log-normale sont utilisées dans ce travail pour la recherche des estimateurs non biaisés des espérances (moyennes estimées) des 11 variables du sol. Le rejet de ces lois de distribution, dans ce cas, pour le Lg, Sf, Sg et K s'explique soit par le fait que l'hypothèse de stationnarité n'est pas vérifiée (Vauclin, 1982), soit du fait de l'interaction de plusieurs effets sur ces variables d'après Burrough (1983). A défaut de la non exploitation d'autres lois de distribution comme la loi gamma, on suppose une distribution normale de ces dernières par appel au théorème central de la limite. Ce dernier suppose une distribution normale d'une moyenne lorsque la taille de l'échantillon est supérieure ou égale à 30. Ceci permet d'estimer dans ce cas les moyennes, les écart-types et les coefficients de

variation de ces variables comme dans le cas de la distribution normale. Seule la transformation logarithmique de la CE répond à une distribution normale, permettant ainsi de déterminer son estimateur non biaisé de son espérance au niveau de la parcelle d'étude. Il semble ainsi, d'après cette analyse que la correspondance entre distribution et variable n'est pas aussi simple comme l'a démontré Gascuel-Odoux (1987).

L'approche factorielle multiparamétrique représentée d'abord par l'analyse des composantes principales (ACP) ne montre pas de groupements distincts pouvant servir à la constitution des ensembles de sondages pédologiques cohérents. En majeure partie, les sondages du sol se concentrent avec une proximité élevée au milieu du plan factoriel formé par la 1^{ère} et la 2^{ème} composantes principales. Cette situation exprime sur le terrain une couverture pédologique continue où il est difficile d'établir des limites géographiques franches entre un type de sol et un autre sans imposer un seuil de variation (Legros, 1996).

L'ACP révèle que les variables du sol CE, MO et de façon moindre le I_0 s'opposent le plus aux variables K, pH_{eau} et CaCO_3 total. Cette relation inversement proportionnelle entre ces deux groupements de variables constitue la meilleure différenciation qui existe entre les 348 sondages même si elle est de moindre effet, compte tenu de la faible variance de 20.90 %, expliquée par la 1^{ère} composante principale.

La corrélation, d'une part, positive entre la salinité (CE) et l'indice d'instabilité structural (I_0) et d'autre part, négative entre la CE et l'indice de percolation (K) pourrait être le fait du pourcentage élevé de sodium échangeable (ESP) qui augmente parfois avec la salinité, tout particulièrement lorsque les teneurs en argile sont élevées (Daoud, 1999), comme dans le cas des sols étudiés. La stabilité structurale de ces sols se trouve influencée beaucoup plus par le test de percolation (K) que par l'indice d'instabilité (I_0). Ceci est expliqué par une valeur plus élevée du coefficient de corrélation au carré de la variable K par rapport à celle de (I_0) dans l'espace factoriel représenté par la première composante principale. Henin al. (1958) ont déjà souligné cette réponse plus positive du test de percolation (K) par rapport à celui de l'indice d'instabilité structurale (I_0), dans des conditions salines.

Paradoxalement, le rôle de la MO sur la stabilité structurale trouvé dans ce travail, ne répond pas à ce qui a été attendu au vu de la bibliographie consultée (Monnier, 1965 ; Tisdall & Oades, 1982 ; Churchman & Tate, 1987 ; Chenu, 1989 ; Jouany & al., 1992 ; Haynes, 1993 ; Le Bissonnais & Le Souder, 1995 ; Le Bissonnais & Arrouays, 1997). En fait, c'est surtout par rapport à sa nature et à son stade d'évolution que son action pourrait être expliquée (Gallali, 1980 ; Tisdall & Oades, 1982 ; Mellouhi, 1989 ; Le Bissonnais & Le Souder, 1995 in Douaoui & al., 2001).

Le CaCO_3 total réagit positivement à la dégradation de la stabilité structurale du sol dans la région. Ceci est déduit à la fois par son opposition à I_0 et à son alignement au K. Le rôle positif du CaCO_3 total dans l'amélioration de la stabilité structurale est indéniable (Harris & al., 1966 ; Mbogwu & Bazzoffi, 1998). Le calcium a pour effet de flocculer l'argile et par conséquent d'améliorer la stabilité structurale (Keren, 1989 ; Halliwell & al., 2001).

La différenciation par ces variables, de tous les sondages du sol, correspond aux deux plus hauts niveaux de résolution d'acquisition des données. Il est approximativement

de 2 sondages par hectare pour le transect et d'un sondage pour deux hectares pour le plan de quadrillage.

L'historique cultural ancien des parcelles ainsi que l'activité agricole qui prévaut actuellement ont, sans aucun doute, une part de contribution dans cette différenciation des sols par ces variables à cette échelle élevée d'acquisition des données. La zone d'étude fait déjà partie d'un ancien périmètre irrigué. Des parcelles, autrefois enrichies en matière organique et sur-irriguées, sont aujourd'hui abandonnées. Elles présentent des niveaux de salinité élevés. C'est sans doute, là aussi où beaucoup de sondages du sol requièrent de faibles valeurs de l'indice de percolation avec une accumulation élevée en matière organique. Ces parcelles sont généralement situées à l'intérieur des champs céréaliers, bien distinctes par leur aspect de surface caractérisé par la formation d'une croûte de battance alternant le plus souvent avec une structure très fine ; résultat d'une dégradation très avancée de la structure du sol.

La classification ascendante hiérarchique (C.A.H) complémentaire à l'ACP permet de ressortir les groupements de sondage du sol. Rappelons que le but principal de la CAH, comme c'est le cas de toute classification est de répartir une population en groupes d'observations homogènes, chaque groupe étant bien différencié des autres (Roux, 1985). L'espace factoriel de constitution des groupes de sondage du sol dans la CAH est de meilleure variance par rapport au plan factoriel constitué par les deux premiers axes des composantes principales. Il reste, cependant, faiblement représentatif de l'information sur la variance totale bien qu'il additionne une part de l'information issue du 3^{ème} axe des composantes principales. L'ensemble des trois premiers axes représentent 50.70 % de l'information sur la variance totale.

L'indice de Ward utilisé dans la CAH constitue le critère d'agrégation des groupements de sondage les plus proches dans cet espace euclidien. Rappelons que l'indice de Ward représente le gain d'inertie intra-classe entre deux étapes successives de partition. Plus ce gain est important, entre deux étapes successives de partition, meilleure est la dissemblance entre les groupements de sondage. Une faible décision dans le choix des cinq classes de sondage est admise relativement à l'existence d'un saut légèrement élevé de cet indice, entre la 343^{ème} et la 344^{ème} partition. Cette situation reflète combien l'individualisation des groupes de sondage du sol est faible malgré la possibilité qu'a la CAH de travailler dans un espace relativement de meilleure variance. Ces cinq classes issues de la partition représentent les groupements de sondage nommés D, E, F, G et J.

Le test de position (Kruskal & Wallis) révèle l'existence d'une différence significative dans la distribution de chaque variable du sol, entre les cinq classes de sondage obtenues par la CAH, au seuil de signification égal à 5%. Autrement dit, pour toutes les variables du sol, l'hypothèse d'appartenance des cinq classes de sondage à un même milieu homogène est rejetée à un niveau de signification de 5%. Ceci incite à une illustration graphique, telle suggérée par Becker & Chambers (1984), dans le but de ressortir les classes de sondage à distribution différente de celles à distribution analogue pour chacune des variables du sol.

Trois classes de sondage se distinguent notablement dans l'ensemble par certaines

variables du sol. On reconnaît la classe J par ses faibles valeurs de pH_{eau} , CaCO_3 total et K et élevées en CE et MO. De même, la classe D se distingue par ses valeurs élevées en pH_{eau} et K et faibles en lo. On distingue également, la classe G beaucoup plus, par ses faibles teneurs en A et Lf et élevées en Sf et Sg.

De ce fait, il apparaît que la salinité du sol constitue un paramètre de différenciation entre les classes de sondage J et D. Elle implique dans cette différenciation la stabilité structurale du sol exprimée beaucoup plus par l'indice de percolation K, la MO et l'indice d'instabilité (I_0). Les teneurs élevées en MO dans la classe de sondages J seraient le résultat de sa faible décomposition, en conditions de salinité élevée (Gallali, 1980). La classe J est beaucoup plus représentée dans sa majorité dans la partie sud-ouest ; le reste a tendance à se répartir en de petites plages au nord-est et près du centre. Elle représente sur le terrain, les sols salés, abandonnés, couverts par une végétation halophyte à dominance de *Suaeda Fructosa* représentant une densité de recouvrement très variable dans l'espace. Les sols de cette classe se manifestent en surface par une structure défloculée, poudreuse, surmontée parfois par une mince pellicule glacée. Ces sols sont dénommés par Boulaine (1957) « Solontchak vifs défloculés ».

Il existe cependant, des situations où il est impossible de différencier entre certaines classes de sondage par certaines variables du sol. Cette situation de distribution analogue touche les classes de sondage F, E et G par les variables du sol pH_{eau} et K. Il en est de même pour les classes de sondage F, E, D et G en relation avec les variables CE et MO. Le CaCO_3 total ne permet pas de distinguer entre les classes F, D et G. La variable argile (A) rend impossible la distinction entre les classes de sondage F, E, D et J. On ne différencie pas non plus, par le Lf, d'une part, entre les classes de sondage F et E et d'autre part, entre les classes de sondage D et J. Le même résultat a été observé pour le Lg quant aux classes de sondage F et E d'une part et les classes de sondage D, G et J, d'autre part. Le Sf enregistre la même situation pour l'ensemble des classes de sondage E, D et J de même pour le Sg par rapport aux classes de sondage F, E, D et J. Enfin, le I_0 empêche la différenciation entre les classes de sondage F, E, G et J et les classes de sondage F, E, D et G.

A ce stade de différenciation entre les classes de sondage, il en ressort que la distinction entre celles-ci par une variable unique n'est pas évidente. Les différenciations notables qui distinguent certaines classes de sondage dans l'ensemble ne sont que partielles et spécifiques à certaines classes de sondage. De même, pour la ressemblance entre certaines d'entre elles.

L'analyse de la variance monofactorielle devient l'élément clé dans la recherche des variables du sol les plus discriminantes des cinq classes de sondage issues de la CAH. L'ordre discriminant des 11 variables du sol est fait dans l'hyper-plan représenté par les trois premières composantes principales.

La première place discriminante revient au Sf. Elle est due beaucoup plus à la classe de sondage G qui se distingue par des teneurs plus élevées en Sf. La deuxième place discriminante revient à l'A. Cette variable permet de différencier également la classe G par ses plus faibles teneurs en A. La CE occupe la troisième place discriminante. Elle permet de différencier la classe J par ses valeurs plus élevées en CE. La quatrième place

discriminante est occupée par K. La classe de sondage D en est la cause principale d'après ses valeurs élevées en K. Il est également important de signaler que le Sg et le I₀ constituent les variables du sol les moins discriminantes dans l'ensemble d'après les résultats de l'analyse de la variance monofactorielle. On reconnaît, tout de même, d'après l'illustration graphique portée dans le chapitre VI, la classe de sondage G par ses teneurs élevées en Sg et la classe de sondage D par ses faibles valeurs en I₀.

L'analyse factorielle discriminante (AFD) proprement dite révèle, par contre, une meilleure discrimination et individualisation des cinq classes de sondage par la contribution de l'ensemble des variables du sol. Elle met en évidence, à partir de cet hyper-plan, des niveaux de dissemblance entre les cinq classes de sondage prises deux à deux. La distance de Mahalanobis est son critère de différenciation. De ce fait, il ressort dans l'ensemble une meilleure distinction et d'une façon décroissante entre les couples de classes de sondage du sol E-G, G-J, D-G. Par contre, les couples de classes de sondage E-J et F-E sont les moins discriminants.

En définitif, la CAH fournit une aide qui n'est pas négligeable dans cette étude. Elle se révèle un outil parfaitement adapté pour des données issues directement des distances euclidiennes relatives à l'ACP. La qualité de l'individualisation des classes de sondage du sol découlant de la CAH est révélée par l'AFD. Elle se traduit par le pourcentage de sondages bien classés égal à 83.04. Autrement dit, sur les 348 sondages de sol constituant les cinq classes issues de la CAH, 59 d'entre eux sont concernés par une affectation dans l'AFD. En fait, la constitution des classes de sondage se voit tributaire des deux méthodes : la première est opérationnelle et constitutive et la seconde est qualitative et discriminative. La CAH constitue d'abord les classes de sondages du sol par fusion rapprochée et l'AFD reconstitue ces classes à partir des points virtuels qui seraient les barycentres des points de sondages des classes formées au préalable lors de la CAH. On les qualifie de points de sondage type en référence au profil type. Ces points de sondage type comme les profils types, n'ont pas d'existence réelle sur terrain.

A l'instar des travaux de recherche sur la variabilité intra-unité (Wilding & al., 1964, 1965 ; Mc Cornack & Wilding, 1969 ; Burrough & al., 1971 ; Beckett & Burrough, 1971 ; Beckett & Webster, 1971 ; Bascomb & Jarvis, 1976 ; Ragg & Henderson, 1980 ; Wang, 1982 ; Edmonds & Lentner, 1986 ; Marsman & De Gruijter, 1986 ; Thomas & al., 1989 ; Walter, 1990), un calcul de l'écart-type (σ) après une CAH et une AFD permet, dans ce cas, d'évaluer la variation de la dispersion de chaque variable (Cransser & al., 1989) au sein de chaque unité du sol suite à un changement d'affectation de 58 sondages du sol entre les cinq classes. La CE, MO, Lf, Lg, Sf, Sg et K ont vu une baisse de leur variabilité dans les classes de sondage du sol F, E, D et G. C'est aussi le cas de la variable pH_{eau} relativement aux classes de sondage F et G, du CaCO_3 total pour les classes de sondage F, E et G, de l'A pour les classes de sondage du sol F et E et de I₀ relativement aux classes de sondage du sol F, D et G. Par contre, la classe de sondages J présente dans l'AFD des écarts-types plus élevés pour toutes les variables du sol à l'exception de la CE où la variabilité est plus faible par rapport à celle issue de la CAH.

La cartographie par sections est employée à la fin de la CAH et de l'AFD pour ressortir la distribution spatiale de chacune des classes de sondage du sol dans le milieu d'étude. La représentation spatiale dévoile une répartition fragmentaire des classes en de

grandes et petites plages, allant jusqu'aux surfaces élémentaires représentées par un seul secton. L'exemple des surfaces élémentaires est donné par les classes de sondages F et D et d'une façon moindre par la classe de sondage G. Ces sondages du sol peuvent constituer sur terrain des cas particuliers et ne peuvent être représentatifs de leur environnement immédiat ; tel est l'inconvénient engendré par la cartographie par secton. Elle nous prive de la sorte d'un atout de la cartographie pédologique qui consiste à ne pas s'appuyer uniquement sur des observations ponctuelles du sol (sondages pédologiques), nécessairement limitées, mais à tenir compte également des relations existantes entre les variations du sol comme l'aspect de surface (Walter, 1992). Cet inconvénient est évité par l'emploi de la géostatistique qui propose d'identifier la valeur d'un secton en tenant compte des valeurs des sectons voisins.

Dans l'étude géostatistique multiparamétrique, l'espace factoriel discriminant des cinq classes de sondage du sol issu de l'AFD est reproduit partiellement. La part de reproduction individuelle des cinq classes de sondage du sol dans chaque axe discriminant ou variable linéaire discriminante est différente. Elle l'est surtout dans la nature de la variable du sol, d'où la nécessité d'étudier la répartition spatiale de ces classes de sondage dans la zone d'étude pour chaque axe discriminant.

La géostatistique révèle dans l'analyse variographique des trois variables discriminantes une importante variance locale (effet pépité élevé), notamment pour le 2^{ème} axe discriminant, bien que le plan d'échantillonnage complémentaire représenté par les deux transects a pour rôle de minimiser celle-ci, comme le préconise Atteia & al. (1994). Gajem & al. (1981) et Burrough (1983) attribuent cette situation soit à des hétérogénéités spatiales existant à une échelle plus détaillée que celles prises en compte par l'échantillonnage complémentaire, soit aux erreurs de mesures lors de l'analyse des variables du sol au laboratoire.

Après une interpolation par krigeage à voisinage glissant, la reprise des valeurs des limites de localisation de chacune des classes de sondage du sol sur chacun des axes discriminants dévoile des zones de transition à étendues variables. Dans leur ensemble, les zones de transition occupent une surface bien plus importante que celle occupée partiellement par le restant des unités homogènes apparaissant individuellement. La géostatistique effectue une bonne synthèse sur le gradient spatial des trois variables linéaires discriminantes. Elle donne sur les cartes des résultats comparables à une décomposition du gradient en un ensemble de surfaces qui se succèdent à des degrés croissants dans les zones de transition. Voltz (1986) aboutit par analyse factorielle au même constat dans l'étude de la variabilité spatiale des propriétés physiques du sol en milieu alluvial.

La cartographie de la première variable linéaire discriminante représente une évolution des constituants du sol A, CaCO_3 total, Sf, Sg entre les cinq classes de sondage dans la zone d'étude. Alors que les teneurs en A et MO baissent progressivement lors du passage de l'unité J à l'unité G, celles du CaCO_3 total, Sf et Sg augmentent. Cette variable discriminante se manifeste par une réduction importante de l'espace d'individualisation des cinq classes de sondage du sol. Pourtant, c'est à partir de celle-ci qu'on discrimine le mieux entre ces classes compte tenu de son pouvoir discriminant élevé, exprimé par une valeur propre égale à 61.7 %.

La cartographie de la deuxième variable linéaire discriminante offre un espace de discrimination des classes de sondage beaucoup plus par les variables du sol I_0 , K, pH_{eau} et CE. L'évolution de la salinité et de la dégradation du sol entre les classes de sondage se fait essentiellement en fonction de cette variable linéaire. Alors que les variables du sol CE et I_0 augmentent en allant de l'unité D vers l'unité G, la variable du sol K baisse.

La cartographie de la troisième variable linéaire discriminante représente une évolution de la salinité et des constituants du sol Lf, Lg et MO. Lorsque la teneur en Lf augmente de l'unité J vers l'unité F, celles du Lg et de la MO diminuent. Il en est de même pour la CE qui tend à diminuer dans ce sens là.

Il en ressort, tout de même, de cette approche partielle sur la répartition des cinq classes de sondage du sol dans la zone d'étude diverses possibilités d'examiner la variabilité spatiale selon la nature discriminante des variables du sol. Egalement, bien que cette approche partielle nous prive d'une meilleure individualisation de ces classes de sondage, celle-ci fait apparaître de larges zones de transition entre les unités du sol.

Le plan d'échantillonnage complémentaire représenté par les deux transects avec un pas d'échantillonnage réduit à 50 m permet de garantir les prévisions d'une précision prescrite (Webster, 1985 ; Finke & Stein, 1994) ; celle d'une meilleure estimation de la variable linéaire régionalisée. Autrement dit, le pas réduit vient pour renforcer l'estimation en tout point de la variable linéaire dans le krigeage ordinaire à voisinage glissant. Le faible champ d'interpolation dans ce krigeage, est compensé par le nombre de points réels gagnés à l'intérieur de ce dernier.

Dans une optique de validation croisée, à l'instar des travaux de plusieurs auteurs (Creutin & Obled, 1982 ; Van Kuilenburg & al., 1982 ; Laslett & al., 1987 ; Leenaers & al., 1990 ; Voltz & Webster, 1990 ; Voltz & Goulard, 1994 ; Odeh & al., 1994 ; Bourennane & al., 1996), 35 valeurs réelles prises dans chaque variable linéaire discriminante ont été estimées par krigeage ordinaire à voisinage glissant. Le test des rangs relatif aux cas des échantillons associés par paires (test de Wilcoxon, 1945) confirme la bonne approximation par ce krigeage des trois variables linéaires discriminantes au seuil de signification $\alpha = 0.05$. La meilleure approximation est obtenue par la deuxième variable linéaire discriminante d'après les faibles valeurs de ses indices EM et RMSE, plus proches de zéro. Elle est suivie de la troisième puis de la première variable linéaire discriminante. Il est important de signaler que l'emploi de l'indice RMSE est plus sûr pour arriver à cette conclusion. Cet indice, contrairement à l'indice EM est indemne d'erreurs de compensation qui pourraient exister lors de la sommation des biais négatifs et des biais positifs d'estimation pour chacune des variables linéaires discriminantes.

Le pédologue cartographe serait sûrement indécis tant dans le niveau de distinction entre les unités homogènes que dans leurs éventuels discriminations par les variables du sol dans le champ spatial d'investigation. Il se contente durant la prospection sur terrain d'un choix instantané de facteurs pédologiques pour la discrimination entre les unités dans la parcelle d'étude. L'analyse factorielle, en général, se place dans ce travail comme une alternative à cette carence. Elle corrige et complète l'approche du pédologue cartographe dans le sens où elle constitue les unités homogènes en exploitant,

ultérieurement, les données de base récoltées par ce dernier sur le terrain. Ceci l'amène également à orienter le choix de ses variables en fonction du poids discriminant de ces dernières. C'est justement l'un des résultats que ce travail révèle.

Du niveau de prospection constituant les 348 sondages du sol à l'échelle d'observation donnant lieu aux cinq unités homogènes, il y a un changement de la variable discriminante dans la parcelle d'étude. En effet, de la salinité qui discrimine entre tous les sondages on passe vers le sable fin pour différencier entre les cinq unités du sol. Par ce changement de la variable du sol, il s'avère donc que le critère de discrimination est étroitement lié au niveau d'investigation. Ceci est un second résultat que révèle ce travail.

Il est vrai que le Sf soit la variable du sol la plus discriminante d'après l'AFD. Ceci elle le doit à l'unité G qui en est la cause principale vu son taux élevé en Sf dans l'ensemble. Les autres unités E, D et J ne peuvent être distinguées par cette variable, comme le montre l'illustration graphique de la distribution des cinq classes de sondages du sol issues de l'AFD. C'est cette critique là qu'on voudrait apporter à cette méthode factorielle. Il s'agit de mettre le doigt sur le fait qu'elle n'a pas été réellement discriminante entre les cinq unités homogènes du sol.

La prospection pédologique, telle entreprise par le pédologue cartographe, aurait sans doute piétinée sur les réelles limites entre les unités homogènes du sol, encore pourrait-il les distinguer, dans la zone d'étude. Ceci vient du fait que les limites de transition ne sont pas, toujours, facilement décelables sur terrain. En dépit de nombreux travaux sur la précision des limites entre les unités cartographiques (Vink, 1963 ; Webster, 1973 ; Beckett & Bie, 1975 ; Campbell, 1977 ; Legros, 1978 ; Chrisman, 1982 ; Feeny, 1988 ; Ruellan, 1989 ; Lagacherie & al., 1994 ; Cazemier, 1999), la géostatistique conçoit dans cette étude cette carence et permet de ressortir les limites de transition et de les quantifier. En définitif, elle offre une perspective séduisante pour la description spatiale de la couverture pédologique.

Au terme de ce travail, on a été amené à évaluer le niveau de pureté pour chacune des approches cartographiques. La mesure de la pureté a pour intérêt de jauger l'apport de chaque méthode cartographique pour une différenciation optimale des unités cartographiques.

Les résultats montrent une amélioration des indices de pureté, en passant de la cartographie pédologique classique à la cartographie pédologique factorielle, sous réserve, de la différence d'échelle de travail des deux approches cartographiques.

Marsman & Gruijter (1986) montrent que dans des études cartographiques au 1/10000 la pureté moyenne et la pureté partielle varient pour la première entre 59 % et 62 % et pour la seconde entre 48 % et 84 %. Ce niveau n'est pas atteint dans l'ensemble, surtout pour la pureté moyenne. Ceci n'est probablement que la conséquence du niveau de partition choisi pour l'obtention des cinq classes de sondages du sol dans la CAH. Plus le niveau de partition est faible, plus la dispersion des sondages autour de la moyenne est élevée pour donner lieu à des puretés faibles. Comme pour l'échelle de prospection pédologique, ceci laisse supposer que les indices de pureté sont étroitement liés aux niveaux de partitions des unités du sol lors de l'implication de l'analyse factorielle dans la

cartographie pédologique.

Si la pureté partielle est calculée dans ce travail sur la base de la moyenne pour chacune des approches cartographique, elle reste dépourvue d'objectivité dans la cartographie pédologique factorielle. Cette dernière, la cartographie factorielle, devrait se référer au centre de gravité de chaque unité du sol, autrement dit aux horizons de référence pour le calcul de la pureté partielle. Ce qui n'a pas été le cas. C'est encore un argument pour expliquer la faiblesse des niveaux de pureté pour chacune des méthodes cartographiques factorielles.

Dans l'analyse géostatistique, les indices de pureté des unités du sol obtenus sont faibles compte tenu des valeurs proposées par Marsman & Gruijter (1986), particulièrement au niveau des unités de transition. Les unités de transition ne sont pas issues d'une partition ; partition qui constitue les groupements par critères de ressemblance entre les sondages du sol dans l'espace factoriel discriminant. Elles ne sont que le résultat de l'intersection entre deux ou plusieurs unités du sol dans l'espace factoriel discriminant indépendamment de leurs critères de ressemblance.

Conclusion générale

Un des objectifs essentiels de ce travail est de déterminer le meilleur niveau d'individualisation et donc de différenciation des unités homogènes du sol. Pour ce faire, l'horizon de surface est pris comme exemple en le prenant comme un élément fondamental de l'organisation spatiale de la couverture pédologique.

Une parcelle d'une superficie de 506.25 ha, située au sein de la plaine du Bas-Cheliff, a servi de champ d'investigation pour ce travail. L'espace factoriel issu de l'analyse des composantes principales (ACP) est le point de départ pour la recherche de groupements de sondage les plus cohérents, relativement à un certain nombre de variables du sol choisies en fonction de leur influence sur les propriétés physiques et physico-chimiques des sols de la région.

Les résultats statistiques obtenus par l'ACP n'ont pas été très conséquents, vu le faible pourcentage de la variance expliqué par les plans des variables ou les plans des sondages du sol. A cela, s'ajoute la forte densité de regroupement de sondage au milieu des plans rendant ainsi, plus difficile la distinction ou la séparation des groupes de sondage les plus cohérents.

La classification ascendante hiérarchique (CAH) complémentaire à l'ACP vient comme une alternative à cette carence dans la constitution des groupements de sondage les plus cohérents. Néanmoins, les faibles écarts entre les indices de niveau (D_Q) dans la CAH confirment la difficulté de distinguer entre les groupements de sondage. Tout de même, la décision de retenir un nombre de cinq classes de sondage du sol est admise par le saut relativement élevé de l'inertie intra-classe, entre la 343^{ème} et la 344^{ème}

partition.

L'AFD constitue une étape complémentaire dans ce travail. Par son pouvoir d'allier l'ensemble des variables du sol pour la différenciation entre les cinq classes de sondage issues de la CAH. De celle-ci, il ressort des niveaux de dissemblance entre les cinq classes de sondage du sol par référence aux sondages types pris comme référence par l'AFD. Les variables du sol, dont les plus discriminantes, comme la salinité du sol (CE), la matière organique (MO), le limon fin (Lf), le limon grossier (Lg), le sable fin (Sf), le sable grossier (Sg) et l'indice de percolation (K) ont connu une baisse de leurs variabilités dans les quatre unités sur les cinq retenues. Ceci est aussi l'un des résultats escompté par cette approche de cartographie numérique.

L'exploitation, par partie, de cet espace factoriel discriminant par la géostatistique, après une confirmation de la bonne approximation de chacune d'elles, permet de quantifier les aires de transition entre les cinq unités homogènes du sol. Elle met, également, en valeur la variation progressive des variables du sol lors du passage d'une unité homogène à une autre. C'est ainsi que l'exploitation de la part discriminante la plus élevée (61.7 % de la variance totale) met en relief une baisse progressive des taux d'argile et de matière organique et une augmentation progressive du CaCO_3 total et du sable fin lors du passage de l'unité J vers l'unité G. Par ce résultat, il devient clair que le tracé entre les unités du sol au moyen de la cartographie classique est loin de pouvoir mettre en évidence l'information sur la quantification et la caractérisation d'une limite entre deux unités du sol.

Les classes de sondage du sol ainsi obtenues se voient soumises à une évaluation de leurs niveaux de pureté partielle et moyenne. Le calcul de ces puretés à différentes étapes de la démarche cartographique adoptée révèle une amélioration des niveaux de puretés en allant des unités homogènes du sol obtenues par la CAH vers celles obtenues par l'AFD. Il est à retenir tout de même que ces niveaux de pureté restent en deçà des taux fixés par la bibliographie. On suppose que ces faibles niveaux de pureté sont la conséquence du faible nombre de classes retenu. Pour la géostatistique, le calcul de ces puretés a connu une baisse surtout au niveau des unités de transition. Cette baisse des puretés serait liée au fait que les unités de transition ne soient pas soumises au critère de ressemblance imposé par la CAH.

Abréviation

- Variables du sol
 - A : argile,
 - Lf : limon fin,
 - Lg : limon grossier,
 - Sf : sable fin,
 - Sg : sable grossier,
 - pH_{eau} : potentiel hydrogène,
 - CE : conductivité électrique,

-
- MO : matière organique,
 - CaCO_3 total : calcaire total,
 - I_0 : indice d'instabilité du sol,
 - K : indice de percolation.

- Paramètres statistiques

- min : valeur minimale,
- max : valeur maximale,
- m : moyenne,
- σ : écart-type,
- σ^2 : variance,
- CV : coefficient de variation,
- γ : dissemblance,
- RMSE : racine carrée de l'erreur quadratique moyenne,
- EM : erreur moyenne,
- ACP : analyse des composantes principales,
- CAH : classification ascendante hiérarchique,
- AFD : analyse factorielle discriminante,
- P_{ik} : pureté partielle,
- $P_{m k}$: pureté moyenne.

Références bibliographiques

- AFES., 1995. Référentiel Pédologique 1995. (Baize D., Girard M.C., coord). INRA., Paris, 332 p.
- Arrouays D. (1987). Cartographie des sols et comportements agronomiques. Comparaison des données de cartographie et d'enquêtes agronomiques en vue de la thématization d'une carte de sols. *Sci. Sol*, 25, 1, 43-58.
- Atteia O., Dubois J. P. & Webster R. (1994). Geostatistical analysis of soil contamination in the Swiss Jura. *Envir. Pollut.*, 86, 315-327.
- Aubert G. (1976). Les sols sodiques en Afrique du Nord. *Ann. de l'INA d'Alger*, vol.VI, no1, 185-196.
- Baise D. (1986). Guide des analyses courantes en pédologie. INRA édition, 170 p.
- Baise D. (1986). Couvertures pédologiques, cartographie et taxonomie. *Science du Sol*. 24 (3), 227-243.
- Banfield C.F. & Bascomb C. L. (1976). Variability in three areas of the Denchworth soil mapunit. II Relationships between profiles using laboratory measurements and field observation. *J. Soil Sci.*, 27, 438-450.
- Bascomb C.L. & Jarvis MG. (1976). Variability in three areas of the Denchworth soil map unit I : Purity of map unit and property variability Within it. *Journal of Soil Science*, 27, 420-437.
- Bavaud F. (1999). Mesures BTS et interpolation. (Guide d'utilisation de Variowin et du

- Surfer). 49 p.
- Becker R.A. & Chambers J.M. (1984). S: an interactive environment for data analysis and graphics. Wadsworth advanced book program. Belmont, California, 549 p.
- Beckett P.H.T. & Burrough P.A. (1971). The relation between cost and utility in soil survey. IV- Comparison of the Utilities of Maps produced by Different Survey Procedures and different scales. J. Soil Sci., Vol. 22(4), 466-480.
- Beckett P.H.T. & Webster R. (1971). Soil variability a Review. Soil and fertilizers, 34, 1-15.
- Beckett P.H.T. & Bie S.W. (1975). Reconnaissance for soil survey. I. Pre-survey estimates of the density of soil boundaries necessary to produce pure mapping units. J. Soil Sci., 26, 144-154.
- Becket R.A., Darricau H., Delhomme J. P. & De Marsily G. (1981). Analyse stochastique des propriétés des systèmes poreux naturels hétérogènes. Rapport DGRST, ENSMP, Fontainebleu, 115 p.
- Benbouali E. (1999). Variabilité spatiale de la salinité des sols du périmètre de Hmadna (Bas-Chélif). Mém., Mag., INA d'Alger, 102 p.
- Benbouali E., Douaoui A., Bettahar N. & Berrokia A. (2003). Détermination des niveaux de dégradation des sols du périmètre de Hmadna (Bas Cheliff) par application de l'approche géostatistique multiparamétrique. Ann. Rec. Agro. n° 12, INRA, Alger, 25-35.
- Benziane A. (1983). Problèmes géologiques et hydrogéologiques de la Sebkhia de Benziane (Nord de l'Algérie). Thèse Doc. 3^{ème} cycle, Uni. Claude Bernard - Lyon I, 218 p.
- Beucher-Darricau H. (1981). Apport géostatistique du passage des données de terrain aux paramètres des modèles en hydrogéologie. Thèse de Doc. Ing., E.N.S.M.P., Fontainebleu, 118 p.
- Boiffin J. (1984). La dégradation structurale des couches superficielles du sol sous l'action des pluies. Thèse de Doc.-Ing. INA PG, 320 p.
- Boulaine J. (1957). Etude des sols des plaines du Chélif. Thèse d'état de l'Université d'Alger, 582 p.
- Bourennane H. (1997). Etude des lois de distribution spatiale des sols de Petite Beauce. Application à la cartographie d'un horizon par couplage de méthodes morphométriques et géostatistiques. Thèse de l'Université d'Orléans, 217 p.
- Bourennane H., King D., Chéry P. & Bruand A. (1996). Improving the kriging of a soil variable using slope gradient as external drift. European Journal of Soil Science, 47, 473-483.
- Burrough P.A., Beckett P.H.T. & Jarvis M.G. (1971). The relation between cost and utility in soil Survey. I à III. J. Soil Sci. 22(3), 554-571.
- Burrough P.A. & Webster R. (1976). Improving a reconnaissance soil classification by multivariate methods. J. Soil Sci., 27, 554-571.
- Burgess T.M., Webster R. & Mc Bratney A. B. (1981). Optimal interpolation and isarithmic mapping of soil properties. IV. Sampling strategy. J. Soil Sci., 32, 643-659.

-
- Burrough P. A. (1983).-a- Multiscale sources of spatial variation in soil. I The application of fractal concepts to nested levels of soil variation in soil Survey. *J. Soil Sci.*, 34, 577-597 et 599- 620.
- Burrough P. A. (1983).-b- Multiscale sources of spatial variation in soil. II A non Brownian fractal model and its application in Soil Survey. *J. Soil Sci.*, 34, 577-597 et 599-620.
- Burrough P. A. (1983).-c-. Problems of superimposed effects in statistical study of spatial variation in soil. *Agric. Water Manage.*, 6 (2/3), 123-143.
- Cailliez R. & Pages J.P. (1976). Introduction a l'analyse des données. S.M.A.S.H., 9 rue Duban, 75016 Paris.
- Campbell J. B. (1977). Variation of selected properties accross a soil boundary. *Soil Sci. Soc. Am. J.*, 41(3), 578-582.
- Campbell J. B. (1978). Spatial variation of sand content and pH within single contiguous delineation of two soil mapping units. *Soil Sci. Soc. Amer. J.* 42, 460-464.
- Cazemier D. (1999). Utilisation de l'information incertaine dérivée d'une base de données-sols. Application à la cartographie des propriétés hydriques à l'échelle régionale. Thèse Doc. ENSA Montpellier. 165 p. 12 annexes.
- Chenu C. (1989). Influence of fungal polysaccharide scleroglucan, on clay microstructures. *Soil Biology and Biochemistry*, 21, 299-305.
- Chéry P. (1995). Variabilité de l'épaisseur de la couverture de sol. Conséquences pour le bilan hydrique hivernal d'un secteur de Petite Beauce. Thèse de l'Université de Nancy I.
- Chrisman N. R. (1982). A theory of cartographic error ant its measurement in digital bases. *AUTO-CARTO 5*, International symposium, 159-168.
- Churchman G. J. & Tate K. R. (1987). Stability of aggregates of different size grades in allophonic soil from volcanic ash in Zealand. *J. of Soil Sci.*, 38, 19-27.
- Chrisman N.R.& Lester G. M. (1991). A diagnostic test for error in categorical maps. *AUTO- CARTO 10*, Baltimor, vol. 6, pp 330-348.
- Cipra J. E., Bidwell O. W., Whitney D. A. & Feyerherm A. M. (1972). Variation with distance in selected fertility measurements of Western Kansas Ustoll. *Soil Sci. Amer. J.*, Vol 36, 111-118.
- CPCS (1967). Classification des sols : commission de pédologie et de cartographie des sols. *Doc. Multi.*, INRA Grignon, 96 p.
- Creutin J.G. & Obled C. (1982). Objective analyses and mapping techniques for rainfall fields : an objective comparison. *Water Ressources Research*, 18, 413-431.
- Cressie N. (1984). Towards resistant geosyayistics. In : G. Vrely et al. (Editors), *Geostatistics for National Ressources Characterization*, I. Proc. NATO advanced study Institue, NATO ASI Series : C. Mathematical and physical Science. Reidel, Dordrecht, 21-44.
- Cressie N. (1991). *Statistics for Spatial Data*. John Wiley et Sons, New York, 900 p.
- Cressie N. & Hawkins D. M. (1980). Robust estimation of the variogram. I. *Match. Geol.*, 12, 115-125.
-

- Daoud Y. (1983). Contribution à l'étude de la dynamique des sels dans un sol irrigué du périmètre du Haut-Cheliff (Algérie). Thèse de Doc-Ing. ENSA Université de Rennes, 194 p.
- Daoud Y. (1999). Influence des conditions salines sur les propriétés physiques des sols des plaines du Chellif. Séminaires National sur la salinisation des terres Agricoles en Algérie, Chlef le 01 et 02 Juin 1999.
- Davis B. M. & Borgman L. (1982). A note on the asymptotic distribution of the sample variogram. *Math. Geol.*, 14: 643-653.
- Delhomme J.P. (1978). Kriging in the hydrosciences. *Advances in Water Resources*, 1, 251- 266.
- Derdour H. (1981). Contribution à l'étude de l'influence du taux de sodium échangeable sur le comportement des sols aux compactage. Thèse Mag. INA d'Alger, 46 p.
- Diday F. (1971). La méthode des nuées dynamiques. *Rev. Stat. Appl.*, 19 (2), 19-34.
- Diday F. (1972). Optimisation en classification automatique et reconnaissance des formes. *Rev. Fr. Inf. Rech. Opér.*, 6^{ème} année, 61-95.
- Douaoui A. (1993). Contribution à l'étude des sols des plaines du Chélif : Variabilité spatiale es quelques propriétés physico-chimiques des sols du moyen Chélif, indices cartographiques. *Mém. Mag.*, INA d'Alger, 123 p.
- Douaoui A., Benbouali E. & Gaouar D. (2001). Les facteurs intrinsèques de dégradation de la structure de surface des sols du Bas Chélif. *Ann. Agro. INA d'Alger. Vol. 22, n° 1 et 2*, 27-47.
- Douaoui A., Gascuel-Odoux C. & Walter C. (2004). Infiltrabilité et érodibilité de sols salinisés de la plaine du Bas-Cheliff (Algérie). *Mesures au laboratoire sous simulation de pluie. Etude et Gestion des Sols*, 11, 379-392.
- Douaoui A. (2005). Variabilité spatiale de la salinité et sa relation avec certaines caractéristiques des sols de la plaine du Bas-Cheliff. Thèse Doc. d'Etat Es Sciences Agronomiques. I.N.A. Alger, 261 p.
- Durand J. H. (1958). Les sols irrigables. Etude pédologique. SES., Alger, 198 p.
- Edmonds W.J. & Lentner M. (1986). Statistical evaluation of the taxonomie composition of three mapping units in virginia. *Soil Sci. Soc. Of Amer. J.*, Vol. 49(2), 394-401.
- Etude de l'avant-projet détaillé des extensions de Guerouaou et de la Sebket Benziane et du réaménagement du Bas-Chélif.
- Rapport I- B. Analyse des ressources en sol. Volume 2 : Cartes. Juillet 1990. Sir M. McDonald et Partners Ltd en association avec W.S. Atkins et BNEDER.
- Feeny V. (1988). Méthode d'analyse du contenant des unités cartographiques. DEA fédérale de pédologie, INPG, 57 p.
- Finke P.A. & Stein A. (1994). Application of disjunctive cokriging to optimize fertilizer additions on a field scale. *Geoderma*, 247-263, this issue.
- Freeze H.A. (1975). A stochastic-conceptual of one dimensional groundwater flow in on uniform homogeneous media. *Water. Res.*, 11, 725-741.
- Freycon V. # Sebastien B. (1991). La théorie des variables régionalisées. Répartitions

- spatiales, méthodologie et traitement statistique. Séminaire Montpellier. 9-11. Septembre 1991.
- Gajem Y.M., Warrich A. W. & Myers D. E. (1981). Spatial dependence of physical properties of a typic torrifluent soil. *Soil Sci. Soc. Am. J.*, Vol 45, 709-715.
- Gallali T. (1980). Transfert sels-matière organique en zone arides méditerranéennes : Thèses Univ. Nancy I, 202 p.
- Gascuel-Oudou O.C. (1984). Application de la géostatistique à l'étude de la variabilité spatiale des propriétés hydriques du sol. Thèse de Doct. Ing, ENSM Fontainebleu et ENSA-INR Rennes, 235 p.
- Gascuel-Oudou O. C. (1987). Variabilité spatiale des propriétés hydriques, méthodes et résultats, cas d'une seule variable : revue bibliographique. *Agronomie*, 7 (1), 61-71.
- Girard M.C. (1983). Recherche d'une modélisation en vue d'une représentation spatiale de la couverture pédologique. Thèse Doct. d'état- Sols. 430 p.
- Girard M.C. & King D. (1988). Un algorithme interactif pour la classification des horizons de la couverture pédologique. *Science du sol*. 26 (2), 81-103.
- Goulard M., Voltz M. & Monestiez P. (1987). Comparaison d'approches multivariées pour l'étude de la variabilité spatiale des sols. *Agronomie*, 7 (9), 657-665.
- Guckert A., Chone T. & Jacquin F. (1975). Microflore et stabilité structurale des sols. *Rev. Ecol. Biol. Sol*, 12, 211-223.
- Haddadj D. (1970). Influence des méthodes culturales sur la dégradation structurale des sols du périmètre du Haut Chellif. Thèse ing. INA d'Alger, 86 p.
- Hajrasulika S., Baniabbassi N., Metthey J. & Nulson D. R. (1980). Spatial variability of soil sampling for salinity studies in Southwest Iran. *Irri. Sci*, Vol 1, 197-208.
- Halliwell D. J., Barlow K.M. & Nash D.M. (2001). A review of the effect of wastewater sodium on soil physical properties and their implications for irrigation systems. *Aust. J. of Soil Res.*, 1259-1267.
- Harris R. F. & al (1966). Dynamics of soil aggregation in : N. C.Brady (Ed), *Advances in Agronomy*. Academic Press, New York, NY 18, 107-169.
- Haynes R. J. (1993). Effect of sample pretreatment on aggregate stability measured by wet sieving or turbidimetry on soils of different cropping history. *J. Soil Sci*, 44, 261-270.
- Henderson R. & Ragg., J.M. (1980). A reappraisal of soil mapping in an are of southern Scotland. Part.II The use fulness of some morphological properties and of a discriminant analysis in distinguishing between. The dominant taxa of four mapping units. *J. Soil Sci.*, 31, 573-580.
- Hénin S. # Monnier G. (1956). Evaluation de la stabilité de la structure du sol. In : VI congré international DCI. *Sol AISS Paris*, Vol, B, 49-52.
- Hénin S., Monnier G. & Combeau A. (1958). Méthode pour l'étude de la stabilité structurale des sols. *Ann. Agro.* n° 9, 73-92.
- Hénin S., Gras R. & Monnier G. (1969). Le profil culturale du sol. L'état physique du sol et ses conséquences agronomiques. Masson et Cie, 2^{ème} édition. Masson , Paris,

332 p.

- Hohl R. (1997). L'interpolation spatiale, guide d'utilisation du logiciel Surfer® version 8, Ins. de Géog., Uni. de Lausanne, non publié.
- Hotteling H. (1933). Analysis of a complex of statistical variables into principal components. *J. of Educational Psychology* 24, 417-441, 495-520.
- Iris J. M. (1986). Analyse et interprétation de la variabilité spatiale de la densité apparente dans trois matériaux ferrallitiques. *Science du sol*, Vol. 24, 245-256.
- Issaks E. H. & Srivastava R. M. (1989). *Applied Geostatistics*. Oxford University Press, New York, 561 p.
- Jamagne M. (1967). Bases et techniques d'une cartographie des sols. *Ann. Agro.* 18 (Noh.S), 142 p.
- Jouany C., Chenu C. & Chassin D. (1992). Détermination de la mouillabilité des constituants du sol à partir de mesures d'angles de contacts : revue bibliographique ; *Science du sol*, 30, 33-47.
- Journel A.G. & Huigbregts C. J.(1978). *Mining geostatistics*. Academic Press. New York 600 p.
- Keren R. (1989). Water drop kinetic energy effect on water infiltration in calcium and magnesium soils. *Soil Sci. Soc. Amer. J.* 53: 1624-1628.
- King D. & Duval O. (1989). Traitement informatique des données pédologiques ". La Bibliothèque Logos". *Rev. Sci. Sol*, 27 (1), 77-80.
- Krige D.G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. of the Chem., Metallurgical and Mining Soc. of South Africa*, 52, 119-139.
- Lagacherie P., Andrieux P. & Bouzigues R. (1994). Fuzziness and uncertainty of soil boundaries: from reality to coding in GIS. In "Spatial Conceptual Models for Geographic Objects with undertermined Boundaries" (P.A. Burrough and A. Frank, eds.), 275-287. Taylor & Francis, Baden Austria.
- Laslett G.M., Mc Bratney A.B., Pahl P.J. & Hutchinson M.F. (1987). Comparaison of several Spatial prediction methods for soil pH. *J. Soil Sci.*, 38, 325-341.
- Le Bissonnais Y. (1988). Analyse des mécanismes de dégradation et de la mobilisation des particules de terre sous l'action des pluies. Thèse doct. de Science de la terre. Univ. Orléans, 196 p.
- Le Bissonnais Y. & Le Souder C. (1995). Mesure de la stabilité structurale des sols pour évaluer leur sensibilité à la battance et à l'érosion. *Etude et Gestion des Sols*, 2, 1, 43-56.
- Le Bissonnais Y. & Arrouays D. (1997). Aggregate stability and assessment of soil crustability and erodibility : application to humic loamy soils with various organic carbon contents. In *Eur. J. of Soil Sci.*, march 1997, 39-48.
- Leenaers H., Okx J.P. & Burrough P.A. (1990). Employing elevation data for efficient mapping of soil pollution on floodplains. *Soil Use and Management*, 6, 105-114.

-
- Legros J.P. (1978): Recherche et contrôle numérique de la précision en cartographie pédologique. -1- Précision dans la délimitation des sols. Ann. Agro. 6, 499-519.
-2- Précision dans la délimitation des unités. Ann. Agro. 6, 583-601.
- Legros J. P. (1996). Cartographie des sols : de l'analyse spatiale à la gestion des territoires. Presses polytechniques et universitaires romandes, Lausanne.
(Gérer l'environnement) 10 p.
- Le Souder C. (1990). Effet d'un conditionneur minéral sur la formation des croûtes superficielles du sol sous l'action des pluies. Mode d'action du conditionneur sur la Stabilité. Thèse la stabilité. Thèse doct. INA PG, 181 p.
- Luxmoore R. J. & Shama M.L. (1980). Runoff responses to soil heterogeneity : experimental and simulation comparison for two contrasting watersheds. Water Ressource. Res. Vol. 16 (4), 675-684.
- Mallouhi N. (1989). Etude de la dynamique de la matière organique dans les sols affectés par la salinité. Tropiculture, 7, 3, 90, 90-97.
- Marsman B. A. & Gruijter De J.J. (1986). Quality of soil maps. A comparaison of soil survey methods in a sandy area. Soil Survey Papers n^o 15, Stiboka, Wageningen.
- Martin D. (1976). Analyse multivariable de données analytiques de sols ferrelitiques du Centre Cameroun. Cah. ORSTOM, Ser. Pédo., XIV (1), 3-11.
- Matheron G. (1965). Les variables régionalisées et leur estimation. Masson, Paris 305 p.
- Mbagwu J.S.C. & Bazzoffi P.(1998). Soil characteristics related to resiste of breakdown of dry soil aggregates by water-drops. Soil et tillage research 45, 133-145.
- Mc Bratney A. B. & Webster R. (1981). -a- The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. Computers et Geosciences. Vol. 7 (4), 335-365.
- Mc Bratney A.B., Webster R., Mc Laren R.G. & Spiers R.B. (1982). Regional variation of Extractible copper and cobalt in the topsoil of South-East Scot. Agro., 2, 969-982.
- Mc Cornack D.E. & Wilding L.P. (1969). Variation of soil properties within mapping units of soil with contrasting substrata in N.W. Ohio. Pros. Soil Sci. Soc. Am. 33, 587-593.
- Meirvene M. V. # Hofman G. (1989). Spatial variability of soil mitrate-nitrogen after potatoes and ist change during winter. Plant and soil 120 : 103-110.
- Meylan P. (1986). Régionalisation des données entachées d'erreurs de mesure par krigeage. Application à la pluviométrie. Hydro. Continent, V. I, n^o 1, 25-34.
- Monnier G. (1965). Action des matières organiques sur la stabilité structurale des sols. Thèse doct. , Paris , 140 p.
- Monnier G. # Stengel P. (1982). Structure et état physique du sol, Ed. Techniques, Paris (Coll.Théchniques Agricoles, Fasc. 1140 et 1141).
- Myers D. E. (1985). Some aspects of robustness. Sci. Terre, 24 : 63-79.
- Myers D. E. (1991). On variogram estimation. In : E Dudwicz et al. (Editors), the Frontiers of statistic Scientific Theory and Industrial Application. Proc. Of OCOSCO- I, Vol.II. American Sciences Press, 261-281.
-

- Norris J. M. & Loveday J. (1971). The application of multivariate analysis to soil studies. II-The allocation of soil profiles to established groups : a comparison of soil survey and computer methods. *J. Soil Sci.*, 22 (4), 395-400.
- Nortcliff S. (1978). Soil variability and reconnaissance soil mapping. A statistical study in Norfolk. *J. Soil Sci.*, 29, 403-418.
- Odeh I.O.A., Mc Bratney A.B. & Chittleborough D.J. (1994). Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*, 63, 197-214.
- Olivier M.A. (1984). Soil variation in the myre forest : its elucidation and measurement. P.H.D Thèses, University of Birmingham, 230 p.
- Pannatier Y. (1994). Variowin 2.1. Geostatistical Software Programms. Pannatier Institute of Mineralogy and Petrography. Uni. of Lausanne. Springer. Switzerland.
- Pannatier Y. (1996). Variowin, Software for Spatial Data Annalysis in 2D, Springer-Verlag, New York.
- Philippeau G. (1986). Comment interpréter les résultats d'une analyse en composantes principales?. *Service des études statistiques. STAT-ITCF.*, 61 p.
- Ragg J.M. & Henderson R. (1980). A reappraisal of soil mapping in an area of southern Scotland . Part I: The reliability of four soil mapping units morphological variability of their dominant taxa. *J. Soil Sci.*, 31(3), 559-572.
- Rambal S. (1978). Analyse en composantes principales des variables du stock d'eau du sol. *Bull. Groupe Français d'Humidimétrie neutronique*, 3, 74-81.
- Rayner J. H. (1966). Classification of soils by numerical methods. *J. Soil Sci.*, Vol. 17, 79-82.
- Richards LA (Ed.) (1954). Reclamation of saline and alkali soils. US Salinity Laboratory, Handbook n° 60, USDA, Washington, 176 p.
- Roux M. (1985). Algorithmes de classification. Edition Masson, Parie, 151 p.
- Ruellan A. (1989). Découvrir le sol. CNEARC., Montpellier, 45 p
- Ruellan A., Dosso M. & Fritsch E. (1989). Etude des lois de distribution spatiale des sols de petite Beauce, thèse de doct. De l'Université d'Orléans, 229 p.
- Saidi D. (1992). Etude expérimentale de la stabilité structurale des sols des plaines du Cheliff, thèse de Mag., INA, Alger, 138 p.
- Saidi D., Douaoui A., Le Bissonnais Y. & Walter Ch. (1999). Sensibilité des couches de surface de la plaines du Cheliff (Algérie) à la dégradation structurale. *Etude et Gestion des Sols*, 6, 1, 15-25.
- S. N. REPAL. (1952). Le bassin néogène du Chéllif. Monographies régionales, 1^{er} Série : Algérie, n° 16. XIX^{ème} congrès géologique international, juin 1952, 55 p.
- Thomas P. J., Baker J.C. & Simpson T.W. (1989). Variability of the Cecil Map unit in Appomattox Country, Virginia. *Soil Sci. Amer. J.*, Vol. 53, 1470-1474.
- Thorntwaite C.W. (1948). An approach toward a rational classification climates. *Geogr. Rew.* 38 (1), 55-94.
- Tisdall J.M. & Oades J.M. (1982). Organic matter and water-stable aggregates in soil. *J.*

-
- Soil. Sci., 33, pp 141-163.
- Tomassone R. (1976). Présentation générale des méthodes d'analyse multidimensionnelle. Ann. de l'INA, Vol. VI spécial, 61-99.
- Tomassone R. (1988). Comment interpréter les résultats d'une analyse factorielle discriminante?. INA-PG., Math. Biométrie, STAT-ITCF., 52 p.
- Trangmar B.B., Yort R. S. & Uchara G. (1985). Application of geostatistics of spatial studies of soil properties. Advances in Agronomy, Vol. 38, 45-93.
- Van Kuilenburg J., De Griujter J.J., Marsman B. A. & Bouma J. (1982). Accuracy of interpolation between point data on soil moisture capacity, compared with estimates from mapping units. Geoderma, 27, 311-325.
- Vauclin M. (1982). Méthodes d'étude de la variabilité spatiale des propriétés d'un sol. ED. INRA, Publ. : les colloques de l'INRA, n° 15 , 9-43.
- Veregin H. (1989). Error modelling for the map overlay operation. In accuracy of spatial databases. Taylor et francis, 3-18.
- Vink A.P.A. (1963). Planning of soil surveys in land Development. Publication 10, International Institute for Land Reclamation and Improvement, Wageningen.
- Voltz M. (1986). Variabilité spatiale des propriétés physiques en milieu aluvial. Essai de cartographie quantitative des paramètres hydrodynamiques. Thèse doct-ing, ENSAM, 198 p.
- Voltz M., Bornand M. & Goulard M. (1989). Estimation spatiale des paramètres hydrodynamiques du sol à l'échelle d'une plaine alluviale. In Influence à l'échelle régionale des couvertures pédologiques sur les bilans hydriques et minéraux, G. Vachaud (ed.), Piren, Paris, 81-1.
- Voltz M. & Webster R. (1990). A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J. Soil Sci., 41, 473- 490.
- Voltz M. & Goulard M. (1994). Spatial interpolation of soil moisture retention curves. Geoderma, 62, 109-123.
- Wagenet R. J. & Jurinak J. J. (1978). Spatial variability of soluble salt content in a mancos shale watershed. Soil Sci. Vol 116, n° 6, 342-349.
- Walter C. (1990). Estimation des propriétés du sol et quantification de leur variabilité à moyenne échelle. Université de Paris VI n° SDS 436, 172 p.
- Walter C. & Arousseau P. (1992). Un élément de la précision des cartes pédologiques : la variabilité des propriétés du sol au sein des unités cartographiques. in gestion de l'espace rural et système d'information géographique (Buche, king et Lardon, Eds), INRA Edition, 277-293.
- Wang C. (1982). Variability of soil proprieties in relation to size of map unit delineation. Can J. Soil Sci., 62, 657-662.
- Ward J.H. (1963). Hierarchical grouping to optimize an objective function. Journal of the Amer. Stat. Ass., 58, 236.
- Warrich A. W. & Nielson D.R. (1980). Spatial variability of soil properties in the fields. In
-

- Hillel D. (ed), Application of soil physics. Academic Press, Inc. New York, 319-344.
- Webster R. (1973). Automatic soil boundary location from transect data. *Math. Geol.* 5: 27-37.
- Webster R. (1977). Quantification and numerical methods in soil classification and survey, *Monographs on Soil Survey*, Clarendon Press, Oxford, 269 p.
- Webster R. (1985). Quantitative spatial analysis of soil in the field. In: B.A. Stewart (Editor), *Advances in Soil Science*, 3. Springer, New York, 1-70.
- Webster C. (1990). Estimation de propriétés du sol et quantification de leur variabilité à moyenne échelle. *Cartographie pédologique et géostatistique dans le Sud de l'Ille et Vilaine*. Thèse de l'Université de Paris VI.
- Webster R. & Burrough P.A. (1972). -a- Computer-based soil mapping of small areas sample data. I Multivariate classification and ordination. *J. Soil Sci.*, 23 (2), 210-221.
- Webster R. & Burrough P.A. (1972). -b- Computer-based soil mapping of small areas sample data. II Classification smoothing. *J. Soil Sci.*, 23 (2), 222-234.
- Webster R. & Burrough P.A. (1974). Multiple discriminant analysis in soil survey. *J. Soil Sci.*, 25 (1), 120-134.
- Webster R. & Burgess T.M. (1983). Spatial variation in soil land the role of kriging. *Agronomy and Water Management*, 6, 111-122.
- Wilcoxon F. (1945). Individual comparison by ranking methods. *Biometrics* 1, 80-83.
- Wilding L.P., Jones H. B. & Schafer G. M. (1964). Morley and Blont Soils : statistical summary of certain physical and chemical properties of some selected profiles from Ohio. *Proc. Soil Sci. Soc. Am.* 28, 674-9.
- Wilding L.P., Jones H. B. & Schafer G. M. (1965). Variation of soil morphological properties within Miami, Celina and Gosby mapping units in West-Central Ohio. *Soil Sci. Soc. Amer. Proc.* Vol. 29, 711-717.
- Williot B. (1995). Variabilité spatiale et risques d'erreurs dans l'analyse des horizons holorganiques forestiers. *Etude et Gestion des Sols*, 2, 1, 73-84.
- Wopereis M.C., Gascuel O.C., Bourrie G. & Soignet G. (1982). Spatial variability of heavy metals in soil on a one hectare scale. *J. Soil Sci.*, 146, 113-118.