

**Rapport de stage pour l'obtention
De la 1^{ère} année de Master**

**Méthodes de spatialisation de la résistivité
électrique des sols à l'échelles de micro-
parcelles d'essais variétés**



**Laura Gay
Mai – Août 2017**

Maîtres de stage : Hocine Bourennane¹ et Marie-Noël Mistou²

¹INRA Orléans – Unité de recherche de Sciences du Sol

²UMR Agronomie Grignon

Remerciements

Je tiens tout d'abord à remercier Catherine Hénault, Directrice de l'Unité de Recherche Science du sol de l'INRA d'Orléans pour m'avoir permis de réaliser mon stage au sein de cette unité.

Je remercie particulièrement Hocine Bourennane, pour ses conseils et pour le temps qu'il a consacré pour mon encadrement et aux questions que j'ai pu lui poser. Je le remercie également pour les connaissances en géostatistique que j'ai pu acquérir à son contact.

Je remercie ensuite Marie-Noël Mistou, pour ses conseils et le temps qu'elle a mis à disposition pour répondre à mes questions.

Je tiens également à remercier toutes les personnes qui ont contribué de près ou de loin à mon intégration et au bon déroulement de ce stage : Thomas, Mehdi, Margaux, Aurélie, Manon, Amélia, Marie, Sarah, Sébastien, Charlène, Nicolas, Laëtitia, Romain, Xavier...

Et enfin, un grand Merci à l'ensemble de l'équipe UR Sols et InfoSols pour la bonne ambiance de travail, entrecoupée de pauses, de pots et de sorties, dans laquelle j'ai pu effectuer mon stage de manière agréable à Orléans et au sein de l'INRA particulièrement.

Table des matières

Résumé / Abstract	3
Présentation de l'organisme d'accueil	5
I. Introduction	6
II. Site d'étude et campagne de mesure	7
II.1. L'unité expérimentale d'Epoisses	7
II.2. Mesure de la résistivité électrique du sol.....	8
III. Traitement des données : méthodes	9
III.1. Traitement préliminaire des données	9
III.2. Approche développée pour la spatialisation de la résistivité électrique	10
III.2.1. Analyse variographique	10
III.2.2. Krigeage ordinaire	11
III.2.3. Indicatrice de résistivité électrique des micro-parcelles.....	11
III.2.4. Analyse factorielle discriminante (AFD)	12
III.2.5. Analyse de la variance (ANOVA).....	12
IV. Résultats et discussions	13
IV.1. Spatialisation de la résistivité électrique au niveau des micro-parcelles	13
IV.2. Typologie des micro-parcelles par l'indicatrice de la résistivité	15
IV.3. Validation de la typologie des micro-parcelles.....	17
IV.3.1. Résultats de l'Analyse Factorielle Discriminante (AFD).....	17
IV.3.2. Résultats de l'analyse de la variance (ANOVA)	20
IV.3.3. Comparaisons des différents scénarii	20
V. Conclusion	23
Bibliographie	25
Liste des figures	26
Liste des tableaux	26
Liste des annexes	26
Annexes	a

Résumé

La résistivité électrique des sols a été choisie pour décrire l'hétérogénéité spatiale des sols des micro-parcelles d'essais variétés. Les mesures de résistivité ont été enregistrées à l'automne 2011 pour trois profondeurs (0.5m, 1m et environ 2m) appelées voies de mesures et on observe un nombre de mesures par micro-parcelle faible. Aussi, la résistivité a été spatialisée pour chaque voie de mesure par krigeage ordinaire (KO). Pour caractériser l'hétérogénéité spatiale des sols des micro-parcelles par rapport à la résistivité, les valeurs estimées par KO ont été classées en fonction de leur position par rapport à la moyenne (scénario 1) et à la médiane (scénario 2). Cette discrétisation a conduit à définir deux types de zones (zones avec la résistivité inférieure à la moyenne et inversement) pour chaque paramètre de position. La typologie établie a été validée par une analyse factorielle discriminante (AFD), qui a montré que la résistivité électrique des sols est une variable discriminante des deux types de zones établies. Quelle que soit la voie de mesure et quel que soit le scénario, la probabilité de classement erroné est inférieure à 5%. Par ailleurs, la moyenne des valeurs de la résistivité dans les deux types de zones sont statistiquement différentes (ANOVA : test de Tukey). Les mesures de résistivités des trois voies spatialisées par KO ont aussi été synthétisées par une analyse en composante principale (ACP). Les procédures de discrétisation (les deux scénarii) et de validation (AFD et ANOVA) appliquées à la première composante issue de l'ACP montrent des résultats similaires à ceux obtenus en considérant les voies de mesure de manière individuelle. Enfin, la comparaison des deux scénarii montre un taux de coïncidence compris entre 93.5 % et 96 % selon la voie et le scénario considérés. On peut donc indifféremment considérer la moyenne ou la médiane pour discrétiser la surface des micro-parcelles dans ce contexte d'étude.

Mots clés : Résistivité des sols ; Krigeage ordinaire ; Analyse factorielle discriminante ; Analyse de la variance ; Analyse en composante principale.

Abstract

The electrical resistivity of soils was used to describe the spatial heterogeneity of micro plots soils. The resistivity measurements were recorded at the autumn 2011 for three soil depths (0.5m, 1m and approximately 2m). The number of resistivity measurements by micro plot is low. The resistivity was thus mapped for each depth by ordinary kriging (OK). To describe the spatial heterogeneity of soils in micro-plots relative to resistivity, OK estimates were classified according to their position at the mean (scenario 1) and the median (scenario 2). This classification has led to the definition of two types of areas (resistivity lower to the mean and conversely) for each position parameter. The established typology was validated by a factorial discriminant analysis (FDA), which showed that resistivity is a discriminant variable of the two types of areas established. Regardless the depth of measurement and the scenario considered, the probability of misclassification is less than 5%. Moreover, the mean resistivity values in the two types of areas are statistically different (ANOVA: Tukey test). OK estimates were also summarized through a principal component analysis (PCA). The classification procedure (both scenarios) as well as that of validation (FDA and ANOVA) applied to the first component achieved by PCA show similar results to those obtained by considering separately the resistivity measurements of three depths. Finally, the comparison of the two scenarios shows a coincidence rate ranges from 93.5% to 96% as function of resistivity measurements depth and the scenario considered. Therefore, we can indifferently consider the mean or the median to discretize the surface of the micro plots in the context of this study.

Keywords: Soils resistivity; Ordinary kriging; Factorial discriminant analysis; Analysis of variance; Principal component analysis.

Présentation de l'organisme d'accueil

L'Institut National de la Recherche Agronomique (INRA), est le premier institut de recherche agronomique en Europe et le deuxième dans le monde. C'est un établissement public à caractère scientifique et technologique (EPST) qui dépend à la fois du Ministère de l'Agriculture et de l'agroalimentaire, ainsi que du Ministère de l'enseignement supérieur et de la recherche.

L'INRA a été fondé en 1946, à la sortie de la seconde guerre mondiale, avec comme objectif de rendre la France autosuffisante pour son alimentation. Pour cela, il doit réussir à allier science et technologie, afin de les mettre au service du développement de l'agriculture et de rendre cette dernière plus productive, spécialisée et régionalisée. Cet objectif fût atteint au début des années 1960 et les missions de l'INRA se sont peu à peu diversifiées (forêts et milieux aquatiques, agroalimentaire, biologie et biotechnologies, environnement, développement durable et climat). Aujourd'hui, l'INRA poursuit trois axes de recherche principaux : l'alimentation, l'agriculture et l'environnement. L'objectif est désormais de développer une agriculture à la fois durable, adaptée aux besoins nutritionnels de l'Homme et compétitive, dans le but de « nourrir durablement le monde ». L'INRA se regroupe ainsi en plusieurs centres régionaux dans l'ensemble de la France.

Le centre INRA Val de Loire est un centre régional implanté sur 3 sites : Orléans, Tours et Bourges et mène des recherches autour de 3 axes :

- Biologie animale intégrative, santé animale et publique et gestion durable des productions animales,
- Biologie intégrative des arbres et organismes associés, gestion durable des écosystèmes forestiers,
- Dynamique des sols et gestion de l'environnement.

L'unité de recherche Science du Sol (UR SOLS), dans laquelle j'effectue mon stage, est basée sur le site d'Orléans et appartient au pôle de recherches « Dynamique des sols et gestion de l'environnement », qui abrite également l'unité de service InfoSol. Les enjeux abordés par ce pôle sont particulièrement importants pour le développement durable, puisque les sols participent à fois à la production agricole et à la protection de l'environnement.

L'UR SOLS a pour mission principale de développer des travaux de recherche en Science du Sol, sur les interactions entre les propriétés physiques des sols et leurs fonctionnements hydriques et biogéochimiques, en prenant en compte la diversité des sols.

Elle s'appuie sur des compétences d'observations, de mesures spatiales in situ, d'expérimentation et de modélisation.

Ses travaux contribuent à une meilleure préservation des sols, à une meilleure gestion de la ressource en eau et à la protection de la qualité de l'air. Ils s'inscrivent dans la dynamique scientifique locale, nationale et internationale.

I. Introduction

Le sol est la partie meuble de la lithosphère, aussi appelé « épiderme des terres émergées ». C'est un écosystème situé à un carrefour entre l'atmosphère, l'hydrosphère, la biosphère et la lithosphère : c'est-à-dire un écotone. Ce dernier a de multiples fonctions, telles que la production agricole et forestière, en fournissant le support des cultures, ainsi que les nutriments et l'eau nécessaires aux plantes. C'est également un élément constitutif du paysage, un régulateur des échanges dans les écosystèmes et un système épurateur des eaux par la filtration/fixation, voire dégradation des éléments potentiellement polluants (éléments traces, pesticides). Une bonne connaissance du sol, de ses caractéristiques et de sa variabilité s'avère alors essentielle pour développer une agriculture respectueuse de l'environnement.

Dans ce contexte, la résistivité électrique des sols peut être une variable d'intérêt dans l'amélioration de la connaissance du sol. En effet, cette dernière peut être utilisée pour approximer la variabilité spatiale et temporelle de nombreuses propriétés physiques du sol, telles que la structure ou bien la teneur en eau, par exemple. De plus, cette méthode n'étant pas destructive et étant très sensible, elle offre un outil très intéressant pour décrire les propriétés de sub-surface, sans avoir à réaliser beaucoup de sondages pédologiques. Ainsi, les prospections électriques ont pour objectif de déterminer la distribution de la résistivité pour le volume de sol prospecté. Un courant électrique est ainsi généré artificiellement puis fourni au sol. Ce sont ensuite les différences de potentiels en résultant qui sont mesurées. Les modèles de différence de potentiel fournissent alors des informations sur la forme des hétérogénéités de sub-surface et sur leurs propriétés électriques (Samouëlian *et al.*, 2005).

L'INRA (UR SOLS et UMR Agronomie) est ainsi partenaire d'un projet CASDAR intitulé *Optirès*, coordonné par le GEVES (Groupe d'Etude et de contrôle des Variétés Et des Semences) et qui a pour but d'optimiser les réseaux d'essais variétés par des méthodes permettant, en particulier, de mieux contrôler l'hétérogénéité spatiale dans un essai. En effet, les réseaux d'évaluation des variétés sont parfois critiqués pour ne pas représenter suffisamment la diversité des situations de production dans lesquelles les variétés seront utilisées. Ils sont bien souvent réalisés dans des milieux homogènes garantissant un minimum de réussite de l'expérimentation. Par ailleurs, le risque d'hétérogénéité augmente avec la taille de l'essai, créant ainsi des difficultés lorsque l'on cherche à tester un grand nombre de variétés dans un même essai. Enfin, cette hétérogénéité est encore mal prise en compte par le dispositif statistique censé la contrôler, avec parfois d'importantes incertitudes, voire même l'absence complète ou partielle des résultats d'une variété dans un essai.

Dans ce contexte, le projet *Optirès* aura comme finalité de permettre, dans un premier temps, l'amélioration de la précision des essais et d'en valoriser un plus grand nombre, tout en maintenant la qualité des résultats finaux issus des regroupements. Dans un second temps, il aura pour but d'élargir le référencement des variétés à des situations agropédologiques plus nombreuses, plus variées et plus proches des conditions rencontrées par l'agriculteur.

La résistivité des sols a été choisie pour décrire l'hétérogénéité spatiale des sols des micro-parcelles. Mon stage a ainsi pour objectif de spatialiser cette résistivité électrique des sols à l'échelle de micro-parcelles d'essais variétés et de proposer une méthodologie pour caractériser l'hétérogénéité des sols grâce à la résistivité électrique. Ce rapport présente la méthodologie mise en place pour atteindre cet objectif.

II. Site d'étude et campagne de mesure

II.1. L'unité expérimentale d'Epoisses

Le site expérimental de l'INRA d'Epoisses est situé sur la commune de la Bretenière (Côte-d'Or, Bourgogne), à 10 km au sud-est de Dijon (Figure 1). Le domaine compte 120 hectares de terre, subdivisés en de nombreuses parcelles expérimentales qui accueillent différentes cultures (maïs, blé, orge, betterave...), dans le cadre de recherches en agroécologie dans des conditions réelles de cultures en plein champ.

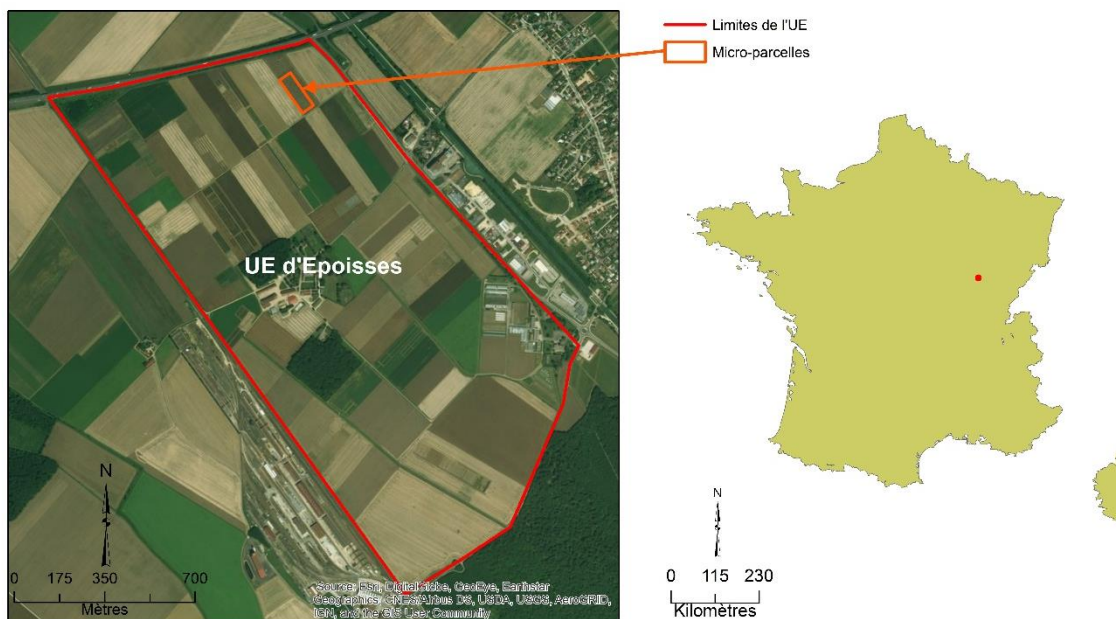


Figure 1 : Localisation du site d'étude, domaine expérimental de l'INRA d'Epoisses

Le climat qui caractérise cette région est un climat de type semi-continentale, avec une pluviométrie assez importante, de l'ordre de 760 mm par an (Météo France, station de Dijon). Le site d'étude se situe dans la vallée de l'Ouche, large plaine alluviale s'étalant sur une largeur de 3 à 8 km. On y retrouve des dépôts alluvionnaires anciens et récents provenant de l'écoulement de l'Ouche et de la Tille, de texture argilo-limoneuse (Chrétien, 1976 ; InfoTerre, 2017).

D'après la carte pédologique de Dijon au 1/100 000^{ème} (Chrétien, 1976), le site d'étude présente deux types de sols : des rendzines brunifiées de faibles épaisseurs (20 à 40 cm) et des sols bruns calciques plus épais (entre 50 et 100 cm).

Il a également été mis en évidence la présence d'une nappe souterraine à 0,5 voire 1,5 m de profondeur, s'écoulant dans la direction Nord-Ouest / Sud-Est.

II.2. Mesure de la résistivité électrique du sol

Une campagne de terrain a été menée au cours de l'automne 2011, dans le but de mesurer la résistivité électrique des sols sur l'unité expérimentale (UE) d'Epoisses. L'appareil utilisé pour réaliser ces mesures fut un ARP (Automatic Resistivity Profiler) et elles ont été géo-référencées grâce à un GPS différentiel.

La mesure de la résistivité électrique nécessite en général 4 électrodes : 2 électrodes généralement appelées A et B qui injectent le courant et 2 autres électrodes appelées M et N, qui enregistrent les différences de potentiels en résultant.

Le dispositif ARP (Figure 2) utilisé ici est un système multipôle tracté, qui a été développé par le centre de recherche géophysique de Garchy (CNRS et Paris VI). Il permet de mesurer la résistivité électrique du sol à l'échelle de la parcelle de façon rapide, grâce à un système de tractage par véhicule (Panissod et *al.*, 1997 ; Ndjiba Mitombo, 2013). Il implique 1 dipôle émetteur et 3 dipôles récepteurs en forme de roues dentées qui sont des sondes métalliques : le dipôle émetteur AB permet d'injecter le courant dans le sol et la série des 3 dipôles récepteurs (M1N1, M2N2 et M3N3) permet de mesurer la différence de potentiel électrique. L'espace entre la sonde émettrice et la sonde réceptrice est de 0.5 m pour le premier dipôle récepteur, de 1 m pour le deuxième et de 2 m pour le troisième. Pour chaque point de mesure, il y a donc 3 valeurs de résistivité apparente mesurées. Ces 3 valeurs correspondent à 3 profondeurs d'investigation différentes (respectivement 0.5 m, 1 m et environ 2 m), appelées « voies de mesure » (V1, V2 et V3). De plus, les mesures de résistivité sont intégratrices, c'est-à-dire que la mesure de la voie 3 intègre le volume prospecté par les voies 1 et 2. De même pour la voie 2, qui intègre le volume prospecté par la voie 1. L'intervalle moyen de mesure le long des profils est de 0.2 m et la distance séparant 2 profils est d'environ 6 m dans ce travail.

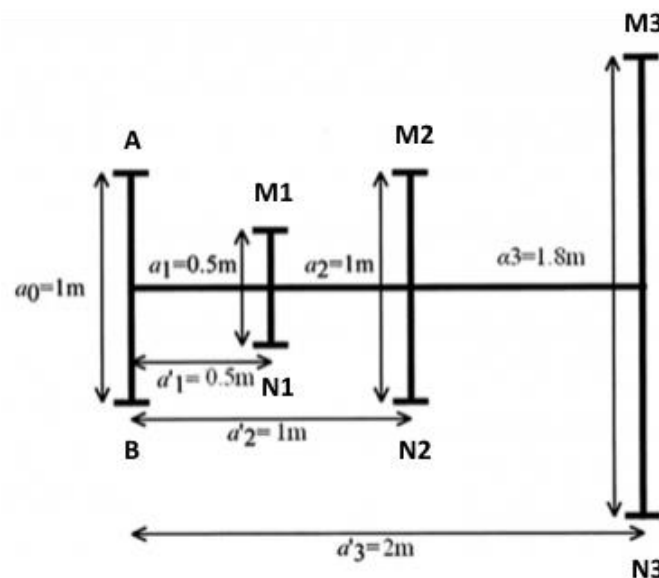


Figure 2 : Schéma du dispositif ARP (Source : Papadopoulos *et al.*, 2009)

III. Traitement des données : méthodes

III.1. Traitement préliminaire des données

Afin de bien comprendre et maîtriser les données de résistivité électrique acquises lors de la campagne de mesure de novembre 2011, une première analyse de ces données a été effectuée. A partir du fichier de forme de la zone d'étude des micro-parcelles, ainsi que du fichier texte réunissant les valeurs de résistivité mesurées pour l'UE du site d'Epoisses, les statistiques descriptives ont été calculées et plusieurs cartes ont été réalisées sur le logiciel ArcGis®.

Tout d'abord, une première carte (cf. Figure 1) montrant la position et la taille de la zone d'étude des micro-parcelles par rapport à l'étendue de l'unité expérimentale a été réalisée. Cette carte permet de s'apercevoir que les micro-parcelles représentent une partie relativement petite de l'UE d'Epoisses et qu'elles se trouvent exclusivement sur sa partie nord.

De ce fait et dans le but d'alléger les calculs, le fichier contenant les valeurs de résistivité électrique a été limité à la seule étendue de la zone d'étude des micro-parcelles. De plus, chaque valeur de résistivité électrique s'est vue liée à la micro-parcelle à laquelle elle appartient, ce qui permet donc de calculer les statistiques descriptives de chaque micro-parcelle (résultats non présentés dans ce rapport).

Enfin, ces différentes actions permettent de connaître le nombre de valeurs de résistivité électrique mesurées par micro-parcelle (Figure 3) et leur répartition au sein de cette dernière (Figure 4).

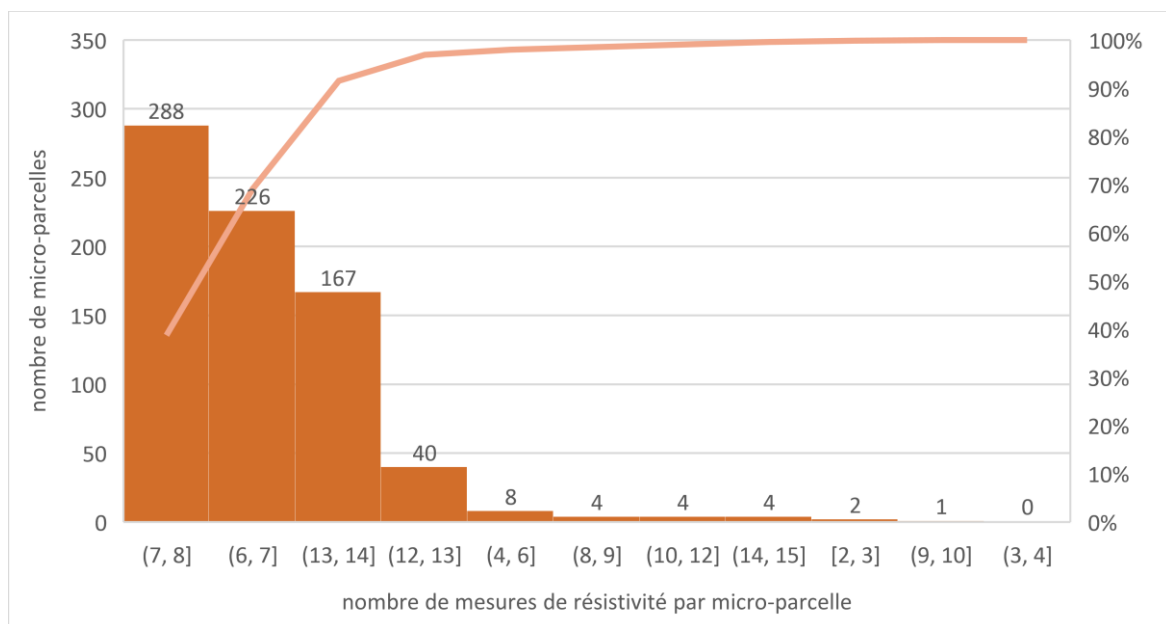


Figure 3 : Fréquence du nombre de mesures de résistivité électrique par micro-parcelle

On remarque alors qu'il y a très peu de mesures par micro-parcelles, avec environ 6 à 8 mesures pour 70% des micro-parcelles (Figure 3) et que ces dernières sont réparties de façon linéaire, souvent sur une seule trainée électrique et parfois deux pour les micro-parcelles ayant le plus de mesures (Figure 4 et Annexe 1).

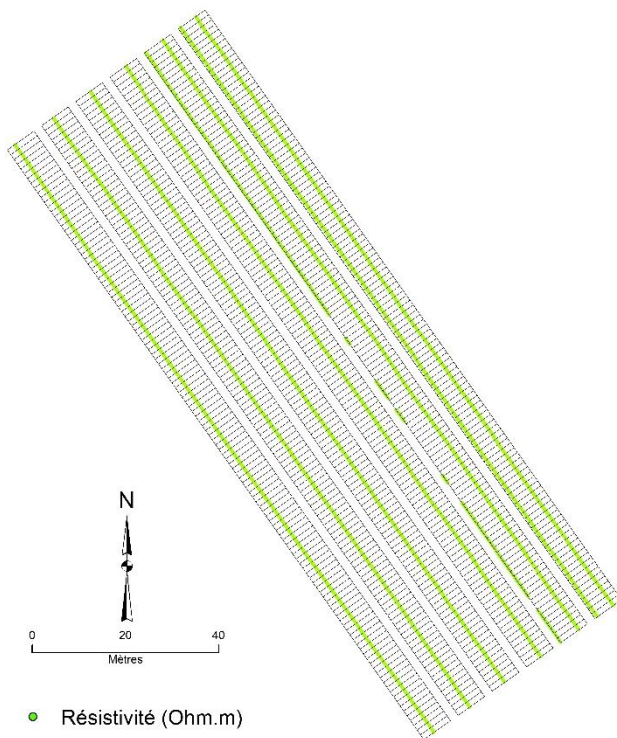


Figure 4 : Profils de résistivité électrique des micro-parcelles sur l'UE d'Epoisses

Suite à cette première analyse, il est apparu que le nombre de mesures de résistivité électrique par micro-parcelle est très peu élevé et que leur répartition n'est pas homogène sur l'ensemble de celle-ci. C'est pourquoi une spatialisation des données par krigeage ordinaire a ensuite été réalisée. Pour cela, une zone tampon de 20 mètres autour du périmètre des micro-parcelles a été créée (Annexe 2), permettant d'avoir un nombre de mesures de résistivité électrique assez important et de limiter les effets de bord. En effet, le nombre de mesures contenues dans les 744 micro-parcelles est de 6 603, alors qu'il s'élève désormais à 20 408 dans la zone tampon.

III.2. Approche développée pour la spatialisation de la résistivité électrique

Afin d'estimer de nouvelles valeurs de résistivité électrique, une analyse variographique, ainsi qu'une estimation par krigeage ordinaire ont été réalisées pour chaque voie de mesure sur le logiciel de géostatistique Isatis.

Cette méthode de géostatistique permet d'obtenir de nouvelles valeurs de résistivité grâce à des interpolations utilisant les données mesurées et ainsi des cartes de la variable interpolée.

III.2.1. Analyse variographique

Le variogramme est un outil central de la géostatistique (Oliver et Webster, 2014), qui permet de savoir si une variable est structurée dans l'espace géographique ou bien si sa distribution est aléatoire.

Le variogramme expérimental est une mesure de dissemblance entre n paires d'observations formées à partir de N points de mesures.

Il est défini par l'équation suivante :

$$\gamma^*(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{\alpha=1}^{n(\mathbf{h})} [Z(X_{\alpha} + \mathbf{h}) - Z(X_{\alpha})]^2 \quad (1)$$

Avec : $n(\mathbf{h})$ = nombre de paires d'observations reliées par un vecteur \mathbf{h} ,

Z = valeur prise par la variable aux points X_{α} .

Une fois le variogramme expérimental construit, l'étape suivante consiste à ajuster une courbe aux valeurs expérimentales : c'est le variogramme théorique. Cette courbe doit décrire

les principales caractéristiques du variogramme expérimental. Pour cela, plusieurs modèles sont disponibles, dont les plus utilisés sont les modèles sphérique et pépitique (Oliver et Webster, 2014). L'effet de pépité correspond à la variation non capturée par l'échantillonnage (mesures).

III.2.2. Krigeage ordinaire

Le krigeage est une technique géostatistique permettant, à partir de données dispersées, d'estimer de nouvelles valeurs. Il a l'avantage de tenir compte, contrairement aux méthodes statistiques classiques, des distances entre les données, des distances entre les données et la cible (le point pour lequel on veut estimer la mesure) et de la structure spatiale des données. Le krigeage ordinaire (KO) est une des méthodes d'interpolation la plus couramment utilisée pour interpoler de nouvelles valeurs à partir d'un semi de points.

III.2.3. Indicatrice de résistivité électrique des micro-parcelles

Le but ici est d'attribuer à chaque mesure de résistivité électrique spatialisée, l'indicatrice 0 ou 1 en fonction : (1) de sa position par rapport à la moyenne de la résistivité de la voie de mesure à laquelle elle appartient, et (2) de sa position par rapport à la médiane. Pour cela, chaque point de mesure de résistivité est comparé à la valeur moyenne puis à la valeur médiane de la résistivité électrique, comme présenté par la Figure 5. Si la valeur de résistivité du point de mesure est inférieure à la moyenne ou à la médiane, on lui attribue l'indicatrice 0 et si cette valeur est supérieure ou égale à la moyenne ou à la médiane, on lui attribue l'indicatrice 1.

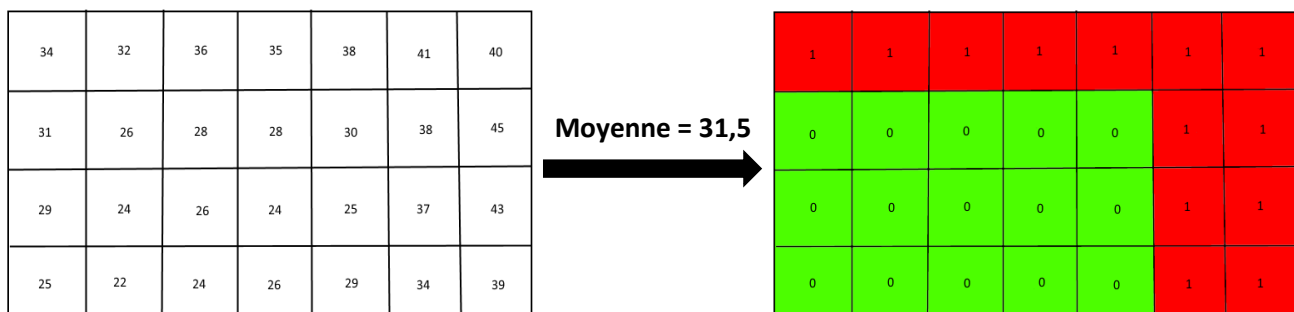


Figure 5 : Schéma explicatif de la méthode appliquée pour attribuer l'indicatrice de résistivité (exemple par rapport à la moyenne)

Le même protocole est appliqué au facteur F1 issu de l'Analyse en Composante Principale (ACP), conduite sur les trois voies de mesures de la résistivité électrique. En effet, l'ACP est une méthode permettant de projeter les observations depuis l'espace à p dimensions des p variables, vers un espace à k dimensions (où $k < p$), afin de conserver un maximum d'informations sur les 2 ou 3 premières dimensions, pour pouvoir les représenter plus facilement. Pour cela, le pourcentage de la variabilité totale associé aux principales variables doit être suffisant. Dans notre cas, il y a trois variables, qui sont les trois voies de mesure de la résistivité électrique et le facteur F1 représente 95% de la variabilité totale des trois voies de mesure de la résistivité. C'est pourquoi ce facteur sera également testé pour spatialiser la résistivité globale des micro-parcelles.

Ce traitement permet ainsi de classer les valeurs de résistivité en deux catégories distinctes. Il importera ensuite de valider ou non ce classement grâce, notamment, à une analyse factorielle discriminante (AFD) et par une analyse de la variance (ANOVA).

III.2.4. Analyse factorielle discriminante (AFD)

L'analyse factorielle discriminante est une méthode statistique explicative et prédictive, dont l'objectif est d'étudier la relation entre une variable qualitative et un ensemble de variables quantitatives (ici une seule : la résistivité électrique). Trois objectifs principaux peuvent être attribués à l'analyse discriminante :

1. déterminer les variables les plus discriminantes au regard de catégories spécifiques,
2. déterminer la catégorie d'un individu sur la base de ses caractéristiques,
3. valider une classification ou bien choisir entre différentes classifications pour déterminer laquelle est la plus pertinente.

L'analyse discriminante vient à posteriori de la classification et peut être considérée comme une extension du problème de régression où la variable est qualitative. Les données se composent de n observations divisées en k classes ou catégories et sont décrites par p variables. Traditionnellement, on peut distinguer 2 aspects de l'analyse discriminante :

1. un aspect **descriptif** qui consiste à trouver des combinaisons linéaires de variables qui séparent de la meilleure façon les catégories k et donnent une représentation graphique qui reflète bien cette séparation,
2. un aspect **décisionnel** où apparaît un nouvel individu et pour lequel nous connaissons les valeurs des prédicteurs. Il faut alors décider à quelle catégorie il devrait être affecté (problème de classification).

Deux modèles d'AFD sont possibles en fonction de l'hypothèse fondamentale de départ. Si on suppose que les matrices de covariance sont identiques, on est dans le cas d'une analyse discriminante factorielle linéaire. A l'inverse, si on suppose que les matrices de covariance sont différentes pour au moins deux catégories, on est alors dans le cas d'un modèle quadratique. C'est le test de Box qui permet de vérifier cette hypothèse (Tabachnick et Fidell, 1996 ; Bourennane *et al.*, 2014).

III.2.5. Analyse de la variance (ANOVA)

L'ANOVA, ou analyse de la variance, a pour but de comparer des moyennes sur plusieurs populations, à partir d'échantillons aléatoires et indépendants prélevés dans chacune d'elles. Elle est utilisée ici afin de comparer les moyennes de la résistivité des deux classes 0 et 1, afin de déterminer si elles sont statistiquement différentes ou non. L'hypothèse nulle (H_0) à vérifier est que tous les échantillons ont la même moyenne. H_0 est rejetée si le F de Fisher obtenu à partir des mesures est supérieur ou égal à sa valeur critique (ou valeur théorique).

IV. Résultats et discussions

IV.1. Spatialisation de la résistivité électrique au niveau des micro-parcelles

Dans un premier temps, un variogramme expérimental est estimé à partir des données de résistivité électrique. Ce variogramme est ensuite ajusté par un variogramme théorique (Figure 6). Cette étape est réalisée pour chaque voie de mesure de résistivité et pour le facteur F1 issu de l'ACP.

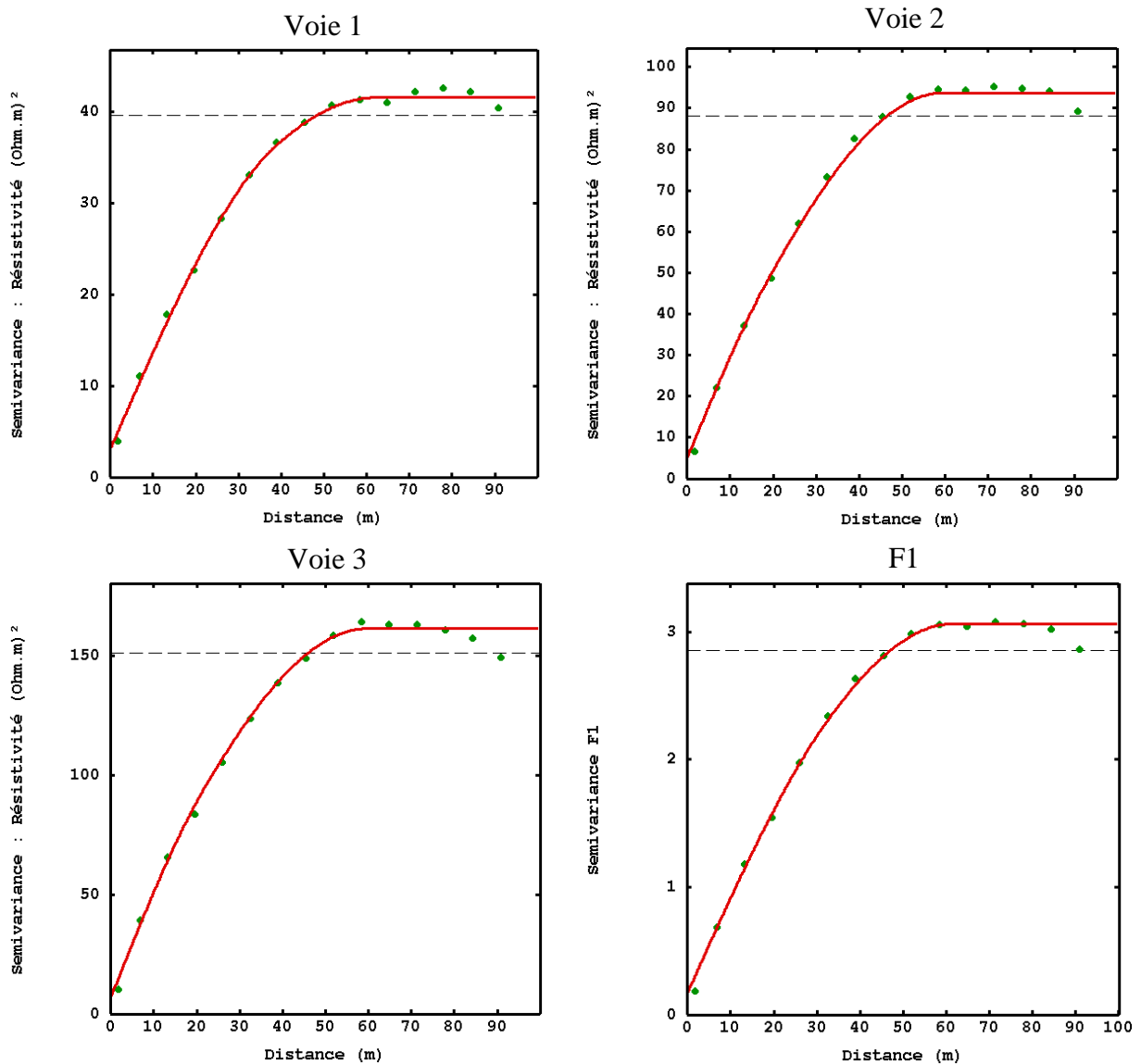


Figure 6 : Variogrammes expérimentaux (points verts) et modèles théoriques ajustés (courbes rouges) pour les 3 voies de mesure de la résistivité électrique et le facteur F1 de l'ACP

Ces variogrammes expérimentaux sont estimés sur une distance de 100 m et avec un pas de 6.5 m, puis ajustés à l'aide de trois structures : un effet de pépité (S1) et deux modèles sphériques (S2 et S3), dont les caractéristiques sont présentées dans le Tableau 1.

Tableau 1 : Coefficients des variogrammes théoriques ajustés aux variogrammes expérimentaux

	Unité	Voie 1	Voie 2	Voie 3	F1
Effet de pépité	($\Omega.m$) ²	3	5	7	0.17
Portée	m	39.06	18.71	21.62	31.38
Palier	($\Omega.m$) ²	10.18	5.16	14.08	0.21
Portée	m	62.36	59.75	59.63	61.76
Palier	($\Omega.m$) ²	28.30	83.59	140.23	2.68

La résistivité électrique du sol est ainsi spatialisée pour chaque voie de mesure par krigeage ordinaire sur une grille de mailles régulières de 0.04 m², en utilisant les variogrammes expérimental et théorique correspondants. Les statistiques descriptives sont présentées dans le Tableau 2 pour le jeu de données original et dans le Tableau 3 pour les valeurs estimées de la résistivité électrique.

Tableau 2 : Statistiques des valeurs de résistivité mesurées pour chaque voie de mesure

	Effectif	Unité	Minimum	Maximum	Moyenne	Variance
Voie 1	20408	$\Omega.m$	15.59	64.76	31.36	39.59
Voie 2	20408	$\Omega.m$	17.63	81.42	37.05	88.18
Voie 3	20408	$\Omega.m$	19.57	106.12	51.14	151.18

Tableau 3 : Statistiques des valeurs de résistivité estimées par KO sur une grille régulière de résolution 0.04m² pour chaque voie de mesure (micro-parcelles)

	Effectifs	Unité	Minimum	Maximum	Moyenne	Variance
V1_KO	162500	$\Omega.m$	21.77	52.55	32.97	20.90
V2_KO	162500	$\Omega.m$	20.01	64.71	39.49	55.92
V3_KO	162500	$\Omega.m$	20.56	91.97	53.46	110.18

Les valeurs estimées par KO sont proches de celles mesurées. On peut néanmoins remarquer que l'amplitude des valeurs est plus faible que pour le jeu de données original, avec des valeurs minimales plus élevées et des valeurs maximales plus faibles. Cela est dû à l'estimateur utilisé (KO), qui sous-estime les valeurs fortes et qui surestime les valeurs faibles.

On obtient ainsi une carte (Figure 7) qui décrit le signal électrique pour chaque voie de mesure de la résistivité électrique. Les trois cartes obtenues montrent des structures très similaires, avec une répartition des zones de faible et de forte résistivité électrique plutôt semblable pour les trois voies de mesure.

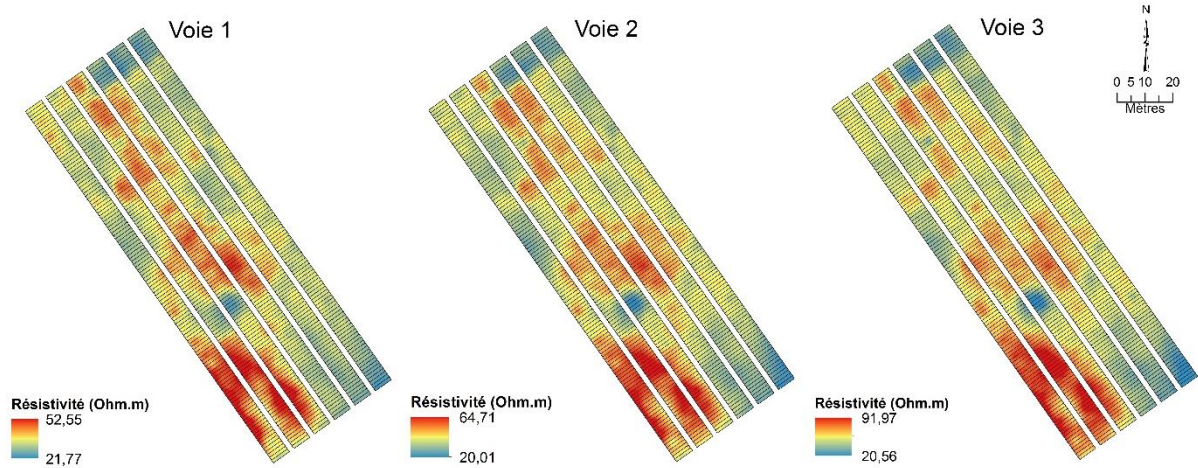


Figure 7 : Cartes des résistivités électriques estimées par KO

IV.2. Typologie des micro-parcelles par l'indicateur de la résistivité

Après avoir attribué à chaque valeur de résistivité électrique l'indicateur 0 ou 1 comme expliqué au paragraphe III.2.3, trois cartes sont obtenues pour chaque scénario (scénario 1 : par rapport à la moyenne de la voie de mesure ; scénario 2 : par rapport à la médiane de la voie de mesure), indiquant deux types de zones par construction (Figure 8 et Figure 9).

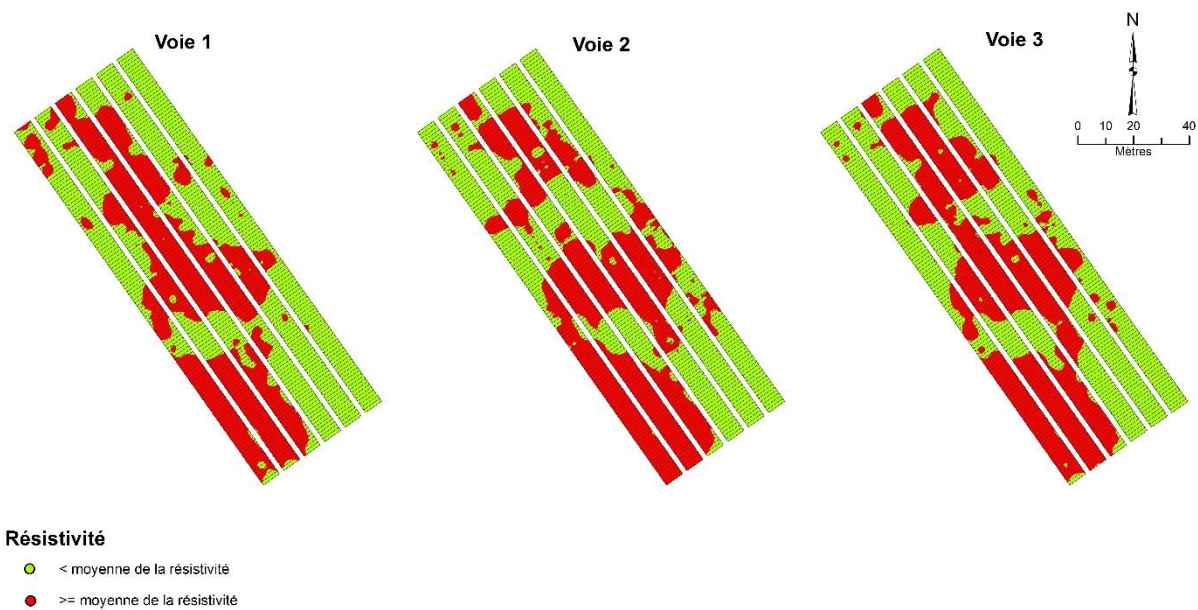


Figure 8 : Indicateur de résistivité électrique par rapport à la moyenne de la voie de mesure (scénario 1)

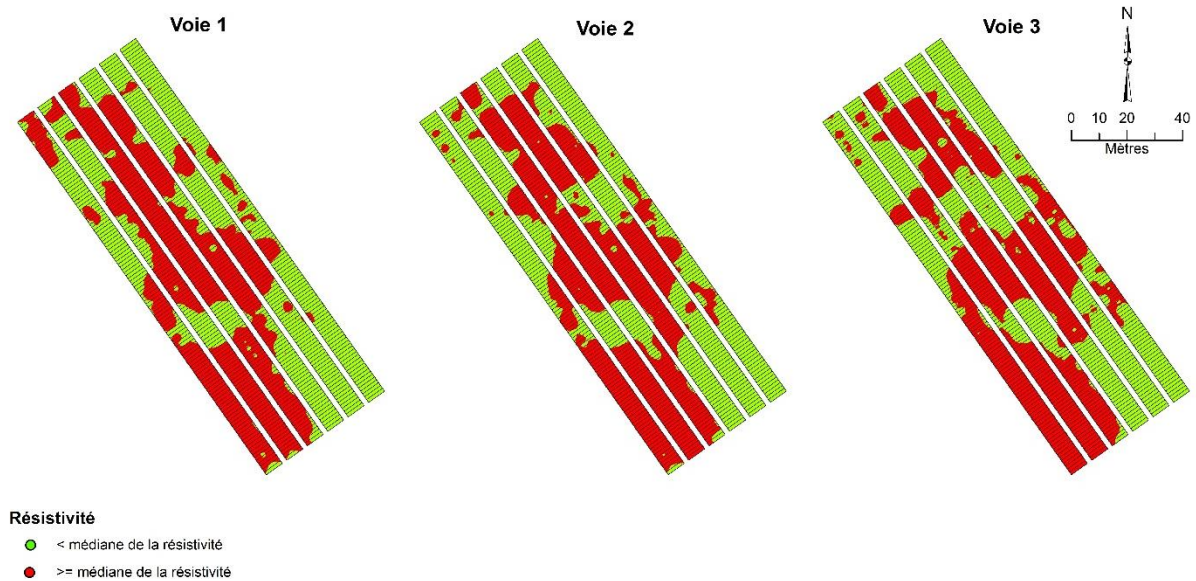


Figure 9 : Indicatrice de résistivité électrique par rapport à la médiane de la voie de mesure (scénario 2)

Grâce aux classements mis en place, on observe une discrétisation de la surface d'étude en deux zones assez continues dans l'espace géographique. De plus, il semble y avoir une répartition des zones semblable à celle observée sur la Figure 7, avec les zones de forte résistivité électrique correspondant aux zones 1 (résistivité supérieure ou égale à la moyenne ou à la médiane) et les zones de faible résistivité électrique correspondant aux zones 0 (résistivité inférieure à la moyenne ou à la médiane).

La résistivité électrique est utilisée dans le projet Optirès en tant que variable explicative de la variation des rendements de différentes variétés. On cherche donc à savoir si certaines variétés ne seraient pas toujours sur des zones défavorables ou favorables, ce qui fausserait l'interprétation des résultats de rendements. Les classements présentés ont donc pour objectif de mettre en évidence des hétérogénéités dans le terrain des micro-parcelles, pour pouvoir, par la suite, déterminer si la résistivité électrique (qui exprime la variation des propriétés du sol) peut expliquer la variabilité du rendement.

La Figure 10, quant à elle, présente les deux cartes obtenues pour l'indicatrice de résistivité par rapport à la moyenne et à la médiane de la composante principale F1 issue de l'ACP. On obtient également pour les deux scénarii, deux types de zones par construction, avec la même répartition observée Figure 8 et Figure 9.

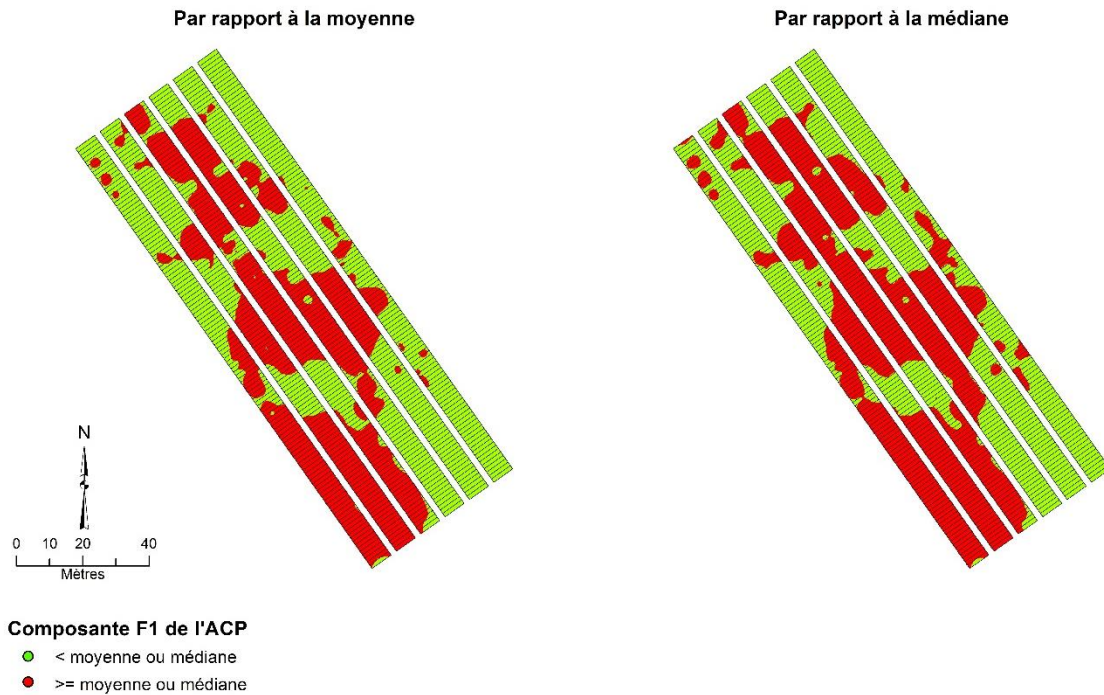


Figure 10 : Indicatrice de résistivité par rapport à la moyenne et à la médiane de la composante principale F1 de l'ACP

IV.3. Validation de la typologie des micro-parcelles

Dans cette partie, seuls les résultats de la voie de mesure 1 (V1) sont présentés, ainsi que les résultats du facteur F1 de l'ACP et ce pour les deux scénarii (S1 et S2). Les résultats pour les voies de mesure 2 et 3 (V2 et V3) sont disponibles en Annexe 3 et 4.

IV.3.1. Résultats de l'Analyse Factorielle Discriminante (AFD)

L'AFD utilisée dans ce travail a pour objectif de valider la typologie établie. L'idée est de voir si le classement réalisé peut être retrouvé grâce à une AFD. Dans cette dernière, on se pose la question de savoir si la résistivité électrique estimée par KO peut retrouver le même classement qui a été établi sur la base de la moyenne et de la médiane.

Deux modèles d'AFD sont possibles, sur la base d'une supposition fondamentale. Si on suppose que les matrices de covariances sont identiques, le traitement se fera avec une analyse factorielle discriminante linéaire. A l'inverse, si on suppose qu'elles sont différentes pour au moins deux groupes, alors le traitement se fera avec un modèle quadratique. C'est le test de Box (Tableau 4 et Tableaux 1 des Annexe 3 et 4) qui permet d'évaluer cette hypothèse (Bourennane *et al.*, 2014). Soit l'hypothèse H_0 : les matrices de covariance intra-classe sont égales et l'hypothèse H_1 : les matrices de covariance intra-classe sont différentes. Etant donné que pour toutes les AFD effectuées la p-value calculée est inférieure au niveau de signification $\alpha = 0.05$, l'hypothèse H_0 doit être rejetée et l'hypothèse H_1 retenue. Les résultats de ce test confirment donc que les matrices de covariance ne sont pas égales.

Tableau 4 : Test de Box (Approximation asymptotique du khi²)

	V1 – S1	V1 – S2	F1 – S1	F1 – S2
-2Log(M)	17220,083	22926,236	6385,689	8394,811
Khi ² (Valeur observée)	17219,974	22926,095	6385,649	8394,760
Khi ² (Valeur critique)	3,841	3,841	3,841	3,841
DDL	1	1	1	1
p-value	< 0,0001	< 0,0001	< 0,0001	< 0,0001
alpha	0,05	0,05	0,05	0,05

Les coefficients des fonctions de classement (Tableau 5 et Tableaux 2 des Annexe 3 et 4) peuvent être directement utilisés pour déterminer à quelle catégorie doit être assignée une observation, sur la base des valeurs prises pour les différentes variables explicatives. Une observation est ainsi assignée à la catégorie pour laquelle la fonction de classement est la plus élevée (Bourennane *et al.*, 2014).

Tableau 5 : Fonctions de classement

<i>(a) VI – S1</i>			<i>(c) F1 – S1</i>		
	0	1			
Constante	92,41	-57,67	Constante	-0,54	-2,02
R1	6,10	3,00	F1	-0,92	1,57
R1×R1	-0,10	-0,04	F1×F1	-0,89	-0,51
<i>(b) VI – S2</i>			<i>(d) F1 – S2</i>		
	0	1			
Constante	100,953	-53,334	Constante	-0,74	-1,68
R1	6,756	2,817	F1	-1,15	1,38
R1×R1	-0,115	-0,039	F1×F1	-0,93	-0,49

La matrice de confusion, quant à elle, résume l'information concernant les reclassements d'observations et donne le pourcentage de classement correct des observations pour chaque classe et le pourcentage correct total. Ces pourcentages correspondent au rapport du nombre d'observations bien classées sur le nombre total d'observations. Les résultats de la classification montrent ainsi que dans l'ensemble, les pourcentages de « bons classements » sont supérieurs à 95% pour les trois voies de mesure en ce qui concerne les scénarii 1 et 2 (Tableau 6a, Tableau 7a et Tableaux 3a et 4a des Annexe 3 et 4), ainsi que pour le facteur F1 de l'ACP (Tableau 8a et Tableau 9a). Enfin, la validation croisée systématique consiste à exclure systématiquement de l'échantillon l'observation qui doit être classée, afin d'éviter qu'elle n'influence trop sa prévision (Bourennane *et al.*, 2014). Dans les cas présentés ici, la validation croisée systématique donne les mêmes résultats que pour les échantillons initiaux (Tableau 6b, Tableau 7b, Tableau 8b, Tableau 9b et Tableaux 3b et 4b des Annexe 3 et 4).

Tableau 6 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 1 – scénario 1)

Catégorie	0	1	Total	% correct
<i>(a) Matrice de confusion pour l'échantillon d'estimation</i>				
0	91802	0	91802	100,00
1	5814	64884	70698	91,78
Total	97616	64884	162500	96,42
<i>(b) Matrice de confusion pour les résultats de la validation croisée</i>				
0	91802	0	91802	100,00
1	5814	64884	70698	91,78
Total	97616	64884	162500	96,42

Tableau 7 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 1 – scénario 2)

Catégorie	0	1	Total	% correct
<i>(a) Matrice de confusion pour l'échantillon d'estimation</i>				
0	81176	0	81176	100,00
1	4594	76730	81324	94,35
Total	85770	76730	162500	97,17
<i>(b) Matrice de confusion pour les résultats de la validation croisée</i>				
0	81176	0	81176	100,00
1	4594	76730	81324	94,35
Total	85770	76730	162500	97,17

Tableau 8 : Matrices de confusion résultant de l'analyse factorielle discriminante (F1 – scénario 1)

Catégorie	0	1	Total	% correct
<i>(a) Matrice de confusion pour l'échantillon d'estimation</i>				
0	89857	0	89857	100,00
1	7873	64770	72643	89,16
Total	97730	64770	162500	95,16
<i>(b) Matrice de confusion pour les résultats de la validation croisée</i>				
0	89857	0	89857	100,00
1	7873	64770	72643	89,16
Total	97730	64770	162500	95,16

Tableau 9 : Matrices de confusion résultant de l'analyse factorielle discriminante (F1 – scénario 2)

Catégorie	0	1	Total	% correct
<i>(a) Matrice de confusion pour l'échantillon d'estimation</i>				
0	80728	0	80728	100,00
1	5074	76698	81772	93,79
Total	85802	76698	162500	96,88
<i>(b) Matrice de confusion pour les résultats de la validation croisée</i>				
0	80728	0	80728	100,00
1	5074	76698	81772	93,79
Total	85802	76698	162500	96,88

IV.3.2. Résultats de l'analyse de la variance (ANOVA)

Le zonage établi étant validé par l'AFD, on examine ici si la résistivité moyenne dans les deux zones est statistiquement similaire. Le Tableau 10 présente les résultats du test de Tukey. L'hypothèse nulle H_0 est que la moyenne des zones 0 est égale à la moyenne des zones 1. La probabilité associée à H_0 pour chaque test réalisé est < 0.0001 , ce qui signifie que l'on prend un risque $< 0.01\%$ en concluant que les moyennes des zones 0 et 1 sont significativement différentes. Les mêmes résultats sont obtenus pour les voies 2 et 3 (Tableaux 5 des Annexes 3 et 4).

Tableau 10 : Test de Tukey – Analyse des différences entre les modalités avec un intervalle de confiance à 95%

Modalité	Moyennes estimées	Erreur standard	Borne inférieure (95%)	Borne supérieure (95%)	Groupes
V1 – S1					
1	37,037	0,011	37,016	37,058	A
0	29,829	0,009	29,811	29,848	B
V1 - S2					
1	36,461	0,010	36,441	36,481	A
0	29,463	0,010	29,443	29,483	B
F1 – S1					
1	1,550	0,003	1,544	1,556	A
0	-0,517	0,003	-0,522	-0,511	B
F1 – S2					
1	1,414	0,003	1,408	1,421	A
0	-0,613	0,003	-0,619	-0,607	B

Plusieurs ré-échantillonnages ont été réalisés, afin de vérifier si pour des échantillons plus petits, les moyennes entre les zones 0 et 1 sont toujours significativement différentes. Des échantillonnages aléatoires sans remise ont été réalisés avec 32 500, 15 000, 2 000, 500, 200, 150 et 50 individus. Les ANOVA réalisées sur ces échantillons donnent toutes les mêmes résultats : les moyennes entre les deux zones sont statistiquement différentes l'une de l'autre, quelle que soit la taille de l'échantillon.

IV.3.3. Comparaisons des scénarii

Comme observé sur la Figure 8 et la Figure 9, la distribution des zones « rouges » et « vertes » semble varier légèrement selon le scénario appliqué. Pour chaque voie de mesure, les deux scénarii ont donc été comparés. Pour cela, les cartes ont été superposées l'une sur l'autre et les coïncidences et les différences ont été observées. C'est-à-dire qu'une variable de comparaison a été créée, afin de comparer la variable contenant la valeur 0 ou 1 pour chaque scénario et pour la même voie de mesure. La nouvelle variable indique donc si la valeur de résistivité est inférieure à la moyenne et à la médiane de la résistivité pour les deux scénarii (valeur 0 : zones vertes), si elle est supérieure ou égale pour les deux scénarii (valeur 1 : zones rouges) ou bien si le classement est différent d'un scénario à l'autre (valeur 2 : zones blanches). Une carte (Figure 11) est ainsi obtenue.



Figure 11 : Comparaison des deux scénarii mis en place pour chaque voie de mesure de la résistivité électrique

On observe très peu de zones blanches sur les cartes, ce qui signifie que le taux de coïncidence entre les deux scénarii est très élevé et ce pour les trois voies de mesure de la résistivité. En effet, comme le montre le Tableau 11, on note un pourcentage de coïncidence de 93,5 % pour la voie 1, 95,6 % pour la voie 2 et 96,0 % pour la voie 3.

Tableau 11 : Tableau récapitulatif des pourcentages de coïncidences et de différences entre les scénarii 1 et 2

	Effectifs			
	Voie 1	Voie 2	Voie 3	F1
0	81176	81200	81245	80728
1	70698	74203	74726	72643
2	10626	7097	6529	9129
% coïncidences	93,5	95,6	96,0	94,4
% différences	6,5	4,4	4,0	5,6

Effectif total = 162 500

Ces résultats peuvent s'expliquer par l'absence de valeurs extrêmes, ce qui implique que les valeurs de la moyenne et de la médiane sont très proches l'une de l'autre. On peut ainsi dire que les mesures de résistivité possèdent une distribution symétrique.

Le même exercice a été effectué pour comparer les deux scénarii appliqués au facteur F1 de l'ACP (Figure 12). Comme précédemment, on observe très peu de zones blanches et donc un taux de coïncidence très élevé, de 94,4 % (Tableau 11).

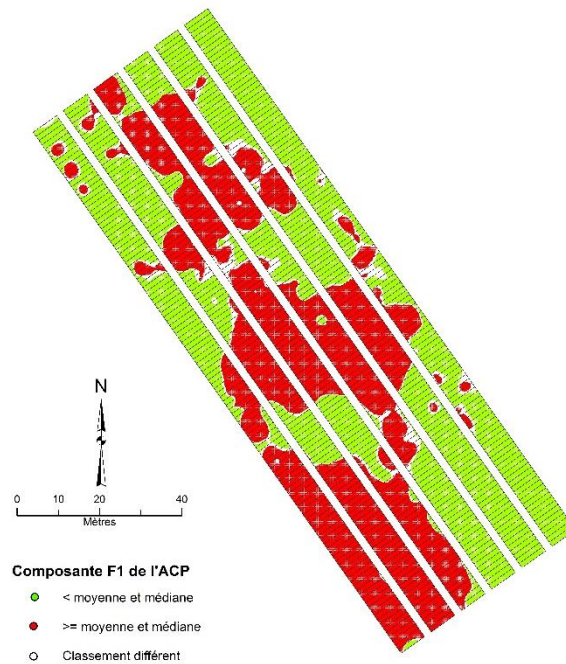


Figure 12 : Comparaison des deux scénarii mis en place pour le facteur issu de l'ACP sur les mesures de résistivité électrique

Les mesures étant symétriques, les classements par la moyenne et la médiane sont très similaires (classement pour les voies de mesure et pour F1) et dans le contexte de cette étude, les deux scénarii peuvent être considéré indifféremment pour discrétiser la surface des micro-parcelles. Cependant, sur des mesures non symétriques, les résultats de ces deux classements seraient sûrement très différents. On peut imaginer que dans un tel cas, la médiane pourrait être plus appropriée.

V. Conclusion

Le travail présenté dans ce document s'inscrit dans le cadre d'un projet CASDAR intitulé Optirès. Ce dernier a pour but d'optimiser les réseaux d'essais variétés par des méthodes permettant de mieux contrôler l'hétérogénéité spatiale dans un essai.

Dans ce mémoire, la résistivité électrique des sols a été choisie pour décrire l'hétérogénéité spatiale des sols des micro-parcelles. L'objectif était donc de spatialiser la résistivité électrique du sol à l'échelle de micro-parcelles d'essais variétés, puis de proposer une méthodologie pour caractériser l'hétérogénéité des sols grâce à cette variable.

Les mesures de résistivité électrique proviennent d'une prospection réalisée à l'automne 2011 sur des parcelles englobant les micro-parcelles concernées par ce travail. Ainsi, la résistivité électrique des sols a été enregistrée pour trois profondeurs (0.5m, 1m et environ 2m) appelées voies de mesures. Un traitement préliminaire des données a révélé que le nombre de mesures de résistivité par micro-parcelle est faible (6 à 8 mesures pour 70 % des 744 micro-parcelles) et qu'elles sont réparties de façon linéaire quelle que soit la voie de mesure. Aussi, la résistivité a été spatialisée pour chaque voie de mesure par krigeage ordinaire (KO) sur une grille de mailles régulières de 0.04 m². Pour cela, une analyse variographique a été réalisée pour chaque voie de mesure de la résistivité, avec l'estimation du variogramme expérimental, puis son ajustement par un variogramme théorique.

Dans un premier temps, pour caractériser l'hétérogénéité spatiale des sols des micro-parcelles par rapport à la résistivité, les valeurs estimées par KO ont été classées en fonction de leur position par rapport à la moyenne (scénario 1) et à la médiane (scénario 2). Cette discrétisation a conduit à définir deux types de zones (zones avec la résistivité inférieure à la moyenne ou la médiane et inversement) dans les micro-parcelles pour chaque paramètre de position. La typologie établie a été validée par une analyse factorielle discriminante (AFD). Cette dernière a montré que la résistivité électrique des sols est une variable discriminante des deux types de zones établies. Quelle que soit la voie de mesure et quel que soit le scénario, la probabilité de classement erroné est inférieure à 5%. Par ailleurs, la moyenne des valeurs de la résistivité dans les deux types de zones sont statistiquement différentes (ANOVA : test de Tukey).

Dans un second temps, les mesures de résistivités spatialisées par KO ont été synthétisées par une analyse en composante principale (ACP). La première composante (F1) issue de l'ACP, qui explique 95 % de la variabilité contenue dans les mesures de résistivité des trois voies de mesure, a été discrétisée selon les scénarii décrit ci-dessus. L'AFD et le test de Tukey appliqués au zonage obtenu à partir de F1 montrent des résultats similaires à ceux obtenus en considérant les voies de mesure de manière individuelle.

L'influence de la taille de l'échantillon sur les résultats de l'ANOVA a été testée par un tirage aléatoire de plusieurs échantillons dans les données estimées. Les résultats ont montré que quelle que soit la taille de l'échantillon (même pour des échantillons de très petite taille), les moyennes des résistivités entre les zones établies sont significativement différentes.

Enfin, la comparaison des deux scénarii montre un taux de coïncidence compris entre 93.5 % et 96 % selon la voie et le scénario considérés. Ce résultat indique que la distribution des mesures de résistivité est symétrique et dans le contexte de cette étude, on peut indifféremment considérer la moyenne ou bien la médiane pour discrétiser la surface des micro-parcelles.

Il faut néanmoins souligner que l'étendue des valeurs de résistivité est faible d'un point de vue géophysique, ce qui peut poser la question de la pertinence du classement établi. Cependant, d'un point de vue statistique, les moyennes de la résistivité électrique dans les deux zones sont statistiquement différentes, y compris pour des échantillons de petite taille.

En perspective du travail réalisé dans ce mémoire, il serait intéressant de confronter le zonage établi à des propriétés des sols telles que la texture et la profondeur. Par ailleurs, la stabilité du zonage établi peut être examinée en considérant les autres campagnes de mesures de résistivité réalisées sur le site d'étude.

Bibliographie

Articles scientifiques :

Bourennane, H., Couturier, A., Pasquier, C., Chartin, C., Hirschberger, F., Macaire, J.J. & Salvador-Blanes, S., 2014. Comparative performance of classification algorithms for the development of models of spatial distribution of landscape structure. *Geoderma*, 219 – 220: 136 – 144.

Bourennane, H., Hirschberger, F., Chartin, C. & Salvador-Blanes, S., 2017. Spatial filtering of electrical resistivity and slope intensity: Enhancement of spatial estimates of a soil property. *Journal of Applied Geophysics*, 138: 210-219.

Bourennane, H., Nicoulaud, B., Couturier, A., Pasquier, C., Mary, B. & King, D., 2012. Geostatistical filtering for improved soil water content estimation from electrical resistivity data. *Geoderma*, 183-184: 32-40.

Oliver, M.A. & Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena*, 113: 56 – 69.

Panissod, C., Dabas, M., Jolivet, A. & Tabbagh, A., 1997. A novel mobile multipole system (MUCEP) for shallow (0–3m) geoelectrical investigation: the 'Vol-de-canards' array. *Geophysical prospecting*, 45: 983-1002.

Papadopoulos, G. N., Tsokas, N. G., Dabas, M., Yi, M.J. & Tsourlos, P., 2009. 3D Inversion of Automated Resistivity Profiling (ARP) Data. *ArcheoSciences, revue d'archéométrie*, suppl. 33, p. 329-332

Samouëlian, A., Cousin, I., Tabbagh, A., Bruand, A. & Richard, G., 2005. Electrical resistivity survey in soil science: a review. *Soil & Tillage Research*, 83: 173-193.

Mémoire d'étudiant :

Ndjiba Mitombo, D., 2013. Valorisation d'une prospection électrique pour établir une carte de l'épaisseur du sol, Mémoire de 2ème année de Master GPRE Sol, Eau & Environnement, Université de Lorraine, INRA Orléans.

Cartes :

Chrétien, J., 1976, Carte pédologique de la France à 1/100 000. Feuille de Dijon. INRA. Editions carte et notice. 218p

Ouvrages :

Tabacknick, B.G. & Fidell, L.S., 1996. *Using Multivariate Statistics*. Harper Collins, New York.

Sites internet :

InfoTerre. Visualiseur InfoTerre [En ligne]. Disponible sur : <<http://infoterre.brgm.fr/viewer/MainTileForward.do>> (consulté le 15/06/2017).

Institut National de la Recherche Agronomique (INRA). INRA - Centre Val de Loire [en ligne]. Disponible sur : <<http://www.val-de-loire.inra.fr>> (consulté le 15/05/2017).

Météo France. Site Officiel de Météo-France [en ligne]. Disponible sur : <<http://www.meteofrance.com/accueil>> (consulté le 19/06/2017).

XLSTAT. Analyse Factorielle Discriminante (AFD) [en ligne]. Disponible sur : <<https://www.xlstat.com/fr/solutions/fonctionnalites/analyse-factorielle-discriminante-afd>> (consulté le 24/05/2017).

Liste des figures

Figure 1 : Localisation du site d'étude, domaine expérimental de l'INRA d'Epoisses _____	7
Figure 2 : Schéma du dispositif ARP (Source : Papadopoulos et al., 2009) _____	8
Figure 3 : Fréquence du nombre de mesures de résistivité électrique par micro-parcelle _____	9
Figure 4 : Profils de résistivité électrique des micro-parcelles sur l'UE d'Epoisses _____	10
Figure 5 : Schéma explicatif de la méthode appliquée pour attribuer l'indicatrice de résistivité (exemple par rapport à la moyenne) _____	11
Figure 6 : Variogrammes expérimentaux (points verts) et modèles théoriques ajustés (courbes rouges) pour les 3 voies de mesure de la résistivité électrique et le facteur F1 de l'ACP _____	13
Figure 7 : Cartes des résistivités électriques estimées par KO _____	15
Figure 8 : Indicatrice de résistivité électrique par rapport à la moyenne de la voie de mesure (scénario 1) _____	15
Figure 9 : Indicatrice de résistivité électrique par rapport à la médiane de la voie de mesure (scénario 2) _____	16
Figure 10 : Indicatrice de résistivité par rapport à la moyenne et à la médiane de la composante principale F1 de l'ACP _____	17
Figure 11 : Comparaison des deux scénarii mis en place pour chaque voie de mesure de la résistivité électrique _____	21
Figure 12 : Comparaison des deux scénarii mis en place pour le facteur issu de l'ACP sur les mesures de résistivité électrique _____	22

Liste des tableaux

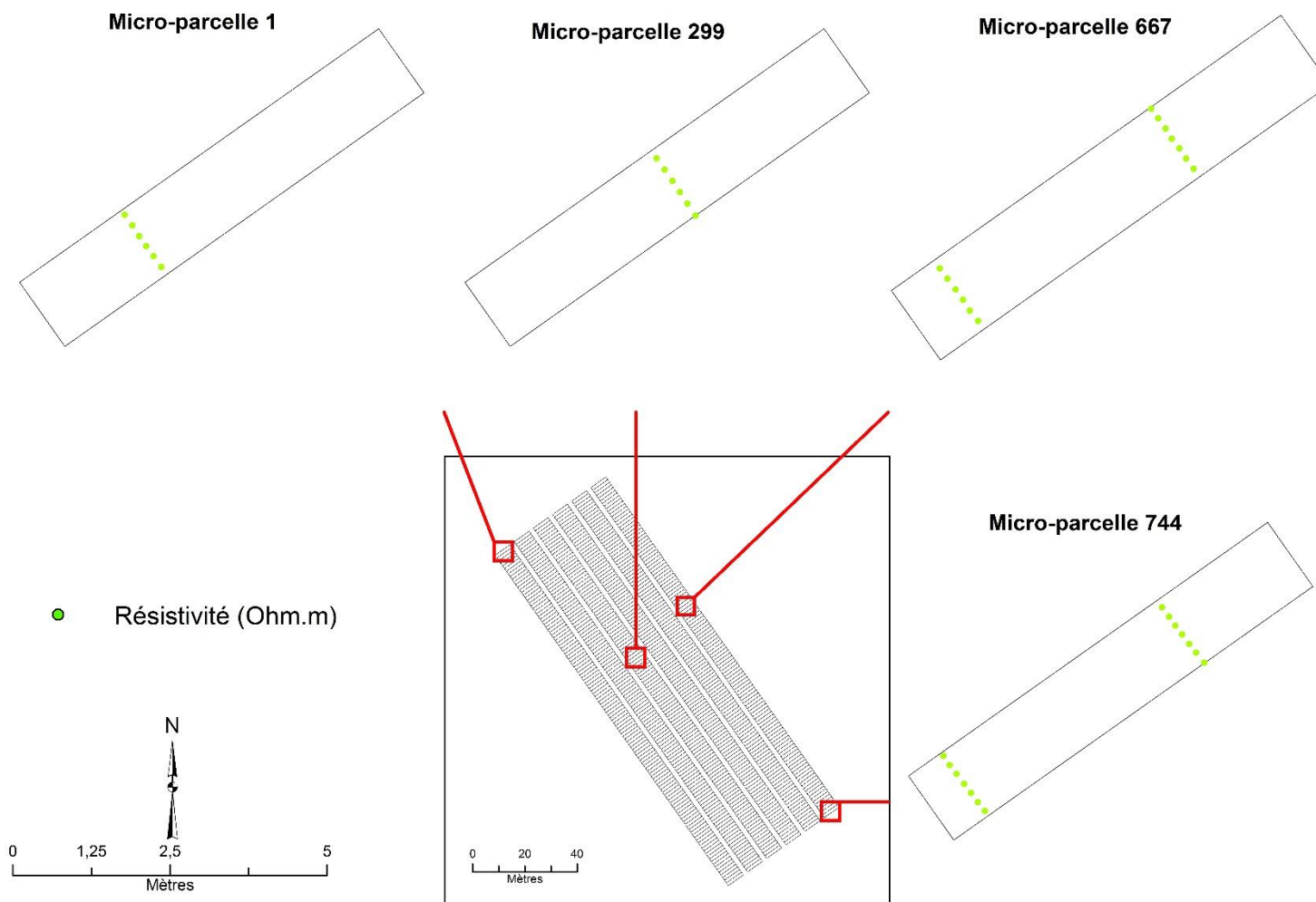
Tableau 1 : Coefficients des variogrammes théoriques ajustés aux variogrammes expérimentaux _____	14
Tableau 2 : Statistiques des valeurs de résistivité mesurées pour chaque voie de mesure _____	14
Tableau 3 : Statistiques des valeurs de résistivité estimées par KO sur une grille régulière de résolution 0.04m ² pour chaque voie de mesure (micro-parcelles) _____	14
Tableau 4 : Test de Box (Approximation asymptotique du khi ²) _____	18
Tableau 5 : Fonctions de classement _____	18
Tableau 6 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 1 – scénario 1) _____	19
Tableau 7 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 1 – scénario 2) _____	19
Tableau 8 : Matrices de confusion résultant de l'analyse factorielle discriminante (F1 – scénario 1) _____	19
Tableau 9 : Matrices de confusion résultant de l'analyse factorielle discriminante (F1 – scénario 2) _____	19
Tableau 10 : Test de Tukey – Analyse des différences entre les modalités avec un intervalle de confiance à 95% _____	20
Tableau 11 : Tableau récapitulatif des pourcentages de coïncidences et de différences entre les scénarii 1 et 2 _____	21

Liste des annexes

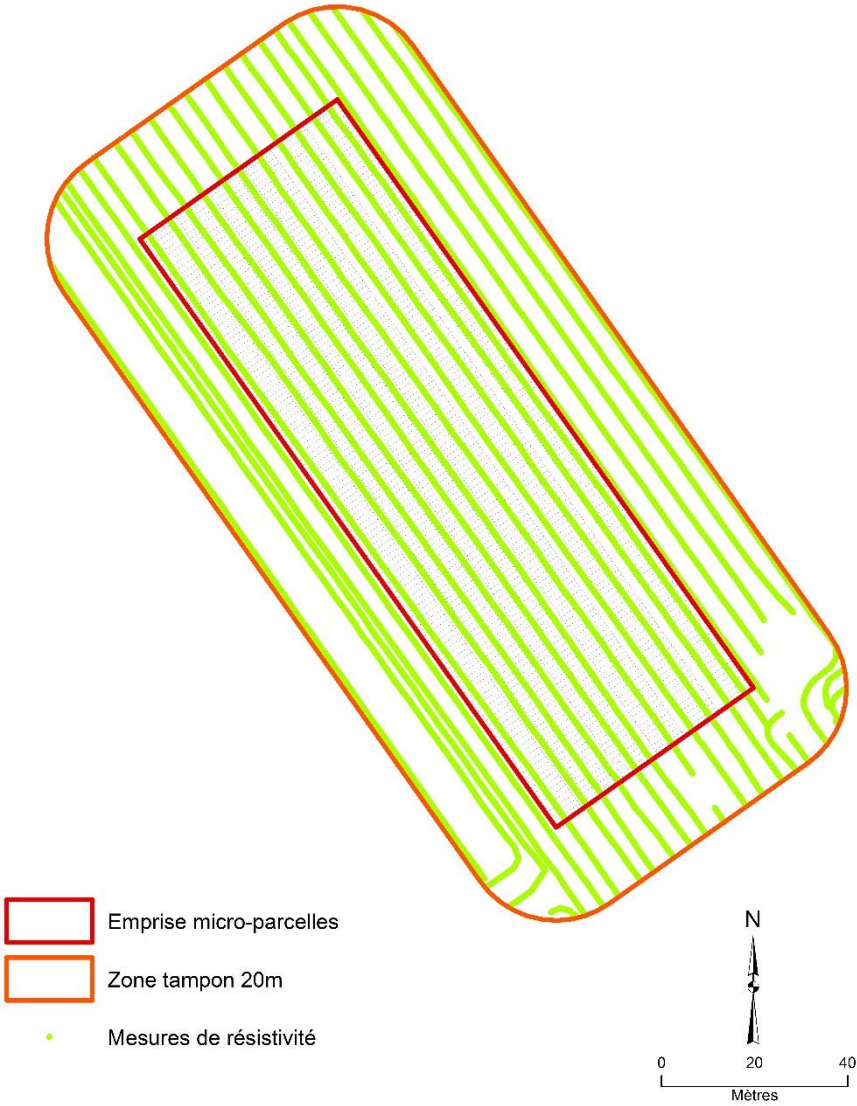
Annexe 1 : Exemples des situations des mesures disponibles de la covariable par micro-parcelle _____	a
Annexe 2 : Zone tampon de 20 m établie autour de la zone d'étude des micro-parcelles _____	b
Annexe 3 : Résultats pour le scénario 1 _____	c
Annexe 4 : Résultats pour le scénario 2 _____	e

Annexes

Annexe 1 : Exemples des situations des mesures disponibles de la covariable par micro-parcelle



Annexe 2 : Zone tampon de 20 m établie autour de la zone d'étude des micro-parcelles



Annexe 3 : Résultats pour le scénario 1

Tableau 1 : Test de Box (Approximation asymptotique du khi²)

	Voie 2	Voie 3
-2Log(M)	4310,217	3285,956
Khi ² (Valeur observée)	4310,190	3285,935
Khi ² (Valeur critique)	3,841	3,841
DDL	1	1
p-value	< 0,0001	< 0,0001
alpha	0,05	0,05

Tableau 2 : Fonctions de classement (scénario 1)

(a) Voie 2	0	1
Constante	35,61	-40,58
R2	1,97	1,66
R2×R2	-0,03	-0,02
(b) Voie 3		
Constante	29,26	-34,97
R3	1,16	1,04
R3×R3	0,01	-0,01

Tableau 3 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 2 – scénario 1)

Catégorie	0	1	Total	% correct
<i>(a) Matrice de confusion pour l'échantillon d'estimation</i>				
0	88297	0	88297	100,00
1	5374	68829	74220	92,76
Total	93671	68829	162500	96,69
<i>(b) Matrice de confusion pour les résultats de la validation croisée</i>				
0	88297	0	88297	100,00
1	5374	68829	74220	92,76
Total	93671	68829	162500	96,69

Tableau 4 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 3 – scénario 1)

Catégorie	0	1	Total	% correct
<i>(a) Matrice de confusion pour l'échantillon d'estimation</i>				
0	87874	0	87874	100,00
1	7360	67366	74726	90,15
Total	95134	67366	162500	95,47
<i>(b) Matrice de confusion pour les résultats de la validation croisée</i>				
0	87874	0	87874	100,00
1	7360	67366	74726	90,15
Total	95134	67366	162500	95,47

Tableau 5 : Test de Tukey – Analyse des différences entre les modalités avec un intervalle de confiance à 95%

Modalité	Moyennes estimées	Erreur standard	Borne inférieure (95%)	Borne supérieure (95%)	Groupes
(a) V2 – S1					
1	45,843	0,017	45,809	45,877	A
0	34,154	0,016	34,123	34,185	B
(b) V3 – S1					
1	61,935	0,026	61,885	61,985	A
0	46,237	0,024	46,190	46,283	B

Annexe 4 : Résultats pour le scénario 2

Tableau 1 : Test de Box (Approximation asymptotique du khi²)

	Voie 2	Voie 3
-2Log(M)	6143,601	3924,249
Khi ² (Valeur observée)	6143,563	3924,225
Khi ² (Valeur critique)	3,841	3,841
DDL	1	1
p-value	< 0,0001	< 0,0001
alpha	0,05	0,05

Tableau 2 : Fonctions de classement

Voie 2	0	1
Constante	-36,525	-38,039
R2	2,042	1,576
R2×R2	-0,030	-0,017
Voie 3		
Constante	-29,282	-33,653
R3	1,171	1,010
R3×R3	-0,013	-0,008

Tableau 3 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 2)

Catégorie	0	1	Total	% correct
Matrice de confusion pour l'échantillon d'estimation :				
0	81200	0	81200	100,00
1	4104	77196	81300	94,95
Total	85304	77196	162500	97,47
Matrice de confusion pour les résultats de la validation croisée :				
0	81200	0	81200	100,00
1	4104	77196	81300	94,95
Total	85304	77196	162500	97,47

Tableau 4 : Matrices de confusion résultant de l'analyse factorielle discriminante (voie 3)

Catégorie	0	1	Total	% correct
Matrice de confusion pour l'échantillon d'estimation :				
0	81245	0	81245	100,00
1	5001	76254	81255	93,85
Total	86246	76254	162500	96,92
Matrice de confusion pour les résultats de la validation croisée :				
0	81245	0	81245	100,00
1	5001	76254	81255	93,85
Total	86246	76254	162500	96,92

Tableau 5 : Test de Tukey – Analyse des différences entre les modalités avec un intervalle de confiance à 95%

Modalité	Moyennes estimées	Erreur standard	Borne inférieure (95%)	Borne supérieure (95%)	Groupes
(a) V2 – S2					
1	45,259	0,017	45,227	45,292	A
0	33,717	0,017	33,684	33,750	B
(b) V3 – S2					
1	61,224	0,025	61,176	61,273	A
0	45,686	0,025	45,637	45,734	B